FROM STEM TO STERN: A REVIEW AND TEST OF STEREOTYPE

THREAT INTERVENTIONS ON WOMEN'S MATH

PERFORMANCE AND MOTIVATION

by

Justin Paul Chase

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Psychological Science

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2012

APPROVAL

of a thesis submitted by

Justin Paul Chase

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Dr. Jessi L. Smith

Approved for the Department of Psychology

Dr. Colleen Moore

Approved for The Graduate School

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

Justin Paul Chase

April 2012

DEDICATION

This thesis is dedicated to Rita Currier. May her encouragement in the pursuit of education reach on toward future generations.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

viii

LIST OF FIGURES

ABSTRACT

How do interventions vary in their ability to attenuate women's performance and motivational detriments in science and math triggered by gender related stereotype threat? This project was designed to review the interventions which have been empirically demonstrated to reduce stereotype threat, and then to test the relative effectiveness of these interventions on college women's math performance and motivation. In the phase one, a literature review was conducted to identify interventions which successfully reduced stereotype threat toward women in science, technology, engineering, and mathematics (STEM). Interventions were then categorized into five types based on their common theoretical framework. Phase two was designed to experimentally test the relative effectiveness of the five types of interventions compared to women who did not receive any intervention and men. Although all interventions enhanced women's attitudes toward STEM (all $ps < .05$), results showed that value affirmation (a writing activity that focuses on self-values important to the target) was the only intervention to significantly improve both performance and motivation of women under stereotype threat. Implications for stereotype threat interventions and the manner in which they impact STEM test performance and motivation are discussed.

A REVIEW OF STEREOTYPE THREAT

Introduction

Science, technology, engineering, and mathematics (STEM) are crucial to the economic and social health of the United States. According to the U.S. Labor Department, 9 out of 10 of the fastest growing occupations projected for 2018 will require significant mathematics or science preparation. Concern about America's ability to stay competitive in these fields has led to calls by President Obama to strengthen these areas (National Academy of Science, 2007). For example, in a 2011 White House press release, President Obama announced his "*Educate to Innovate*" campaign to promote higher achievement in science and math over the next decade. In the press release, President Obama identifies that an overarching priority for improving STEM performance is to expand STEM education and career opportunities for underrepresented groups, such as women. There are well over 400 articles exploring the possible causes of women's underrepresentation in STEM fields, ranging from biological to social factors (Ceci, Williams, & Barnett, 2009). Certainly there is a range of possible reasons why women remain a minority in STEM, and the solutions are hotly debated and range from investing money and resources to encourage participation to changing institutions to be more female-friendly. One central theme in the literature is the important role of social norms and gendered expectations for success as barriers to women pursing or persisting in STEM. For example, cultural associations, which imply that men are more proficient in STEM domains, negatively influence women's motivation and aspirations to pursue

careers (Correll, 2004). Indeed, there is a stereotype that women are not skilled in math and science (Maccoby & Jacklin, 1974; Levine & Ornstein, 1983; Swim, 1994; Good, Aronson, & Inzlicht, 2003; Good, Rattan, & Dweck, 2009) and even if women themselves do not endorse such stereotypes, they are often still victimized by it (Gonzales, Blanton & Williams, 2001; Inzlicht & Ben-Zeev, 2000; Spencer, Steele, & Quinn, 1999; Foschi, 1996; Steele, 1997; Lovaglia, Lucas, Houser, Thy, & Markovsky, 1998). Such gender stereotypes most often favor men in STEM settings (Ceci & Williams, 2007; Smith & Johnson, 2006) and as reviewed below, can negatively influence women's performance as well as motivation, once women reach college (Smith & White, 2002; Spencer et al., 1999; Inzlitch, & Ben-Zeev, 2000). As Dr. Schiebinger's quote suggests, if institutions are trying to recruit and retain more women, then it is the institution which must implement changes to make the norms and environment of STEM workplaces impartial toward gender. How can this be accomplished? How can the gender stereotypes regarding women's performance and motivation be attenuated? The goal of the current project was to review and test the relative effectiveness of peer-reviewed gender-stereotype threat interventions designed to assist women's motivation and performance in STEM fields.

<u>The Importance of Women in STEM</u>

Scientific innovation has produced roughly half of all the U.S. economic growth over the last 50 years (NSF, 2008). However, in polls of the top corporations of the leading Fortune 1000 STEM companies, 63% of all surveyed acknowledged that the

underrepresentation by women and minorities in STEM threatens America's ability to remain competitive in STEM (Bayer, 2010). Additionally, that same survey reported that 44% believed that an increase in diversity would foster global competition and provide solutions to economic problems currently facing the United States.

This potential economic gain is not simply derived from increasing productivity by doubling a workforce, but also by eliminating the harmful side effects which have risen from the exclusion of women in scientific research. For example, between 1997 and 2000, 10 drugs were withdrawn from the US market due to life threatening effects, with four of these drugs exhibiting negative consequences which were especially severe for women (Wald & Wu, 2010). Researchers noted that testing of the drugs was primarily conducted on male lab animals and men, which failed to properly test negative consequences for women. This finding makes two meaningful suggestions. First, because women are lacking in science, scientific innovations are being developed which do not consider their needs. Second, this cost which has arisen from gender bias is not just measured in terms of dollars, but rather can be measured in terms of human life.

In addition to improving the economic welfare of the nation, increasing the number of women in STEM will also help establish economic security for women. For example, in 2009, the average starting salary for a bachelor's degree recipient entering non-STEM fields such as marketing ($42,000) and accounting ($48,500) was much lower than recipients of degrees in STEM such as computer science ($61,500) and mechanical engineering ($59,000) (National Association of Colleges and Employers, 2009). While occupational segregation accounts for the vast majority of the wage gap (AAUW

Educational Foundation, 2007); women in science and engineering tend to earn more

than women in other sectors of the workforce. Given the growth of an economy is linked

to innovation, it stands to reason that by attracting and retaining a greater number of

women in STEM, the US will increase its innovation, creativity, and competitiveness so

that it can remain an economic power throughout the 21st century. Finally, in civilized

societies, women should be granted the same privileges, resources, and support that men

are afforded, as it is socially just. From an equity view, the power and prestige associated

with science careers is a fundamental reason to address the fact that women remain

underrepresented in STEM. When women are scarce from academic fields which train

students in STEM, those same students are deprived of the educational and motivational

contributions of women. This creates a cycle of social injustice, as numerous studies

suggest that role models of underrepresented groups are imperative for the success of

future generations (Fort, 2005; Nettles & Millett, 2006). According to Fox (2001),

science is a professional claim to "authoritative knowledge," which implies that if women

are not brought into STEM, they are at a disadvantage in shaping the future of the United

States. Even if STEM fields are not inherently set up to discourage women, the norms

which favor men are still damaging. As a result, the institution should elect changes

which promote success and motivation among women, and favor gender neutrality

throughout the workplace.

Given that incorporating women into STEM has a number of benefits, it is

troubling that according to the US Department of Labor (2009), women make up 46.8%

of the workforce, and are expected to account for 51.2% of the increase in total labor

force through 2018; however, men are more likely to major in STEM fields in college, outnumbering women nearly 2:1 in all ethnic categories (Higher Education Research Institute, 2007; NSF, 2009). Women are less likely to be hired than men for academic STEM jobs (Bentley & Adamson, 2003; Ginther & Kahn, 2006), and perhaps as a cause or consequence, a smaller percentage of qualified women apply for these positions (National Research Council, 2009). Even when women do pursue STEM careers, studies examining employees in STEM fields suggest that women have higher attrition rates when compared to both male counterparts and women in other fields (Hewlett et al., 2008; Simard, Henderson, Gilmartin, Schiebinger, & Whitney, 2008). Women who enter STEM majors in college tend to be well qualified considering that both men and women are equally likely to have earned high grades in prerequisite math and science classes during high school (Brainard & Carlin, 1998; U.S. Department of Education, National Center for Education Statistics, 2000, Vogt, Hocevar, & Hagedorn, 2007). Yet, these capable women turn away from STEM majors altogether, or leave them earlier in their college careers (Seymour & Hewitt, 1997), suggesting that women may perceive a threat in these domains.

Why is it that there are fewer women in STEM fields relative to men? The answer to this question is vast and multifaceted, and research has sought to identify whether women and men differ in ability for STEM domains (for review see Hyde, Fennema, & Lamon, 1990). In a review of No Child Left Behind data sets from 10 states, the difference in average math performance between men and women was shown to be non-significant in a general school population (Hyde et al., 2008). Furthermore, girls and

women often have better grades in math and science, while taking the same high level math courses as men (U.S. Department of Education, National Center for Education Statistics, 2007). These data suggest women are, on average, just as capable as men when entering STEM domains. This implies that any discrepancies in performance between men and women are likely due to stereotype contextual factor (such as gender stereotypes), as opposed to a biological factor. Finally, some may argue that women, overall, have a lower interest in STEM because they are more drawn to helping professions (Bressler & Wendell, 1980; Boughn & Letini, 1999; Correll, 2001); which may explain why women, who do show an interest in STEM, typically seek out careers in biology, which have a direct connection with providing assistance for others (NSF, 2008; Turner & Lapen, 2005; Eisenhart, 2008; Plant, Baylor, Doerr, & Rosenberg-Kima, 2009). However, such responses to explain the unequal ratio of men to women do not take important situational factors into consideration, for the notion that women are free to simply "choose" an academic or career path dismisses the role that social norms and stereotypes play in restricting people's choices (Eagly & Steffen, 1984; Margolis & Fisher, 2002; Su, Rounds, & Armstrong, 2009).

If women are capable of performing well in STEM relevant classes, and can have equal interest in STEM, why is there an underrepresentation of women in most STEM domains? One possible contributing explanation is that women, even those highly skilled at STEM, seem to be the victim of a "threat in the air" (Steele, 1997) that can lead to underperformance and low motivation, leading to disidentifying with the domain completely (Crocker & Major, 1989). When a gender stereotype is "in the air," it is said

to result in *stereotype threat* which refers to the concern or worry that a stigmatized

individual experiences; the worry that they are at risk of confirming a negative stereotype

about their group (Steele & Aronson, 1995).

<div align="center">

Gender Stereotypes in STEM and
the Experience of Stereotype Threat

</div>

One deeply embedded socially perceived gender difference is that men perform

better in math and science domains, whereas women excel at English and reading (Eagly

& Wood, 1991; Eccles, Jacobs, & Harold, 1990; Levine & Ornstein, 1983; Maccoby &

Jacklin, 1974; Swim, 1994; Hyde, Fennema, Ryan, Frost, & Hopp, 1990; Smith & White,

2002; Walton & Cohen, 2003; Nosek, Banaji, & Greenwald, 2002; Keifer &

Sekaquaptewa, 2007). Knowledge of this negative gender stereotype about women's

math and science abilities are reported by age groups ranging from children in elementary

school (Ambady, Shih, Kim, & Pittinsky, 2001), to adults (Nosek et al., 2002). As a

cause or consequence, women are underrepresented in college majors and professional

careers which place an emphasis on science skills, such as engineering and computer

science (Strenta, Elliot, Adair, Scott, & Matier, 1993; Ramist, Lewis, & McCamley-

Jenkins, 1994). These gender stereotypes can be transmitted through the media, parents,

peers, and teachers (Bandura, 1977: 1997; Eccles, Jacob, & Harold, 1990). Ultimately,

such gender stereotypes influence the way children play, the books they read, how they

are treated in class, and the careers they ultimately pursue (Meece, Eccles, Kaczala, Goff,

& Futterman, 1982; Eccles, 1987; Biernat, 1991; Hewitt & Seymour, 1991). Even if

women do not endorse the stereotypes, the expectation of what is valued for a given

social group is enough to influence judgments and decisions about that group and among those group members (Foschi, 1996; Steele, 1997; Lovaglia et al., 1998; Allen & Smith, 2011). What is more, gender stereotypes seem to work against those who care most about achieving and success (Smith, Sasone, & White, 2007; Pronin, Steele, & Ross, 2003), as well as those women who identify most with their gender (Kaiser & Hagiwara, 2011). Nevertheless, the reach of stereotype threat on women's experiences and expectations for math and science is far and wide, and can even be detrimental to those women who do not identify highly in the stereotyped domain (Osborne & Walker, 2006; Keller, 2007).

Given that a pervading stereotype about women's negative performance in STEM exists, it is important to develop ways in which the stereotype threat can be made less threatening and ideally eliminated. Changing a stereotype is difficult at the individual level and nearly impossible at the societal level (Allport, 1954; Devine, 1989). As a result, researchers have attempted to develop models which will combat the relatively low number of women compared to men. The dominant model for increasing the number of women in STEM fields has been the pipeline model (Berryman, 1983). This model, which has focused on recruitment, has helped more women enter STEM careers, and while women now make up a more substantial percentage of graduate students in these fields (NSF, 2008), if the current rate of recruitment stays steady, the gender gap will not close until the year 2149 (Glazer-Raymo, 1999: 2001).

Yet, even if the rate of recruitment stayed consistent, recruitment alone would not suffice, as the pipeline is leaking, as attrition rates remain high for the few women who do enter STEM domains, and increasing the rate of recruitment does not address the need

to patch the fissures (Chesler, Barabino, Bhatia, & Richards-Kortum, 2010). As a result, it is necessary to develop a greater understanding of stereotypes and how they may threaten an individual, so that framework can be implemented to foster improvements in performance and motivation. Stereotype threat is a form of social identity threat. According to social identity threat theory, people strive to maintain a positive perception of their groups and collectives (Tajfel & Turner, 1986). If these positive perceptions are challenged, a sense of threat arises, which may manifest negative emotions or reinforcing behaviors which align with the group norms (Walton & Cohen, 2007). Essentially, people may treat others (Baumeister & Bushman, 2008), and even themselves (Hall & Schmid-Mast, 2008; Smith & Lewis, 2009; Logel et al., 2009; Allen & Smith, 2011), in a manner which is consistent with that group's norms. Social identity threats are instances in which people believe that the groups to which they belong have been evaluated negatively, whereas stereotype threat refers to the individual's concerns about confirming their group's negative reputation (Derks, Inzlicht, & Kang, 2008). When taking a STEM relevant test, for example, a woman may be concerned she will confirm the negative gender stereotype, which ironically leads her to underperform on the test and report lower motivation for STEM related tasks (Osborne, 1995; Keller, 2002; Keller & Dauenheimer, 2003; Cadinu, Maass, Rosabinca, & Kiesner, 2005).

This *stereotype threat* experience has been replicated in a variety of domains for a wide range of different groups including race (Frantz, Cuddy, Burnett, Ray, & Hart, 2004), age (Hess & Hinson, 2002) and even socioeconomic status (Croizet & Claire, 1998). Stereotype threat is so pervasive that even middle class white men, who are the

socially dominate group in America, are vulnerable to the stereotype that people who are Asian are better at math (Aronson et al., 1999; Smith & White, 2002) or people who are Black are better at athletics (Stone, Lynch, Sjomerling, & Darley, 1999). In these studies, white men who take a test worry about confirming negative stereotypes about whites, and perform significantly worse in the stereotype threat conditions. However, gender stereotype threat may be unique compared to other forms of stereotype threat, due to physical and social power differences between men and women; and the permeation of male-female relationships in everyday life (Fiske & Stevens, 1993). For example, one study found that children with Williams Syndrome, a rare genetic disorder, recognized gender stereotypes but lacked formation of racial stereotypes (Santos, Meyer-Lindenberg, & Deruelle, 2010). This implies that from a cognitive standpoint, gender stereotypes have a different biological root than racial stereotypes. Additionally, researchers have found differences between gender and racial stereotypes in impression formation, such that, when asked to evaluate a potential employee, women were evaluated entirely on an attached photograph, whereas candidates of difference races were evaluated on both a photograph and trait description (Branscombe & Smith, 1990). Finally, a meta-analysis of stereotype threats has suggested that women experience the greatest performance deficit to stereotype threats when those threats are ambiguous or subtle (e.g., implied in a given context), whereas those who are exposed to a stereotype threat based on race and ethnicity experience a larger performance deficit when the threat is explicitly stated (Nguyen & Ryan, 2008), also suggesting a different basis for the application of the threats faced based on the identity that is threatened.

Consequences of Stereotype Threat

Stereotype threat has been shown to produce numerous negative consequences ranging from eating behavior to increased aggression (for review, see Inzlicht & Sonia, 2010); however, stereotype threat is primarily examined under the context of performance deficits on high stakes tests (Brown & Josephs, 1999; Cadinu, Mass, Frigerio, Impagiliazzo, & Latinotti, 2003; McFarland, Lev-Arey, & Ziegert, 2003; Beilock, Jellison, Rydell, McConnell, & Carr, 2006; Cole, Matheson, & Anisman, 2007; Koch, Müller, & Sieverding, 2008; Goff, Steel, & Davies, 2008). Performance is an importation variable to examine as much of one's ability to succeed in a career is extrapolated from tests which are said to measure ability, and are often used as benchmarks to decide progress and promotions, despite an overall lack of predictive validity (Hunter & Hunter, 1984). However, detriments in performance are likely to lead to less job satisfaction (Judge, Thoresen, Bono & Patton, 2001), and may explain why women who enter STEM are more likely to quit compared to male counterparts. Although research has examined the potential motivational deficits of stereotype threat (Cohen, Steele, & Ross, 1999; Smith et al., 2007), and its relationship with performance (Jamieson & Harkins, 2007), results have been mixed, and interventions have yet to address the importance of motivational constructs (for a review see Aronson et al., 2009). It is not enough to tackle the problems women face in STEM unilaterally, because although recruitment is important, it is meaningless without decreasing attrition. If the pipeline is leaking (Blickenstaff, 2005), research must uncover the best way to facilitate performance while improving motivation.

The most commonly studied consequences of stereotype threat is reduced performance on challenging high stake tests when negative stereotypes about a person's group abilities are relevant. Under stereotype threat, performance declines for the target group and this poor performance has been observed in both laboratory (Steele & Aronson, 1995) and classroom settings (Cole et al., 2007) when the stereotype threat is activated either subtly (Smith et al., 2007; Inzlicht & Ben-Zeev, 2000; Stone & McWhinnie, 2007) or blatantly (Smith & White, 2002; Rosenthal & Crisp, 2006; Cadinu et al. 2003). For example, after reading an article that articulated the debate about gender differences in math ability, women performed poorly on GRE math items compared to women who took the test when the gender stereotype was rendered irrelevant to the test at hand (Smith & White, 2002) and compared to men whose performance was unaffected by the gender stereotype (Spencer et al., 1999).Gender stereotype threat has also been shown to reduce women's performance on tests related to engineering (e.g., Crisp, Bache, Maitner, 2009), science (Good, Aronson, & Harder, 2008) and computer science (e.g. Smith, Morgan, & White, 2005; Cheryan, Plaut, Davis, & Steele, 2009). In one study examining the impact of stereotype threat on performance, men and women engineering students were asked to take a reasoning test which was framed as either "diagnostic" or "diagnostic but gender fair" (Régner et al., 2010). In addition to the reasoning test, working memory capacity was measured, allowing researchers to inspect whether or not individual differences in working memory moderated performance based effects derived from stereotype threat. Researchers found that engineering women under stereotype

threat performed more poorly on the exam, especially if they were low in working

memory capacity.

Declined performance has also been observed on state-wide standardized tests

upon college level entry (Good et al., 2003). Specifically, stereotype threat impacts a

woman's math performance (Nguyen & Ryan, 2008), and may account for up to 20 points

on the math portion of the SAT (Walton & Spencer, 2009). Although 20 points may

appear to be insignificant on a test worth 800 points, the average male score on the SAT

math portion was 33 points higher than that of females (Wainer & Steinberg, 1992),

suggesting that stereotype threat could account for two thirds of the SAT math difference.

Although women are earning similar grades in classes, these high stakes tests at the

college level may not accurately represent the ability of these women, suggesting that

social factors may have lead to performance deficits. For example, Spencer and

colleagues (1999) recruited men and women who both had strong math backgrounds and

similar abilities, as measured by course grades, to take a math section of the Graduate

Record Exam (GRE). When the women participants were told that men outperformed

them, they tested significantly worse than when no gender difference was stated. This

implies that if the gender differences in performance were in fact due to sex-linked

differences in ability, women would have performed worse regardless of a stereotype

threat. Considering undergraduate college students, research has suggested that both SAT

and ACT scores tend to under predict women's future GPA (Breland & Griswold, 1982;

Sawyer, 1986). Gender bias, in the form of predicting future GPA, has also been noted on

the Miller Analogies Test (House & Keeley, 1995) and even the GRE (Kaczmarek &

Franco, 1986). Stereotype threat effects have been replicated on tests which are considered "pure" measures of ability and reasoning (Brown & Day, 2006), suggesting that stereotype threats can even hinder performance on exams which have been deemed "culturally fair" and are not reliant on reading skills or background knowledge. Given the importance of these tests to the admissions process, it is important to understand ways in which this stereotype threat can be alleviated.

Poor performance notwithstanding, even women who excel in classes which are necessary for success in STEM fields, do not often choose college majors relating to these areas suggesting a motivational competent to stereotype threat effects (Betz & Hackett, 1981; LeFevre, Kulak, & Heymans, 1992; Stangor & Sechrist, 1998; Lubinski & Benbow, 2006). An important need to delineate between performance and motivation arises, as both constructs are related but research shows that even when an individual performs well, their motivation to continue working in a given domain may be influenced by stereotypes (e.g., Seymour & Hewitt, 1997; Bleeker & Jacobs, 2004). Although low performance can impact motivation, the notion that motivation can be harmed when performance is high may help explain the low attrition rates in STEM fields. For example, women's interest in computer science was lower when a stereotype threat was present (Smith et al., 2007). Drawing from the stereotyped task engagement process (STEP) model (Smith, 2004), Smith and colleagues (2007) suggested that this lack of interest was linked to the goals participants adopted. Specifically, they found that participants with higher achievement motivation were more likely to adopt performance-

avoidance goals when a stereotype threat was triggered. Thus, even if women perform well, motivation can still wane (Eccles, 1994; Seymour, 1995).

Under stereotype threat, women's motivation can be pulled toward non-stereotyped domains. For example, when men and women viewed stereotypic television advertisements, women (but not men) who viewed stereotypical ads were less likely to attempt math items, and instead attempted more verbal problems (Davies, Spencer, Quinn, & Gerhardstein, 2002). Certainly, there are a number of ways in which motivation can be conceptualized. Although for the purposes of this project, we are particularly emphasizing interest in the task and field, stereotype threat effects also negatively affect other motivational variables such as effort exertion and attributional style. Further, women who endorse stereotypic beliefs which highlight negative abilities in a given domain may lead to a decrease of motivation in relevant domains (Schmader, Johns, & Barquissau, 2004). For example, in one study, college students were asked to find a document on the internet and save it to a memory stick, while being told that women typically perform less well than men, or that men perform less well than women (Koch et al., 2008). In reality, the memory stick was rigged to fail, preventing students from finishing their task. Women under stereotype threat were more likely to attribute their failure to internal factors, as opposed to men who attributed failure to internal and external causes. In conjunction with the findings presented by Schmader and colleagues, these results suggest that women may attribute a failure toward the self, which can hinder motivation.

Although few studies have examined the motivational consequences of interventions, relative to performance, motivational constructs are important to consider as they are mechanisms of personal agency where the motivational experience is most proximally related to career and college major pursuit and persistence (Sansone & Smith, 2000; Renninger, Sansone & Smith 2004). Indeed, the USA suffers from a STEM motivational crisis of sorts, whereby both men and women are not inclined to report enjoying math and science (Ames, 1992; Hu & Kuh, 2002; Legault, Green-Demers, & Pelletier 2006; Fielding-Wells & Makar, 2008), and STEM majors are often among the lowest enrollments on college campuses (HERI, 2001; 2011). Achievement motivation, or the general will-to-succeed (Eccles, Wigfield, & Schiefele, 1998), does not always coincide with interest. As a result, intrinsic motivation may be best measured by assessing both involvement and interest (Thoman, Smith & Silva, 2011). Interest is not just a measure of positive feelings toward a domain, but rather takes into account stored knowledge, feelings that influence engagement, questioning, and the activity levels of groups (Renninger et al., 2004). Those with greater interest in science domains tend to have a greater capacity to find meaning and identify their own questions (Renninger, Ewen, & Lasher, 2002), as well as utilize more effective learning strategies (Csikzentmihayli, Rathunde, & Whalen, 1993; Renninger & Hidi, 2002; Lipstein & Renninger, 2006; Barron, 2006) and seek answers in systematic ways (Renninger, 2000; Engle & Conant, 2002). Finally, some research has even found that normalizing effort by stressing that everyone needs to work hard in order to achieve success, results in higher self-reports of belonging and future interest in STEM (Smith, Lewis, Hawthorne, &

Hodges, under review). These findings suggest that motivation may be just as, if not more, important to a women's experience in STEM, than performance (Eccles et al., 1983, Sorby, 2000: 2008; Stout, Dasgupta, Hunsinger, & McManus, 2011).

Additionally, research has noted that women who excel in mathematics experience a decline in both interest and motivation as they graduate from high school (Malpass, O'Neil, & Hocevar, 1999; Reis & Park, 2001). It has even been suggested that the skewed gender ratio of STEM experts may reduce motivation to pursue careers in science, engineering, and technology (Stout et al., 2011). Beliefs about the stereotypes which color STEM careers as masculine domains may lower women's interest and undermine the belief that they are able to excel in these fields (Smith et al., under review). In a test of a series of factors which may predict consideration of math and science majors, studies have found that while confidence and self-efficacy may play important roles for males and minorities, only feelings of interest predicted this consideration for women, with women showing greater interest in non-math and science fields (Post, Stewart, & Smith, 1991). Indeed, research has long suggested that fostering a positive attitude toward STEM is necessary for increasing future interest and performance in related domains (Stodolsky, 1985; Eccles & Jacobs, 1986; Fennema, 1989; Gierl & Bisanz, 1995). Collectively, this data suggests that stereotype threat may lead women to lose interest in STEM, even if they manage to overcome performance decrements triggered by stereotype threat. As such, the current project will examine both performance and motivational outcomes of stereotype threat interventions in an effort to examine the relative effectiveness of the identified interventions.

Theory-based stereotype threat interventions are informed by an understanding of the mechanisms involved; documenting *how* stereotype threat effects do their damage to performance. In order to inoculate women for stereotype threat, the mediating mechanisms of such threats must be understood before interventions can be designed.

The Mediating Mechanisms of Stereotype Threat

The hundreds of peer reviewed articles on stereotype threat reiterate that a variety of situations could arise which causes these threats, but despite the robustness of this effect, the specific mechanisms which cause stereotype threat are still somewhat uncertain (for review see Wheeler & Petty, 2001; Smith, 2004). Originally, Steele and Aronson (1995) cited distraction, narrowed attention, anxiety, self-consciousness, and over effort may play a role in the observed results. Since then, research has found mixed support for all of these proposed mechanisms and other variables as well (Spencer et al., 1999; Stangor, Carr, & Kiang, 1998; Stone, 2002; Smith & White, 2002; Keller & Dauenheimer, 2003; O'Brien & Crandall, 2003; Beilock, Rydell, & McConnell, 2007). Smith's 2004 model provides an account for multiple mediators while suggesting that goal adoption is key (Brodish & Devine, 2005; Ryan & Ryan, 2005). The STEP model takes a phenomenological account of how stereotype threat feels, as opposed to cognitive accounts (Schamder, Johns, & Forbes, 2008), which argue that a physiological stress response, tendency to actively monitor performance, and efforts to suppress negative thoughts and emotions are interrelated. These mechanisms are thought to combine in such a way that executive resources commonly needed to perform well on a variety of

cognitive and social tasks are taxed. While these models provide a basis for understanding stereotype threats via performance or motivation, both postulate a theoretical framework which allows one to understand the multiple routes by which stereotype threat causes damage. By understanding the causes of the damage, interventions can be designed to attenuate the detriments of stereotype threat.

Although both accounts focus on different facets of stereotype threats, it is important to note that the context under which the threat develops could be important. Some studies have examined threats which were blatantly stated (Spencer et al., 1999; Smith & White, 2002), whereas others used a more subtle manipulation (Brown & Pinel, 2003; Beilock et al., 2007). While both cues can create stereotype threats, there could be differences in the way these threats impact performance or motivation (Smith, 2004; Stone & McWhinnie, 2008). For example, according to a meta-analysis of stereotype threat triggers, subtle threats produce a larger effect than blatant threats for women (Nguyen & Ryan, 2008). Although it is unclear why women experience a greater detriment from subtle threats, many studies have replicated stereotype threat using blatant manipulations (Keller, 2002; McIntyre et al., 2003; Ambady et al., 2004; Rosenthal & Crisp, 2006; Dar-Nimrod & Heine, 2006; Thoman et al., 2008).

To date, numerous interventions have been developed in order to combat the threats derived from these stereotypes, yet research has yet to examine the relative effectiveness of each intervention (Quinn & Spencer, 2001; Good et al., 2003; Stricker & Ward, 2004; Schimel, Arndt, Banko, & Cook, 2004; Huguet & Régner, 2007), prompting a need for a focused direction within the intervention literature so that policy makers and

educators are empirically informed as to how to allocate limited resources efficiently. The groundwork to fix the leaky pipeline is in place thanks to the numerous interventions which empirically attenuate stereotype threat, but the field has been waiting for an attempt to critically examine the groundwork and comparatively test the effectiveness of each intervention.

Project Overview

Although there have been many attempts to explain why there are fewer women in STEM relative to men, and why women who do pursue STEM careers have higher attrition rates (for review see AAUW, 2010), the negative impact of stereotype threat on women's performance and motivation for STEM cannot be disputed. Current efforts to fix the numerous problems brought about by this crisis are insufficient (Glazer-Raymo, 1999: 2001), as they have primarily focused on recruitment (Berryman, 1983), but even those which have looked at lowering attrition and improving performance through interventions have not directly assessed how to improve both performance and motivation, let alone whether they have improved the overall attitude of individuals toward the threatened domain (see Yeager & Walton, 2011). Additionally, research has not informed scientists how these interventions are similar, different, and which may be the most effective intervention at curtailing the negative consequences that arise from gender stereotypes in STEM.

As a result, the aim of this project was twofold, interventions which have been previously identified as reducing stereotype threat in STEM domains would be identified,

and they tested for their relative effectiveness in reducing performance and motivational deficits. First, in phase one, an extensive literature review was conducted as a means to identify stereotype threat interventions aimed at improving the performance and or motivation of college aged women in STEM. These interventions were then coded for theoretical similarities and were collapsed into five subtypes. Next, in the second phase, a prototypical example was selected for each of the five interventions, and original authors were contacted so that materials could be implemented in the same manner that they were in the original studies. Finally, a mildly explicit stereotype threat (Schmader, 2002) was embedded in experimental instructions to create a standardized situation for all participants, who were then assigned to one of seven conditions (five interventions, or two controls). Based on the literature review, it was hypothesized that all interventions would improve performance relative to women who did not receive an intervention, yet it is unclear which intervention may best attenuate stereotype threat. Secondly, it was important to examine whether or not intrinsic motivation, which tends to be naturally low for STEM domains (Ames, 1992; Hu & Kuh, 2002; Legault, Green-Demers, & Pelletier 2006; Fielding-Wells & Makar, 2008), and has only been examined in one study (Thoman et al., 2008), could be improved by any of the interventions. Intrinsic motivation may be exclusive of performance and could explain why equally capable women do not pursue STEM careers (Seymour & Hewitt, 1997).

PHASE ONE: A LITERATURE REVIEW OF INTERVENTIONS

Methods

Review Selection Procedure

PsychInfo was used as the means to identify articles published between November 1995 – June 2011. Criteria for inclusion in the review were as follows: 1) indexed with the key word "stereotype threat". This keyword was used to search within abstracts, which produced 488 articles. Given that many stereotype threat articles examine race based threats (for review see Steele et al., 2002) the next criteria was 2) the article be focused on women. As such, "*gender*" was then added as a secondary key term to search the abstracts, yielding a total of 134 entries, which was then refined by English only selections of peer reviewed journals, limiting articles to 77. Additionally, another search was run using "*women*" as the secondary key term, yielding 158 entries, which was refined to 108 when the non-English non-peer reviewed journals were excluded. Finally, a third search using "*sex*" as the secondary key term reported 32 entries, which was refined to 21 English peer reviewed journal articles, creating a total of 217 articles to review. To detect any duplicates, the titles of these 217 articles were inspected in order to exclude duplicate entries, as some articles appeared in all three searches, excluding 79 articles. Article abstracts for the remaining 139 peer reviewed English journal articles were then read to determine the sample used in each study. Once each sample was recorded, our next inclusion criteria was implemented 4) which included studies that tested interventions on college aged women, reducing the total number of articles to 93.

The final criteria, 5) which sought studies which experimentally manipulated stereotype threat and significantly improved either performance or motivation was implemented, leaving 15 journal articles which consisted of 28 studies (note that 30 studies were published within the 15 journal articles, but two of the three studies from Schimel et al., 2004 were excluded for not manipulating stereotype threat).

Coding of Stereotype Threat Intervention Studies

First, the content analysis examined the theoretical framework for each intervention (Table 1, see pages 48-49), categorizing them into five subgroups (deemphasizing threatened group identity, providing external attributions for difficulty, encouraging value affirmation, providing role models, and encouraging an incremental mindset). Next, the methodology of each intervention was examined. Of the interventions (Table 1), 1 experiment (3%) measured motivation, whereas all 28 (100%) measured performance. Additionally, whether a women under stereotype threat only or a men comparison group was used as a control was recorded. Results found that 15 (53%) compared the intervention to only women under stereotype threat. The content analysis also revealed that 23 experiments (82%) examined stereotype threats in mathematics. Finally, the type of threat each experiment utilized was recorded (Table 2, see pages 50-51), with results showing that 14 (50%) used subtle threats, 12 (43) used blatant threats, and 2 (7%) used mildly explicit threats.

Results

Overview

Stereotype threat effects have been demonstrated in a variety of studies using numerous tests. Yet, establishing a problem is merely the first step to uncovering viable solutions. Over the past decade, research has focused primarily on reducing or eliminating women's stereotype threat performance deficits using various methods. As of yet, no "silver bullet" intervention, one known to be most effective in reducing stereotype threat, exists (Shapiro & Neuberg, 2007); likely because of the vast sea of literature on the topic. Using the review selection criteria above, the literature of PsychInfo indexed gender stereotype threat interventions for women in college yielded 15 distinct experimental articles, consisting of 28 individual studies which were collapsed into five types of interventions based on their common theoretical framework: deemphasizing threatened group identity, providing external attributions for difficulty, encouraging value affirmation, providing role models, and encouraging an incremental mindset. The following section provides a detailed account of each intervention category and an overview of each of the 28 studies (Table 1 & 2, see pages 48-51).

Intervention Category 1: Deemphasizing
Threatened Group Identity

With five articles across nine empirical intervention experiments, the most researched method of reducing stereotype threat for college aged women is to deemphasize threatened group identities. The theoretical underpinnings of this intervention category were first evident in the landmark stereotype threat paper by Steele

and Aronson (1995), who suggested that individuals may be more likely to show performance decrements under conditions which create stereotype threats when they strongly identify with the domain in question (e.g., women who strongly identified in STEM domains may be more susceptible to the negative consequences of stereotype threat). Seeking to test whether the strength's of one's group identification may play a role in stereotype threat activation; empirical research found that women under stereotype threat who report gender as a central component of their identity performed less well than men, and women who did not view gender as a crucial part of their identity (Schmader, 2002).

As empirical data established the negative consequences of threatened group identities (Schmader, 2002), researchers theorized that one way to reduce the negative consequences of stereotype threat would be to have participants focus on the self as opposed to a group to which they belong so that they do not focus on the portion of their group identity that was threatened (Ambady et al., 2004). In two experiments, Ambady and colleagues (as depicted as ID 1 in Table 1) examined how distancing one from threatened group identities provided a buffer against stereotype threat. In the first study, women completed a task which triggered stereotype threat by subconsciously priming gender with words related to women (stereotype threat) or not related to women (control). Participants were then asked to either indicate their favorite food, movie, book, hobbies, and briefly discuss their best and worst traits, or answer a series of questions which did not pertain to the self. Participants then completed a challenging math test. Results suggested that women who were primed with their gender group identity performed less

well than the control, but only when they had not deemphasized the threatened group identity. Study two replicated the effects of the first study. These studies show that this deemphasis may act as a buffer against stereotype threat, because it affords women a means to distance the self from group level identities which have been linked to the stereotype in the air.

Extending the finding that focusing on multiple identities may attenuate stereotype threat, Gresky and colleagues (2005) (ID 2, Table 1) sought to explore the effects of multiple identities on performance in a stereotyped domain. Participants were recruited to take a GRE math test and explicitly told that men usually outperform women on math tests. However, before taking the math test, some women were given a task to deemphasize their threatened group identity. Instead of focusing on traits, researchers asked participants to create self-concept maps, which pending experimental condition, varied in their complexity. Results found that women who made simple maps, which reflected the person's most basic characteristics such as one's humor, performed significantly worse on the math test compared to men, and just as poorly as women who received no intervention. However, results found that women who were highly identified in math, and who completed complex self-concept maps, which reflected a complete description of the person, performed just as well as men. Although Ambady et al., (2004) suggested that encouraging women to think of themselves in terms of their valued and unique characteristics may attenuate stereotype threat, Gresky and colleagues argue that individuals must consider themselves as complex and multi-faceted individuals to reduce such vulnerability.

Given that all people have multiple identities (albeit some have been shown to be more or less important), how might the difference in which these identities are highlighted influence the detriments of stereotype threat? In an experiment examining how different highlighted social identities could moderate vulnerability to stereotype threat, participants completed a series of questionnaires designed to highlight different aspects of their social identities that may or may not be associated with stereotypical gender differences in spatial abilities (McGlone & Aronson, 2006, ID 3, Table 1). Participants then completed a mental rotation test, and results suggested that women performed worse when the pre-test survey highlighted gender identity, but best when the survey highlighted academic identity. Conversely, men performed best when their gender identity was made salient and worst when their academic identity was highlighted. Consistent with previous studies, the results suggested that highlighting social identities can moderate performance if they differ in implications for the tested domain.

If focusing on different aspects of social identities improved one group's performance, while hindering another, how can researchers encourage this deemphasis when individuals experience threatened social identities in domains in which a variety of groups may underperform? One such tactic that could be employed involves blurring intergroup boundaries (Rosenthal & Crisp, 2006; Rosenthal et al., 2007). Rosenthal and Crisp (2006) (ID 4, Table 1), asked participants to think of items which men and women have in common (category overlap) or they completed no task (control). When exposed to category overlap, women not only had less preference for stereotypical feminine careers (study one), but also performed better than women in the control on a math test

(study two). Of importance, a third study was conducted which manipulated whether the task preceded the stereotype threat manipulation. Results suggested that performance also improved when the category overlap task preceded the threat manipulation, leading Rosenthal and Crisp to conclude that this intervention might halt stereotype threat from forming.

In a follow-up study (Rosenthal et al., 2007 ID 5, Table 1), researchers examined whether emphasizing similarities between groups might reduce the harm of stereotype threat by changing performance expectancies. In the first experiment, men and women were asked to generate similarities between the groups physical attributes, academic characteristics, non-academic characteristics, or were not asked to generate any similarities (control). Students were then asked to imagine that they were the only member of their gender group in a class of eight others. Participants were then asked for their perceptions of everyone's performance, and women who generated similarities regarding academic characteristics of men and women had higher performance expectancies in math compared with the other conditions. Study two not only replicated study one, but also had participants take a math exam at the end of the study. A mediation analysis suggested that performance expectations partially accounted for the relation between condition and performance. This intervention suggested that generating similarities between men and women in academic settings raises performance expectations, which in turn improves actual performance of women.

Thus far the literature review has focused on the interventions selected based on our inclusion criteria. Although these articles are most relevant to the primary objectives

of this study, research deemphasizing threatened group identity has uncovered important findings for a variety of domains. For example, research has demonstrated that simply moving standard demographic questions about ethnicity and gender to the end of a test resulted in higher performance for women taking an advanced placement (AP) calculus test (Stricker & Ward, 2004). Although these effects were somewhat modest, with an effect size of d = 0.20, a reanalysis of the data with less conservative significant criteria (using $\eta \geq .05$ as opposed to $p < .05$ given the use of the conservative post hoc Bonferroni correction), suggested that if testing organizations implemented shifting the demographic questions to the end of the AP calculus test, an additional 4,700 women would receive credit annually. Additionally, experiments have also demonstrated that this intervention may successfully attenuate ethnic based stereotypes (Shih, Bonam, Sanchez, & Peck, 2007). Specifically, it was hypothesized that biracial individuals may suppress racial stereotypes in conditions in which race is salient. Four studies found that emphasizing the social construction of race, and deemphasizing its biological basis, reduced vulnerability to stereotype threat. These studies speak to the robustness and generalizability that deemphasizing threatened group identities has in real world application, as well as its ability to reduce threats derived from different stereotypes.

Collectively, the five articles grounded in the selection criteria spanning nine studies suggested that performance deficits under stereotype threat can be alleviated by prompting women to distance the self from their threatened group identity. As demonstrated in Table 1, nine of the studies (100%) examined performance, but none have measured motivation. Additionally, five of the nine studies (56%) compared their

intervention to both women who had no intervention, and men. Also, seven (78%) of the studies measured stereotype threat in math, whereas one study (11%) looked at spatial reasoning, with the remaining study (11%) examining threat on careers in general. Finally, as seen in Table 2, seven (78%) of the studies used a subtle threat as opposed to a blatant threat.

Intervention Category 2:
<u>Encouraging Self-Value Affirmations</u>

Consisting of two articles with five empirical intervention experiments, self-value affirmation is founded on that notion that people are motivated to maintain integrity of the self (Steele & Liu, 1983; Steele, 1988; Tesser, 1988). Self-affirmation involves encouraging individuals to consider aspects of the self that are not related to the threatened domain (McGlone & Aronson, 2006). Individuals often belong to many different groups (Cooley, 1902; Rose, 1996) and may share values with both in-group and out-group members. Self-affirmation consists of having individuals think about aspects of themselves that they value or believe are important aspects of their esteem, which can in turn protect damaged self perception that has arisen from the threat (Schimel et al., 2004), allowing individuals to focus on the task at hand, or adopt optimal strategies to complete the task (Stone & McWhinnie, 2008). By fulfilling the need to protect self-integrity in the face of threat, self-worth can be restored when their image of self-integrity is threatened. If the ultimate goal of the self is to protect an image of its adequacy, then it stands to reason that encouraging self-value affirmations would enable

people to deal with threatening events and information without resorting to defensive biases.

The first such study to empirically examine whether or not self-value affirmation would reduce stereotype threat explored whether different kinds of self-value affirmation differ in protecting individuals from stereotype threat (Schimel et al., 2004, ID 6, Table 1). In experiment one, participants were randomly assigned to complete sentence fragments referring to the self to provide an intrinsic (e.g., "being a ____ makes me feel ____") an extrinsic (e.g., "when I am a successful ____ I receive ____") or no affirmation (control; e.g., "watching television is a good way to ____") account. After the affirmation manipulations, students completed a mental math task. Results suggested that students who affirmed the self did not self-handicap as much before engaging in the math task relative to the other two conditions. Additionally, those individuals who self-affirmed performed best on the math test. Experiment two replicated experiment one, except that participants were given information about a "Quantitative Examination" that was either an indication of their "mathematical intelligence" (stereotype threat) or involved, "testing problems for future research" (control). Results of exam performance indicated that women who intrinsically self-affirmed while under stereotype threat excelled on the math test compared with the control condition. However, women who extrinsically self-affirmed performed worse on the exam under stereotype threat compared with the control condition. Finally, in the third experiment, participants completed a task designed to provide either intrinsic or extrinsic affirmation. Women then received either positive or no feedback from an ostensible evaluator. All participants then completed a task designed

to reflect the accessibility of thoughts about rejection by others. Results found that women who received no feedback had more thoughts about rejection if they had extrinsically self-affirmed than if they had intrinsically self-affirmed. Women who received positive feedback showed lowered accessibility of thoughts about rejection. Collectively, these experiments suggested that intrinsic value affirmation reduces concern regarding performance, which subsequently improves performance for women under stereotype threat.

A related study, which sought to directly examine the consequences of value affirmation on women operating under stereotype threat, involved recruiting participants who had a strong history of math performance (Martens et al., 2006, ID 7, Table 1). Participants were told that they would take a test which measured math ability (stereotype threat) or that researchers sought to gain a better understanding of individual impressions of problems (control). Participants were then randomly assigned to self-value affirm by selecting an important characteristic they possessed and writing about a time that it had been important to them. Results found that math ability for men did not differ if they had self-value affirmed; however, women who had self-value affirmed performed better than women in the control condition. Experiment two replicated the findings of the first experiment in a different domain (spatial ability). These findings suggested that value affirmation may help reduce the negative performance consequences of stereotype threat.

The previously outlined studies suggest that encouraging self-value affirmations has reduced stereotype threat for college aged women entering STEM domains; however, a slew of studies have examined how robust such an intervention is. For example,

research has showed that Whites who were given the opportunity to affirm their commitment to being nonracist were less likely to respond in a stereotypic fashion to an implicit measure of racial associations (Frantz, Cuddy, Burnett, Ray, & Hart, 2004). Additionally, research has demonstrated that African-American seventh grade students at racially-diverse schools who were asked to self-value affirm by indicating values that were important to them by writing an essay indicating why those values were important, performed 0.3 grade points better during a semester compared to who had not undergone the intervention. This study found that through the implementation of one 15 minute intervention at the start of the school year, students earned higher grades at the end of their first semester (Cohen, Garcia, Apfel, & Master, 2006). Finally, observational studies have also been conducted in the realm of gender threats of STEM domains (Miyake et al., 2010). Results found that self-values affirmation not only reduced male-female performance and learning differences on overall course work, but was reflected in the elevation of a full letter grade for those women who underwent the intervention relative to those who had not. These studies suggest that value-affirmation is a viable solution to stereotype threat in a variety of domains, and naturalistic settings.

The two articles outlined in the inclusion criteria with five experiments have demonstrated that encouraging self-value affirmation reduced the negative impact of stereotype threat. As seen in Table 1, four (80%) of the studies measured performance, but none have experimentally measured motivation. Additionally, three experiments (60%) only compared the intervention to women under threat who did not receive an intervention. In regards to domains tested, three of the studies (60%) measured threats in

eliminated, because the fear of the target confirming the stereotype is no longer there; the role model has debunked the stereotype (Blanton, Crocker, & Miller, 2000).

In a series of experiments, Marx and Roman (2002) (ID 8, Table 1) examined whether a competent role model could attenuate stereotype threat. In the first experiment, highly math-identified participants were given a difficult diagnostic math test by a male or female experimenter who was said to have developed the math test (high competence). Results indicated that men did not differ in performance when the gender of the experimenter was changed, but women performed significantly better when the test was administered by a competent woman. Experiment two replicated the findings of the first experiment, but added varying degrees of experimenter competence. Results found that women performed less well when the woman administering the test was not viewed as competent, whereas men performed better in the presence of this non-competent woman. Finally, in the third study, women participants were led to believe that they would interact with a same gendered experimenter with high or low competence. Women exposed to the high-competence experimenter not only performed better on their math exam, but also had elevated appraisals of their own math ability, and showed a higher level of state self-esteem compared to women exposed to the low-competence experimenter. These results indicated that role models who challenge a stereotypic assumption of ability in a given domain can eliminate the negative effects of stereotype threat on women's math performance.

Although interacting with a role model may produce a reduction in stereotype threat, a viable role model may not always be present. As a result, McIntyre, Paulson, and

Lord (2003) (ID 9, Table 1) sought to determine if exposing women to something as simple as another woman's accomplishments would be sufficient in alleviating stereotype threat effects. In experiment one, participants were told that men tended to outperform women in math, but the empirical evidence supporting this claim was mixed. Participants were told that they could participate in studies later throughout the semester, but that only woman could participate because, "women produce more reliable and valid data, comprehend the task requirements better, and produce better results in all types of psychological experiments" (women superior) or were told nothing about the gender based restrictions (control). Results showed that women in the control condition performed worse on the math task compared with those who were told women were superior. In experiment two, participants were told that they would participate in two studies, one focusing on "developing stimulus materials for future experiments" and the other involving finding solutions to math problems. The experimenter then induced stereotype threat, read and critiqued four biographical essays either about four successful women (positive role models for women) or about four successful corporations (control). Math task performance varied as a function of gender and exposure to positive role models, yet men did not differ based on the number of essays they read. Collectively, these results suggested that women can overcome performance decrements under gender stereotype threat in mathematics when reminded of the accomplishments of other women.

Competent role models are able to combat stereotype threat, regardless of whether the role model is presently interacting with students (Marx & Roman, 2002), or learned about from a brief information sheet in which their accomplishments were listed

(McIntyre et al., 2003). However, the precise relationship between exposure to positive role models and facilitation of performance under stereotype threat remained unclear. In order to shed light on this lingering question, McIntyre, Lord, Gresky, Ten Eyck, Frye, and Bond (2005) (ID 10, Table 1) conducted an experiment in which the relationship of the amount of exposure and magnitude of the attenuating effect could be ascertained. Participants were asked to participant in two studies, one of which focused on developing stimulus materials for future experiments, and the other on finding math solutions. When the tasks were introduced, the experimenter explained that women perform worse than men on math tests, and explained that they would first read biographical sheets describing women who were successful in various fields. Results found that women performed worse than men when they read no essays about successful women, but that there was no difference in performance if women read even just one essay about a successful woman. Additionally, women's performance did improve as the number of essays they read increased, but to a lesser degree with each additional article. Men's performance was not influenced by the number of role models they were exposed to. This study suggested that exposure to positive role models attenuated stereotype threat as a function of the amount of exposure one has to competent role models.

The previously outlined studies were selected based on our inclusion criteria, but a wealth of literature exists for this intervention, consisting of numerous studies which did not meet our inclusion criteria. For example, research has found that imagining female role-models helped middle school girls perform well in single-sex settings even under conditions where the task description typically produces stereotype threat (Huguet

& Régner, 2007). Additionally, research has noted that self-esteem can be boosted among African Americans under conditions that produce threat if a competent role model is provided who disconfirms negative stereotypes (Blanton, Crocker, & Miller, 2000), which was supported by findings from Marx and Goff (2005), who found that providing in-group role models can reduce stereotype threat and bolster performance. Still yet, naturalistic examples have demonstrated that having African Americans think of Barrack Obama reduced stereotype threat for students completing a standardized verbal test, and that making President Obama's accomplishments salient improved verbal scores on a GRE exam (Aronson, Jannone, McGlone, & Johnson-Campbell, 2009; Marx, Ko, & Friedman, 2009). These experiments suggest that providing competent role models can be applied to ethnic based threats, and observational settings.

Providing a competent role model reminds women that the stereotypes they may face throughout a variety of domains do not need to be ascribed to. By providing examples of women who have excelled in these domains, women are able to overcome performance decrements. Furthermore, the more exposure to these role models that women have, the more their performance tends to increase. The three articles from our inclusion criteria (Table 1), which presented results from six studies, all (100%) measured performance, whereas none measured motivation. Five (83%) of the studies have compared the interventions effectiveness to both women under threat who did not receive an intervention, as well as men. Additionally, five (83%) of the studies examined threats in math with one (17%) study measuring threats in spatial reasoning. Finally, as

noted in Table 2, three of the studies (50%) measured their intervention effectiveness

using a subtle stereotype threat.

Intervention Category 4:
Providing External Anxiety Attributions

Consisting of three articles which conducted seven empirical studies, providing

external anxiety attributions is rooted in attribution theory (Heider, 1958; Kelley, 1967).

This theory suggests that by providing external anxiety attributions, individuals are

immune to the anxiety created by threats which would otherwise distract them from

focusing on the task at hand. Providing external anxiety attributions may be especially

effective for women faced with a STEM stereotype as research has suggested that women

tend to internalize failure (blame the self) when faced with stereotype threats, whereas

men are more likely to believe external factors caused the failure (Koch, Müller,

Sieverding, 2008). As a result, this theory claims that stereotype threat can be diminished

by providing individuals with explanations regarding why anxiety and distraction are

occurring that have not implicated the self or validated the stereotype. Indeed, these

experiments have suggested that stereotype threat can be overcome by overriding feelings

of threat-induced anxiety by misattributing the anxiety to something external.

In one of the earliest experiments testing whether or not providing an external

anxiety attribution to explain the stereotype threat reduced threat detriments (Ben-Zeev,

Fein & Inzlicht, 2005, ID 11, Table 1), participants were told that they would take a

difficult math test, and told that gender differences existed (stereotype threat), or no such

differences had been noted (control). After the test, participants were asked to complete

either an easy task (write their names as many times as they could in 20 seconds) or a novel difficult task (write their name backwards as many times as they could in 20 seconds). Results found that women in the stereotype threat condition wrote significantly more forward names, but fewer backward names, than in the control condition. Men showed no effects of the stereotype threat manipulation. These results were interpreted to mean that women under stereotype threat experience increased physiological arousal. In the second study, researchers were interested in disarming the arousal noted in the first experiment. Female participants were asked to complete a math test in either the presence of two other women (control) or two men (stereotype threat). Before taking the test, the participants were made aware of a "subliminal noise generator" which produced a noise that, pending condition, had no discernible effects, or that it might elevate arousal, nervousness, and heart rate. Results found that women in the stereotype threat condition performed better on the math test when they were told that the generator may cause anxiety. Researchers inferred that stereotype threat was reduced because women had a means to explain their increased arousal; any anxiety induced arousal they felt could be attributed to the subliminal generator and not their worry or concern about confirming the negative stereotype.

Another example of using an external anxiety attribution for anxiety to reduce women's stereotype threat taught students about stereotype threats and its negative impact (Johns et al., 2005, ID 12, Table 1). Participants were recruited to complete a difficult math test which was described as a problem solving task (control), math test (stereotype threat), or a math test with an informative brief lecture of stereotype threat

(intervention). Specifically, women in the intervention condition were told that if they are feeling anxiety, it may be due to negative stereotypes which may be widely known, but have nothing to do with their actual ability. When women were taught about stereotype threat, they performed equally as well as men (who were unaffected by the intervention), and significantly better than the control group of women who did not learn about stereotype threat and its effects. These results suggest that teaching students about stereotype threat could be a viable means for reducing its negative consequences.

Teaching an individual about stereotype threat could reduce the negative consequences which may arise from it (Johns et al., 2005). However, this intervention may not prevent such a threat from arising, but rather allow individuals to function within its presence. As a result, research shifted focus to examine the role of one of the proposed mechanisms of stereotype threat, emotional regulation (Johns, Inzlicht, & Schmader, 2008, ID 13, Table 1). It has been proposed that stereotype threat provokes a variety of emotional responses, and to control these emotional responses, working memory resources are required, depriving the individual under threat of the cognitive resources needed for optimal task performance. In the first experiment, participants were asked to complete a math task in the presence of two men (stereotype threat) or two women (control). Participants then completed tasks designed to measure working memory capacity, and attention to threat relevant stimuli. For the attention task, women were randomly assigned to a vague description of the task - so that the task would function as a measure of the accessibility of threat emotions - or were told it would measure whether they were experiencing anxiety. Results indicated that women under threat showed lower

working memory capacity and those who were given a vague description while under

threat showed a greater attention to threat relevant stimuli. Those who were told about the

nature of the task showed a reduced attention to threatening stimuli compared with the

control condition. Results compared to the control condition, which focused on varying

the strategies used for emotional regulation under stereotype threat, found that effective

emotional regulation could preserve capacity and ensure successful performance. In

experiment three, previous findings were replicated, but participants were told that the

anxiety they experience under threat could even help them. Results suggested that

participants who were told about anxiety performed significantly better than those who

completed the test but anxiety was not mentioned. These experiments suggested that the

process of emotional regulation in response to stereotype threat can tax cognitive

resources, but these resources may be freed when an individual is taught about the

anxiety which may be caused by stereotypes.

Providing reasons as to why failure occurs which does not validate the stereotype

has been examined on a variety of dimensions. Much of the research has examined how

these attributions arise in the face of stereotypes, and their negative impact on judgments,

responsibility, deservingness, punishment, self-esteem, and motivation both in and out of

the classroom (for review see Reyna, 2000). Additionally, research has demonstrated that

providing middle school children with mentors who emphasize external attribution for

academic struggle, helps improve performance such that, girls did not significantly differ

in math performance when they received the intervention (Good, Aronson, & Inzlicht,

2003). Finally, one component that may set providing external anxiety attributions apart

from other interventions is that it has been replicated to assist all students in a general sense. For example, research has found that having students write down their thoughts about an upcoming test, reduced their anxiety about the test, as well as improved student exam scores (Ramirez & Beilock, 2011). Collectively, these studies assert that providing external attributions has a positive impact in classroom settings.

Research on external anxiety attributions has repeatedly found that by providing an external reason which explains why anxiety may be present, without validating the threat, has improved performance. All of the seven studies (Table 1) identified by our inclusion criteria (100%) have been shown to improve performance, but none have measured motivation. Of these studies, four (57%) have found that the intervention is effective compared to women who were only exposed to stereotype threat, and men. All seven (100%) of the studies consisted of threats based in math domains. Finally, as seen in Table 2, three (43%) of the studies were effective in combating a blatant threat, with two shown to having been shown to successfully attenuate the impact of both subtle (28%) and mildly explicit threats (28%).

Intervention Category 5: Emphasizing
an Incremental View of Ability

The least researched area of stereotype threat interventions in the gender and STEM literature consists of two articles which conducted three studies, emphasizes an incremental view of intelligence. The foundation for the theory is derived from differences in how people view intelligence (Dweck & Leggett, 1988; Dweck & Sorich, 1999). Some individuals hold that intelligence is fixed, asserting that ability levels are

innate and cannot be changed over time (entity theorists). Alternatively, others view

intelligence as malleable, suggesting that it can be increased through effort (incremental

theorists). Gender based stereotype threats assert that ability is fixed such that women are

innately worse in STEM domains (Good, Rattan, & Dweck, 2009). As a result,

individuals with this entity orientation were more likely to experience (Sawyer & Hollis-

Sawyer, 2005) and be negatively influenced by stereotype threat (Goff et al., 2008). As a

result, it has been proposed that by emphasizing the importance of effort for success,

stereotype threat can be reduced (Lawrence, Crocker, & Dweck, 2006).

If default mindset established by a stereotype promotes an entity view, changing

this mindset would be key in improving performance and motivation (Dar-Nimrod &

Heine, 2006, ID 15, Table 1). In the first study which empirically tested whether an

incremental mindset can improve performance, participants completed a math test,

followed by a reading comprehension test, and then a second math test. The content of

the reading test served as the experimental manipulation, which depending on condition,

acknowledged gender differences in math performance but attributed the differences to

being rooted in genetics (entity) or effort (incremental) theories, or a control condition

which did not offer an explanation. Performance on the math test was lowest for women

who read essays which attributed math differences to genetic (and thus fixed, entity)

factors, and those who did not receive an explanation as to the gender differences.

Experiment two replicated the results of the first experiment, using a different design.

The results from these studies suggested that anchoring gender differences under entity

theory harmed performance, whereas attributing gender differences to effort attenuated the negative consequences for performance.

Elaborating on the findings which asserted that an emphasis on effort could reduce stereotype threat, Thoman et al., (2008) (ID 15, Table 1) conducted an experiment which speculated that stereotype threat regarding women's math performance might be based on the implication that women differ from men in inherent math ability as opposed to effort. In this study, participants completed a difficult math test under three conditions. Students were randomly assigned to either read an experimentally crafted scientific article which emphasized that gender differences in math performance were "innate" (entity threat) or existed because men exerted more effort (incremental threat), or did not read an article and went straight to the test. Although students in the incremental threat condition completed fewer items, they answered a greater percentage of test items correctly compared to both the control and entity-threat conditions. These results suggested that performance deficits could be ameliorated if differences in effort were emphasized other those of ability. Additionally, Thoman et al., (2008) included pre and post test measures of motivation. Although no differences were found on the pre or post exam subjective measures of motivation and evaluation of the exam, significant simple correlations with the ratio of correct problems were noted, such that, participants performed better when they expected to do well.

Although incremental theory research is limited on college aged women who face stereotype threats in STEM domains, a wide body of literature exists for how views of intelligence can impact stereotype threats. One such study found that malleability of

intelligence was related to test scores and stereotype threat beliefs, with those who naturally believed that intelligence was fixed (entity theorists) performing worse and demonstrating a higher degree of stereotype threat on cognitive tests (Sawyer & Hollis-Sawyer, 2005). Additionally, as noted in the previous section, Good et al., (2003) not only provided external anxiety attributions for difficulty, but also designed their intervention to promote an incremental mindset, which found that emphasizing the malleability of intelligence through effort significantly improved grades. Such findings have also been replicated for threats derived from other stereotypes. Similar findings have also shown that when women learn new mathematic techniques in classrooms with an entity perspective, they perform less well when exposed to a stereotype relative to when no stereotype is present, but when an incremental perspective is emphasized, no differences between stereotype threat and non-threat conditions existed. For example, research has noted that by having students read essays which encourage incremental orientations (i.e. intelligence can grow with work and effort) were less likely to confirm negative racial stereotypes (Aronson, Fried, & Good, 2002; Goff, Steele, & Davies, 2008). These studies suggest that beliefs about the malleability of intelligence can reduce the consequences of stereotype threat through many domains.

Studies which encouraged an incremental over entity mindset were able to significantly improve performance when an individual faced stereotype threat. Of the three studies identified by our inclusion criteria (Table 1), all measured performance (100%), whereas motivation was only measured in one (33%). Of importance, all studies (100%) only compared the interventions effectiveness against women who did not

receive an intervention, and all (100%) measured math based threats. Finally, all (100%) studies have only been tested against blatant threats (Table 2).

In Summary

Although there are many effective means for reducing stereotype threat in college women, it is not known the extent to which one intervention is better than another, if all are equally effective, or if some interventions are primarily useful for enhancing performance but not motivation and attitudes. In conducting this literature review, 15 distinct studies with 28 experiments were located. Between these 28 studies, a mix of methodologies were employed ranging from how stereotype threat was manipulated (subtle v. blatant) to what control group was used for comparison (men - no threat v. women - threat only). While this range may impose some limitations on the manner in which the results of this study can be generalized, it also strengthens this project's ability to make assertions about the target sample (i.e. women in STEM). Thus, the aim of phase two was to hold constant the manipulation of stereotype threat and to provide both comparison groups to test the relative effectiveness of the five interventions outlined above.

Table 1. *Intervention collapsed by subtypes.*

TYPE	AUTHORS	ID	YEAR	EXP	PERFORMANCE	MOTIVATION	COMPARED	DOMAIN
Emphasize Individuation	*Ambady, Paik, Steele, Owen-Smith, & Mitchell*	01	2004	01	YES	NO	CONTROL	MATH
				02	YES	NO	CONTROL	MATH
	Gresky, Ten Eyck, Lord, & McIntyre	02	2005	03	YES	NO	CONTROL / MEN	MATH
	McGlone & Aronson	03	2006	04	YES	NO	CONTROL / MEN	SPATIAL
	Rosenthal & Crisp	04	2006	05	YES	NO	CONTROL / MEN	CAREERS
				06	YES	NO	CONTROL	MATH
				07	YES	NO	CONTROL	MATH
	Rosenthal, Crisp, & Suen	05	2007	08	YES	NO	CONTROL / MEN	MATH
				09	YES	NO	CONTROL	MATH
Value Affirmation	*Schimel, Arndt, Banko, & Cook*	06	2004	10	YES	NO	CONTROL	MATH
				11	YES	NO	CONTROL	MATH
				12	NO	NO	CONTROL	AFFIRM
	Martens, Johns, Greenberg, & Schimel	07	2006	13	YES	NO	CONTROL / MEN	MATH
				14	YES	NO	CONTROL / MEN	SPATIAL
Provide Role Models	*Marx & Roman*	08	2002	15	YES	NO	CONTROL / MEN	MATH
				16	YES	NO	CONTROL / MEN	MATH
				17	YES	NO	CONTROL	MATH
	McIntyre, Paulson, & Lord	09	2003	18	YES	NO	CONTROL / MEN	SPATIAL
				19	YES	NO	CONTROL / MEN	MATH
	McIntyre, Lord, Gresky, Ten Eyck, Frye, & Bond Jr.	10	2005	20	YES	NO	CONTROL / MEN	MATH

Table 1 Continued

External Attributions	*Ben-Zeev, Fein, & Inzlicht*	11	2005	21	YES	NO	CONTROL / MEN	MATH
				22	YES	NO	CONTROL / MEN	MATH
	Johns, Schmader, & Martens	12	2005	23	YES	NO	CONTROL / MEN	MATH
	Johns, Inzlict, & Schmader	13	2008	24	YES	NO	CONTROL / MEN	MATH
				25	YES	NO	CONTROL	MATH
				26	YES	NO	CONTROL	MATH
				27	YES	NO	CONTROL	MATH
Incremental Mindset	*Dar-Nimrod & Heine*	14	2006	28	YES	NO	CONTROL	MATH
				29	YES	NO	CONTROL	MATH
	Thoman, White, Yamawaki, & Koishi	15	2008	30	YES	YES	CONTROL	MATH

Note: Control denotes a women no intervention comparison group.

Table 2. *Design and threat type of interventions by subtype.*

TYPE	ID	YEAR	EXP	DESIGN	THREAT
Individ uation	01	2004	01	2 (stereotype threat: gender primed / not primed) x 2 (questionnaire: individuated or nonindividuated)	SUBTLE
			02	2 (stereotype threat: gender primed / not primed) x 2 (questionnaire: individuated or nonindividuated)	SUBTLE
	02	2005	03	3 (map: few nodes / many nodes / no map) x 2 (gender: male / female) x 2 (math identification: high / low)	BLATANT
	03	2006	04	2 (gender: female / male) x 3 (social identity: gender salient / student at private school salient / Northeastern resident salient)	SUBTLE
	04	2006	05	2 (gender: female / male) × 2 (order: female first vs. male first) × 2 (career rating: female vs. male)	SUBTLE
			06	Non-Factorial 3 (Task: overlap / baseline / difference)	SUBTLE
			07	Non-Factorial 4 (baseline / threat / overlap-threat / threat-overlap)	BLATANT
	05	2007	08	2 (gender: male / female) x 2 (test domain: math / English) x 4 (condition: baseline vs. physical vs. non-academic vs. academic) x 2 (target score expectancies: self vs. others)	SUBTLE
			09	Non-Factorial 4 (condition: baseline vs. physical vs. non-academic vs. academic)	SUBTLE
Value Affirm ation	06	2004	10	Non-Factorial 3 (value affirmation: intrinsic / extrinsic / neutral)	
			11	2 (stereotype threat: threat / no threat) × 2 (value affirmation: intrinsic / extrinsic)	SUBTLE
			12	2 (self-affirming description: intrinsic / extrinsic) × 2 (feedback: positive / none)	
	07	2006	13	2 (gender: male / female) x 2 (stereotype threat: reasoning / math ability) x 2 (value affirmation: important attribute / unimportant attribute)	SUBTLE
			14	2 (gender: male / female) x 2 (affirmation: threat + affirm / threat no affirm)	BLATANT
Provid e Role Models	08	2002	15	2 (gender: female, male) x 2 (test description: diagnostic, nondiagnostic)	SUBTLE
			16	4 (social group: student, friend, gender, family) x 2 (test description: diagnostic, nondiagnostic)	SUBLTE
			17	3 (social comparison valence: positive, negative, control) x 2 (test description: diagnostic, nondiagnostic)	SUBTLE
	09	2003	18	2 (gender: female / male) x 2 (instructions: women are better at following instructions / nothing)	BLATANT
			19	2 (gender: female / male) x 2 (essay: successful women / successful companies)	BLATANT
	10	2005	20	2 (gender: female / male) x 5 (biographies read: 0, 1, 2, 3, 4)	BLATANT
Extern al	11	2005	21	2 (stereotype threat: Read differences / no differences) x 2 (Misattribution: task difficult / task easy)	BLATANT

Anxiety Attributions			22	2 (stereotype threat: minority / majority status) x 2 (Misattribution: noise generator / no generator)	SUBTLE
	12	2005	23	2 (gender male / female) x 3 (test description: problem solving / math test/ or teaching intervention)	BLATANT
	13	2008	24	2 (stereotype threat: presence of men / presence of women) x 2 (task description: vague / anxiety)	SUBLTE
			25	Non-Factorial 3 (Test information: Stereotype threat alone / Stereotype threat with suppression / Stereotype reappraisal)	MILDLY EXPLICIT
			26	Non-Factorial 2 (Stereotype alone / Stereotype with anxiety explanation)	MILDLY EXPLICIT
			27	2 (Race: White / Latino) x 2 (Information: Anxiety mentioned / Anxiety Reappraised)	BLATANT
Incremental Mindset	14	2006	28	2 (Sex differences: genetic / experiential) x 2 (stereotype threat: no gender differences / gender differences)	BLATANT
			29	2 (Sex differences: genetic / experiential) x 2 (stereotype threat: no gender differences / gender differences)	BLATANT
	15	2008	30	Non-Factorial 3 (view: ability / effort / control)	BLATANT

Note: Blank fields denote that stereotype threat was not manipulated / measured in the specific experiment

PHASE TWO: AN EMPIRCIAL TEST OF INTERVENTIONS

Methods

Participants

A total of 151 students (129 women, 22 men, 88% Caucasian) were recruited

from the Montana State University psychology subject pool through Sona. 14 students

were excluded from the analysis for failure to follow instructions or errors with the

computer program, leaving 137 participants (117 women, 20 men). All participants were

awarded course credit in exchange for participation. Participants were run in single-sex

groups of up to four participants at a time. Female participants were randomly assigned to

one of 6 conditions: deemphasizing threatened group identity, self-value affirmations,

providing role models, external attributions, encouraging an incremental mindset, and no

intervention threat only conditions, whereas male participants were assigned to the

gender stereotype-present condition only resulting in a total of 7 conditions.

Procedure

Upon arriving in the lab, participants (same sex groups of one to four) were

greeted by an experimenter and asked to sign an informed consent. Only 14 (10%) of the

137 participants were run by a woman experimenter, with all interventions having a range

from one to three sessions run by a female experimenter. Although an attempt to maintain

random gender assignment for experimenter was made, as research has suggested that

gender of the experimenter can influence participants perceptions of stereotype threat

(see Marx et al., 2002) limitations on availability restricted the ability to keep the gender

of the experimenter evenly distributed among conditions (see Table 3, page 63). All

participants were told that the study was designed to examine "how a variety of

psychological factors may influence perceptions of an academic domain" and were lead

to believe that a number of exams were under investigation. Next, all participants

completed a math identification measure to get a baseline of participants experience with

math before they receive the experimental manipulation (Smith & White, 2002). After

participants filled out this survey (which was embedded within a larger set of items so as

not to arouse suspicion), they were told that they have been "randomly assigned to

complete a math exam" (in fact, all participants were given the math exam). Before

participants began their test, they listened to a recorded powerpoint presentation which

gave them instructions about the purpose of their test. Unknown to participants, this

powerpoint presentation served as the activation of the gender stereotype in math in

which participants were told that their "performance on this test will be used to help us

establish performance norms for men and women" and that their "performance will give

us a clear idea of your natural mathematical ability." All participants (male and female)

received this information, and for women, this information served as a mildly explicit

stereotype threat (for men, this same information was presented, but we hesitate to call it

stereotype threat given that the stereotype is positive toward men). Borrowed from

Schmader (2002), this mildly explicit threat was chosen as it seemed to be the best

compromise between the subtle and explicit threats used in past stereotype threat

intervention research (for review see Nguyen & Ryan, 2008). After activating the

stereotype, participants received the experimental condition instructions which explained

that they were either in the writing, listening, reading, or relaxation task pending the

methodology used to implement the intervention (described below). However, despite the

name of the task, all participants were told that the purpose of the task was to obtain

background information on the types of tasks individuals may do before taking an exam.

In all but one case (Ambady et al., 2004) original materials were provided by the original

authors of the intervention studies. In the one case where we were unable to obtain

original materials, we carefully followed the protocol outlined in the published study.

Following the experimental condition, all participants took a 15 minute math

exam, filled out a survey packet containing measures of intrinsic motivation, attitudes

toward STEM, and items pertaining to perceived difficulty of the intervention (described

below). Finally, participants were thanked for their time and fully debriefed.

Experimental Conditions

Deemphasizing Group Identity Threat   Following the procedure utilized by

Ambady et al., (2004), participants were told that they had been assigned to the "writing"

condition, in which the purpose was to collect background information about the type of

tasks participants may be involved in prior to taking an exam. First, participants were

given a paper based survey which asked them to indicate their favorite food, movie,

book, and hobbies and listed their three best and worst qualities. This procedure was

meant to draw attention away from the gender group (which was threatened by the

stereotype activation) and instead focus students on their unique individual qualities.

Although Ambady et al. (2004) was selected as the methodology for this intervention, it

should be noted that five other interventions were identified as alleviating stereotype

threat using similar methodology (Gresky et al., 2005; McGlone& Aronson, 2006;

Rosenthal & Crisp, 2006; Rosenthal et al., 2007; Schamder, 2002). As seen in Table 3

(page 63), this intervention took an average of 3.52 minutes to administer, making it the

second longest intervention to implement. A one-way ANOVA followed up with LSD

post-hoc tests found that this intervention took significantly more time ($p < .001$) than

providing external anxiety attributions or role models, although the test of homogeneity

of variance was violated ($p < .05$). This intervention was rated as relatively easy to

implement ($M = 2.75$) by participants and did not differ in perceived difficulty from any

other condition (see figure 1).

Figure 1. Deemphasizing Group Identity Threat
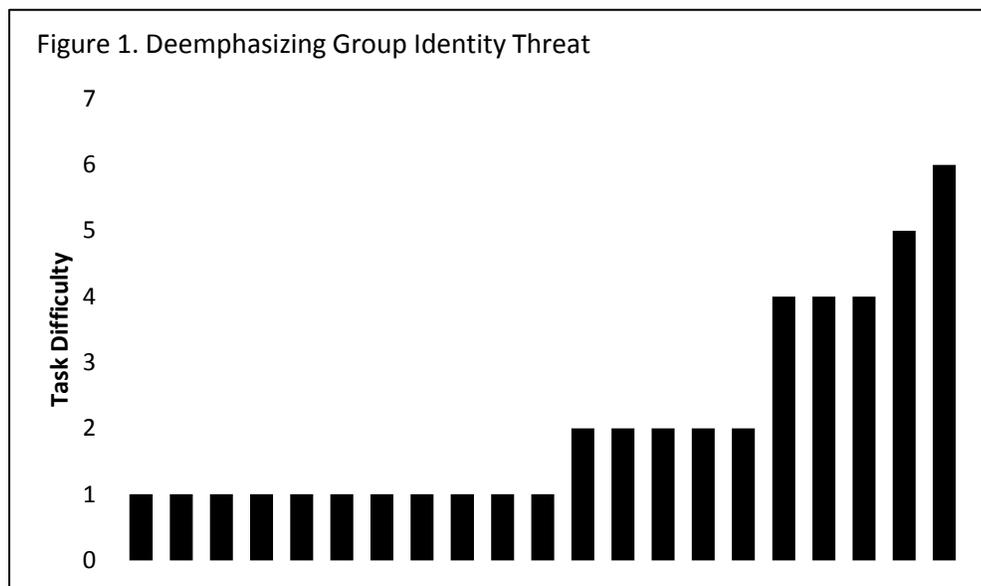
Encouraging Self-Value Affirmation   Following procedures by Martens et al.,

(2006; see also Schimel et al., 2004), participants were told that they had been assigned to

the "writing" condition, in which the purpose was to collect background information

about the type of tasks participants may be involved in prior to taking an exam. First, participants were given a paper based question prompt and asked to rank order five characteristics (creativity, humor, physical attractiveness, social skills, and relations with friends and family) by degree of importance on a scale of 1 (least important) to 5 (most important). Next, participants were then asked to "write why your most valued characteristic (item you ranked as 1) is personally important to you." Finally, participants were asked to write about "a time in which this characteristic had been particularly important to you". This intervention allowed participants to maintain integrity of the self, by considering values which are important to them, which have been successfully used, so that self-worth can be replenished. As seen in Table 3 (page 63), this intervention took an average of 4.07 minutes to administer, making it the longest intervention tested, taking significantly ($p < .001$) more time than both external anxiety attributions and providing a role model. This intervention did not significantly differ in a rating of difficulty compared to all other conditions, and was seen as relatively easy ($M = 2.11$) to implement (see figure 2).

Figure 2. Self-Value Affirmations
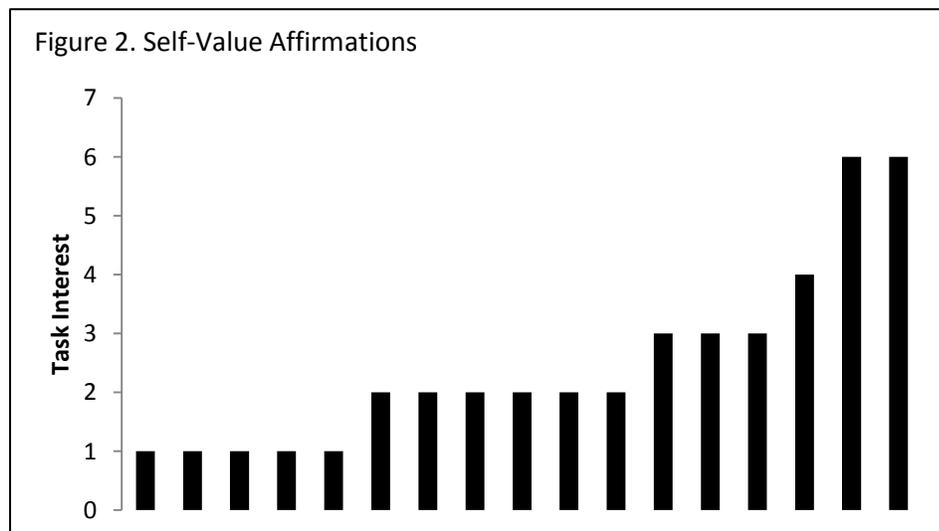
Providing Competent Role Models   Following procedures by Marx and Roman (2002) and McIntyre et al., (2003; 2005), participants were told that they had been assigned to the reading condition, which purpose was to collect background information on reading comprehension, and were warned to read the article carefully as they would be tested for their memory of its contents later in the study. First, participants were asked to read a brief (five paragraphs) experimentally created student newspaper article (Marx et al., 2002; 2005) which portrayed an undergraduate women from the university majoring in math and education. Throughout the article, the role model professes her interest in math ("I have always enjoyed math!", "taking classes in math has really helped me…"), and demonstrates her competence by saying that she has excelled in a variety of math classes such as linear algebra, advanced calculus, and math logic. The article was printed, scanned, and placed into a powerpoint presentation to give it added authenticity. Once on the screen that contained the article, the computer required students to stay on the screen containing the article for at least one minute to ensure that they read the article. By learning about a highly competent in-group member in a threatened domain, stereotype threat is presumably reduced, as the stereotype itself has been disproven. Methodology among Marx and Roman (2002) and McIntyre et al., (2003; 2005) only differed in whether the role model was present, or read about. Participants read about a role model, as opposed to being in the presence of a role model, because it was relatively simple and cost effective to implement. As seen in Table 3 (page 63), providing role models took an average of 1.71 minutes to complete, making it the shortest intervention tested. This intervention took significantly less time ($p < .001$) to complete compared to all

conditions except for providing external anxiety attributions. This intervention and did

not significantly differ in a rating of difficulty ($M = 2.73$) compared to all other

conditions (see figure 3).

Figure 3. Role Model
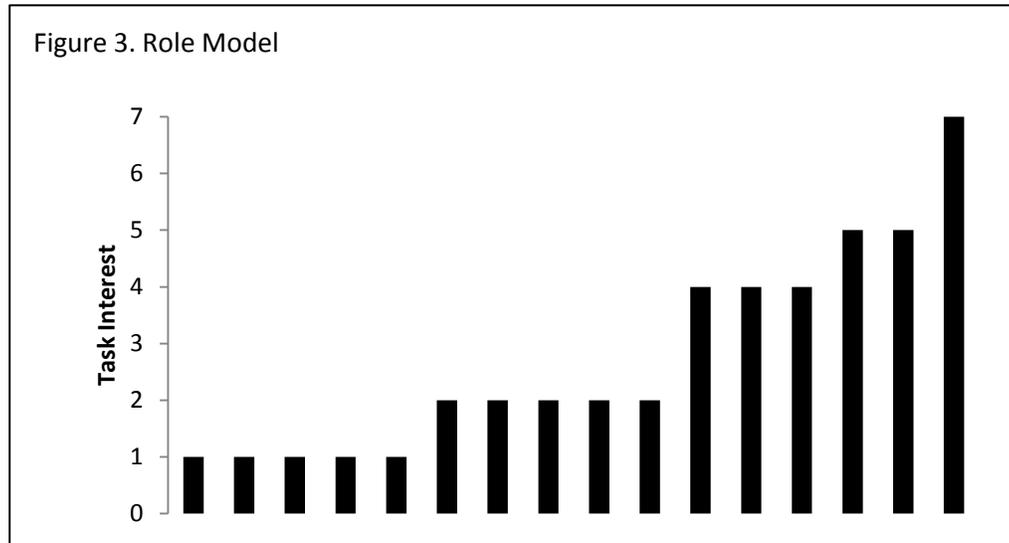
Providing External Anxiety Attributions   Following procedures by Johns et al.,

(2003) and Ben-Zeev et al., (2005) participants were told that they were assigned to the

"listening condition," which purpose was to collect background information about the

type of tasks in which participants may be involved prior to taking an exam. Participants

were first given a short lecture which provided an external reason as to why gender

performance differences on math exams may exist ("Women sometimes do worse than

men on math tests, not because they have less ability, but because there are negative

stereotypes that suggest women should do worse because they aren't "wired" for math

like men.") which did not validate the stereotype ("...it's important that to keep in mind

that if you feel anxious while taking this test, that this anxiety could be a result of these

negative stereotypes that are widely known in society and have nothing to do with your

actual ability to do well on the test"). This intervention was delivered through a

powerpoint slideshow which contained narrated text on each slide. Of importance, the

methodology of the intervention demonstrated by Johns et al., (2003) was selected over

Ben-Zeev et al., (2005) as teaching individuals about stereotype threat had a higher

degree of ecological validity compared to excusing anxiety through a subliminal noise

generator. As seen in Table 3 (page 63), providing external anxiety attributions took the

fourth longest amount of time to implement, with an average of 2.19 minutes, taking

significantly ($p < .001$) less time than self-value affirmation, and deemphasizing

threatened group identity. This intervention did not differ in difficulty ratings ($M = 2.07$)

compared to other interventions (see figure 4).

Figure 4. External Anxiety Attribution
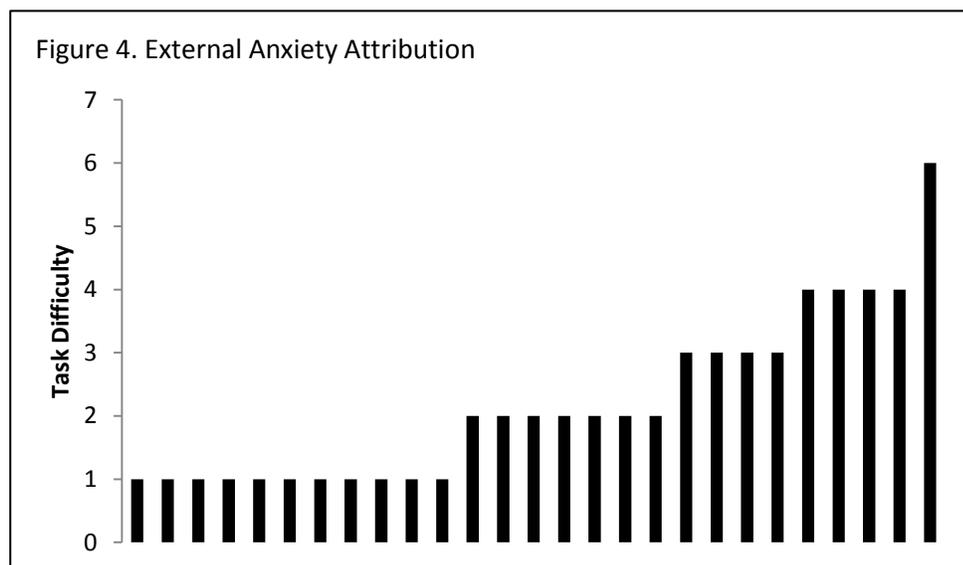
<u>Emphasizing Incremental View of Intelligence</u>   Based on the procedure executed

by Thoman et al., (2008; see also Dar-Nimrod & Heine, 2006), participants were told

they were assigned to the "listening condition," which purpose was to collect background

information about the type of tasks participants may be involved in prior to taking an

exam. On a powerpoint slideshow, participants received narrated text which they were

told was from a published scientific journal (Thoman et al., 2008). This experimentally

constructed article explained that if gender differences exist it is only because of a

"difference in effort toward mathematics," and that "any decrement in mathematical

ability existed because of greater 'mathematical effort.'" This intervention tells students

that by exerting more effort, they will be able to be successful in mathematics,

overcoming any obstacles which existed due to gender bias. As noted in Table 3 (page

63), this intervention took an average of 3.02 minutes to implement, making it the third

longest intervention, taking significantly ($p < .001$) more time than providing role

models. This intervention was viewed as the as easy to implement ($M = 2.90$) as the other

interventions (see figure 5).

Figure 5. Incremental Mindset
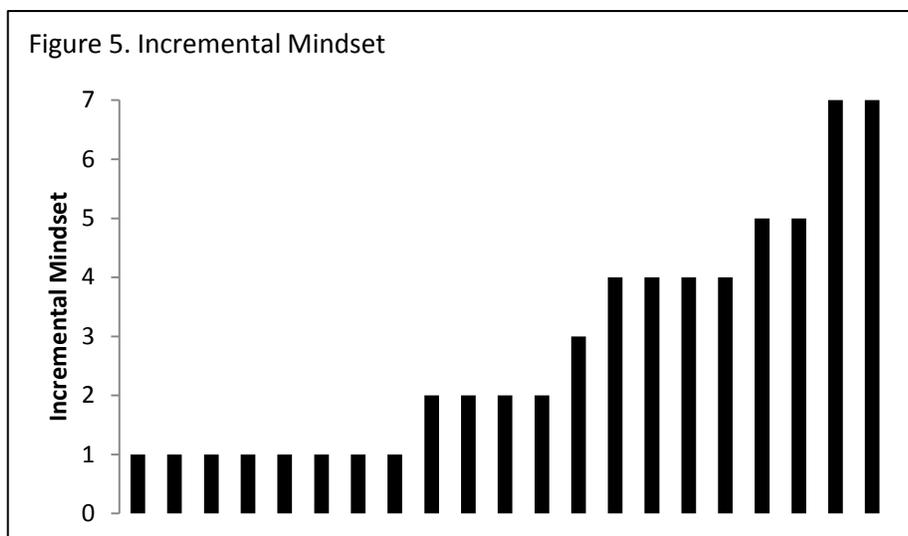
No intervention – Stereotype Threat Only   Women randomly assigned to the

stereotype threat no intervention condition were provided with the stereotype activation

information, and were told that they have been assigned to the "relaxation condition,"

which purpose was to collect background information about the type of tasks participants

may be involved in prior to taking an exam. This consisted of participants sitting in front

of a black computer screen, until the experimenter used a powerpoint clicker which

allowed the participant to the go to the next screen. To ensure that time between the

intervention and the test was not a factor, participants in this stereotype threat only

condition were yoked to the participant who took the longest amount of time in the same

session. Difficulty of this condition ($M = 2.80$) did not differ compared to any of the

interventions (see figure 6).

Figure 6. No Intervention - Stereotype Only
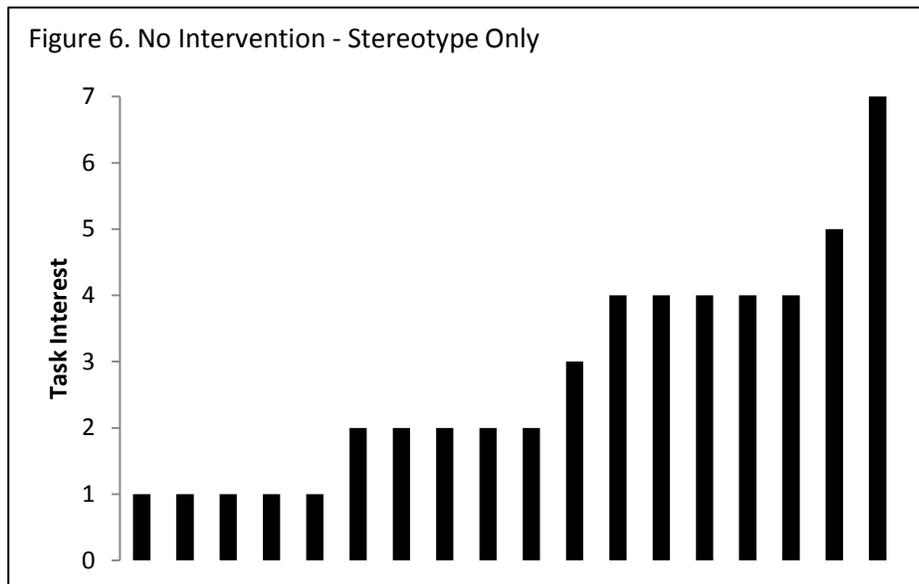
No Intervention – Men   Men were run in the identical manner as women in the *stereotype threat no intervention* condition. When asked to relax, men were given the same amount of time between the stereotype information and the exam as the woman in the *stereotype threat no intervention* condition immediately before them. Difficulty of this condition ($M = 2.55$) did not differ compared to any of the interventions (see figure 7).

Figure 7. No Intervention - Men
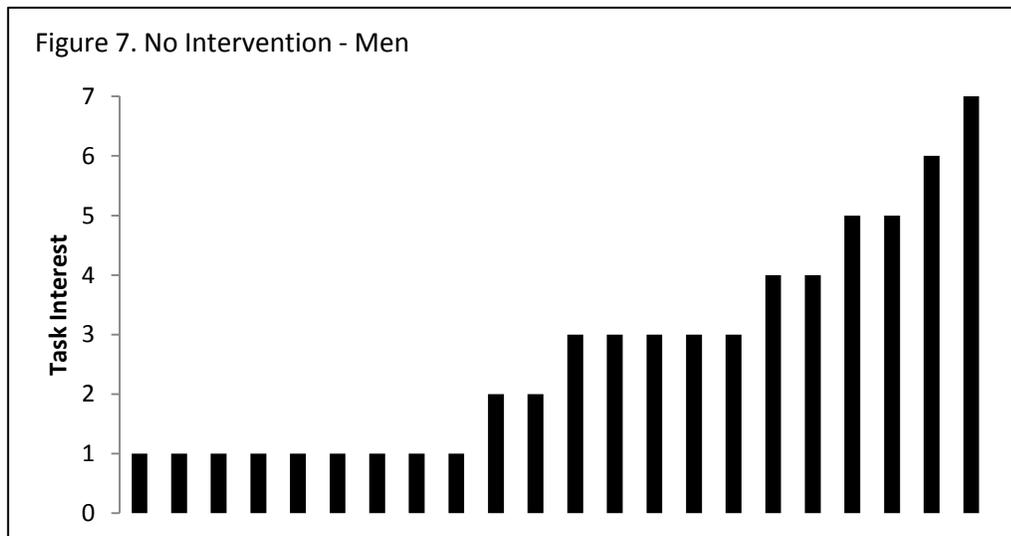
Table 3. *Implementation table for intervention ratings.*

	Sample	Intervention Time		Intervention Difficulty		Exam Difficulty		Exp Gender	
	n	*(Mins.)*		*M*	*(SD)*	*M*	*(SD)*	*W*	*M*
		M	*(SD)*						
Women (Threat Only)	15	3.46^{ac}	(1.94)	2.80	(1.82)	4.33	(1.50)	2	17
Self-Value Affirm	19	4.07^{ac}	(1.59)	2.11	(1.32)	3.95	(1.55)	2	25
External Anxiety	27	2.19^{bd}	(1.41)	2.07	(1.47)	4.44	(1.52)	3	18
Deemphasize	20	3.52^{ac}	(1.86)	2.75	(1.55)	4.10	(1.48)	3	17
Role Models	15	1.71^{d}	(0.41)	2.73	(1.79)	4.00	(1.36)	2	13
Incremental Mindset	21	3.02^{bc}	(0.61)	2.95	(1.71)	3.62	(1.71)	1	14
Men (Gender Prime)	20	3.58^{ac}	(1.86)	2.55	(2.21)	3.70	(1.63)	1	19

Note: All means within rows not sharing a subscript are significantly different at $p < .05$ using Fisher's least significant difference post hoc test. All *d*'s and *g*'s were computed by comparing the condition to women who did not receive an intervention.

<center>Measures</center>

Math Identification (DIM)

Participants completed a math identity measure (Smith et al., 2007) which consisted of 9-items ("Being good at math is important to who I am.") that ensured random sampling was maintained, and that math identification did not differ by condition, as some research suggests that those who have a lower math identification may experience less stereotype threat (Lesko & Corpus, 2006). All responses were collected on a likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Primary Dependent Measures

Math Test Performance   Following protocol by numerous stereotype threat researchers (Aronson et al., 1999; Spencer et al., 1999; Walsh, Hickey, & Duffy, 1999; Quinn & Spencer, 2001; Keller, 2002; Marx & Roman, 2002; Schmader, 2002; Smith & White, 2002; Pronin et al., 2004; Stricker & Bejar, 2004; Gresky et al., 2005; Johns et al., 2005; Cullen, Waters, & Sackett, 2006; Lesko & Corpus, 2006; Beaton, Tougas, Rinfret, Huard, & Delisle, 2007; Keifer & Sekaquaptewa, 2007), the measure of performance consisted of one computerized math test (modeled after Marx et al., 2005), which consisted of 20 multiple choice questions presented on a powerpoint presentation. A math test was chosen because research has demonstrated that it is the math component of STEM domains which activates the stereotype threat (Smith et al., 2005; 2007; Shapiro & Williams, 2011). Participants received two pieces of scratch paper, and a pencil to record their responses onto a bubble sheet. Participants were instructed to raise their hand if they

needed additional scratch paper. Once participants started the exam, a timer in the computer counted down from 15 minutes. When their time was up, a mild alarm was sounded by the computer and narration instructed students to take off their headphones and wait for further instructions. Four of the items on the math test were dropped due to their significant negative correlation with the other items, although patterns of results were the same when including these items. The remaining 16 items were used to compute a total "percent correct" for each participant, and had satisfactory internal consistency (α = 0.64).

Intrinsic Motivation   The extent to which each intervention enhanced (or not) experiences of intrinsic motivation during test taking was examined using a mean composite score of 6 items (e.g. "I would describe this task as very interesting" and "while working on this task I lost track of time"), that indexed task interest and task involvement (Smith et al., 2007; Sansone, Wiebe, & Morgan, 1999). This measure used a 1 (extremely unlikely) to 7 (extremely likely) scale, indicating that the higher the score, the more intrinsic motivation participants experienced. The scale had a sufficient internal consistency (α = 0.65), and an average score was calculated.

Student Subjective Science Attitude Change   To measure general attitude positivity toward science domains, the 28-item SSSAC scale by Stake & Mares (2001) was employed. To generalize findings toward STEM, questions were duplicated so that they asked about perceptions of a career in both science and math ("I would enjoy a career in math"; "I would enjoy a career in science"). All responses were collected on a

likert scale ranging from 1 (not at all) to 7 (a great deal) and averaged into a mean score.

The scale had a high degree of internal consistency ($\alpha = 0.91$).

Secondary Measures

Feelings of Belonging   The 10 item scale was revised to include items related to both math and science, making the revised version a total of 20 items ("People in STEM domains accept me") was implemented to test to what degree participants felt accepted in STEM (Leary, Kelly, Cottrell, & Schreindorfer, 2007). Items were rated on a likert scale from 1 (strongly disagree) to 5 (strongly agree) the scale had a high degree of reliability ($\alpha = 0.93$), and a total mean level of belonging was calculated.

Intervention Time   To test how long each participant took in each intervention condition, time spent engaging with the intervention was kept with a stop watch by the experimenter and recorded in minutes. In the case of more than one participant, multiple stop watches were used. The experimenter began the stop watch after the participant received their intervention instructions, and ended it when the participant clicked the "next" button on the computer.

Intervention Difficulty   The perceived difficulty of the intervention task was assessed by each participant with one question ("I think this is a hard task") on a likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

Results

Analysis Overview

Separate one way analysis of variances (ANOVAs) were conducted on the three primary dependent measures. In order to interpret the overall significant differences, Fisher's least significant differences (LSD) were computed (Hayter, 1986), with significance determined at $p < .05$. To measure the strength of the difference between all conditions compared to the *no intervention – stereotype threat only*, Cohen's d and Hedge's g were computed. When examining the F- tests for ANOVAs, effect sizes can range from small (d = 0.20), to medium (d = 0.50), to large (d = 0.80) (Cohen, 1992; Wilkinson, 1999).

Math Identification

A one way ANOVA was computed on a 9 item scale ($\alpha = 0.93$), measuring pre-test math identification. Results indicated that the ANOVA was not significant $F(7,149) = 0.454$, $p > .80$, suggesting participants in each condition did not differ in math identification.

Performance

To examine if the presence vs. absence of any intervention improved performance, a planned contrast was performed in which all interventions were weighted equally (dummy coded as +1), and the pooled differences against intervention women only comparison group (-5). Results yielded a significant contrast effect $t(130) = 2.14$, $p < .05$, suggesting that women who received the deemphasizing group identity threat, self-

value affirmation, external anxiety attributions, role models, and an emphasis on an incremental mindset interventions any intervention performed significantly better than those women in the *no intervention – stereotype threat only*.

In order to determine whether interventions differed between one another in their ability to attenuate stereotype threat, a one way ANOVA showed the percentage of answers correct on the math exam significantly differed by condition $F(6,130) = 2.76$, $p <$ .05[1] (see Table 4, page 71). A LSD post hoc analysis indicated that women in the *no intervention – stereotype threat only* condition ($M = 33.43$, $SD = 12.62$) performed significantly worse than men ($M = 55.75$, $SD = 20.65$) with a large effect size difference $d = 1.28$ replicating past stereotype threat effects on performance. Importantly, Table 4 shows that women in the self-value affirmation intervention ($M = 47.47$, $SD = 17.07$) and external attribution intervention ($M = 47.78$, $SD = 20.10$) performed significantly better than women under stereotype threat who did not receive an intervention ($M = 34.33$, $SD = 12.62$). In addition, in these two intervention conditions, women performed similarly well compared to men ($M = 55.75$, $SD = 20.65$). As noted in Table 4, the self-value affirmation intervention resulted in a large difference in improved performance for women compared to women in the no intervention stereotype threat only condition ($d = 0.86$) whereas the external attribution intervention resulted in a moderate improved difference in performance ($d = 0.77$). The other interventions did not significantly improve performance compared to women in the no intervention stereotype threat only

---

[1] An analysis of covariance (*ANCOVA*) using intervention time was computed for performance $F(1,122) = 0.90$, $p = 0.35$. Results found that time did not alter the reported patterns.

condition (although in some cases the effect size difference was still moderately large, see Table 4).

Intrinsic Motivation

A one way ANOVA was computed to examine mean intrinsic motivation, and found significant group differences $F(6,130) = 4.62$, $p < .001$[2] (See Table 4). A LSD post hoc analysis indicated that women in the *no intervention – stereotype threat only* condition ($M = 2.69$, $SD = 1.03$) did not differ from men ($M = 2.65$, $SD = 0.90$) in intrinsic motivation. However, as documented on Table 4, women in the role model ($M = 3.58$, $SD = 0.73$), self-value ($M = 3.73$, $SD = 0.76$), and deemphasize threatened group identities ($M = 3.42$, $SD = 0.90$) had significantly higher intrinsic motivation relative to both men and women controls. No other interventions differed significantly from women whom did not receive an intervention. Effect size analysis revealed that value-affirmation provided the largest group difference relative to women in the *no intervention – stereotype threat only* condition ($d = 1.19$), followed by providing a role model ($d = 1.01$), and deemphasizing threatened group identities ($d = 0.77$).

Attitude Toward STEM

A one way ANOVA found significant differences between conditions in attitude toward STEM, $F(6,92) = 3.53$, $p < .01$[3] (see Table 4). LSD post hoc analysis revealed that women in the *no intervention – stereotype threat only* condition ($M = 3.10$, $SD =$

---

[2]An ANCOVA using intervention time was computed for intrinsic motivation $F(1,122) = 1.82$, $p = 0.17$. Results found that time did not alter the reported patterns.

[3] An ANCOVA using intervention time was computed for attitude toward stem $F(1,122) = 0.63$, $p = 0.44$. Results found that time did not alter the reported patterns.

1.01) had a significantly less positive attitude toward STEM than men ($M = 4.04$, $SD = 1.05$). As indicated by Table 4, all interventions significantly improved attitude toward STEM compared to women who did not receive an intervention. Incremental mindset had the largest effect size ($d = 1.98$), with all other interventions also showing a large effect size as well.

Feelings of Belonging

A one way ANOVA found no significant differences in feelings of belonging in STEM between conditions $F(6,111) = 1.41$, $p > .05$. These results suggest that interventions did not contribute toward feelings of belonging in STEM domains

Table 4. *Mean ratings of performance, intrinsic motivation and future task motivation.*

	Performance (%)				Intrinsic Motivation				Attitudes Toward STEM			
	M	*(SD)*	*d*	*g*	*M*	*(SD)*	*d*	*g*	*M*	*(SD)*	*d*	*g*
Women (Threat Only)	33.43^c	(12.62)	---	---	2.69^b	(1.03)	---	---	3.10	(1.01)	---	---
Self-Value Affirm	47.47^{ab}	(17.06)	0.88	0.86	3.73^a	(0.76)	1.19	1.16	4.23^a	(1.18)	1.95	1.01
External Anxiety	47.78^{ab}	(20.10)	0.77	0.75	2.86^b	(0.79)	0.19	0.19	4.15^a	(1.14)	0.96	0.93
Deemphasize	45.15^{abc}	(16.59)	0.74	0.72	3.42^{ac}	(0.90)	0.77	0.75	4.45^a	(0.61)	1.73	1.68
Role Models	42.60^{abc}	(16.65)	0.57	0.56	3.58^a	(0.72)	1.01	0.99	4.02^a	(1.19)	0.86	0.83
Incremental Mindset	39.90^{bc}	(13.65)	0.43	0.42	2.95^{bc}	(1.05)	0.25	0.25	4.72^a	(0.64)	1.98	1.92
Men (Gender Prime)	55.75^a	(20.65)	1.28	1.25	2.65^b	(0.90)	0.04	0.04	4.04^a	(1.05)	0.93	0.90

Note: All means within rows not sharing a subscript are significantly different at *p < .05* using Fisher's least significant difference post hoc test. All *d*'s and *g*'s were computed by comparing the condition to women who did not receive an intervention.

DISCUSSION

<u>Overview</u>

As predicted, any intervention improved performance compared to the women in the no intervention – stereotype threat only condition, but upon closer inspection, the self-value affirmation intervention ($d = 0.88$) resulted in the largest improvement in math performance, and also had the greatest positive impact on the intrinsic motivational experience during the test taking experience. Moreover, the self-value affirmation intervention also improved attitudes toward STEM. Although all interventions improved attitude toward STEM, and other interventions did significantly increase performance (external anxiety attribution), and intrinsic motivation (role models and deemphasizing the self), self-value affirmation is the only intervention which improves both. Yet, as noted in Table 3 (page 63), the self-value affirmation intervention took the longest time to implement, which may imply that if time is limited and assurances cannot be made that the intervention is implemented correctly, other interventions may be easier to carry out (e.g., role models, or external anxiety attributions) and still result in at least modest improvement in either performance or motivation, however, of importance, time had no effect on the significant effects noted for all of our dependent measures. Collectively, these results suggest that in the presence of a mildly explicit stereotype threat (Schmader, 2002), self-value affirmations are the best intervention for improving performance, motivation, and increasing positive attitudes toward STEM, for college aged women who have experienced such a threat in STEM domains.

Some research suggests that stereotype threat effects are stronger among women who are more identified with the domain under study (Lesko & Corpus, 2006; Eriksson & Lindholm, 2007), although other scholars argue that all women are highly susceptible to stereotype threat experiences in STEM domains (Smith & White, 2002; Keller, 2002; 2007). The data from this found that math identification did not differ by condition, suggesting that self-value affirmations may help everyone, not just those who have high domain identification in the presence of a domain specific threat. The findings of this thesis are important at both a theoretical and application level. At a theoretical level, these results suggest that although stereotype threat may be multifaceted in the manner through which it harms performance (Smith, 2004; Schamder et al., 2008), self-value affirmations has the greatest impact on attenuating the threat. As a result, this research implies that understanding the mechanisms which explain how stereotype threat harms self-worth may be paramount in uncovering the cognitive and/or affective underpinnings of the threat's effect (Tesser, 1988; Tesser, 2000). From an application standpoint, this research suggests that policy makers and administrators implement self-value affirmation interventions in order to have the greatest impact on reducing the detriments stereotype threat has on performance, while fostering intrinsic motivation, and creating positive attitudes toward STEM.

The goal of this project was to determine the best way to reduce stereotype threat so that gender diversity within STEM domains could be increased. These results suggest that by implementing self-value affirmation interventions in college classrooms, women will experience the greatest boost in performance and motivation, which could have long

lasting effects (Cohen et al., 2006) that would eventually lead to women pursuing STEM

college majors (Sansone & Smith, 2000), and ultimately seeking out a STEM career

(Correll, 2004). Currently, men outnumber women in STEM (HERI, 2007; NSF, 2009),

and those women who do enter STEM, tend to leave at a higher rate compared to their

male counterparts (Hewlett et al., 2008; Simard et al., 2008). These deficits exist because

stereotype threats, which negatively impact performance and motivation, are embedded

in society. In order to increase the diversity of women in STEM, it is necessary to

implement a self-value affirmation intervention, as our results suggest that it was the only

intervention which attacked all three detriments imposed by stereotype threat. Although

these results recommend that self-value affirmations be used in STEM settings in which

college aged women experience threat, it should be noted that stereotype threat effects are

difficult to replicate, and interventions are only effective when the original

methodologies, which are grounded in psychological theory, are followed (Stoet &

Geary, 2012). This is especially important when considering how incorrectly

implementing self-value affirmations may actually backfire (Jacks & O'Brien, 2004),

making individuals more resistant to change, and increase a person's commitment to the

issue and their identity (Sherman & Cohen, 2006).

The effectiveness of self-value affirmations has a long history of support

throughout psychological literature (Steele, 1988; Major, Spencer, Schmader, Wolfe, &

Crocker, 1998; Sherman & Cohen, 2006), with research suggesting that such an

intervention is not only adaptive (Taylor & Brown, 1988), but good for mental and

physical health (Armor & Taylor, 2002). Self-value affirmation interventions may

produce the largest improvements in both performance and motivation, as it is aimed at altering students' threat/stress responses to perceived threats, implying that it may be psychologically farther downstream compared to other processes which different interventions may rely on to attenuate threat (Martens et al., 2006; Sherman & Cohen, 2006). In addition to improving performance in stereotyped domains (Aronson, Fried, & Good, 2002; Frantz et al., 2004; Schimel et al., 2004; Cohen et al, 2006; Martens et al., 2006; Miyake et al., 2010), research has found self-value affirmation to minimize prejudice to protect self-worth (Crocker & Major, 1989; Adams, Tormala, & O'Brien, 2006), allowing motivation to be fostered.

Although self-value affirmation has the potential to assist individuals across a variety of threatening situations, it is also worth noting that moderating variables and qualifying conditions do exist. For example, culture seems to have an important impact on self-value affirmation, as studies have demonstrated that the intervention has not produced significant differences for Asian cultures (Heine, Lehman, Markus, & Kitayama, 1999; Kitayama, Snibbe, Markus, & Suzuki, 2004; Hoshino-Browne et al., 2005). Additionally, research has found that one way people respond to threats is by affirming the self in an alternative domain, but that those with fewer affirmational resources (i.e. lower self-esteem), and those with insecure views of the self, tended to be less resilient to threat (Steele, Spencer, & Lynch, 1993; Jordan, Spencer, Zanna, Hoshino-Browne, & Cornell, 2003). Yet another important variable which can influence the strength of self-value affirmation is one's identity centrality and salience (Boninger, Krosnick, & Berent, 1995). For example, research has suggested that highlighting one of

several potential identities, such as being a student and being an athlete, can improve task performance (Yopyk & Prentice, 2005), asserting that the more multifaceted an individual is able to consider the self, the easier it is for them to self-affirm (Baysu, Phalet, & Brown, 2011). Given that individuals may vary in the strength of their commitment to a specific source of identity, individuals with less dynamic views of the self may see the world in black and white terms, making them less likely to self-value affirm (Correll, Spencer, Zanna, 2004; Zuwerink & Devine, 1996).

Laboratory Based Interventions in Educational Settings

As noted throughout this thesis, numerous laboratory experiments have inspected the negative impact stereotype threat has on performance. Although research has generally reported that negative stereotypes harm performance (Spencer et al., 1999) whereas positive stereotypes enhance performance (McGlone & Aronson, 2006), many considerations must be made when extracting the methods used in a laboratory setting, and applying them to educational environments. Although it has been 17 years since the first stereotype threat experiment was published (Steel & Aronson, 1995), and nearly 500 articles of stereotype threat have been published since, it is reasonable to ask whether stereotype threat emerges in naturalistic settings, and whether or not interventions can be implemented successfully. The practical relevance of lab-based studies have become a source of debate among scholars (Falk & Heckman, 2009), and admittedly, whether an experimental is artificial or realistic, conducted with college age freshmen or pre-teens, looking at the immediate or long term effects, no single study is able to definitively establish the degree to which a particular intervention is useful in a real world setting.

Early field experiments in stereotype threat have focused on rigorously testing the

practicality of stereotype threat interventions in classrooms (Keller & Dauenheimer,

2003; Kellow & Jones, 2007; Good et al., 2008) and have yielded favorable results,

which map on to findings produced by laboratory experiments. However, when applying

laboratory research to field settings, certain limitations exist.

First, as noted in the methodology of this thesis, participants were run in single

sex groups. This precaution was taken in order to control for environmental factors,

which have been shown to impact stereotype threat (Inzlicht & Ben-Zeev, 2000).

Although having this level of control helps cut down on experimental noise, it is not

always a practical option in educational settings. Research has demonstrated that when

placed in threatening environments (math class), women experience a greater degree of

threat if the majority of the class is comprised of men. This caveat becomes especially

problematic when one considers that field experiments which have tested interventions

also tested the interventions in single sex groups (Huguet & Régner, 2007). Although one

may suggest that policy segregates math related courses by gender, this may create an

additional burden on educational institutions, and especially in domains in which the

gender gap is prominent (Ceci & Williams, 2007), may not even be a possible option.

Therefore, the results from this thesis, which suggest that self-value affirmations are the

best, can only be generalized to single sex environments.

In addition to the gender of the student, the gender of the teacher can be of great

importance. For example, research has shown that when an individual who is

administering a math test is a woman, she might be viewed as a role model, if she has

demonstrated her competence by creating the test (Marx & Roman, 2002). In this study, an attempt was made to assure that the gender of the experimenter was randomly assigned throughout all conditions. However, due to limitations of available experimenters, the data from this thesis was primarily collected by a male experimenter (see Table 3, page 63). This is an important limitation, as this study cannot exclude the possibility that gender of the experimenter may have not only influenced the results, but creates an important practical consideration for interpreting this data in a real world setting. For example, this research suggests that self-value affirmations are the best interventions, when tested in the presence of a male administrator. However, it is unclear whether or not these results would have emerged if the individual administering the test had been a woman. Although it is reasonable to predict that if a female experimenter was present, she may have been viewed as a role model, which could have increased the performance and motivation for any of the interventions (McIntyre et al., 2003), it is important to note that these results cannot attest how men's performance may have been influenced in the presence of a competent woman.

Another important consideration derived from the limitations of this study has to do with the sample tested. Although intervention studies have found similar results of an intervention's effectiveness in a variety of high school and middle school settings (Good et al., 2003; Huguet & Régner, 2007), these results can only suggest that self-value affirmations are only the most effective intervention for college age students. Relative to those in high school, college students tend to demonstrate a higher level of achievement, as competitive admission procedures by accredited universities (Duffy & Goldberg,

1998), coupled with the fact that public school is compulsory until at least the age of 14 (NCES, 2008), suggest that studies which use none college age students may include participants who have a different range of beliefs and educational experiences. Additionally, research suggests that college aged samples tend to have difference socioeconomic ranges that may typically be found among younger samples (Hossler, Schmit, & Vesper, 1999). Given these factors, it is possible that the results of this study, which suggest that self-value affirmations are most effective, may not reign true among younger samples.

Finally, the results of this thesis also suggest that college aged men and women tend to have low intrinsic motivation for STEM domains. Results yielded that only self-value affirmations, providing a role model and deemphasizing group identity threat yield a positive impact toward intrinsic motivation, but given the previously mentioned limitations regarding gender and age, it is possible that younger populations, who receive interventions from women, may differ in their intrinsic motivation toward STEM domains. Although it is argued that there is a national crisis for motivation in STEM courses, research has also suggested that motivation tends to be lower in the lab, relative to naturalistic environments (Schmit & Ryan, 1997). Collectively, very little research has examined motivation, and it is important for future research to not only measure how these interventions may impact this variable, but also, how stereotype threats in the classroom may harm motivation. Field experiments are needed to assess how students naturally interact with threats which emerge in their environment, and whether they are motivated to disconfirm the stereotype (Forbes & Schmader, 2010), withdraw effort as a

means of self-handicapping (Keller, 2002), or decide to avoid the domain altogether (Cheryan et al., 2009). Researchers should aim future efforts at understanding how motivation and future motivational items can be influenced by traditional background factors, such as family income, and state and national standardized exam scores (Walton & Spencer, 2009).

In conclusion, although the results of this study are limited to the specific environment crafted by the experimenter, it is important to consider that experiments, whether they are conducted in the lab or field, share some important similarities. For example, laboratory experiments often test stereotype threat interventions on students, which is often the exact study population of interest (Harrison & List, 2004). Additionally, characteristics of lab experiments which are generally scrutinized also happen to exist as key features in the classrooms and testing centers where many of these field experiments take place (Schmader & Stone, 2008). Although there are many variables to consider, field experiments have attempted to provide guidance for ways in which educators can implement the methodology in the classroom, and begin to put the scientific community's knowledge to practical use. Future studies are needed to ascertain whether or not empirical findings demonstrated in the laboratory can be fully extended to educational settings, and pending the findings, which factors may be of importance in explaining why results may not fully map on to those demonstrated in lab environments.

General Limitations and Future Directions

Although it is the contention of this thesis to interpret the results of this experiment such that self-value affirmation is the best intervention for improving performance, intrinsic motivation, and attitude toward STEM, several limitations exist. First, it should be noted that this intervention was tested within the context of improving immediate performance and motivation for college aged women under stereotype threat related to STEM domains, in a controlled lab setting. This limitation has several important implications. Given that men did not undergo an intervention, it is difficult to say which of these interventions could impact the stereotype lift men experience in the face of a stereotype which favors their performance (Walton & Cohen, 2003). As a result, future research should examine each interventions impact on men to ensure that implementing the intervention will not boost performance of women at the cost of men. Additionally, although this study directly tested the interventions effectiveness for college aged women in STEM, self-value affirmation interventions have been successfully used to improve the performance of racial minorities (Frantz et al., 2004), those in middle school (Cohen et al., 2006), and have even been successfully implemented in classroom environments, showing an improvement across an entire semester (Miyake et al., 2010). This suggests that, although this study cannot directly speak to the interventions relative effectiveness compared to other interventions in different domains, much research reaffirms that self-value affirmations are a generally effective intervention. Future research should test whether there are domain or stereotype

specific factors which may impact the overall effectiveness of an interventions ability to reduce stereotype threat.

Additionally, in this project, the manner of stereotype threat induction was held constant across interventions, but it is certainly the case that the way in which stereotype threat is triggered might reduce/enhance the efficacy of a given intervention. For example, Stone and McWhinnie (2008) found that subtle and blatant threats differ in how they create performance deficits. Subtle threats create a sense of distraction as participants are curious as to whether or not a threat exists, whereas blatant threats are said to prompt participants to adopt strategies which may be more cautious, and subsequently less effective, as participants are afraid of confirming the negative stereotype. As noted in Table 2 (pages 50-51), the design and threat type differed by study. Most studies in our inclusion criteria manipulated stereotype threat using a subtle threat (50%), some studies used blatant (42%), whereas others used mildly explicit (8%). Although mildly explicit was selected as a middle ground between threats, the type of threat which the interventions in this study were tested against could theoretically impact the interventions effectiveness. Future research should explore how interventions may differ in their ability to reduce stereotype threat based on how the threat harms performance, and whether or not these differences in how the threat is delivered impacts intrinsic motivation.

Furthermore, this study selected one intervention from each of the five subtypes. Although the studies within each group tended to have similar methodology, it is important to note that slight differences in methodologies between the subtypes could

influence the strength of the intervention. For example, in providing external anxiety attributions, participants were taught about the dangers of stereotype threat (Johns et al., 2003). However, had anxiety been externalized using different methods, such as providing participants with a black box which ostensibly emits a subliminal tone that could cause anxiety (Ben-Zeev et al., 2005), the observed results may have differed. Although it is difficult to predict which intervention may be most effective, it is necessary to note that these methodological differences could have had some impact on the results. Additionally, had participants been provided with an intervention which combines several of the subcategories (see Good et al., 2003), such as external anxiety attributions (the second shortest intervention, which improved performance) along with received role model information (the shortest intervention, which improve motivation), a compounding effect may have been noted, such that both performance and motivation could have been significantly improved in a fashion similar to self-value affirmations, or even better. As a result, future research should examine how the methodologies of stereotype threat interventions may prove to be more or less effective within each subcategory, as well as how these interventions could be merged to create a greater improvement to performance and motivation.

Finally, the results of this study suggest self-value affirmation is the best intervention at improving intrinsic motivation, but that intrinsic motivation is naturally low for both women under threat, and men. This was surprising, as previous research has suggested that low intrinsic motivation may cause domain disindentification (Crocker & Major, 1989), which could explain why so few women seek out STEM careers (Correll,

2004), and why those women who do choose STEM careers are more likely to leave their jobs relative to men (Seymour & Hewitt, 1997; Bleeker & Jacobs, 2004). However, several possible explanations may account for this finding. First, women in STEM may not experience a reduction in intrinsic motivation until they receive negative feedback about their performance. Research has noted that when women receive feedback affirming the stereotype (i.e. they performed less well than men) intrinsic motivation suffers (Schmader et al., 2004). However, this account seems unlikely given that motivation is closely linked to career and college major pursuit (Sansone & Smith, 2000; Renninger et al., 2004), and until taking the SATs, gender differences in performance do not emerge in STEM domains (Hyde et al., 2008). Additionally, it could be theorized that the men in this study experienced stereotype lift (Walton & Cohen, 2003), which may have acted as feedback which was perceived as an extrinsic reward, subsequently reducing men's intrinsic motivation (Lepper, Green, & Nisbett, 1973). However, very little is known about the motivational underpinnings of stereotype lift (Walton & Cohen, 2003), and feedback, such as receiving a grade on a test, is external information which generally does not undermine intrinsic motivation (Deci & Ryan, 2000). However, regardless if women's intrinsic motivation was higher due to a lack of feedback, or if men's motivation was lowered due to stereotype lift, studies have suggested that American students generally report that they do not enjoying math and science (Ames, 1992; Hu & Kuh, 2002; Legault, Green-Demers, & Pelletier 2006; Fielding-Wells & Makar, 2008). This finding is important because it suggests that intrinsic motivation may be low for everyone, regardless of gender. Future research should examine whether

everyone has naturally low intrinsic motivation for STEM, and whether lift negatively impacts the motivation of men, or women's motivation is only harmed when negative feedback is received.

Conclusion

This research is the first paper to provide empirical evidence which suggests that stereotype threat interventions differ in their effectiveness in curtailing performance and motivational deficits. This novel finding encourages researchers to reevaluate how stereotype threat is combated, by considering efficiency. The findings provide researchers, educators, and policy makers with evidence designed to help select which intervention may be most appropriate to use. This research proposes that self-value affirmation is the best at curtailing negative consequences related to performance and intrinsic motivation for college aged women who face a mildly explicit stereotype threat in STEM domains, but notes that it also takes the longest to implement. Stereotype threat is an insidious construct which has taken root in the fabric of our society. By taking action to reduce this threat, attrition rates will decrease for women, performance will equalize with men, and innovation and economic prosperity will be encouraged. The time to act is now. Women must be motivated to enter STEM domains, and performance detriments caused by stereotype threat must be reduced to increase efficiency in developing new technologies, as well as promoting social justice. Women have been hindered by stereotype threat for decades, and we must now implement a self-value affirmation intervention to ensure that this cycle of injustice stops. By putting the results of this study into practice we can increase diversity in STEM majors and careers.

REFERENCES

AAUW Educational Foundation. (2007). *Behind the pay gap*, by J. G. Dey & C. Hill. Washington, DC: Author.

AAUW Educational Foundation. (2010). *Why so few*, by C. Hill, C. Corbett, & C. Rose. Washington, DC: Author.

Adams, G., Tormala, T. T., & O'Brien, L. T. (2006). The effect of self-affirmation on perceptions of racism. *Journal of Experimental Social Psychology*, *42*, 616-626.

Allport, G. W. (1954). The nature of prejudice. Cambridge, MA: Perseus Books

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, *40*, 401-408. doi: 10.1016/j.jesp.2003.08.003

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. (2001, September). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, *12*, 385–90. doi: 10.1111/1467-9280.00371

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*, 261-271. doi: 10.1037/0022-0663.84.3.261

Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.). *Heuristics and biases: The psychology of intuitive judgment*. (pp. 334-347). New York: Cambridge University Press.

Aronson, J., Jannone, S., McGlone, M., & Johnson-Campbell, T. (2009). The Obama Effect: An experimental test. *Journal of Experimental Social Psychology*, *45*, 957-960. doi: 10.1016/j.jesp.2009.05.006

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the Effects of Stereotype Threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*, 113-125. doi: 10.1006/jesp.2001.1491

Bandura, A. (1977). Social Learning Theory. New York: General Learning Press.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: W.H. Freeman.

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, *49*, 153-224. doi: 10.1159/000094368

Baumeister, R. F., & Bushman, B. J. (2008). *Social psychology and human nature* (1st ed.) Belmont, CA: Thomson Wadsworth.

Bayer Corporation (2010). Bayer Facts of Science Education XIV: Female and Minority Chemists and Chemical Engineers Speak about Diversity and Underrepresentation in STEM. Retrieved June 13, 2011, from http://www.womenscolleges.org/files/u1/BayerFactsofScienceEducation.pdf.

Baysu, G., Phalet, K., & Brown, R. (2011). Dual identity as a two-edged sword: Identity threat and minority school performance. *Social Psychology Quarterly*, *74*, 121 143. doi: 10.1177/0190272511407619

Beaton, A., Tougas, F., Rinfret, N., Huard, N., & Delisle, M. N. (2007). Strength in numbers? Women and mathematics. *European Journal of Psychology of Education*, *22*, 291-306. doi: 10.1007/BF03173427

Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of in-group versus out-group social comparison on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, *36*, 519–530. doi: 10.1006/jesp.2000.1425

Beilock, S. L., Jellison, W. A., Rydell, R. J., McConnell, A. R., & Carr, T. H. (2006). On the causal mechanisms of stereotype threat: Can skills that don't rely heavily on working memory still be threatened? *Personality and Social Psychology Bulletin*, *32*, 1059–1071. doi: 10.1177/0146167206288489

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

Bentley, J. T., & Adamson, R. (2003). *Gender differences in the careers of academic scientists and engineers: A literature review* (NSF 03-322). Arlington, VA: National Science Foundation.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, *41*, 174-181. doi: 10.1016/j.jesp.2003.11.007

Berryman, S.E. (1983). *Who will do science? Trends, and their causes in minority and female representation among holders of advanced degrees in science and mathematics*. A special report. New York, NY: Rockefeller Foundation. (ERIC Document Reproduction Service No. ED245052)

Betz, N.E. & Hackett, G. (2006). Career self-efficacy theory: Back to the future. *Journal of Career Assessment*, *14*, 3-11. doi: 10.1177/1069072705281347

Biernat, M. (1991). Gender stereotypes and the relationship between masculinity and femininity: A developmental analysis. *Journal of Personality and Social Psychology*, *61*, 351-365. doi: 10.1037/0022-3514.61.3.351

Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of in-group versus out-group social comparison on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, *36*, 519–530. doi:10.1006/jesp.2000.1425

Bleeker, M. M. & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology*, *96*, 97-109. doi: 10.1037/0022-0663.96.1.97

Boughn S, Letini A. (1999). Why Do Women Choose Nursing? *Journal of Nursing Education*, *38*, 156-161.

Brainard, S. G., & Carlin, L. (1998). A six-year longitudinal study of undergraduate women in engineering and science. *Journal of Engineering Education*, *87*, 369 75. doi: 10.1109/FIE.1997.644826

Branscombe, N.R., & Smith, E.R. (1990). Gender and racial stereotypes in impression formation and social decision-making processes. *Sex Roles*, *22*, 627-647. doi: 10.1007/BF00288239

Breland, H. M. & Griswold, P. A. (1982). Use of a performance test as a criterion in a differential validity study. *Journal of Educational Psychology*, *74*, 713–721. doi: 10.1037/0022-0663.74.5.713

Bressler, M, & Wendell, P. (1980). The sex composition of selective colleges and gender differences in career aspirations. *Journal of Higher Education*, *51*, 650-663. doi: 10.2307/1981170

Brodish, A. B., & Devine, P. G. (2005). The dynamics of prejudice, stereotyping, and intergroup relations: Intrapersonal and interpersonal processes. *Social Psychological Review*, *7*, 54-70.

Brown, R. P., & Day, E. A. (2006). The difference isn't black and white: Stereotype threat and the race gap on Raven's Advanced Progressive matrices. *Journal of Applied Psychology*, *91*, 979-985. doi: 10.1037/0021-9010.91.4.979

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*, 246-257. doi: 10.1037/0022-3514.76.2.246

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

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, *33*, 267–285. doi: 10.1002/ejsp.145

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? *Psychological Science*, *16*, 572-578. doi: 10.1111/j.0956-7976.2005.01577.x

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, *135*, 218–61. doi:10.1037/a0014412

Ceci, S. J. & Williams, W. M. (2007). *Why aren't more women in science? Top researchers debate the evidence*. Washington, DC: APA Books.

Cheryan, S., Plaut, V. C., Davies, P., & Steele, C. M. (2009). Ambient belonging: How stereotypical environments impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*, 1045-1060. doi: 10.1037/a0016239

Chesler, N.C., Barabino, G.A., Bhatia, S.N., Richards-Kortum, R.R. (2010). The pipeline still leaks and more than you think: A status report on gender diversity in biomedical engineering. *Annals of Biomedical Engineering*, *38*, 1928-1935. doi: 10.1007/s10439-010-9958-9

Cohen, J (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. doi: 10.1037/0033-2909.112.1.155

Cohen, G.L., Garcia, J., Apfel, N., Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307-1310. doi: 10.1126/science.1128317

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin*, *25*, 1302-1318. doi: 10.1177/0146167299258011

Cole, B., Matheson, K., & Anisman, H. (2007). The moderating role of ethnic identity and social support on relations between well-being and academic performance. *Journal of Applied Social Psychology*, *37*, 592-615. doi: 10.1111/j.1559-1816.2007.00176.x

Cooley, C.H. (1902). Human Nature and the Social Order. New York: Scribner's.

Correll, S.J. (2001). Gender and the career choice process: The role of biased self assessments. *American Journal of Sociology*, *106*, 1691–1730. doi: 10.1086/321299

Correll, S.J. (2004). Constraints into preferences: Gender, status and emerging career aspirations. *American Sociological Review*, *69*, 93–133. doi: 10.1177/000312240406900106

Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology*, *40*, 350-356. doi: 10.1016/j.jesp.2003.07.001

Crisp, R. J., Bache, L. M. & Maitner, A. T. (2009). Dynamics of social comparison in counter stereotypic domains: Stereotype boost, not stereotype threat, for women engineering majors. *Social Influence*, *4*, 171-184. doi: 10.1080/15534510802607953

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, *96*, 608-30.

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. Gilbert, S. Fiske, and G. Lindzey (Eds.), *The Handbook of Social Psychology*, Vol. 2 (4th ed.). (pp. 504-553). Boston: McGraw-Hill.

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

Csikszentmihali, M., Rathunde, K., and Whalen, S. (1993). *Talented teenagers: The roots of success and failure*. New York: Cambridge University Press.

Cullen, M.J., Waters, S. D., and Sackett, P. R. (2006). Testing stereotype threat theory predictions for math majors and non-majors by gender. *Human Performance*, *19*, 421-440.

Danaher, K., & Crandall, C. S. (2008). Stereotype threat in applied settings re-examined. *Journal of Applied Social Psychology*, *38*, 1639-1655. doi: 10.1111/j.1559-1816.2008.00362.x

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science*, *314*, 435. doi: 10.1126/science.1131100

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. doi: 10.1177/014616702237644

Davis, D.E., & Austin, H.S. (1990). "Life cycle, career patterns and gender stratification in academe: Breaking myths and exposing truths", in Lie, S.S. and O'Leary, V.E. (eds.), *Storming the tower: Women in the academic world*, 89-107, New York, Nichols/GP Publishing.

Deci, E.L, & Ryan, R.M. (2000). The what and why of goal pursuits: Human needs and the self determination of behavior. *Psychological Inquiry*, *11*, 227-268. doi: 10.1207/S15327965PLI1104_01

Derks, B., Inzlicht, M., & Kang, S. (2008). The neuroscience of stigma and stereotype threat. *Group Processes & Intergroup Relations*, *11*, 163-181. doi: 10.1177/1368430207088036

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

Dweck, C., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–73. doi: 10.1037/0033-295X.95.2.256

Dweck, C. S., & Sorich, L. (1999). Mastery-oriented thinking. In C. R. Snyder (ed.), *Coping* (pp. 232-251). New York: Oxford University Press.

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. doi: 10.1037/0022-3514.46.4.735

Eagly, A.H., & Wood, W. (1991). Explaining sex differences in social behavior: A social role perspective. *Personality and Social Psychology Bulletin*, *17*, 306-315.

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, *18*, 585–609. doi: 10.1111/j.1471-6402.1994.tb01049.x

Eccles (Parsons), J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Perspective on achievement and achievement motivation* (pp. 75–146). San Francisco: W. H. Freeman.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, *46*, 186-201. doi: 10.1111/j.1540-4560.1990.tb01929.x

Eccles, J.S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Ed.), *Handbook of child psychology: Volume 3 - Social, emotional, and personality development* (5th ed.). New York: Wiley.

Eisenhart, M. (2008). *We can't get there from here: High school girls consider engineering*. Presentation for a Women in Engineering ProActive Network (WEPAN) national webcast.

Elliot, A.J., Murayama, K., & Pekrun, R. (2011). A 3 x 2 achievement goal model. *Journal of Educational Psychology*, *103*, 632-648. doi: 10.1037/0022-3514.80.3.501

Engle, R.A. and Conant, F.R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners' classroom. *Cognition and Instruction*, *20*, 399-483. doi: 10.1207/S1532690XCI2004_1

Falk, A., & Heckman, J.J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326, 535-538. doi: 10.1126/science.1168244

Fielding-Wells, J. & Makar. K. (2008). *Using mathematical inquiry to engage student learning within the overall curriculum*. Paper presented at the 11th International Congress of Mathematics Education, Monterrey, Mexico.

Fenigstein, A., Seheier, M. E, & Buss, A. H. (1975). Public and private self consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*, 522-527. doi: 10.1037/h0076760

Fiske, S. T., & Stevens, L. E. (1993). What's so special about sex? Gender stereotyping and discrimination. In S. Oskamp & M. Costanza (Eds.), *Gender issues in contemporary society* (Vol. 6., pp. 173–196). Newbury Park, CA: Sage.

Fox, M.F. (2001). Women, science, and academic: Graduate education and careers. *Gender and Society*, *15*, 654-666. doi: 10.1177/089124301015005002

Forbes, C.E., & Schmader, T. (2010). Retraining implicit attitudes and stereotypes to distinguish motivation from performance in a stereotype threatening domain. *Journal of Personality and Social Psychology*, *99*, 740-754. doi: 10.1037/a0020971

Fort, D.C. (2005). *A hand up: Women mentoring women in science*. Washington, DC: The Association for Women in Science.

Foschi, M. (1996). Double standards in the evaluation of men and women. *Social Psychology Quarterly*, *59*, 237–54. doi: 10.2307/2787021

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

Ginther, D. K., & Kahn, S. (2006). *Does science promote women? Evidence from academia 1973–2001* (NBER Working Paper W12691). Cambridge, MA: National Bureau of Economic Research.

Glazer-Raymo, J. (1999, 2001). *Shattering the myths: Women in academe*. Baltimore: Johns Hopkins University Press.

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

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

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, *29*, 17–28. doi: 10.1016/j.appdev.2007.10.004

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology*, *24*, 645–62. doi: 10.1016/j.appdev.2003.09.002

Good, C., Rattan, A., & Dweck, C. S. (2009). Why do women opt out? Sense of belonging and women's representation in mathematics. Unpublished paper, Baruch College, Stanford University.

Gresky, D. M., Ten Eyck, L. L., Lord, C. G., & McIntyre, R. B. (2005). Effects of salient multiple identities on women's performance under mathematics stereotype threat. *Sex Roles*, *53*, 703-716. doi: 10.1007/s11199-005-7735-2

Hall, J. A., & Schmid Mast M. (2008). Are women always more interpersonally sensitive than men? Impact of goals and content domain. *Personality and Social Psychology Bulletin*, *34*, 144-155. doi: 10.1177/0146167207309192

Harrison, G.W., & List, J.A. (2004). Field experiments. *Journal of Economic Literature*, 42, 1009-1055.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley & Sons.

Heine, S. J., Lehman, D. R., Markus, H., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, *106*, 766-794.

Hoshino-Browne, E., Zanna, A., Spencer, S., Zanna, M., Kitayama, S., & Lackenbauer, S. (2005). On the cultural guises of cognitive dissonance: The case of Easterners and Westerners. Journal of Personality and Social Psychology, 89, 294-310.

Hess, T. M., Hinson, J. T. (2006). Age-related variation in the influences of aging stereotypes on memory in adulthood. *Psychology and Aging*, *21*, 621-625. doi: 10.1037/0882-7974.21.3.621

Hewitt, N.M., & Seymour, E. (1991). Factors contributing to high attrition rates among science, mathematics, and engineering undergraduate majors. Boulder, CO: University of Colorado.

Hewlett, S. A., Buck Luce, C., Servon, L. J., Sherbin, L., Shiller, P., Sosnovich, E., & Sumberg, K. (2008). *The Athena Factor: Reversing the brain drain in science, engineering and technology*, Harvard Business Review Research Report. Boston: Harvard Business Publishing.

Higher Education Research Institute. (2001; 2007; 2010, January). *Survey of the American freshman: Special tabulations*. Los Angeles, CA: Author.

Hossler, D., Schmit, J., & Vesper, N. (1999). *Going to college: How social, economic, and educational factors influence the decisions students make*. Johns Hopkins University Press, Baltimore, MD.

House, J. D. and E. J. Keeley (1995). Differential prediction of adult student performance from miller analogies test scores. *Journal of Genetic Psychology*, *157*, 501-504. doi: 10.1080/00221325.1996.9914882

Hu, S., and Kuh, G. D. (2002). Being (dis)engaged in educationally purposeful activities: The influences of student and institutional characteristics. *Research in Higher Education*, *43*, 555–575.

Huguet, P. & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, *99*, 545-560. doi: 10.1037/0022-0663.99.3.545

Hyde, J. S., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*, 139–155. doi: 10.1037/0033-2909.107.2.139

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect. *Psychology of Women Quarterly*, *14*, 299–324. doi: 10.1111/j.1471-6402.1990.tb00022.x

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, *11*, 365–71. doi: 10.1111/1467-9280.00272

Inzlicht, M. & Schmader, T. (2011). Stereotype threat: Theory, process, and application. New York: Oxford University Press.

Jacks, J. Z., & O'Brien, M. E. (2004). Decreasing resistance by affirming the self. In Knowles, E. S. & Linn, J. A. (Eds.), *Resistance and persuasion* (pp. 235-257). Mahwah, NJ: Lawrence Erlbaum Associates.

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, *93*, 544-564. doi: 10.1037/0022-3514.93.4.544

Javidan, M., Bemmel, B., Devine, K.S., & Dastmalchian, A. (1995) Superior and subordinate gender and the acceptance of superiors as role models. *Human Relations*, *48*, 1271-1284. doi: 10.1177/001872679504801102

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*, 691-705. doi: 10.1037/a0013834

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, *16*, 175–79. doi: 10.1111/j.0956-7976.2005.00799.x

Jordan, C. H, Spencer, S. J., & Zanna, M. P. (2003). "I love me...I love me not:" Implicit self-esteem, explicit self-esteem, and defensiveness. In S. J. Spencer et al. (Eds.), *Motivated social perception: The Ontario symposium* (Vol. 9, pp. 117-145). Mahwah, NJ: Lawrence Erlbaum Associates.

Kaczmarek, M., and Franco, J.N. (1986). Sex Difference in Prediction of academic Performance by the Graduate Record Examination. *Psychological Reports*, *59*, 1197-1198. doi: 10.2466/pr0.1986.59.3.1197

Kaiser, C. R., & Hagiwara, N. (2011). Gender identification moderates social identity threat effects on working memory. *Psychology of Women Quarterly*, *35*, 243-251. doi:10.1177/0361684310384102

Keifer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*, 825-832. doi: 10.1016/j.jesp.2006.08.004

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. doi: 10.1023/A:1021003307511

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology*, *77*, 323-338. doi: 10.1348/000709906X113662

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. doi: 10.1177/0146167202250218

Kelley, H.H. (1967). Attribution Theory in Social Psychology. *Nebraska Symposium on Motivation*, *15*, 192-238.

Kellow, T. J., & Jones, B. D. (2007). The effects of stereotypes on the achievement gap: Reexamining the academic performance of African American high school students. *Journal of Black Psychology*, *34*, 94-120. doi: 10.1177/0095798407310537

Kitayama, S., Snibbe, A.C., Markus, H.R., & Suzuki, T. (2004). Is there any "free" choice? Self and dissonance in two cultures. *Psychological Science*, *15*, 527-533. doi: 10.2307/40064011

Koch, S. C., Müller, S. M., & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. Computers & Education, 51, 1795 1803. doi: 10.1016/j.compedu.2008.05.007

LeFevre, J., Kulak, A. G., & Heymans, S. L. (1992). Factors influencing the selection of university majors varying in mathematical content. *Canadian Journal of Behavioural Science*, *24*, 276-289. doi: 10.1037/h0078742

Legault, L., Green-Demers, I., & Pelletier, L.G. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of educational psychology*, *98*, 567-582.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, *28*, 129–137.

Lesko, A. C., & Corpus, J. H. (2006). Discounting the difficult: How high math identified women respond to stereotype threat. *Sex Roles*, *54*, 113-125. doi: 10.1007/s11199-005-8873-2

Levine, D.U., & Ornstein, A.C. (1983). Sex differences in ability and achievement. *Journal of Research and Development in Education*, *16*, 62-72.

Lipstein, R. & Renninger, K. A. (2006). "Putting things into words": The development of 12-15-year-old students' interest for writing. In Boscolo, P. & Hidi, S. (Eds.), *Motivation and writing: Research and School Practice* (pp. 113- 140). New York: Kluwer Academic/Plenum.

Logel, C., Walton, G.M., Spencer, S.J., Iserman, E.C., von Hippel, W., Bell, A.E. (2009). Interacting with sexist men triggers social identity threat among female engineers. Journal of Personality and Social Psychology, 96, 1089-1103. doi: 10.1037/a0015703

Lovaglia, M. J., Lucas, J. W., Houser, J. A., Thye, S. R., & Markovsky, B. (1998). Status processes and mental ability test scores. *American Journal of Sociology*, *104*, 195–228. doi: 10.1086/210006

Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math-science pipeline. *Current Directions in Psychological Science*, *1*, 61–66. doi: 10.2307/20182131

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Major, B., Spencer, S. J., Schmader, T., Wolfe, C. T., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, *24*, 34-50. doi: 10.1177/0146167298241003

Margolis, J., Fisher, A., & Miller, F. (2002). *Caring about connections: Gender and computing*. Pittsburgh, PA: Carnegie Mellon University, School of Computer Science.

Martens, A., Johns, M., Greenberg, J., & Schimel (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, *42*, 236-243. doi: 10.1016/j.jesp.2005.04.010

Malpass, J. R., O'Neil, H. F., & Hocevar, D. (1999). Self-regulation, goal orientation, self efficacy, worry, and high-stakes math achievement for mathematically gifted high school students. *Roeper Review*, *21*, 281-228. doi: 10.1080/02783199909553976

Marx, D. M., Ko, S. J., & Friedman, R. A. (2009). The "Obama Effect": How a salient role model reduces race-based performance differences. *Journal of Experimental Social Psychology*, *45*, 953-956. doi: 10.1016/j.jesp.2009.03.012

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*, 1183-1193. doi: 10.1177/01461672022812004

Marx, D. M., Stapel, D. A., & Muller, D. (2005). We can do it: The interplay of construal orientation and social comparisons under threat. *Journal of Personality and Social Psychology*, *88*, 432-446. doi: 10.1037/0022-3514.88.3.432

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. doi: 10.1207/S15327043HUP1603_2

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*, *27*, 486-493. doi: 10.1016/j.appdev.2006.06.003

McIntyre, R. B., Lord, C. G., Gresky, D. M., Ten Eyck, L. L., Frye, G. D. J., & Bond, C. F., Jr. (2005). A social impact trend in the effects of role models on alleviating women's mathematics stereotype threat. *Current Research in Social Psychology*, *10*, 116–36.

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*, 83–90. doi: 10.1016/S0022-1031(02)00513-9

Meece, J. L., Eccles, J. S., Kaczala, C. M., Goff, S. B., & Futterman, R. (1982). Sex differences in math achievement: Towards a model of academic choice. *Psychological Bulletin*, *91*, 324–348. doi: 10.1037/0033-2909.91.2.324

Miyake, A., Kost-Smith, L.E., Finkelstein, N.D., Pollock, S.J., Cohen, G.L., & Ito, T.A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science*, *330*, 1234-1237.

Morgan, S. L., & Mehta, J. D. (2004). Beyond the laboratory: Evaluating the survey evidence for the disidentification explanation of black-white differences in achievement. *Sociology of Education*, *77*, 82–101. doi: 10.1177/003804070407700104

Mueller, C. M., and Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*, 33-52. doi: 10.1037/0022-3514.75.1.33

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. doi: 10.1111/j.1467-9280.2007.01995.x

National Academy of Sciences. (2007). *Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering*. Washington, DC: National Academies Press. National Association of Colleges and Employers (2009). Salary survey.

Nettles, M. T., & Millett, C. M. (2006). *Three magic letters: Getting to Ph.D*. Baltimore: The Johns Hopkins University Press.

National Association of Colleges and Employers. (2009, Fall). *Salary survey*. Author.

National Research Council. (2009). *Gender differences at critical transitions in the careers of science, engineering and mathematics faculty*. Washington, DC: National Academies Press.

National Science Foundation. Division of Science Resources Statistics. (2008). Science and engineering degrees: 1966–2006 (Detailed Statistical Tables) (NSF 08-321). Arlington, VA: Author. Retrieved August 22, 2011, from www.nsf.gov/statistics/nsf08321/pdf/nsf08321.pdf.

National Science Foundation. Division of Science Resources Statistics. (2009). Women, minorities, and persons with disabilities in science and engineering: 2009 (NSF 09-305). Arlington, VA: Author. Retrieved June 23, 2011, from www.nsf.gov/statistics/wmpd.

Nguyen, H.-H. H., & Ryan, A. M. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*, 1314–34. doi: 10.1037/a0012702

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59. doi: 10.1037/0022-3514.83.1.44

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. doi: 10.1177/0146167203029006010

O'Leary, V.E., & Mitchell, J.M. (1990). Women connecting with women: Networks and mentors. In *Storming the tower: Women in the academic world*, ed. Suzanne S. Lie and Virginia E. O'Leary, 58-73. New York: Nicholas Publishing.

Osborne, J. W. (1995). Academics, self-esteem, and race: A look at the assumptions underlying the disindentification hypothesis. *Personality and Social Psychology Bulletin*, *21*, 449-455. doi: 10.1177/0146167295215003

Osborne, J. W., & Walker, C. (2006). Stereotype threat, identification with academics, and withdrawal from school: Why the most successful students of colour might be the most likely to withdraw. *Educational Psychology*, *26*, 563-577. doi: 10.1080/01443410500342518

Phillips, S. D., & Imhoff, A. R. (1997). Career development and counseling for women. *Annual Review of Psychology*, *48*, 31-59.

Plant, E. A., Baylor, A. L., Doerr, C. E., & Rosenberg-Kima, R. B. (2009). Changing middle school students' attitudes and performance regarding engineering with computer based social models. Computers and Education, 53, 209–215. doi: 10.1016/j.compedu.2009.01.013

Post, P., Stewart, M.A., & Smith, P.L. (1991). Self-efficacy, interest, and consideration of math / science and non-math/science occupations among black freshmen. *Journal of Vocational Behavior*, *38*, 179-186. doi: 10.1016/0001-8791(91)90025-H

Pronin, E., Steele, C., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, *40*, 152-168. doi: 10.1016/S0022-1031(03)00088-X

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. doi: 10.1177/0146167204264088

Quinn, D. M., & Spencer, S. J. (2001). The Interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57*, 55-71. doi: 10.1111/0022-4537.00201

Ramirez, G., & Beilock, S.L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science*, *331*, 211-213. doi: 10.1126/science.1199427

Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1994). *Student group diffences in predicting college grades: Sex, language, and ethnic group*s (College Board Report No. 93-1, ETS No. 94.27). New York: College Entrance Examination Board.

Régner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Monteil, J. M.m & Huguet, P. (2010). Individual differences in working memory moderate stereotype threat effects. *Psychological Science*, *21*, 1646-1648. doi: 10.1177/0956797610386619

Reis, S. M.,&Park, S. (2001). Gender differences in high-achieving students in math and science. *Journal for the Education of the Gifted*, *25*, 52–73. doi: 10.4219/jeg -2001-212

Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone and J. M. Harackiewicz (Eds.) *Intrinsic motivation: Controversies and new directions* (pp. 373-404). San Diego, CA: Academic Press.

Renninger, K. A., Ewen, E., & Lasher, A. K. (2002). Individual interest as context in expository text and mathematical word problems. *Learning and Instruction*, *12*, 467-491. doi: 10.1016/S0959-4752(01)00012-3

Renninger, K. A., & Hidi, S. (2002). Interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation* (pp. 173-195). New York: Academic Press.

Renninger, K. A., Sansone, C., & Smith, J. L. (2004). Love of learning. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification* (pp. 161-179). New York: Oxford University Press.

Reskin, B.F. (1978). Sex differentiation and the social organisation of science. In J Gaston (ed.) *The Sociology of science*, Jossey Bass, San Francisco.

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educational Psychology Review*, *12*, 85–110.

Rose, N. (1996). *Inventing our selves: Psychology, power, and personhood*. Cambridge, University Press.

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

Rosenthal, H. E. S., Crisp, R. J., & Suen, M.-W. (2007). Improving performance expectancies in stereotypic domains: Task relevance and the reduction of stereotype threat. *European Journal of Social Psychology*, *37*, 586-597. doi: 10.1002/ejsp.379

Ryan, K. E., & Ryan, A. M. (2005). Psychological processes underlying stereotype threat and standardized math test performance. *Educational Psychologist*, *40*, 53-63.

Santos, A., Meyer-Lindenberg, A., Deruelle, C. (2010). Absence of racial, but not gender, stereotyping in Williams syndrome children. *Current Biology*, *20*, R307–R308. doi:10.1016/j.cub.2010.02.009

Sansone, C., & Smith, J.L. (2000). The "how" of goal pursuit: Interest and self regulation. *Psychology Inquiry*, *11*, 306-309. doi: 10.2307/1449627

Sansone, C., Wiebe, D. J., & Morgan, C. (1999). Self-regulating interest: The moderating role of hardiness and conscientiousness. *Journal of Personality*, *61*, 701-733. doi: 10.1111/1467-6494.00070

Sawyer, R. (1986). Using demographic subgroup and dummy variable equations to predict college freshmen grade average. Journal of Educational Measurement, 23, 131-145. doi:10.1111/j.1745-3984.1986.tb00239.x

Sawyer, T. P. Jr., & Hollis-Sawyer, L. A. (2005). Predicting stereotype threat, anxiety, and cognitive ability test performance: An examination of three models. *International Journal of Testing*, *5*, 225-246. doi: 10.1207/s15327574ijt0503_3

Scheier, M. F., & Carver, C. S. (1985). The Self-Consciousness Scale: A revised version for use with general populations. *Journal of Applied Social Psychology*, *15*, 687-699. doi: 10.1111/j.1559-1816.1985.tb02268.x

Schimel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefits of affirming the intrinsic (vs. extrinsic) self. *Social Cognition*, *22*, 75-99. doi: 10.1521/soco.22.1.75.30984

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38*, 194 201. doi: 10.1006/jesp.2001.1500

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

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

Schmader, T., & Stone, J. (2008). Toward a problem-focused approach to prejudice, *Psychological Inquiry*, *19*, 108-113. doi: 10.1080/10478400802050165

Schmit, M.J., & Ryan, A.M. (1997). Applicant withdrawal: The role of test-taking attitudes and racial differences. *Personnel Psychology*, *50*, 855-876.

Schroeder, D.S. & Mynatt, C.R. (1993). Female graduate students' perceptions of their interactions with male and female major professors. *Journal of Higher Education*, *64*, 555-573. doi: 10.2307/2959993

Sekaquaptewa, D., & Thompson, M. (2002). The differential effects of solo status on members of high- and low-status groups. *Personality and Social Psychology Bulletin*, *28*, 694-707. doi: 10.1177/0146167202288013

Seymour, E. (1995). The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account. *Science Education*, *79*, 437–473. doi: 10.1002/sce.3730790406

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Shapiro, J. S., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, *11*, 107-130. doi: 10.1177/1088868306294790

Shapiro, J.R., & Williams, A.M. (2011). The role of stereotype threats in undermining girl's and women's performance and interest in STEM fields. Sex Roles, 66, 175-183. doi: 10.1007/s11199-011-0051-0

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

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*, 80-83. doi: 10.1111/1467-9280.00111

Shih, M., Bonam, C., Sanchez, D., & Peck, C. (2007). The social construction of race: Biracial identity and vulnerability to stereotypes. *Cultural Diversity and Ethnic Minority Psychology*, *13*, 125-133. doi: 10.1037/1099-9809.13.2.125

Simard, C., Henderson, A. D., Gilmartin, S. K., Schiebinger, L., & Whitney, T. (2008). *Climbing the technical ladder: Obstacles and solutions for mid-level women in technology*. Stanford, CA: Michelle R. Clayman Institute for Gender Research, Stanford University, & Anita Borg Institute for Women and Technology.

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. doi: 10.1006/jesp.1998.1373

Smeby, J. (2000). Same-gender relationships in graduate supervision. *Higher Education*, *40*, 53-67. doi: 10.1023/A:1004040911469

Smith, J.L, Johnson, C. (2006). A stereotype boost or choking under pressure? Positive gender stereotypes and men who are low in domain identification. *Basic and Applied Social Psychology*, *28*, 51-63. doi: 10.1207/s15324834basp2801_5

Smith, J. L. & Lewis, K. L. (2009). Men's interpersonal (mis)perception: Fitting in with gender norms following social rejection. *Sex Roles*, *61*, 252-264. doi: 10.1007/s11199-009-9621-9

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (revision invited at Personality and Social Psychology Bulletin). When trying hard isn't natural: Women's motivation for male-dominated fields as a function of effort expenditure concerns.

Smith, J.L ., Morgan, C.L, & White, P.H. (2005). Investigating a measure of computer technology domain identification: A tool for understanding gender-differences and stereotypes. *Educational and Psychological Measurement*, *65*, 336-355. doi: 10.1177/0013164404272486

Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and achievement motivation. *Journal of Educational Psychology*, *99*, 99-114. doi: 10.1037/0022-0663.99.1.99

Smith, J. L., & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles*, *47*, 179-191. doi: 10.1023/A:1021051223441

Stake, J.E. & Mares, K.R. (2001). Science enrichment programs for gifted high school girls and boys: Predictors of program impact on science confidence and motivation. *Journal of Research in Science Teaching*, *38*, 1065-1088. doi: 10.1002/tea.10001

Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. Journal of Personality and Social Psychology, 75, 1191-1197. doi: 10.1037/0022-3514.75.5.1191

Stangor, C., & Sechrist, G. B. (1998). Conceptualizing the determinants of academic choice and task performance across social groups. In J. K. Swim and C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 105-124). San Diego: Academic Press.

Steele, C.M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowtiz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261-302). New York: Academic Press.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–29. doi: 10.1037/0003-066X.52.6.613

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. doi: 10.1037/0022-3514.69.5.797

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*, 5-19. doi: 10.1037/0022-3514.45.1.5

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology*, *64*, 885-896. doi: 10.1037/0022-3514.64.6.885

Stoet, G., & Geary, D.C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology*, *16*, 93-102. doi: 10.1037/a0026617

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

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology*, *77*, 1213-1227. doi: 10.1037/0022-3514.77.6.1213

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. doi: 10.1016/j.jesp.2007.02.006

Stout, J.G., Dasgupta, N., Hunsinger, M., McManus, M.A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*, 255-270. doi: 10.1037/a0021385

Strenta, A.C., Elliot, R., Adair, R., Scott, J., & Matier, M. (1993). Choosing and leaving science in highly selective institutions. Unpublished report to the Alfred P. Sloan Foundation.

Stricker, L. J., & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, *34*, 665-693. doi: 10.1111/j.1559-1816.2004.tb02564.x

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and Things, Women and People: A Meta-analysis of Sex Differences in Interests. *Psychological Bulletin*, *135*, 859-894. doi: 10.1037/a0017364

Swim, J.K. (1994). Perceived versus meta-analytic effect sizes: An assessment of accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21-36. doi: 10.1037/0022-3514.66.1.21

Tajfel, H. & Turner, J. C. (1986). The social identity theory of inter-group behavior. In S. Worchel & L. W. Austin (Eds.), *Psychology of Intergroup Relations*. Chicago: Nelson Hall.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*,193–210.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 181-227). San Diego, CA: Academic Press.

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, *4*, 290-299.

Thoman, D. B., Smith, J. L., & Silvia, P. J. (2011). Interest and ego depletion: demonstrating the resource replenishment function of interest. Social *Psychological and Personality Science* (*online first*). doi: 10.1177/1948550611402521

Thoman, D. B., White, P. H., Yamawaki, N., & Koishi, H. (2008). Variations of gender math stereotype content affect women's vulnerability to stereotype threat. *Sex Roles*, *58*, 702-712. doi: 10.1007/s11199-008-9390-x

Turner, S. L., & Lapan, R. T. (2005). Evaluation of an intervention to increase non traditional career interests and career-related self-efficacy among middle-school adolescents. *Journal of Vocational Behavior*, *66*, 516–31. doi: 10.1016/j.jvb.2004.02.005

Wald, C., & Wu, C. (2010). Of mice and women: The bias in animal models. *Science*, *26*, 1571-1572. 10.1126/science.327.5973.1571

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456–467. doi: 10.1016/S0022-1031(03)00019-2

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*, 82-96. doi: 10.1037/0022-3514.92.1.82

Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*, 1132–1139. doi: 10.1111/j.1467-9280.2009.02417.x

Wainer, H. & Steinberg, L. S. (1992). Sex differences in performance on the mathematics section of the scholastic aptitude test: A bidirectional validity study. *Harvard Educational Review*, *62*, 323-36

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

Whitley, B. E., & Kite, M. E. (2010). The psychology of prejudice and discrimination. Belmont, CA: Wadsworth.

Wilkinson, L., (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594-604. doi:10.1037/0003-066X.54.8.594

Wood, J.V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*, 231-248. doi: 10.1037/0033-2909.106.2.231

U.S. Department of Education. National Center for Education Statistics. (2000). Entry and persistence of women and minorities in college science and engineering education (NCES 2000-601), by G. Huang, N. Taddese, & E. Walter. Washington, DC: Author.

U.S. Department of Education. National Center for Education Statistics. (2007). *The nation's report card: America's high school graduates*. Results from the 2005 NAEP high school transcript study, by C. Shettle, S. Roey, J. Mordica, R. Perkins, C. Nord, J. Teodorovic, J. Brown, M. Lyons, C. Averett, & D. Kastberg. (NCES 2007-467). Washington, DC: Government Printing Office.

U.S. Department of Education. National Center for Education Statistics. (2008). *The Condition of education*. Results from June 2008, by M. Planty, W. Hussar, T. Snyder, S. Provasnik, G. Kena, R. Dinkes, A. KewalRamani, & J. Klemp. (NCES 2008-031). Washington, DC: Government Printing Office.

U.S. Department of Labor. Bureau of Labor Statistics. (2009). *Women in the labor force: A databook* (Report 1018). Washington, DC: Author.

Vogt, C. M., Hocevar, D., & Hagedorn, L. S. (2007). A social cognitive construct validation: Determining women's and men's success in engineering programs. *The Journal of Higher Education*, *78*, 337–64. doi: 10.2307/4501213

Yeager, D.S. & Walton, G. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*, 267-301. doi: 10.3102/0034654311405999

Yopyk, D. J. A., & 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. doi: 10.1207/s15324834basp2704_5

Zuwerink, J. R., & Devine, P. G. (1996). Attitude importance and resistance to persuasion: It's not just the thought that counts. *Journal of Personality and Social Psychology*, *70*, 931-944.