THE EXPERIENCE OF SHIFTING STANDARDS FOR WOMEN ATHLETES: CONSEQUENCES OF STEREOTYPED FEEDBACK

by

Jill Marie Wagaman

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

of

Master of Science

in

Applied Psychology

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2009

APPROVAL

of a thesis submitted by

Jill Marie Wagaman

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 Division of Graduate Education.

Dr. Jessi L. Smith

Approved for the Department of Psychology

Dr. Richard Block

Approved for the Division of Graduate Education

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Jill Marie Wagaman

April 2009

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

Table Page

ABSTRACT

Women athletes were recruited to investigate the experience of 'shifting standards.' Shifting standards occurs when people use stereotypes (e.g. gender and athleticism) to make a relative judgment about performance or behavior. Whereas past research has examined shifting standards from the perceiver's perspective, the current project investigates shifting standards from the target's perspective. To inform our hypotheses, we relied on stereotype threat literature (Stone et al., 1999) and the motivational model for stereotyped tasks (Smith, Sansone & White, 2007). Using athletics as the stereotyped domain, Study 1 demonstrated that the masculine nature of the domain was important in predicting reactions to shifting standards feedback. In addition, there was a positive relationship between stigma consciousness and domain identification. In Study 2, college students and community participants were recruited using a pre-screen questionnaire assessing domain identification and primary sport played. As a result, self-identified women athletes (n = 77, 15.6% community) were blocked on the masculine nature of the sport, resulting in a 2 (type of sport: masculine, non-masculine) x 3 (feedback: positive, shifting standards, no feedback) between-subjects design. After first engaging in an ambiguous athletic test, participants received the feedback manipulation. Participants were presented with a "word puzzle" and completed a measure of gender stereotype activation. Then, a second unambiguous test was administered to measure performance, and participants completed post-test measures assessing interest, future motivation, and self-esteem. The athletic tasks were ostensibly assessing the sport each woman most identified with, but all participants actually received the same two tests. The proposed relationship between feedback, performance, and motivation was unaffected by type of sport. Explanations are provided for why this relationship was not supported. Instead, stigma consciousness moderated the effect of feedback on performance and motivation. Women high in stigma consciousness receiving shifting standards feedback showed high gender stereotype activation, low performance, and low interest compared to women high in stigma consciousness receiving positive feedback. These results were not due to participants' self-esteem or level of commitment to the test. Theoretical and practical implications for the experience of shifting standards are discussed.

INTRODUCTION

Lisa Leslie and LeBron James are both exceptional basketball players who excel in their sport; both players won a gold medal at the 2008 Summer Olympics. However, when trying to judge their basketball ability, the standard for what is good ability for a woman shifts. Lisa Leslie is good at basketball "for a girl," but she is not good compared to LeBron James. This illustrates that gender prejudice toward women continues to be an issue in the United States, but in much more subtle ways than in the past (Swim, Aikin, Hall & Hunter, 1995; Paul & Smith, 2008). Using a shifting standard is one example of subtle, modern form of prejudice.

Forms of Prejudice

Much research has been conducted on when and how people will be biased toward a group of others (Biernat, 2003; Biernat & Manis, 1994; Biernat, Manis & Nelson, 1991; Biernat & Kobrynowicz, 1999; Kobrynowicz & Biernat, 1997). Prejudice is the negative attitude held about an individual based on their group membership, and discrimination refers to the individuals' actions that result from the prejudiced attitudes (Eagly & Diekman, 2005). Gender is one such avenue that a perceiver may use to make stereotyped judgments. A perceiver is the person who evaluates a group who he or she believes to be devalued. To better understand what forms prejudice regarding gender can take, consider the following discussion on sexism. Sexism is an attitude about someone based on their sex alone and can take three different forms: hostile, ambivalent, and benevolent sexism (Glick & Fiske, 1996; Glick & Fiske, 2001). Hostile sexism is

characterized by negative beliefs about and negative emotional responses to women. Someone high in hostile sexism views women as opponents of men (i.e. women seek power by gaining control over men). In contrast, benevolent sexism is characterized by positive beliefs about and positive emotional responses to women. Someone high in benevolent sexism views women as weak, fragile, and to be cherished. Someone high in both hostile and benevolent sexism is termed "ambivalent sexism," which is characterized by both positive and negative emotional responses (i.e. Women should be cherished and protected, but they manipulate men into behaving this way). However, all three forms of sexism can predict the perceiver treating someone differently.

Recall that prejudicial attitudes about a group results in discrimination when the prejudice is acted upon. For example, discrimination based on one's gender is sometimes evident in the hiring process. In Phelan, Moss-Racusin and Rudman (2008), backlash effects toward agentic women (i.e. women who are independent and ambitious) were investigated. Backlash effects are negative consequences for women who possess leadership qualities, but suffer social and economic outcomes (Rudman, 1998). In this study, confederates served as lab-manager candidates who were interviewed by participants who played the interviewer role. The interviewee was either a man or a woman and followed a communal or agentic interview style. The communal interview followed a traditional women's gender role, whereas the agentic interview did not. Participants then rated the interviewees on competence, social skills, and hireability. In the agentic condition, women were rated lower in social skills than men. Women were also rated as less hireable than men in the agentic condition. Results suggested that

different aspects of hiring criteria were emphasized for men and women. For men, competence and social skills ratings carried equal weight, but for women social skills were emphasized more than competence. Women who expressed agency in their interview were the victims of backlash effects and were also judged against standards that discounted their positive qualities. This suggests that discrimination occurred against agentic women because they were less likely to be hired than communal women.

Further research has examined how career advancement differs for men and women who reach management positions. In one study, men and women executives from the same company were matched by the human resources department on age and career success (i.e. compensation and level of management; Lyness & Thompson, 2000). Barriers to advancement for women included more "lack of culture fit" and less geographic mobility than men. Results from Lyness and Thompson (2000) also showed that women executives value different criteria in the process of career advancement than men, such as a good track record and relationships with higher management. Thus, even though women executives achieved high status in this study, they did so by very different means than men.

Additionally, research shows that a pay disparity still exists between men and women (U.S. Bureau of Labor Statistics, 2008). In 2007, women's median weekly earnings represented only 80 percent of men's median weekly earnings. While this percentage has increased substantially over time (62% in 1979), women are still disadvantaged in salary.

In fact, a meta-analysis of resume studies using masculine and feminine jobs reported that applicants were rated lower when applying for an opposite-sex-typed job (i.e. men in nursing) (Davison & Burke, 2000). Furthermore, when a woman enters a stereotypical domain, she may not "fit" and therefore, suffer a "female disadvantage" such as being seen as socially incompetent and earning less income (Eagly, Makhijani & Klonsky, 1992; Eagly, 2007). In sum, gender is a highly visible cue in the workplace that leads to different treatment for women such as hiring decisions, advancement strategies, and income (Blau & Kahn, 2000; Maume, 2004).

Consequences of Prejudice

Men and women are expected to possess certain gendered qualities. The gender-trait hypothesis suggests that an individual's gender leads to expectations about masculinity and femininity (Huddy & Terkildsen, 1993). For example, gender leads to expectations about who should go into a career. This is known as a prescriptive stereotype, which provides insight into what characteristics group members should have (e.g. women are caring and nurturing; Diekman & Eagly, 2000). An individual can obtain role congruity if they behave in a way that subscribes to the gendered norms (e.g. nurturing woman enters elementary teaching occupation; Glick, Wilk & Perreault, 1995; Cejka & Eagly, 1999).

What happens when a woman enters a stereotypically masculine domain? The gender-role incongruency hypothesis suggests that prejudice may occur because gender and role of domain are incompatible. Prejudice can be blatant, but it may also surface in

subtle forms (Eagly & Diekman, 2005; Smith, Sansone & White, 2007). Recently, Diekman and Hirnisey (2007) investigated age-related stereotypes by assessing hireability of old and young candidates for a stable or dynamic company. This study asked the question "How is the individual evaluated for a specific context?" It was predicted that older adults would show less hireability than younger adults for dynamic jobs, but that younger adults would show equal or less hireability for stable jobs. Using resumes from an ostensibly completed applicant file, participants rated the job candidates on hireability, adaptability (i.e. willingness to learn new skills), and reliability (i.e. conscientiousness, loyalty). The company manipulation described the company as either *dynamic* (opening many new offices and has moved several of their department managers to different locations) or *stable* (doing well with their current sites and has not needed to open additional offices). The age manipulation presented the applicants as older or younger adults and was listed on the resume. Results of Diekman and Hirnisey (2007) showed a fan interaction on hireability, with younger adults being equally likely to be hired for the stable or dynamic company and older adults less likely to be hired for the dynamic company. This interaction was mediated by the perceived adaptability of the candidate. However, both older and younger adults received positive ratings, which further suggests that a subtle form of prejudice occurred (Diekman & Hirnisey, 2007). Similar to gender stereotypes, age-related stereotypes resulted in prejudice because the social group and role were incompatible, and this resulted in a "misfit".

As demonstrated, relying on stereotypes to make judgments can lead to prejudice. But can forms of prejudice predict whether and how subsequent discrimination will

occur? Biernat and colleagues (2009) investigated this question for racism and racial discrimination. Using implicit (i.e. Implicit Association Test [IAT] and lexical decision task) and explicit measures of racial stereotyping (i.e. pro-black and anti-black attitudes questionnaire; Katz & Hass, 1988), the authors measured to what extent participants would assign valued resources to a Black Student Union group. Additionally, Biernat, Collins, Katzarska-Miller and Thompson (2009) developed a new measure which assessed how much a participant used various types of ratings to assess Blacks' intelligence and determine if Blacks and Whites are held to different standards of intelligence. The new measure was utilized as an independent variable to predict funding to the Black Student Union. Results of Biernat et al. (2009) suggested that the only predictor of the Black Student Union funding was the individual difference measure assessing the tendency to hold Blacks and Whites to different standards. Across three studies, there was evidence to support that high levels of implicit prejudice and high likelihood of using different standards predicted less Black Student Union funding. In this study, the tendency to evaluate Blacks and Whites differently resulted from stereotypes about race being implicitly activated (Biernat et al., 2009). As such, other types of prejudice (e.g. sexism), may set the stage for predicting when different judgment standards will be employed.

Using Stereotypes to Inform Judgment Standards

A shifting standard is characterized by members of certain social groups being evaluated differently for the same performance (Biernat et al., 1991). Stereotypes are

used as the basis of information to determine which judgment standard should be employed. Thus, the standard shifts depending who is being evaluated (Biernat & Manis, 1994). Prejudice is evident by applying a shifting standard for similar performance leading to less desirable outcomes (Biernat & Manis, 1994; Biernat, 2003). Shifting standards are characterized by the type of measurement scale used for evaluation (Biernat et al., 1991; Biernat & Manis, 1994), the comparison group being evaluated (Paul & Smith, 2008), outcomes of the evaluation (Phelan et al., 2008; Biernet et al., 2009), and whether the judgment is congruent or incongruent with the stereotype (Hodson, Dovidio & Gaertner, 2002).

Measurement: Objective and Subjective Judgments

The first characteristic of a shifting standard is a discrepancy between an objective and subjective judgment. To assess judgments, two scales of measurement are used in this literature, including objective and subjective ratings. Objective ratings involve absolute judgments that are made *across* category. For example, test scores, grades, monetary (i.e. salary), and rank order are all examples of objective ratings. Alternatively, subjective ratings involve relative comparisons such as Likert-type scales and ratings on a continuum. Shifting standards occurs when a different subjective rating is given for the same objective performance.

To demonstrate shifting standards in the workplace, Biernat et al. (1991) used both subjective and objective ratings for financial success. Participants were asked to make judgments about how financially successful men and women were by looking at various photographs. Some participants made objective judgments (estimates of salary)

whereas others made subjective ratings of financial success (Likert-type ratings). Objective results showed that men were estimated to earn $9,000 more than women. However, subjective results showed that women were rated as more financially successful than men (even though women were expected to earn less).

In Biernat and Manis (1994) participants judged an author's ability to write articles concerning sex-typed topics (i.e. bass fishing and salaries of professional baseball players, cooking nutritious meals and eye makeup). The articles were identical except that the supposed author was either a woman (Joan) or a man (John). Results showed that on an objective scale, women were better at writing about feminine tasks and men were better at masculine tasks. However, the subjective ratings did not show this same bias. So, people used different "anchors" for evaluating what is good for a woman versus what is good for a man. Importantly, the standard depended on the gender of the author and the gender-relevant topic they wrote about.

Comparison Groups

Within-category judgments tend to be made for stereotyped groups (e.g. women judged against other women) rather than between-category judgments (e.g. women judged against men). When a group is stereotyped as "deficient" in a domain, people have lower standards for that group to ensure that stricter objective evidence is required that an individual possesses a counter-stereotypical attribute (Foddy & Smithson, 1989).

Research supporting the notion that different comparison groups are used comes from the 2008 presidential election. Women running for the Presidency are held to different standards than men (Paul & Smith, 2008; Smith, Paul & Paul, 2007). In Paul

and Smith (2008), telephone surveys were conducted to determine how likely a person would be to vote for the following candidates: Hillary Clinton, John Edwards, Elizabeth Dole, Rudy Giuliani, and John McCain. All candidates had similar objective qualifications (e.g. current position, years in office). However, when asked to match up participants head-to-head, the women candidates received a lower percentage of the hypothetical vote. Moreover, when the match-up involved Dole versus Clinton, the percentages for a women candidate increased. This suggests that participants frequently switched their vote to a male candidate when the opponent was a woman, unless the only other choice was a woman (Paul & Smith, 2008). It is evident that shifting standards occurred in this study because Dole and Clinton were objectively qualified, but still did not receive high percentages of votes, unless the match-up was against each other.

Outcome of Evaluation

Another factor of the shifting standards paradigm is the outcome of the evaluation. For example, people rate Lisa Leslie as a good athlete, but if forced to pick one athlete for the team, they will choose LeBron over Lisa. Two types of outcomes can occur including zero-sum behaviors and nonzero-sum behaviors (Biernat, 2003). Zero-sum behaviors are decisions involving resource allocation or rank-order situations, whereas nonzero-sum behaviors can be given to an unlimited number of individuals. Zero-sum behaviors require a cross-category (objective) judgment. On the other hand, nonzero-sum behaviors use a within-category (subjective) judgment. Thus, understanding shifting standards depends upon the frame of reference (across-category or within-category) and the outcome of the evaluation at hand (zero-sum or nonzero-sum).

As suggested in Phelan et al. (2008), agentic women were less likely to be hired than agentic men for a managerial position. This represents a zero-sum outcome because there is only one position to fill, requiring that the participant choose whether to hire the man or the woman. In other words, a judgment that required cross-category reference was needed (Phelan et al., 2008). Similarly in Biernat et al. (2009), the funding allocation for the Black Student Union group was a zero-sum decision. There was a limited amount of money available, and the participant had to decide which student organization would receive funding (Biernat et al., 2009).

Contrast and Assimilation to Stereotypes

The effects of stereotyping can also be explained by assimilation or contrast effects (Biernat, 2003). An assimilation effect occurs when a person is judged in terms of a specific stereotype and no individuating information is provided (e.g. Blacks perform poorly in academics). A contrast effect occurs when a person is judged in counter stereotypical terms when provided with individuating information (e.g. A Black individual performs really well in academics; Biernat & Kobrynowicz, 1999).

In a study investigating shifting standards and race, participants were asked to make college admissions decisions (e.g. accept or deny) based on the candidate's race, entrance exam scores (i.e. ACT) and high school achievement (i.e. GPA). Results showed that when Blacks had "mixed credentials" they were less likely to be recommended for admission than Blacks with strong credentials. In contrast, "mixed" credentials for Whites led to the same recommendation for admission as Whites with strong credentials.

This suggests that when there is ambiguity in the situation, there tends to be assimilation toward the stereotype among high prejudice participants (Hodson et al., 2002).

Shifting Standards Summary

In sum, shifting standards occur when there is a discrepancy between objective and subjective measures, within-category judgments are made, zero-sum outcomes are at stake, and assimilation to stereotypes occurs. The consequences of shifting standards include blatant and subtle forms of discrimination as well as denial of limited resources (Biernat & Manis, 1994; Biernat et al., 2009). But the question remains: How does it feel to be the victim of shifting standards? How would Lisa Leslie feel if she knew her good ability was qualified by her gender? This is one question addressed in the current project. The aim of the current project is to investigate how shifting standards feedback (i.e. you did good for a girl) affects motivational and performance outcomes for women who care about succeeding in (i.e. are highly identified with) athletics.

The current project builds on this line of research to ask how it feels to be the victim of a shifting standard. Much work has been done on characterizing the nature of shifting standards, but the *subjective* experience from the victim's perspective has not yet been investigated. In order to make predictions about how women will experience shifting standards, I draw from related research on patronizing, stereotype threat, and the motivational model for stereotyped tasks.

Responses to Prejudice

A person's response to shifting standards may be multi-faceted, involving performance and motivational consequences. Although positive in nature, shifting standards feedback may activate gender stereotypes (Pinel, 1999). Gender stereotype activation may lead to decreased performance (Stone, Lynch, Sjomeling & Darling, 1999; Stone, 2002). Shifting standards feedback may also reduce interest in the activity both immediately and in the future. (Smith, Sansone & White, 2007). The following discussion considers these potential responses to shifting standards feedback.

Patronizing Feedback

Shifting standards feedback may be viewed as "devalued praise." Devalued praise is often perceived as patronizing (Gervais & Vescio, 2007). Patronizing is specific to situations where someone of high status uses and endorses stereotypes of low-power individuals (Vescio et al., 2003). Patronizing behavior is often motivated by benevolent forms of sexism that result in the victim being devalued (Vescio, Gervais, Snyder & Hoover, 2005). For example, in one study men and women engaged in a stereotypically masculine task (strategic, academic challenge) and then received feedback from a high-power man (Vescio et al., 2005). Women who received patronizing feedback were presented with a computer-generated email message from the male evaluator, "Jimmy John." The message contained high praise, but resulted in a devalued position (being selected as an alternate for the challenge). For women in the patronizing condition, results showed decreased performance on the next task relative to the non-patronizing

conditions. In contrast, men showed increased performance in response to patronizing feedback. Thus, it appears that patronizing feedback leads to lower performance in gender-stereotyped domains.

Subsequent research has examined perceived sexism as a potential mediator to the patronizing feedback/poor performance relationship (Dardenne et al., 2007). In this study, women were asked to complete job hiring tests for a stereotypically feminine job (not described in detail). Feedback was provided by a high-power male recruiter that contained elements of hostile sexism, benevolent sexism, or no feedback. Performance was then measured on a problem-solving test. Results of Dardenne et al. (2007) suggested that although perceived sexism was highest in the hostile sexism condition, women's performance was lowest in the benevolent sexism condition. Given the ambiguity of the benevolent sexism condition, women may have had difficulty creating external attributions for the feedback (Dardenne et al., 2007). Instead, they may have created internal attributions such as negative thoughts about one's ability on the problem-solving test). Taken together, the patronizing research suggests that devalued praise (from a high status other) leads to decreased performance, especially when the feedback is of a benevolent nature.

Stereotype Threat

As patronizing research makes clear, performance decreases in response to devalued praise from a high-status other. Stereotype threat also predicts decreased performance when negative stereotypes about group membership become salient. Stereotype threat is the experience of being reminded of a negative stereotype about

one's group, and then confirming the content of the stereotype (Steele & Aronson, 1995). For example, the stereotype that Whites are not athletic and Blacks are not intelligent was investigated by Stone et al. (1999). In this study, a golf test was framed either as a sports intelligence test or natural athletic ability test. It was predicted that Whites would have lower performance on the golf test framed as a natural athletic ability test, whereas Blacks would perform worse on the golf test framed as a sports intelligence test. Performance was measured as number of strokes taken on a miniature-golf putting test. Results showed that when race was made salient, Blacks did worse on the "athletics intelligence" test relative to the "natural athletic ability test," and Whites did worse on the "natural athletic ability" test relative to the "athletics intelligence test." This shows that performance suffers when reminded about a negative stereotype about one's group.

In addition to stereotypes being explicitly activated, stereotype threat may result from implicit activation as well. Implicit threat occurs when a group feels stigmatized in the situation and relies on previous knowledge about the stereotype, such as the belief that Asian men outperform White men in math (Smith & Johnson, 2006; Kaiser, Vick & Major, 2006; Schmader, Johns & Barquissau, 2004). Inzlicht and Ben-Zeev (2000) showed that women who were outnumbered by men in a testing situation performed poorly on a math test. This suggests that stereotype threat occurred even though the stereotype that women are inferior to men at math was not explicitly mentioned. In addition, Smith and White's results (2002) showed that stereotypes can be activated under explicit stereotype threat and normal testing circumstances (implicit threat). Smith and White (2002) suggested that implicit and explicit threat are equally harmful to

performance. It is clear that stereotype threat effects occur by explicit and implicit activation; however, less is known about *why* this occurs.

What Variables Influence Stereotype Threat?

Past research has investigated the mechanism resulting in low performance in the stereotype threat literature. It is possible that the process operating to produce stereotype threat is cognitive, emotional, motivational, or a combination of these factors (Major & O'Brien, 2005; Schmader, Johns & Forbes, 2008; Smith, Sansone & White, 2007). Smith (2004) outlines a number of potential mechanisms that may be operating to lower performance under stereotype threat. Among others, these include performance avoidance goal adoption (Smith, 2006; Smith, Sansone & White, 2007), lower interest in the task (Smith, Sansone & White, 2007), low working memory capacity (Schmader et al., 2008), and "choking under pressure" (Smith & Johnson, 2006; Beilock & McConnell, 2004). Choking under pressure is defined as a monitoring error where an individual concentrates too much on the task, resulting in decreased performance (Beilock & McConnell, 2004). The quest to find one mediator for stereotype threat has yielded many failed attempts, resulting in partially supported mechanisms (see Smith, 2004 for a review).

Research has recently begun to examine physiological and neurological methods, in an effort to understand the biological basis of a stereotype threatening experience. Using fMRI data, Krendl, Richeson, Kelle and Heatherton (2008) found in one study that stereotype threat vs. no threat activates different brain regions. In this study, women participants engaged in a stereotyped math/gender Implicit Association Test (IAT) or

neutral IAT. Then a fMRI was conducted while participants completed math problems. In the stereotype threat condition, there was no activation in mathematical or working memory brain regions, but instead activation was present in the ventral anterior cingular cortex (emotional area). In the neutral condition, there was activation in the angular gyrus, left parietal, and prefrontal cortex (mathematical and working memory regions; Krendl et al., 2008). The authors point out that these results draw into question whether lower working memory is the mechanism by which stereotype threat leads to poor performance. Instead, these results suggest that there is an emotional component to the stereotype threat experience. It is possible, then, that being aware of one's stigma might result in a negative emotional experience during task engagement.

Stigma Consciousness: Stigma consciousness is defined as the extent to which individuals are chronically self-conscious of their stigmatized status (Pinel, 1999). Stigma consciousness is a distinct concept from group identity (e.g. gender identity) and group consciousness because stigma consciousness does not require feelings of group membership. Stigma consciousness is also empirically distinct from related constructs, as shown by weak or negative correlations with public self-consciousness and modern sexism (Pinel, 1999; Study 2). Gender identity and stigma consciousness may also differ in terms of their development. Gender identity is thought to develop beginning in childhood (see Martin, 1999 for a review), but one way stigma consciousness might emerge is through an emotional reaction that develops following a negative gender-related experience (e.g. sexual harassment; Stockdale, 1998; Bergman et al., 2002). Since sexual harassment leads to negative work-related outcomes, psychological and physical

outcomes for the victim, this experience could lead to elevated stigma consciousness. In fact, Pinel (1999) showed that stigma consciousness was related to sensitivity to sexism. Consequently, being high in stigma consciousness allows the individual to more readily perceive future acts of discrimination, such as sexism (Pinel, 1999; 2004).

Research shows that stigma consciousness moderates stereotype threat effects on performance. For example, in a study conducted by Brown and Pinel (2003), women participants who highly identified with math were categorized as either high or low in gender stigma consciousness. Participants were pre-screened for math identification using the Mathematics Identification Questionnaire (Brown & Josephs, 2000). Under conditions of stereotype threat (or not), participants completed a 20-item math test consisting of items from the math GRE. Women who were high in stigma consciousness and in the stereotype threat condition scored lower on the math test than any of the other conditions (see also Smith, Sansone & White, 2007). Being aware of one's stigma led to higher susceptibility to stereotype threat effects, and thus moderated the relationship between stereotype threat and performance. These results suggest that stigma consciousness is an important factor to consider in stereotype threat experiences.

Stigma consciousness has also been linked to disengagement among stereotyped employees in the workplace. Pinel and Paulin (2005) investigated stigma consciousness among staff workers employed by a university in a naturally occurring setting. A modified version of the Stigma Consciousness Questionnaire (Pinel, 1999) was used to assess stigma consciousness as a staff worker. Participants were contacted initially and after a two year period to assess their level of stigma consciousness and intent to leave

the job. Results of Pinel and Paulin (2005) suggested that staff workers high in stigma consciousness were more likely to feel disrespected, show increased intent to leave their jobs, and left their jobs more often compared to participants low in stigma consciousness (Pinel & Paulin, 2005). Although disengagement can sometimes occur psychologically, results suggested that employees physically disengaged with the stereotyped domain by leaving their jobs following a two year period. These findings suggest that high levels of stigma consciousness set the stage for staff workers to downplay the personal importance of a domain, ultimately disengaging completely (Crocker & Major, 1989; Davies, Spencer, Quinn & Gerhardstein, 2002; Schmader, Major, Eccleston & McCoy, 2001).

Stigma consciousness has negatively affected non-performance outcomes as well. In Smith, Kausar and Holt-Lunstad (2007), women's academic experiences in Pakistan were investigated. Specifically, achievement goal adoption and level of well-being were assessed for women in science (gender-atypical) and non-science (gender-typical) fields of study. Achievement goals are defined as a construct representing the desire to show competence (Harackiewicz & Elliot, 1993). Goals can take a performance or mastery orientation with an approach or avoidance focus (Elliott & McGregor, 2001). Gender stigma consciousness was measured using the Stigma Consciousness Questionnaire (Pinel, 1999). Results showed that women high in gender stigma consciousness in science fields reported higher performance-avoidance goal adoption and lower levels of well-being. This suggests that how Pakistani women perceive stigma affects their experience in stereotyped domains. Whereas most stereotype threat literature has reported immediate performance decrements, this line of research highlights that stereotype threat is also

negatively related to non-performance outcomes (Smith, Kausar & Holt-Lunstad, 2007; Schmader et al., 2004). It is possible that stereotype threat might affect additional outcomes as well, such as level of interest in the task.

Motivational Model for Stereotyped Tasks

The Stereotyped Task Engagement Process model presents a motivational account of stereotype threat effects (STEP; Smith, 2004). The STEP model predicts that individuals under stereotype threat will experience the lowest level of task interest (both immediately and in the future), because of the adoption of avoidance achievement goals. To test this, Smith et al. (2007) reminded women of the computer science-gender stereotype that women are not as good as men in math–related tasks. In Study 1, women participants were assigned to adopt a performance-avoidance goal, performance-approach goal, mastery goal, or no goal. Participants learned about the upcoming task called the computing aptitude assessment tool (CAAT), and then were provided with a "journal article" reminding the women participants that men are superior to women in mathematics. The stereotype threat condition contained a statement that the current task had found gender differences, whereas the stereotype nullified condition stated that the current task had found no such gender differences. Next, participants completed the CAAT task (a computer programming exercise) and reported their interest and future motivation for the task.

Results showed that participants in the stereotype threat condition were less interested in a computer science task when performance-avoidance goals were assigned compared to mastery and performance-approach goals. When no goal was assigned,

participants reported similar interest to the performance-avoidance goal condition, suggesting that avoidance goals might be adopted spontaneously in response to stereotype threat. Study 2 showed that performance-avoidance goals were spontaneously adopted in response to stereotype threat. However, Study 2 also used an achievement motivation distinction and suggested that women low in achievement motivation adopt a different type of goal (performance approach) compared to women high in achievement motivation (performance avoidance) in response to stereotype threat.

Study 3 showed that women high in achievement motivation reported higher interest after assignment to a performance- approach goal than women low in achievement motivation. Women who were low in achievement motivation showed lower interest in the task following performance approach compared to performance avoidance goal assignment. In a meditational analysis, task absorption and perceived competence were both related to increased immediate interest in the task. The results of Study 3 suggest that the most negative effects of stereotype threat occurred for individuals high in achievement motivation. In summary, highly identified women in computer science reported decreased interest and future motivation in the computer science task in the stereotype threat condition. Additionally, participants also showed adoption of performance avoidance goals for the task.

Domain Identification: As suggested above, stereotype threat and the STEP model (Smith, 2004) show that negative effects are most pronounced for people who place high importance on succeeding in the domain (Smith & White, 2001; Smith, Morgan & White, 2005). People who care about a particular field are considered high in domain

identification. Domain identification is defined as the incorporation of a particular domain into one's self concept (Smith & White, 2001) and is positively associated with career pursuit, grades, successful school adjustment, and athletic performance (Smith et al., 2005; Crocker, Major & Steele, 1998; Smith & White, 2001; Stone, 2002; Koestner & Losier, 2002; Osborne, 1997). For example, individuals who identified politics as an important part of their identity were more likely to show political activism by voting (Hodgins & Knee, 2002). Thus, high domain identification is associated with a number of positive outcomes.

Ironically, individuals high in domain identification may also be the most susceptible to stereotype threat effects (Smith & White, 2001; Aronson et al., 1999). For example, Aronson and colleagues (Study 2; 1999) blocked male calculus students on math domain identification, resulting in two groups of highly and moderately identified individuals. Stereotype threat was manipulated by reminding participants that research has shown Asians' math ability to be superior to White student's math ability. As such, participants were told that the current study was being conducted to test whether Asians' superior math ability generalizes to all types of math problems. Using a 2 (Domain identification: high, moderate) x 2 (Condition: stereotype threat, control) design, results suggested that high domain identification in math resulted in higher math performance than moderate domain identification in math for the control condition. This is in line with other research showing the positive benefits associated with high domain identification under normal conditions. Importantly, the performance pattern was reversed under stereotype threat, such that men higher in domain identification performed worse

compared to men with moderate domain identification in math. This suggests that stereotype threat effects are more likely occur to when the domain is a central component of the individual's self-concept.

As this review suggests, stereotype threat effects are found primarily among those high in domain identification (Smith, Sansone & White, 2007; Brown & Pinel, 2003; Aronson et al., 1999, but see Keller, 2002). Thus, level of domain identification must be accounted for when examining stereotype threat effects. One technique used to investigate stereotype effects is to pre-select participants who are high in domain identification (Aronson et al., 1999; Spencer, Steele & Quinn, 1999; Smith, Sansone & White, 2007). Another technique that has been used in previous research involves statistically controlling for domain identification to make all groups statistically equal on domain identification (Smith, 2006; Smith & White, 2002; Brown & Josephs, 1999). The current research will use a combination of these approaches.

Relationship between Interest, Perceived Competence, and Self-Esteem

In addition to interest, perceived competence in the task and self-esteem are important factors in understanding stereotype threat. Thoman and Sansone (in press) investigated the impact of negative stereotyped feedback on interest, perceived competence, and self-esteem. In Study 2, women participants read mystery stories (including a crime and conclusion by a forensic scientist) and then were asked to report which clue led the forensic scientist to their conclusion. The women believed that a man was also completing the same task but was in a different room for the beginning phase of

the experiment.

The experimenter ostensibly graded the answer sheets, arriving at the conclusion that the woman (participant) and the man (confederate) scored the same. To settle the dispute concerning who would be the "Outstanding Group" member, the experimenter ostensibly had a conversation with her male supervisor (John) and the male confederate was always selected. In the overheard conversation, one of four types of feedback was provided including: ambiguous reason, unfair but non-gender biased reason, individual-specific gender-biased reason, and domain gender-biased reason. For the ambiguous condition, no reason was provided. In the unfair but non-gender bias reason, the name on top was selected. In the individual-specific gender-bias reason, the supervisor mentioned a man he knew that excelled on science tasks. Finally, for the domain gender-bias condition, a stereotype regarding all men's and women's science abilities was mentioned. Post-test measures of activity interest, perceived competence, and state self-esteem were taken.

Results of Thoman and Sansone's experiment (in press) suggested that womens' interest decreased significantly when they received a gender-biased reason for not being selected as "Outstanding Group" member. When receiving unfair feedback in general there was a marginal effect for interest, with a trend in the direction of lower interest. When perceived competence was higher for a task, state self-esteem was also higher. This suggests that interest, perceived competence, and self-esteem are all important in the experience of an individual who is the victim of prejudice.

State Self-Esteem

State self-esteem is defined as an individual's evaluation of their personal characteristics and behavioral patterns at a given point in time. State self-esteem is different than trait self-esteem because it can fluctuate and is less stable over time (Crocker & Major, 1989; Crocker, Voelkl, Testa & Major, 1991). For example, an individual with high trait self-esteem might demonstrate temporary signs of low state self-esteem following poor performance. Past research has resulted in mixed findings concerning the effect of negative feedback on self esteem. In a study conducted by Dion (1975) women reported lower self-esteem after experiencing interpersonal rejection from men compared to women. Of note, this feedback was unambiguous and dealt with gender.

More recently, research has shown that self-esteem may not be lowered in response to negative feedback. In a study conducted by Crocker et al. (1991) an ambiguous (blinds down) or unambiguous (blinds up) target provided positive or negative feedback from the next room. Both Blacks and Whites played the role of the participant, but the confederate in the next room was always White. When the window blinds were closed and the participant was unable to attribute the negative feedback to prejudice, then self-esteem was negatively affected. However, when the window blinds were open and the participant could attribute the feedback to the evaluator's racism, self-esteem was not affected.

These findings suggest that, at least for race, there may be self-protective mechanisms preserving self-esteem during ambiguous situations. Additionally, state self-esteem was protected in Thoman and Sansone's work (in press). This mixed finding suggests that self-esteem should be included as a dependent measure in this study. One goal of the current project is to examine the effects of shifting standards feedback on self-esteem.

Differences among Shifting Standards, Stereotype Threat, and Patronizing Research

The goal of the current study is to understand how it might feel to be the victim of a shifting standard when it is attached to positive feedback. Although feedback might seem positive, it can sometimes take the form of benevolent sexism (Dardenne et al., 2007). Stereotype threat is characterized by worrying about confirming a negative stereotype and no feedback is provided regarding individual performance (Stone et al., 1999; Smith, Sansone & White, 2007). During stereotype threat, the threat is "in the air" and lacks interpersonal communication (Steele & Aronson, 1995). In some cases, shifting standards may be perceived as a form of patronizing (when low minimum standards are called to mind for a stereotyped group; Biernat et al., 2009).The patronizing experience is more akin to the shifting standards experience than stereotype threat because it does involve interaction with the person making the judgment (Vescio et al., 2005). In addition, patronizing feedback is characterized by the feedback coming from a high-power source, but it is not only provided for stereotyped tasks. This project will draw

from these literatures to inform how it might feel to experience shifting standards in a stereotyped task.

Shifting Standards for Women in Athletics

Although any stereotyped domain could have been chosen for this research, I am are choosing athletics. According to Hall (2008), athletics are a male-identified domain overall, but media has recently been portraying some women athletes positively. Women have only recently become more accepted in their role as athletes (Parsons & Betz, 2001). Due in large part to the legislation from Title IX in 1972, the opportunity to participate in athletics was extended to women. Title IX states that sex-based discrimination is prohibited and equal opportunity must be available, regardless of sex (U.S. Department of Labor, 2009). However, the decision to include women was not made easily, as opponents of Title IX expressed their concerns. John Fuzak, National Collegiate Athletic Association (NCAA) president, wrote to President Ford in 1975 stating, "The Department of Health, Education and Welfare [contends that] Title IX as expressed could seriously damage, if not destroy, the major men's intercollegiate athletic programs." Nevertheless, in 1980 the NCAA began its initiation of women's programs. Today there are 19 women's sports that hold NCAA sponsored championships (NCAA, 2009).

According to the National Sporting Goods Association (2007), basketball (31 % women), snowboarding (26.5% women), golf (22.6 % women), and ice hockey (18.1 %) are among the sports with the lowest levels of participation for women. Differing levels of participation may be due to how each sport scores on a continuum of masculinity and

femininity (Parsons & Betz, 2001). Parsons and Betz (2001) had participants evaluate the masculine nature of sports using a Likert-type scale. The most masculine sports in this sample included golf, soccer, crew, basketball, and lacrosse whereas the most feminine sports included dance, cheerleading, synchronized swimming, and gymnastics (Parsons & Betz, 2001). Those sports containing masculine qualities might be less desirable for women than those with feminine qualities, presumably because the masculine qualities do not "fit" for women.

Research on shifting standards in sports from the perceiver's perspective shows that stereotypes concerning gender and athleticism are used to inform zero-sum decisions (Biernat & Vescio, 2002). In Biernat and Vescio (2002), participants played the role of "team manager" and made judgments about men and women softball players. Photographs were used to predict both objective and subjective athletic performance. For the objective judgments, participants were asked to pick 13 of the 18 players to be on their team, choose the starting line-up, pick positions (including 3 players benched), and create a batting order. For the objective ratings, participants were asked to estimate batting average and fielding error rate. In addition, participants were asked to rate the objective judgments on subjective scales (i.e. how good is this person at hitting?) and indicate how they would respond if the person hit a single (i.e. zero-sum vs. non-zero sum behavior).

Results showed that a man athlete with the same performance as a woman athlete was rated lower in overall athletic ability. Additionally, zero-sum behaviors (i.e. being picked for the team) favored male athletes, but non-zero sum behaviors (i.e. smiling)

favored women athletes. Thus, it appears that perceivers' judgments were made on a "slippery slope," by applying different standards of athleticism for men and women. That is, what was good for a man was not the same as what was good for a woman. How would finding out that you have been the victim of a shifting standard affect performance and motivation? The current project tested this question to determine if shifting standards for athletics have deleterious effects on performance and motivation for highly-identified women athletes.

Project Overview

Study 1 was conducted to investigate a real-world experience of shifting standards for elite women athletes. Interviews with athletes in male-dominated (snowboarding) versus less male-dominated (skiing) winter sports were conducted. Shifting standards feedback was randomly provided by the experimenter to determine how estimated athletic performance was influenced by the feedback. The goal of this field study was to refine the general hypotheses of the current research and help determine which variables to investigate in Study 2.

The objective of Study 2 was to test the effect of shifting standards feedback on performance and motivation in a more controlled environment. In Study 2, participants were recruited based on their responses to an online pre-screen assessing domain identification in athletics and primary sport played. As a result, self-identified elite women athletes (n = 77) were blocked on the masculine nature of the sport, resulting in a 2 (type of sport: masculine, non-masculine) x 3 (feedback: positive, shifting standards, no

feedback) between-subjects design. Stigma consciousness as a woman athlete was also measured (See Appendix A). The masculine nature of the sports was determined by pilot testing. After first engaging in an ambiguous athletic test (hand grip), participants received the feedback manipulation. Then, a second unambiguous test (balance disk) was administered to measure performance, and participants then completed post-test measures assessing interest, future motivation, and several exploratory measures. It was predicted that motivation and performance would be low among women who played masculine sports and received shifting standards feedback compared to women in masculine sports who received positive feedback. It was also predicted that women high in stigma consciousness who received shifting standards feedback would show low performance and motivation relative to women low in stigma consciousness receiving the same feedback. To further explore the mechanisms underlying the proposed relationship, measures of stereotype activation and goal adoption were also included.

STUDY 1: BRIDGER BOWL FIELD STUDY

Winter sports provide an opportunity to examine women athletes as they naturally participate in their sport. Elite women skiers and snowboarders were interviewed at Bridger Bowl Ski Resort in Montana on a Saturday in February. On the day of data collection, there was a NCAA skiing competition occurring at the same time.

Method

Participants and Design

Participants (n = 28, Mean age = 24.5 years) were randomly recruited near advanced runs in an effort to try and select "elite" athletes. Overall, the sample contained 36 % snowboarders and 64 % skiers. In particular, participants were overall very highly identified with their sport (M = 6.71, SD = 1.92). A one sample t-test showed that participants were highly identified with their sport by testing against the neutral point (5) on the 1 to 9 scale, with higher scores indicating more identification with their sport (t [27] = 4.71, $p < .01$). Participants also had high levels of experience with their sport (M = 12.71 years, SD = 10.20) and reported completing the blue or black (advanced) runs 92.8 % of the time. A one-sample t-test showed that participants enjoyed skiing or snowboarding to a high extent (M = 6.82, SD = 0.55) by testing against the neutral point (5) on a scale from 1 to 9 with higher scores indicating more enjoyment (t [27] = 39.90, $p < .01$). Participants who were interviewed were blocked on type of sport resulting in a 2 (type of sport: skiing, snowboarding) x 2 (type of feedback: shifting standards, positive)

between subjects design. An interview was conducted with participants in exchange for a $5.00 drink coupon for the ski lodge coffee shop.

Procedure

Every third woman was approached, unless participants had already participated. Participants first reported several demographic variables (including domain identification items) and were then asked to report an estimated time from their previous run ("Estimated Time 1"). The experimenter recorded this time and then turned the page of the packet. On the next page there was a mark designed to tell the experimenter whether to respond with shifting standards or positive feedback. Until this time, the experimenter remained unaware of condition assignment. Next, participants were asked several questions including one item assessing stigma consciousness. Specifically, participants were asked "How often are ever judged differently on the slopes, just because you are a woman?" on a 1 (never happens) to 9 (always happens) scale. Finally, participants were asked to predict their time on the next run ("Estimated Time 2").

Shifting Standards Manipulation: Pilot Study 1 was conducted first in the lab to explore the best wording to invoke the shifting standards experience. Participants were asked to self-report which of the following phrases was most familiar to them: good for a girl, good for a woman, or good for a female presented in counterbalanced order. A separate sample of 40 participants (90.0% women, Mean age = 22.9 years) completed the survey. Results showed that the phrase "good for a girl" was the most familiar 85 % of the time. The phrase "good for a woman" only accounted for ten percent of most familiar

responses, whereas "good for a female" accounted for five percent of the responses. By using the most commonly heard phrase, we were able to maintain a high degree of experimental realism in manipulating the shifting standards experience.

Based on the information obtained in Pilot Study 1, participants in the positive feedback condition received the following feedback: "Wow! That's a really good time. You are clearly good at skiing/snowboarding." Participants in the shifting standards condition were given the same feedback, except "for a girl" was added to the end of the statement.

Dependent Variables: A difference score was calculated by subtracting Estimated Time 1 from Estimated Time 2, which served as the main dependent variable and index to estimated performance. Finally, participants were told the nature of the study and asked an open-ended item: "How would you feel/react if someone had just told you: You are good at skiing/snowboarding for a girl?" Quotations from the open-ended item are listed in the End Note.

Results

There was a significant type of sport x feedback condition interaction on predicted time, ($F[1,24] = 5.29$, $p < .05$). Follow-up simple effects tests revealed that when given shifting standards feedback, women snowboarders reported a slower estimated time on the next run ($M = -0.25$ minutes, $SE = 1.29$ minutes) compared to women snowboarders in the positive feedback condition ($M = 4.63$ minutes, $SE = 1.57$ minutes), ($F[1,24] = 4.63$, $p < .05$). No difference in expected time for skiers was found. Thus, it appears that

positive feedback qualified by "good for a girl" had a negative impact on expected performance for elite women in male-dominated sports. Among skiers, shifting standards feedback did not cause a significant change in estimated time on the next run compared to positive feedback.

Relationship between Variables

To test the relationship between variables, a correlation between stigma consciousness and domain identification was calculated. A significant positive correlation was revealed ($r[28]= .56$, $p < .01$), suggesting that both stigma consciousness and domain identification should be examined in Study 2 as individual difference measures.

Discussion

Results of this field study showed that women who participate in male-dominated sports became less confident in their expected performance on the next trial after receiving shifting standards feedback. Given that the shifting standards feedback could have been interpreted as benevolent sexism if it had come from a man (Glick & Fiske, 2001), Study 2 held this variable constant by using only a male evaluator. The field study also suggested that the masculine nature of the sport mattered, evident in decreased perceived competence for snowboarders (male-dominated sport) upon receiving shifting standards feedback but not for skiers (relatively less male-dominated sport). Study 2 blocked participants on the masculine nature of the sport, as obtained from the sign-up process.

Pilot Test 2 was conducted to inform the type of sport variable. The objective of this study was to classify masculine sports and non-masculine sports. A separate set of 31 coders (64.5% women, 80.6% athletes, Mean age= 22.6 years) rated a list of sports generated from the entire pool of participants. Each sport was rated on two scales, including a Likert-type scale and forced choice. The Likert-type scale assessed masculinity and femininity by using a 1 (completely feminine) to 9 (completely masculine) scale. The forced choice option required that each participant choose whether the sport is more masculine or feminine. This coding was use to classify athletes into masculine or non-masculine sports in Study 2 (See Table 1).

Table 1. Complete List of Sports Classified into Masculine or Non-Masculine Categories.

Masculine	Neutral (Non-masculine)	Feminine (Non-masculine)
Kayaking	Badminton	Horseback riding
Football	Hiking	Dance
Rock climbing	Skiing	Gymnastics
Field hockey	Soccer	Softball
Cycling	Swimming	Cheerleading
Snowboarding	Tennis	Equestrian
Baseball	Cross country	Volleyball
Golf		
Rugby		
Kickboxing		
Karate		
Fly fishing		
Basketball		
Lacrosse		
Bowling		
Wrestling		
Rodeo		
Track & field		

Note: Scores ranged from 1 (feminine) to 9 (masculine) and masculinity scores for each sport were entered into a one-sample t-test using 5 (midpoint) as the test value.

End Note

To gain insight into how participants in the relatively male-dominated sport felt after receiving shifting standards feedback, the open ended items concerning women's participation in athletics were examined. When asked how a participant would have felt if "for a girl" was added on, a snowboarder in the positive feedback condition commented, "I would have a negative reaction, and might think it's sexist." When another snowboarder in the positive feedback condition was asked how she felt, she reported, "It's fine that you [a woman experimenter] said it, but coming from a guy it might have been different." These comments suggest that the gender of the evaluator may be an important variable to control for in the shifting standards experience.

STUDY 2: CONTROLLED LABORATORY STUDY

Method

Study 2 was conducted to investigate the subjective experience of shifting standards for participant outcomes (e.g. performance and intrinsic motivation) and processes (e.g. gender stereotype activation, goal adoption). Participants were blocked by the primary sport they identified with (i.e. masculine or non-masculine sport). They were provided with computer-generated feedback concerning the participant's athletic performance on an athletic ability predictor test or computer-generated feedback concerning the athletic equipment. Participants believed that they were being watched through a video camera by a male evaluator in the next room. Participants received performance feedback from the evaluator on their computer through an instant messaging system. The score given to all participants was the same; however, the explanation of the score varied across conditions. People received either positive feedback (good), shifting standards feedback (good for a girl) or no feedback. Next, participants completed a stereotype activation measure for gender. A different (unambiguous) athletic ability test was then administered. Finally, participants reported on their motivational experience using established measures.

Hypotheses

Hypotheses 1a & 1b: The Role of Gender Stereotype Activation

It was predicted that shifting standards feedback would elicit gender stereotype activation, and that this activation would, in turn, be correlated with lower performance and motivation. To test the following predictions, a new measure assessing gender stereotype activation was included.

Type of Sport (1a). It was predicted that women who identified with a masculine sport in the shifting standards feedback condition would show the high gender stereotype activation compared to women receiving the same feedback, but identifying with a non-masculine sport. Based on gender-role incongruency theory (Eagly & Diekman, 2005), it was predicted that gender stereotypes would be activated in response to feedback that reminds participants of their gender role misfit (i.e. woman in a masculine sport).

Stigma Consciousness (1b). It was predicted that among women high in stigma consciousness, shifting standards feedback condition would elicit high gender stereotype activation compared women in the positive and no feedback conditions. This prediction stems from research showing that any references to gender can lead to stereotype-consistent behavior (Major & O'Brien, 2005). However, it was expected that gender stereotype activation would be especially high for women already showing stigma sensitivity.

Hypotheses 2a & 2b: Responses to Shifting Standards Feedback

Type of Sport (2a). It was predicted that women athletes who highly identified with a masculine sport and received shifting standards feedback would experience low immediate and future motivation and poor performance on a subsequent athletic test. Stereotype threat research predicts that motivation and performance will only decrease for a woman athlete who is highly identified with her domain (Stone et al., 1999; Smith, Sansone & White, 2007). Based on stereotype threat research and patronizing research (Vescio et al., 2005; Dardenne et al., 2007), the most harmful shifting standards feedback should be that which discounts a woman's accomplishments in a domain which does not "fit" her gender role. Thus, it was predicted that shifting standards in the "misfit" domain (woman in a male dominated sport) would result in low immediate and future motivation and performance.

Stigma Consciousness (2b). It was predicted that women athletes who were high in stigma consciousness and received shifting standards feedback would experience low motivation and poor performance on a subsequent athletic test. Brown and Pinel (2003) showed that the relationship between stereotype threat and performance was moderated by stigma consciousness. Therefore, it was predicted that shifting standards feedback would result in low immediate interest, future motivation, and performance for those who are most conscious of the stigma associated with being a woman athlete.

Hypotheses 3a & 3b: Responses to Positive Feedback

Type of Sport (3a). Positive feedback was expected to increase motivation and performance to the greatest extent for women in masculine sports. Hypothesis 2a predicts a "shifting standards lift" effect for women in masculine sports (Smith & Johnson, 2006). When a woman receives positive feedback *despite* her involvement with a stereotyped domain, she is expected to believe the feedback. Role congruity theory predicts that a woman in a masculine sport is violating her gender role (Eagly & Diekman, 2005) and *should* receive feedback reflecting that bias. However, if she receives positive feedback the augmenting principle suggests that the positive feedback is attributed to her ability (because it could have been attributed to benevolent sexism but was not) and becomes more valued (Kelley, 1972). Thus, it was predicted that positive feedback would increase performance and motivation, and this effect would be more exaggerated for the "misfit" condition.

Stigma Consciousness (3b*).* It was predicted that positive feedback would increase motivation and performance to the greatest extent for women low in stigma consciousness compared to women high in stigma consciousness. It was also predicted that positive feedback would lead to an increase in perceived competence (Sansone, 1989), which has been positively correlated with increase with test performance in past research (Pinel, 1999). Thus, it was predicted that positive feedback among women athletes low in stigma consciousness would result in the highest levels of immediate and future motivation as well as increased performance.

Hypotheses 4a & 4b: Responses to No Feedback

Type of Sport (4a). The no feedback condition was expected to produce intermediate levels of performance and motivation between the shifting standards and positive feedback. In addition, there was not an expected difference between masculine and non-masculine sports for the no feedback condition. The control condition was included to allow the direction of the effects to be interpreted.

Stigma Consciousness (4b). There were two different possibilities for how women athletes would respond to no feedback. First, no feedback could produce intermediate levels of performance and motivation between the shifting standards and positive feedback. Alternatively, the no feedback condition could produce similar levels of performance and motivation to the shifting standards feedback condition (Smith & White, 2002). Thus, this hypothesis was exploratory because an implicit stereotype about women athletes could be activated in response to no feedback or could produce intermediate levels between shifting standards and positive feedback.

Hypothesis 5: Shifting Standards Feedback Effects on Self-Esteem

There are mixed findings regarding how self-esteem is impacted in response to negative feedback. Recall that according to Dion (1975) state self-esteem should suffer in response to negative feedback. However, given the more recent Crocker and Major (1989) findings, self-esteem should remain unaffected by poor feedback when it can be attributed to an external source (e.g. sexist evaluator). Recall however that the shifting

standards feedback given in this study was positive feedback which has been qualified by a stereotype. As it was unclear what to expect, no directional hypotheses were made for how state self-esteem would be affected by shifting standards feedback.

It was predicted that positive feedback would lead to an increase in perceived competence (Sansone, 1989), which is positively correlated with high self-esteem. It was predicted that positive feedback would increase self-esteem (Thoman & Sansone, in press). However, the no-feedback condition was not expected to result in high levels or low levels of self-esteem.

Expected Relationship Among Variables

Based on past intrinsic motivational research (Smith, Sansone & White, 2007), it was predicted that immediate and future motivation would be positively related. Further, it was predicted that performance as well as both assessments of motivation would all be positive related (Bandura, 1982). It was predicted that gender stereotype activation would be negatively related with performance and interest. Finally, state self-esteem was expected to be positively related to performance and motivation.

Participants and Design

Women college students (n = 65, Mean age = 19.55 years) from a mid-sized Northwestern university and women athletes from the community (n=12, Mean age= 26.08 years) participated in this experiment either for partial course credit or $10.00. Community participants were recruited at the university fitness center, community rock climbing gym, community ice hockey league, and upper division psychology courses.

The college student sample was obtained from introductory psychology courses and the pre-screen for this group occurred using an online recruitment system. The pre-screening for community participants occurred using a paper-and-pencil survey. Only women athletes who considered themselves athletic, currently participated in sports (either competitively or recreationally) and considered working out an important part of who they were and scored at least 3.5 out of 5 on a three-item domain identification measure were recruited to participate. Domain identification items were based on Smith and White (2001) and included: "To what extent do you think athletics are important to the sense of who you are," "How important is it for you to be good at athletics," and "How much do you value being an athlete" to each of which participants responded on a 1 (not at all) to 5 (very much) scale.

Sports in the current study were classified as either masculine (e.g. snowboarding, basketball) or non-masculine (e.g. skiing, volleyball). Though including masculine, neutral, and feminine sports would be a desirable design, the current study lacked the necessary number of participants. As shown in Table 1, the design still allowed for the comparison of a stereotyped task (e.g. women in masculine sports) to a non-stereotyped task (e.g. women in non-masculine sports). Participants were blocked on type of sport, resulting in a 3 (feedback: positive, negative, none) x 2 (type of sport: masculine, non-masculine) between-subjects design.

Procedure

Participants were told that they were participating in a Sports Psychology study in collaboration with the Montana State University Human Performance Lab. They were

further told that the experiment was designed to refine the elements of a test to be used with incoming MSU athletes to gauge their potential athletic ability. In line with this cover story, the lab was set up to resemble a gym setting with athletic equipment including a jump rope, dumbbell weights, hand grips, and balance boards to give the illusion that any athletic test could have been chosen. A fake surveillance camera was mounted on the wall facing the participants' workstations. The room was divided into two sections (with a computer in each) so that two participants were able to complete the experiment at the same time, but participants were not able to see each other.

First, participants were given information on the cover story and completed the informed consent forms. They were told that the athletic-ability predictor test included two components that they believed were specific to their sport: a strength/endurance test and a balance/coordination test. Participants were led to believe that the athletic tests were chosen for them based on the sport they identified in the pre-screen process. In reality, all participants completed the same tests. Participants were provided with photo instructions for the athletic-predictor tests. Participants then took their heart rate and recorded this information. Next, participants engaged in an ambiguous athletic test (the hand grip) with the goal to squeeze the grip as many times as possible in 30 seconds (with dominant hand). A separate pilot study (n = 10) showed that the hand grip test tapped into strength and endurance and was viewed as useful to athletics.

In the next phase, participants received the manipulated performance feedback via an Instant Messenger from the alleged male evaluator in the next room. Immediately after receiving the manipulation, participants were told that they would be given a short

recovery period to return their heart rate to a baseline level. During that time, a word puzzle was provided in which participants unknowingly completed the measure of gender stereotype activation. Participants also completed a pre-experience survey before beginning Test 2, which assessed their level of commitment for the second phase of the athletics test.

Next, Test 2 was given using a balance disk in which participants stood on one foot as long as possible. Pilot testing (n = 10) showed that the balance disk test tapped into balance and coordination and was viewed as useful to athletics. Finally, the remaining dependent measures were administered using Media Lab software, assessing interest, future motivation, and state self-esteem. Participants were probed for suspicion on the manipulation, debriefed, and dismissed.

Independent Variables

Feedback Condition. Participants were randomly assigned to receive feedback via an experimentally created Instant Message system. Specifically, participants who received positive feedback were told "You received a score of 4.175 for execution. Your score is below average. Oops, I was looking at the jump ropes scores! Hold on. Your score is actually good." Participants who received shifting standards feedback were told that "You received a score of 4.175 for execution. Your score is below average. Oops, I was looking at the guys' scores! Hold on. Your score is actually good for a girl." The feedback was designed to be conversational to create a more real-life experience. Participants in the no feedback condition were not given performance feedback. Instead,

they were told "Okay, I am observing your performance from next door for the strength and endurance test. I can see that all of the equipment looks okay and wanted to confirm with you that the camera is working poorperly. Oops, I mean properly! The experimenter will now give you instructions for the next athletic ability predictor test." The instant message was played using an automated PowerPoint show including messages such as "Mark is typing a message" and "You have received a message from Mark." See Appendix B for an example of the last screen shown to participants.

Type of Sport. Pilot Study 2, along with past research on stereotyped tasks (Smith, Sansone & White, 2007), suggested that the type of sport a person identifies with is an important variable to study when trying to understand women athletes' experiences of stereotypes. The primary sport the participant identified with was collected during the prescreening process. These responses were coded for type of sport including masculine (e.g. basketball, snowboarding) or non-masculine (e.g. skiing, volleyball) using criteria established in Pilot Study 2 (see page 30). A total of 39 % of participants in masculine sports participated.

Stigma Consciousness. To assess stigma consciousness, ten items based on Pinel (1999) were included. Items were reworded to reflect stigma consciousness as a "female athlete." Responses were made using a 1 (strongly disagree) to 7 (strongly agree) scale (e.g. I worry about being judged as a "female" athlete). A median split was conducted for stigma consciousness (Median = 3.7), to separate the sample into high and low levels of stigma consciousness (See Appendix A for survey).

Dependent Variables

A list of major and ancillary dependent measures can be found in Table 2.

Table 2. List of Dependent Measures.

Major Measures	Ancillary Measures
Performance (Test 2)- Length in seconds on balance disk	State self-esteem
Motivation: self-ratings (interest, future motivation)	
Gender stereotype activation (word-fragment completion)	Achievement goal adoption

Performance: Test performance was measured for Test 2 using the balance disk. Performance on the test was defined as the length of time (in seconds) the participant was able to stay on the disk using one foot. The experimenter timed the participant using a hand-held stopwatch, but if the participant continued balancing past five minutes they were asked to stop to ensure enough time to complete the experiment.

Interest: The experience of interest is an important component of intrinsic motivation both immediately and in the future (Sansone & Smith, 2000). Interest was assessed as a post-test dependent measure. Following Test 2, four items assessed interest such as "I enjoyed doing this test very much" on a 1 (strongly disagree) to 7 (strongly agree) scale (Smith, Sansone & White, 2007).

Future Motivation: Future motivation was assessed as a post-test dependent measure using three items such as "How willing would you be in the future to do more

tests?" on a 1 (not at all willing) to 7 (very willing) scale (Smith, Sansone & White, 2007).

State Self-Esteem: To assess state self-esteem, participants completed a measure developed by Heatherton and Polivy (1991). The scale included seven items such as "I am confident about my abilities" which was anchored on a 1 (not at all) to 5 (extremely) scale.

Achievement Goal Adoption: The specific type of goal adopted (performance vs. mastery) as well as whether the orientation was approach vs. avoidance may differ based on the feedback received. The stereotype threat experience has been shown to lead to adoption of performance-avoidance goals, or the focus on not failing at a task (PAV; Smith, 2006; Smith, Sansone & White, 2007). To assess performance-avoidance achievement goal adoption, the performance-avoidance goal subscale from the Achievement Goal Questionnaire was included (Elliot & McGregor, 2001). Six items such as "I just wanted to avoid doing poorly on the test" were rated on a 1 (not at all) to 7 (very true of me) scale.

Exploratory Process Measure: Gender Stereotype Activation

To test the processes influencing the relationship between shifting standards feedback, performance, and motivation, a momentary gender stereotype activation measure was developed. This measure was modeled after the Asian race activation measure developed by Gilbert and Hixon (1991), as well as the word-completion task used to measure accessibility of group-relevant vs. irrelevant words (Knowles & Gardner,

2008). The new gender activation measure included 14 stereotype relevant items about women (Bem, 1974; Diekman & Eagly, 2000; Nelson. McEvoy & Schreiber, 1998) such as S E _ _ which could be completed in at least two ways (i.e. sexy or sent). Words were matched as much as possible for word frequency using HAL_log norms (Balota, Yap, Cortese, Hutchison, Kessler et al., 2007). Ten neutral words (Hutchison, Neely, Neill & Walker, 2004) were included to avoid contamination from previous responses. Word fragments were presented in a fixed random order between stereotyped words and neutral words (See Appendix C for survey).

Results and Discussion

Although participants were pre-selected for high domain identification as an athlete, it was important to ensure that domain identification did not differ between conditions. To test this, a 3 (Feedback condition) x 2 (Type of sport) ANOVA on domain identification was conducted. A feedback x type of sport interaction emerged for domain identification ($F[2,71] = 4.72$, $p < .05$), with participants in the positive feedback condition showing significant differences despite random assignment (Masculine sport: $M = 4.10$, $SE = .13$; Non-masculine sport: $M = 4.45$, $SE = .11$). As such, ANCOVAs were conducted in order to control for the participant's level of domain identification (Smith & White, 2002; Spencer et al. 1999; Brown & Josephs, 1999). Thus, domain identification will serve as a covariate in all subsequent analyses. Additionally, two participants were excluded from the analyses due to incomplete data (1) and failed manipulation check (1).

Effects on Gender Stereotype Activation

Type of Sport. To test the effects of feedback condition and type of sport on gender stereotype activation, a 3 (Feedback: shifting standards, positive, none) x 2 (Type of sport: masculine, non-masculine) ANCOVA was conducted using domain identification as a covariate. There was no main effect of type of sport. A main effect emerged for feedback condition ($F[2,66] = 6.68$, $p < .01$, $\eta_p^2 = .17$), such that shifting standards feedback elicited a higher proportion of stereotyped words when compared to positive and no feedback. However, this was qualified by a feedback condition x type of sport interaction ($F[2,66] = 4.04$, $p < .05$, $\eta_p^2 = .11$). In partial support of Hypothesis 1a, follow-up analyses showed that when no feedback was provided, women athletes who identified with masculine sports reported a higher proportion of gender stereotypic words compared to women athletes who identified with non-masculine sports ($F[1,67] = 3.80$, $p = .055$, $\eta_p^2 = .05$).

Stigma Consciousness: Next, a 3 (Feedback: shifting standards, positive, none) x 2 (Stigma consciousness: high, low) ANCOVA using domain identification as a covariate was conducted to test the effects of feedback condition and stigma consciousness on gender stereotype activation. There was a main effect for feedback condition ($F[2,66] = 6.29$, $p < .01$, $\eta_p^2 = .16$), such that shifting standards feedback elicited the highest proportion of gender stereotypic words compared to positive or no feedback. There was a trend for a main effect of stigma consciousness ($F[1,66] = 3.23$, $p = .077$, $\eta_p^2 = .05$). Women athletes high in stigma consciousness tended to report a higher proportion of

gender stereotypic words compared to women athletes low in stigma consciousness. As shown in Table 3, these main effects were qualified by a feedback condition x stigma consciousness interaction ($F[2,66] = 3.13$, $p = .05$, $\eta_p^2 = .09$).

In support of Hypothesis 1b, follow-up simple effect tests showed that gender stereotype activation differed only in the high stigma consciousness condition. Women athletes receiving shifting standards feedback reported a higher percentage of stereotyped words compared to positive or no feedback. Among high stigma conscious participants, shifting standards feedback resulted in higher gender stereotype activation than positive feedback or no feedback. There were no significant differences on gender stereotype activation among low stigma conscious participants.

Summary of Gender Stereotype Activation Results: When no feedback was provided, type of sport did result in higher gender stereotype activation for women in masculine sports compared to women in non-masculine sports. Results suggested that being in a domain in which there was not role congruity led to higher gender stereotype activation than a domain in which there was role congruity.

When stigma consciousness was entered as an independent variable, results suggested that participants high in stigma consciousness receiving shifting standards feedback showed increased gender stereotype activation compared to positive feedback and no feedback. Thus, women athletes who already worried about being stigmatized and received stereotyped feedback had the highest gender stereotype activation. It remained possible that gender stereotype activation was a mechanism for any found effects on

performance and motivation. First, however, it was necessary to test for any main effect and interactive effects among the conditions.

Effects on Performance

Performance data for the balance disk test was highly positively skewed (Skewness= 1.00). To account for the substantial skewness, a data transformation (Log 10 X + 1) was used to normalize the data (Tabachnick & Fidell, 2001). For ease of interpretation, both the transformed and raw data (in seconds) are presented.

Type of Sport: To test the effects of feedback condition and type of sport on performance, a 3 (Feedback: shifting standards, positive, none) x 2 (Type of sport: masculine, non-masculine) ANCOVA was conducted using domain identification as a covariate. There were no main effects of feedback condition or type of sport, and the feedback condition by type of sport interaction was not significant. Hypotheses about type of sport (2a, 3a, 4a) were not supported; performance on the athletic test was not significantly affected by the type of sport played or the performance feedback given.

Stigma Consciousness*:* To test the next set of hypotheses, a 3 (Feedback: shifting standards, positive, none) x 2 (Stigma consciousness: high, low) ANCOVA using domain identification as a covariate was conducted to test the effects of feedback condition and stigma consciousness on performance. There was no main effect of feedback condition on performance. A main effect of stigma consciousness emerged ($F[1,69] = 6.81$, $p < .05$, $\eta_p^2 = .09$), such that women athletes high in stigma consciousness performed worse than

Table 3. Outcome Measures as a Function of Stigma Consciousness and Feedback Condition.[1]

	Stigma Consciousness			
	High SC		Low SC	
Feedback condition	*M*	*SE*	*M*	*SE*
	Gender stereotype activation			
Shifting standards	74.8[a]	3.6	57.7[b]	4.5
Positive	56.5[b]	3.9	54.5[b]	3.6
None	52.0[b]	4.0	53.8[b]	3.7
	Performance			
Shifting standards	73.9[a]†	27.0	120.9[b]†	31.5
Positive	104.5[b]	28.7	84.0[b]	26.7
None	61.3[a]	27.5	123[b]	27.5
	Interest			
Shifting standards	4.00[a]	0.28	5.27[bc]	0.33
Positive	5.77[b]	0.30	5.65[bc]	0.28
None	4.89[c]	0.29	5.08[bc]	0.33
	Future motivation[2]			
Shifting standards	4.94	0.29	5.17	0.34
Positive	5.26	0.31	5.37	0.29
None	4.81	0.30	4.94	0.30
	Self-esteem[2]			
Shifting standards	3.92	0.13	4.27	0.15
Positive	4.25	0.14	4.17	0.13
None	4.19	0.13	4.23	0.13
	Performance-avoidance goal adoption[2]			
Shifting standards	3.91	0.36	3.62	0.42
Positive	3.38	0.38	3.56	0.35
None	3.34	0.36	3.03	0.36

Note. † $p = .066$. SC= Stigma consciousness. Performance scores are in seconds, but analyses were run using a log transformation. Gender stereotype activation reported in percent stereotype words completed. Interest & future motivation (Range = 1-7), Self-esteem (Range = 1-5), PAV goal adoption (Range = 1-7).

[1] Means within each dependent measure not sharing a superscript differ at $p < .05$.
[2] No significant interactions for future motivation, self-esteem, or PAV goal adoption.

women athletes low in stigma consciousness. Importantly, this was qualified by a feedback x stigma consciousness interaction ($F[2,69] = 3.80$, $p < .05$, $\eta_p^2 = .10$).

As seen in Table 3, simple effects follow-up tests showed that when women athletes received shifting standards feedback, having high stigma consciousness resulted in marginally worse performance compared to low stigma consciousness ($p = .066$). In support of Hypothesis 2b, results showed that among high stigma consciousness participants, shifting standards feedback resulted in significantly lower performance compared to participants who received positive feedback ($p < .05$). Further, among high stigma conscious participants, there was no difference in performance after receiving shifting standards feedback or no feedback, suggesting potential "implicit" stereotype activation (Smith & White, 2002). Feedback conditions did not significantly affect performance for low stigma consciousness participants.

To test Hypothesis 3b, simple-effect follow-up analyses were conducted. Hypothesis 3b was not supported; results revealed that when positive feedback was given, participants high and low in stigma consciousness performed statistically similar on the athletic test. Among high stigma consciousness participants, positive feedback resulted in higher performance than no feedback. However, positive feedback did not result in higher performance than shifting standards feedback.

To test Hypothesis 4b, simple-effect follow-up analyses were conducted. When no feedback was given, women low in stigma consciousnesses performed significantly better than women high in stigma consciousness. High stigma conscious participants

performed better when given positive feedback than when no feedback was given. Shifting standards feedback was statistically equal to no feedback.

Summary of Performance Results: In contrast to Study 1 findings, results did not show a significant relationship between feedback condition and type of sport on performance. These findings suggest that type of sport did not significantly alter performance. One possible reason is that the type of sport a woman plays does not influence how women experience performance feedback (but see Study 1). Another possibility for why type of sport did not produce any significant effects might be due to methodological problems in this study (i.e. using a masculine vs. non-masculine distinction instead of masculine vs. feminine). If this is the case, then future research should include masculine, neutral and feminine distinctions. Finally, a third possibility is that type of sport did not matter because participants were highly identified with multiple sports. The sport identification process used for Study 2 only allowed people to choose the sport they "identified with most." However, many participants listed more than one sport that they played competitively (M = 3.31 sports, SD = 1.91). Another possibility is that the ratio of male-dominated vs. female dominated sports may be a more accurate predictor given that many participants in our sample identified with multiple sports. Future research could investigate these possibilities.

When stigma consciousness was used as an independent variable, women athletes who were high in stigma consciousness and received shifting standards feedback showed low performance. Importantly, these same participants performed as well as the participants who also had high stigma consciousness and were given no feedback. This

suggests that an implicit stereotype about women in athletics was activated (Smith & White, 2002), such that receiving no feedback produced a similar performance to the shifting standards feedback among high stigma conscious participants. Only explicit, unqualified, positive feedback resulted in high performance for these women. However, women low in stigma consciousness performed equally well, regardless of the feedback condition.

Effects on Immediate Interest

Type of Sport: To test the effects of feedback condition and type of sport on immediate interest, a 3 (Feedback: shifting standards, positive, none) x 2 (Type of sport: masculine, non-masculine) ANCOVA was conducted using domain identification as a covariate. There were no main effects of feedback condition or type of sport, and the feedback condition x type of sport interaction was not significant. Hypotheses about type of sport (2a, 3a, 4a) were not supported; type of sport played and performance feedback provided did not significantly affect immediate interest.

Stigma Consciousness: Next, a 3 (Feedback: shifting standards, positive, none) x 2 (Stigma consciousness: high, low) ANCOVA using domain identification as a covariate was conducted to test the effects of feedback condition and stigma consciousness on immediate interest. There was a main effect for feedback condition ($F[2, 69] = 7.01$, $p < .01$, $\eta_p^2 = .17$). Follow-up tests showed that shifting standards feedback resulted in the lowest level of interest and positive feedback resulted in the high level of interest. The control condition (no feedback) fell in between shifting standards and positive feedback.

A trend for a main effect of stigma consciousness emerged (F[1,69] = 3.42, p = .069, η_p^2 = .05) such that women athletes low in stigma consciousness tended to report higher interest compared to women athletes high in stigma consciousness. Importantly, this was qualified by a significant feedback condition x stigma consciousness interaction (F[2,69] = 2.98, p = .057, η_p^2 = .08).

As shown in Table 3, simple effect follow-up tests showed that immediate interest varied as a function of stigma consciousness when provided with shifting standards feedback. In support of Hypothesis 2b, results showed that women athletes high in stigma consciousness reported significantly less interest than women athletes low in stigma consciousness. Among high stigma consciousness participants, positive feedback resulted in significantly higher interest than no feedback. Unlike the performance results, in which equal performance was observed in the shifting standards and no feedback conditions, shifting standards feedback resulted in significantly lower interest than no feedback. Similar to the performance results for low stigma conscious participants, different feedback conditions had no effect on interest.

Contrary to Hypothesis 3b, positive feedback did not elicit different levels of immediate interest as a function of stigma consciousness. However, as expected, participants high in stigma consciousness reported the highest level of interest when given positive feedback. Participants' level of interest when receiving no feedback did not significantly differ as a function of stigma consciousness. In support of Hypothesis 4b, participants high in stigma consciousness who received no feedback reported higher

interest than participants who received shifting standards feedback. But these participants on average reported lower interest than those receiving positive feedback.

Summary of Immediate Interest Results: Similar to performance, interest was not significantly affected by type of sport and feedback condition. However, when stigma consciousness was entered as an independent variable, the interest results followed a similar pattern to the performance results. Similar to the results for performance, feedback condition did not influence interest for women low in stigma consciousness. For women high in stigma consciousness in the shifting standards condition, interest levels were low compared to all other conditions. However, the high stigma conscious/shifting standard condition was significantly lower than the high stigma conscious/no feedback condition. Thus, it did not appear that an implicit stereotype about women athletes was operating in the no feedback condition, when examining effects for interest. Nevertheless, it can be determined that for women athletes who are the most sensitive about their stigma, shifting standards feedback had negative effects for both performance and interest.

Effects on Future Motivation

Type of Sport: To test the effects of feedback condition and type of sport on future motivation, a 3 (Feedback: shifting standards, positive, none) x 2 (Type of sport: masculine, non-masculine) ANCOVA was conducted using domain identification as a covariate. There were no main effects of feedback condition or type of sport, and the

feedback condition x type of sport interaction was not significant. Hypotheses about type of sport (2a, 3a, 4a) were not supported.

Stigma Consciousness. Next, a 3 (Feedback: shifting standards, positive, none) x 2 (Stigma consciousness: high, low) ANCOVA using domain identification as a covariate was conducted to test the effects of feedback condition and stigma consciousness on future motivation. There were no main effects of feedback condition or stigma consciousness, and the interaction between feedback condition and stigma consciousness was not significant. Hypotheses about stigma consciousness (2b, 3b, 4b) were not supported.

Summary of Future Motivation Results: Similar to the results for performance and interest, future motivation was not affected by the type of sport played or the performance feedback given. In contrast to the results for performance, interest, and gender stereotype activation, future motivation was not affected by stigma consciousness or the performance feedback given. However, there was a correlation between interest and future motivation ($r = .34$, $p < .01$), suggesting that immediate interest and future interest are related. Thus it is possible that over time, any negative effects on immediate interest could translate to lower levels of future motivation. Certainly, it is also possible that the self-report nature of the future motivation measures may not have been sensitive enough to uncover any effects in this study as has been found in past research (Smith, Sansone & White, 2007).

Effects on Ancillary Variables: State Self-Esteem and Performance-Avoidance Goal Adoption

Type of Sport: To test the effects of feedback condition and type of sport on state self-esteem and achievement goal adoption, two separate 3 (Feedback: shifting standards, positive, none) x 2 (Type of sport: masculine, non-masculine) ANCOVAs were conducted using domain identification as a covariate. There were no main effects of feedback condition or type of sport, and the feedback condition x type of sport interaction was not significant for either of the ancillary variables.

Stigma Consciousness: Next, separate 3 (Feedback: shifting standards, positive, none) x 2 (Stigma consciousness: high, low) ANCOVAs using domain identification as a covariate were conducted to test the effects of feedback condition and stigma consciousness on state self-esteem and goal adoption. There were no main effects of feedback condition or stigma consciousness, and the feedback condition x stigma consciousness interactions were not significant.

Summary of Self-Esteem and Performance-Avoidance Goal Adoption Results: Results suggested that type of sport and feedback condition did not significantly affect state self-esteem or goal adoption. Crocker and Major (1989) predicted that self-esteem would be protected when the performance feedback could be attributed to an external source (i.e. sexism). But, state self-esteem was not significantly affected by stigma consciousness or the performance feedback given. The finding that positive feedback did not enhance self-esteem compared to no feedback is not in line with past research

(Sansone, 1989; Thoman & Sansone, in press). Nonetheless, this pattern of results suggests that the effects on performance and interest cannot be accounted for by differences in self-esteem.

Goal adoption also did not differ by condition. This suggests that effects of shifting standards feedback are different from stereotype threat effects, which have been repeatedly demonstrated to lead to performance-avoidance goal adoption (Smith, 2006; Smith, Sansone & White, 2007).

Relationships Among Variables

Table 4 displays the correlations among the primary and ancillary dependent measures. Because gender stereotype activation was not significantly correlated with performance, stereotype activation could not be tested as a process variable in this study. However, performance was positively related to immediate interest. As predicted, immediate interest was positive correlated with future interest. State self-esteem was positively related to both performance and immediate interest. However, self-reported future interest may not be reliable and future studies might employ a different measure of this variable.

Overall Summary

Results of Study 2 suggested that type of sport did not significantly affect motivation or performance in this study. However, a pattern emerged for women athletes high in stigma consciousness who received shifting standards feedback, such that performance and intrinsic motivation decreased when compared to other feedback

conditions. In contrast, women athletes low in stigma consciousness did not suffer performance or motivational decrements. Although it is possible that an implicit stereotype about women athletes was acting in the no feedback condition, these results were mixed. Gender stereotype activation was highest for the high stigma consciousness/shifting standards feedback condition, which is suggestive that gender stereotype activation could serve as a process variable. However, the correlation between stereotype activation and performance was non-significant, so stereotype activation could not be tested as a process variable in this study.

Table 4. Correlations between Gender Stereotype Activation Performance, Immediate and Future Interest, State Self-Esteem and Performance-Avoidance Goal Adoption.

	1	*2*	*3*	*4*	*5*	*6*
1. Gender stereotype activation	1	-.10	-.27*	-.09	-.12	-.02
2. Performance		1	.24*	.11	.23*	.13
3. Immediate interest			1	.34*	.33*	-.12
4. Future interest				1	.10	.09
5. State self-esteem					1	.16
6. Performance-avoidance goal adoption						1

*Correlation is significant at $p < .05$.

GENERAL DISCUSSION

The purpose of the current project was to investigate the consequences of shifting standards feedback on women athletes. Shifting standards is a judgment informed by a stereotype that results in stereotyped groups being judged against different standards than non-stereotyped groups. Thus, positive feedback that is qualified by gender may be experienced negatively. In the context of athletics, two studies were conducted to examine the experience of shifting standards for women athletes. Specifically, this project manipulated performance feedback given for an athletic test to determine effects on performance and motivational processes and outcomes. Interactive effects of the manipulated feedback with participants' trait level of stigma consciousness were observed on performance and motivation, such that women athletes high in stigma consciousness who received shifting standards feedback experienced high gender stereotype activation, low performance, and low immediate interest compared to women high in stigma consciousness receiving positive feedback.

Specifically, Study 1 recruited skiers and snowboarders to test whether anticipated athletic performance was affected by shifting standards feedback and the masculine nature of the sport. Study 1 results showed that women snowboarders (a more male-dominated sport) reported less confidence in their anticipated future performance after receiving shifting standards feedback compared to positive feedback. However, women skiers (a more female-dominated sport) reported equal confidence in their anticipated future performance when receiving positive or shifting standards feedback. The type of sport distinction was further tested in Study 2, along with the potential

moderating role of stigma consciousness. Stigma consciousness is the extent to which an individual is chronically aware of their stigmatized status. Thus, the goal of Study 2 was to test whether actual athletic performance and self-reported motivation was influenced by chronic stigma consciousness and the nature of the feedback provided.

In Study 2, domain identification as a woman athlete was accounted for by pre-selecting only highly identified women athletes and statistically controlling for domain identification. Construction of the shifting standards feedback included the four defining features shown in past research: disconnect between objective and subjective judgments, between-groups comparison, a zero-sum behavioral component, and tests of stereotype assimilation (Bienat & Manis, 1994; Phelan et al., 2008; Biernat et al., 2009).

Study 2 results showed that women athletes high in stigma consciousness who received shifting standards feedback showed high gender stereotype activation, low performance, and reported low immediate interest compared to women high in stigma consciousness who received positive feedback. When no feedback was provided, similar effects emerged for performance among women athletes low in stigma consciousness. This suggests that the effect of shifting standards feedback may function at both implicit and explicit levels (Smith & White, 2002). However, this explanation should be interpreted with caution, given that the pattern did not replicate for immediate interest and gender stereotype activation. Importantly, effects on performance and motivation were not due to the participants' level of self-esteem following the stereotyped feedback or to level of commitment to the test. These results demonstrate that shifting standards feedback is a unique phenomenon because negative consequences resulted from positive

feedback qualified by gender.

The current research goes beyond stereotype threat effects in two ways. First, shifting standards feedback did not result in performance-avoidance goal adoption in this study. However, women high in stigma consciousness who received shifting standards feedback showed increased gender stereotype activation compared to all other conditions. Second, the information received in a stereotype threat situation is negative, but the feedback provided to the participants in the current research was positive. These findings suggest that the target's experience of prejudice was negatively impacted by positive (but qualified) feedback when sensitivity to a stigmatized identity was heightened (Pinel, 1999). However, a moderated meditational analysis did not support gender stereotype activation as a process variable for interest (which would be predicted by stereotype threat).

The mechanisms that produce stereotype threat and shifting standards appear to be different, and future research could investigate the processes that make each phenomenon distinct. One possibility is the role of anger in response to a shifting standards experience. In response to patronizing feedback, women (but not men) responded with anger (Vescio et al., 2005). As such, it is possible that emotions (e.g. anger) might interfere with the ability to concentrate to produce low performance. However, separate processes may be working to produce low levels of performance and motivation. Gender stereotype activation may prompt differential goal adoption in response to a shifting standards experience. The STEP model (Smith, 2004) suggests that performance-avoidance goals lead to low levels of interest in the task. As such, it is

possible that gender stereotype activation could result in low interest through PAV goal adoption. Although the process by which shifting standards feedback does its damage is not yet understood, it is clear that consequences on performance and motivation are negative.

Thus, the current project combined past research on patronizing feedback, stereotype threat, the motivational model for stereotyped tasks, and shifting standards. Additionally, this study offered a unique contribution to the literature with its focus on positive (but qualified) feedback from an unknown evaluator. Previous research examining stereotype threat effects has focused primarily on the immediate performance outcomes (Steele & Aronson, 1995; Stone et al., 1999). Similarly, patronizing feedback has resulted in low performance in stereotyped domains (Vescio et al., 2005). More recently, long term non-performance outcomes have been included as dependent measures, including motivation and general well-being (Smith, Kausar & Holt-Lunstad, 2007; Schmader et al., 2004). The motivational model for stereotyped tasks suggests that both immediate and future interest in the task is negatively impacted by stereotype threat experiences (Smith, Sansone & White, 2007). Although much research has supported that shifting standards occurs (Biernat & Manis, 1994; Biernat, 2003; Paul & Smith, 2008; Phelan et al., 2008), Study 2 is the first study to our knowledge that demonstrates the effect of a "shifting standards" experience for the target. This study adds to the shifting standards literature in understanding how it feels to be the victim of a shifting standard.

Limitations and Future Directions

Study 2 showed that shifting standards feedback resulted in negative performance and motivational consequences only among women athletes high in stigma consciousness. Women athletes in masculine and non-masculine sports had similar levels of performance and motivation when receiving shifting standards feedback. This suggests that stigma consciousness was an important moderator, but that type of sport was not. A recent study conducted by Smith, Kausar, and Holt-Lunstad (2007) used both stereotyped domain and stigma consciousness as moderators in examining women's academic experiences in Pakistan. This design was not employed in the current research due to low sample size and difficulty recruiting participants who highly identified with masculine sports.

Based on Smith, Kausar and Holt-Lunstad (2007), future work should examine type of sport, stigma consciousness, and shifting standards feedback together, resulting in a 3 (Feedback: shifting standards, positive, none) x 2 (Type of sport: masculine, feminine) x 2 (Stigma consciousness: high, moderate) design. It would be useful to know if these factors interact to alter performance and motivation. On one hand, receiving shifting standards feedback and being high in stigma consciousness in a masculine sport might lead to even lower performance and motivation than either condition alone. For example, a woman ice hockey player high in stigma consciousness may interpret shifting standards feedback more negatively than a woman gymnast high in stigma consciousness because of the addition of a gender role misfit (Eagly & Diekman, 2005) where stigma consciousness already exists. Thus, the effects of shifting standards feedback may be

impacted by both stigma consciousness and stereotyped domain and warrants further investigation.

Future research could examine the motivational and performance consequences of shifting standards feedback compared to negative feedback. An empirical research question is whether shifting standards feedback results in more negative outcomes than explicit, negative feedback. It would be predicted that negative feedback would be as damaging to performance and motivation as shifting standards feedback. This finding would further elucidate how praise that is qualified by one's gender can lead to negative outcomes for highly-identified woman athletes.

Additionally, research on patronizing feedback suggests that "devalued praise" must be given by a person in power to be considered patronizing (Gervais & Vescio, 2007). However, in the current study, it is possible that the confederate was not perceived as an authority figure, even though he was seen wearing a lab coat and served in an evaluator role. For women athletes, a coach would serve as a high power individual, but a team member would serve as a peer. The effects of shifting standards feedback coming from a high-power vs. low-power source may carry different weight. It is predicted that shifting standards feedback coming from a high power source would negatively impact the participants more than from a low power source, especially if the observer was classified as an "expert" in the domain. Future research could test this by manipulating the status of the person providing the performance feedback.

Subsequent studies could also examine the terminology used to invoke a shifting standards experience. In the current study, the phrase "good for a girl" was used.

Although pilot testing supported the notion that this phrase was commonly heard and would lead to high external validity, it is possible that using the word "girl" was patronizing in and of itself. Given that this study is the first to examine shifting standards from the target's perspective, future research could manipulate shifting standards feedback to test the use of the words girl vs. woman vs. female. It is expected that there would be no difference in performance and motivation between feedback conditions that used the phrases: good for a girl, good for a woman, and good for a female. This finding would allow the current research to rule out the possibility that the results presented here are an artifact of the language used to invoke the shifting standards experience.

Past research on stereotyped tasks has used domains where men and women are directly compared with each other, such as men and women working toward the same promotion in the workplace. However, athletics is decidedly unique because it has distinct competitions where women compete against other women and men compete against other men. Thus, the current research can only be generalized to domains where men and women are judged separately. The current study used athletics as a stereotyped domain because of the visible increase in women's athletics since the Title IX legislation. Although the law dictates that men's and women's level of athletic participation must be equal, this does not necessarily mean that the *experience* of men and women athletes is the same. Consider the differences in women's and men's professional basketball salaries. In 2008, the WNBA players' salaries ranged from $34,500 to $97,500 (WNBA, 2008). However, the average NBA player's average salary was $5.2 million and the highest salary possible was $55.63 million (NBA, 2008). Additionally, an analysis of

Olympic television coverage from 1996-2006 suggested that men's sports received more air time (61.6%) compared to women's sports (38.4%) during the Winter Olympics (Billings, 2008). Although the Summer Olympics offered a more equal distribution of air time with men's sports (51.9%) showing a slight advantage over women's sports (48.1%), the women's sports most commonly shown during primetime hours included sports which sexualized the women's body (e.g. gymnastics, diving) and highlighted the "pretty" female (Billings, 2008; Jones, Murrell & Jackson, 1999). Women athletes may become targets of subtle prejudice, such as being the recipient of shifting standards feedback, thereby suffering negative outcomes.

Conclusion

The present research indicated that stigma consciousness moderated the experience of shifting standard feedback on women's gender stereotype activation, interest in the task, and to a lesser extent, performance on the task. Specifically, compared to women low in stigma consciousness, women high in stigma consciousness experienced greater gender stereotype activation, somewhat lower performance, and lower immediate interest in the athletics task. Most importantly, women high in stigma consciousness receiving shifting standards feedback showed higher gender stereotype activation, lower performance, and lower interest compared to women high in stigma consciousness receiving positive feedback. These findings have important implications for understanding how subtle forms of sexism can result in negative consequences for individuals in stereotyped tasks. In this case, performance outcomes were negatively

affected as in previous stereotype threat research, but non-performance outcomes including immediate interest and gender stereotype activation were also negatively influenced by high stigma consciousness and shifting standards feedback. This suggests that the process underlying the experience of shifting standards might be different than stereotype threat.

In sum, subtle forms of sexism can result in negative consequences for the target of shifting standards feedback. Moreover, this research is an initial step to suggest that the experience of shifting standards feels demoralizing, creates high sensitivity to one's gender, and diminishes performance and motivation. In a society that values women's voices and fights for women's rights, this work is a call to action to recognize that modern forms of sexism exist and are deserving of attention to continue working toward gender equity.

REFERENCES CITED

Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C. & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35,* 29-46.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39,* 445-459.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychology, 37,* 122-147.

Beilock, S. & McConnell, A. (2004). Stereotype threat and sport: Can athletic performance be threatened? *Journal of Sport and Exercise Psychology, 26,* 597-609.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162.

Bergman, M., Langhou, R.D., Palmieri, P., Cortina, L. & Fitzgerald, L. (2002). The (un)reasonableness of reporting: Antecedents and consequences of reporting sexual harassment. *Journal of Applied Psychology, 87,* 230-242.

Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist, 58,* 1019-1027.

Biernat, M., Collins, E., Katzarska-Miller, I. & Thompson, E. (2009). Race-based shifting standards and racial discrimination. *Personality and Social Psychology Bulletin, 35,* 16-28.

Biernat, M. & Kobrynowicz, D. (1999). A shifting standards perspective on the complexity of gender stereotypes and gender stereotyping. In Swann, W., Langlois, J., Gilbert, L. & Albino, L. (Eds.), *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence* (pp. 75-106). Washington, DC: American Psychological Association.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66*, 5-20.

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60*, 485-499.

Biernat, M. & Vescio, T. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin, 28,* 66-77.

Billings, A. (2008). Clocking gender differences: Televised Olympic clock time in the 1996-2006 summer and winter Olympics. *Television and News Media, 9,* 429-441.

Blau, F. D., & Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, *14,* 75–99.

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.

Brown, R.P. & Josephs, R.A. (2000). The importance of importance: The mathematics identification questionnaire. Unpublished manuscript: The University of Texas at Austin.

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.

Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin, 25*, 413- 423.

Crocker, J. & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96*, 608-630.

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), Handbook of social psychology (4th ed., pp. 504-553). Boston: McGraw-Hill.

Crocker, J. Voelkl, K., Testa, M. & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 60,* 218-228.

Dardenne, B., Dumont, M. & Bollier, T. (2007). Insidious dangers of benevolent sexism: Consequences for women's performance. *Journal of Personality and Social Psychology, 93,* 764-779.

Davies, P., Spencer, S., Quinn, D. & 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.

Davison, H. & Burke, M. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior, 56,* 225-248.

Diekman, A. & Eagly, A. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin, 26,* 1171-1188.

Diekman, A. & Hirnisey, L. (2007). The effect of context on the silver ceiling: A role congruity perspective on prejudiced responses. *Personality and Social Psychology Bulletin, 33,* 1353-1366.

Dion, K. (1975). Women's reactions to discrimination from members of the same or opposite sex. *Journal of Research in Personality, 9,* 294-306.

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111,* 3–22.

Eagly, A. (2007). Female leadership advantage and disadvantage: Resolving the contradictions. *Psychology of Women Quarterly, 31,* 1-12.

Eagly, A. & Diekman, A. (2005). What is the problem? Prejudice as an attitude-in-context. In Dovidio, J., Glick, P. & Rudman, L. (Eds.), *On the nature of prejudice: Fifty years after Allport*, (pp. 19-35). Malden, MA: Blackwell Publishing.

Elliot, A. & McGregor, H. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80,* 501-519.

Foddy, M. & Smithson, M. (1989). Fuzzy sets and double standards: Modeling the process of ability interference. In J. Berger, M. Zelditch & B. Anderson (Eds.), *Sociological theories in progress: New formulations* (pp. 73-99). Newbury Park, CA: Sage.

Gervais, S. & Vescio, T. (2007). The origins and consequences of subtle sexism. In A. Columbus (Ed.), *Advances in psychology research* (Vol. 49, pp. 137-166). Hauppauge, NY: Nova Science Publishers.

Gibert, D. & Hixon, G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509-517.

Glick, P. & Fiske, S. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70,* 491-512.

Glick, P. & Fiske, S. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender equality. *American Psychologist, 56,* 109-118.

Glick, P., Wilk, K. & Perreault, M., (1995). Images of occupations: Components of gender and status in occupational stereotypes. *Sex Roles, 32,* 565-582.

Hall, R. (2008). Sweating it out: The good news and the bad news about women and sport. In Chrisler, J., Golden, C. & Rozee, P. (Eds.), *Lectures on the Psychology of Women* (pp. 97-115). McGraw Hill: Boston.

Harackiewicz ,J. & Elliot, A. (1993). Achievement goals and intrinsic motivation. *Journal of Personality and Social Psychology, 54,* 904-915.

Heatherton, T. & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60,* 895-910.

Hodgins, H. & Knee, R. (2002). The integrating self and conscious experience. In R. Ryan & E. Deci (Eds.), *Handbook of Self-Determination Research.* Rochester, NY: University of Rochester Press.

Hodson, G., Dovidio, J. & Gaertner, S. (2002). Process in racial discrimination: Differential weighting of conflicting information. *Personality and Social Psychology Bulletin, 28,* 460-471.

Hutchison, K., Neely, J., Neill, T. & Walker, P. (2004). Is unconscious identity priming lexical or sublexical? *Consciousness and Cognition, 13*, 512-538.

Huddy, L. & Terkildsen, N. (1993). The consequences of gender stereotypes for women candidates at different levels and types of office. *Political Research Quarterly, 46*, 503-525.

Inzlicht, M. & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365-371.

Jones, R., Murrell, A. & Jackson, J. Pretty versus powerful in the sports pages. *Journal of Sport and Social Issues, 23*, 183-192.

Kaiser, C., Vick, B. & Major, B. (2006). Prejudice expectations moderate preconscious attention to cues that are threatening to social identity. *Psychological Science, 17*, 332-338.

Katz, I. & Hass, R. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 65,* 893-905.

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.

Kelley, H. (1972). Attribution in social interaction. In E.E. Jones, D.E. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior.* Morristown, NJ: General Learning Press.

Knowles, M. & Gardner, W. (2008). Benefits of membership: The activation and amplification of group identities in response to social rejection. *Personality and Social Psychology Bulletin, 34,* 1200-1213.

Kobrynowicz, D., & Biernat, M. (1997). Decoding subjective evaluations: How stereotypes provide shifting standards. *Journal of Experimental Social Psychology, 33*, 579-601.

Koestner, R. & Losier, G. (2002). Distinguishing three ways of being highly motivated: A closer look at introjection, identification, and intrinsic motivation. In R. Ryan & E. Deci (Eds.), *Handbook of Self-Determination Research.* Rochester, NY: University of Rochester Press.

Krendl, A., Richeson, J., Kelley, W. & Heatheron, T. (2008). The negative consequences of threat: A functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science, 19,* 168-175.

Lyness, K. & Thompson, D. (2002). Climbing the corporate ladder: Do female and male executives follow the same route? *Journal of Applied Psychology, 85,* 86-101.

Major, B. & O'Brien, L. (2005). The social psychology of stigma. *Annual Review of Psychology, 56,* 393-421.

Martin, C. L. (1999). A developmental perspective on gender effects and gender concepts. In W. Swann, J. Langlois, & L. Gilbert (Eds.), *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence (*pp. 45-73). Washington, DC: American Psychological Association.

Maume, D. J., Jr. (2004). Is the glass ceiling a unique form of inequality? Evidence from a random-effects model of managerial attainment. *Work and Occupations*, *31,* 250–274.

National Basketball Association (2008). Retrieved April 2, 2009 from http://www.nba.com.

National Collegiate Athletic Association. (2009). Retrieved April 2, 2009 from http://www.ncaa.com.

National Sporting Goods Association. (2007). 2007 Women's participation. Retrieved January 18, 2009 from http://www.nsga.org.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). Word association, rhyme, and word fragment norms. Retrieved July 10, 2008, from University of South Florida Web site: http://w3.usf.edu/FreeAssociation.

Osborne, J. W. (1997). Race and academic disidentification. *Journal of Educational Psychology,* 89, 728-735.

Parsons, E. & Betz, N. (2001). The relationship of participation in sports and physical activity to body objectification, instrumentality, and locus of control among young women. *Psychology of Women Quarterly, 25,* 209-222.

Paul, D. & Smith, J.L. (2008). Subtle sexism? Examining vote preferences when women run against men for the Presidency. *Journal of Women, Politics and Policy, 29,* 451-476.

Phelan, J., Moss-Racuson, C. & Rudman, L. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly, 32,* 406-413.

Pinel, E. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76,* 114-128.

Pinel, E. (2004). You're just saying that because I'm a woman: Stigma consciousness and attributions to discrimination. *Self and Social Identity, 3*, 39-51.

Pinel, E. & Paulin, N. (2005). Stigma consciousness at work. *Basic and Applied Social Psychology, 27,* 345-352.

Rudman, L. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology, 74,* 629-645.

Sansone, C. (1989). Competence feedback, task feedback, and intrinsic interest: An examination of process and content. *Journal of Experimental Social Psychology, 25,* 343-361.

Sansone, C., & Smith, J.L. (2000). Interest and self-regulation: The relation between having to and wanting to. In C. Sansone & J.M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 341 – 404) San Diego, CA: Academic Press.

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.

Schmader, T., Major, B., Eccleston, C., McCoy, A. (2001). Devaluing domains in response to threatening intergroup comparisons: Perceived legitimacy and the status value asymmetry. *Journal of Personality and Social Psychology*, *80,* 782–796.

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

Smith, J.L. (2004). Understanding the process of stereotype threat: A review of meditational variables and new performance directions. *Educational Psychology Review, 16,* 177-206.

Smith, J.L. (2006). The interplay among stereotype threat, performance-avoidance goals, and women's math performance expectations. *Sex Roles, 54*, 287-296.

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.

Smith, J.L., Paul, D., & Paul, R. (2007). No Place for a Woman: Evidence for Gender Bias in Evaluations of Presidential Candidates. *Basic and Applied Social Psychology, 29,* 225-233.

Smith, J.L., Sansone, C. & White, P. (2007). The Stereotyped Task Engagement Process Process: The role of interest and achievement motivation. *Journal of Educational Psychology, 99,* 99-114.

Smith, J.L., Kausar, R. & Holt-Lunstad, J. (2007). Stigma consciousness in the classroom: A comparison of Pakistani women's motivation and well-being in science and non-science fields of study. *Sex Roles, 57,* 3-13.

Smith, J. L., Morgan, C., & 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.

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

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

Steele, C. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797-811.

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

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.

Stockdale, M. S. (1998). The direct and moderating influences of sexual harassment pervasiveness, coping strategies, and gender on work-related outcomes. *Psychology of Women Quarterly*, *22*, 521–535.

Swim, J. Aikin, K., Hall, W. & Hunter, B. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology, 68,* 100-214.

Tabachnick, B. & Fidell, L. (2001). Using multivariate statistics (4^{th} ed.) Boston: Allyn & Bacon.

Thoman, D. & Sansone, C. (in press). Getting different feedback for similar work: How benefiting or being hurt by biased feedback affects activity interest, self-esteem and future motivation. *Journal of Personality and Social Psychology.*

U.S. Bureau of Labor Statistics (2008). Women in the labor force: A data book. U.S. Department of Labor.

U.S. Department of Labor (2009). Title IX: Educational Amendments of 1972. Retrieved January 16, 2009 from http://www.dol.gov,

Vescio, T. K., Gervais, S., Snyder, M., & Hoover, A. (2005). Power and the creation of patronizing environments: The stereotype-based behaviors of the powerful and

their effects on woman performance in masculine domains. *Journal of Personality and Social Psychology, 88*, 658-672.

Vescio, T., Snyder, M. & Butz, D. (2003). Power in stereotypically masculine domains: A social influence strategy x stereotype match model. *Journal of Personality and Social Psychology, 85,* 1062-1078.

Women's National Basketball Association (2008). Retrieved April 2, 2009 from www.wnba.com.

APPENDICES

APPENDIX A

STIGMA CONSCIOUSNESS QUESTIONNARE

Stigma Consciousness Questionnaire for Female Athletes
(Based on Pinel, 1999)

Instructions: Indicate the extent to which you agree with each item on the following scale.

1	2	3	4	5	6	7
Strongly Disagree			Neither Agree or Disagree			Strongly Agree

1. Stereotypes about female athletes have not affected me personally (reverse)

2. I never worry that my behaviors will be viewed as stereotypical of female athletes. (reverse)

3. When interacting with people who know that I am a female athlete, I feel like they interpret all my behaviors in terms of being a female athlete.

4. Most male athletes do not judge female athletes on the basis of their gender. (reverse)

5. Being a female athlete does not influence how people act around me. (reverse)

6. I almost never think about being a female athlete. (reverse)

7. Most male athletes have a lot more sexist thoughts than they actually express.

8. I often think that male athletes are unfairly accused of being sexist. (reverse)

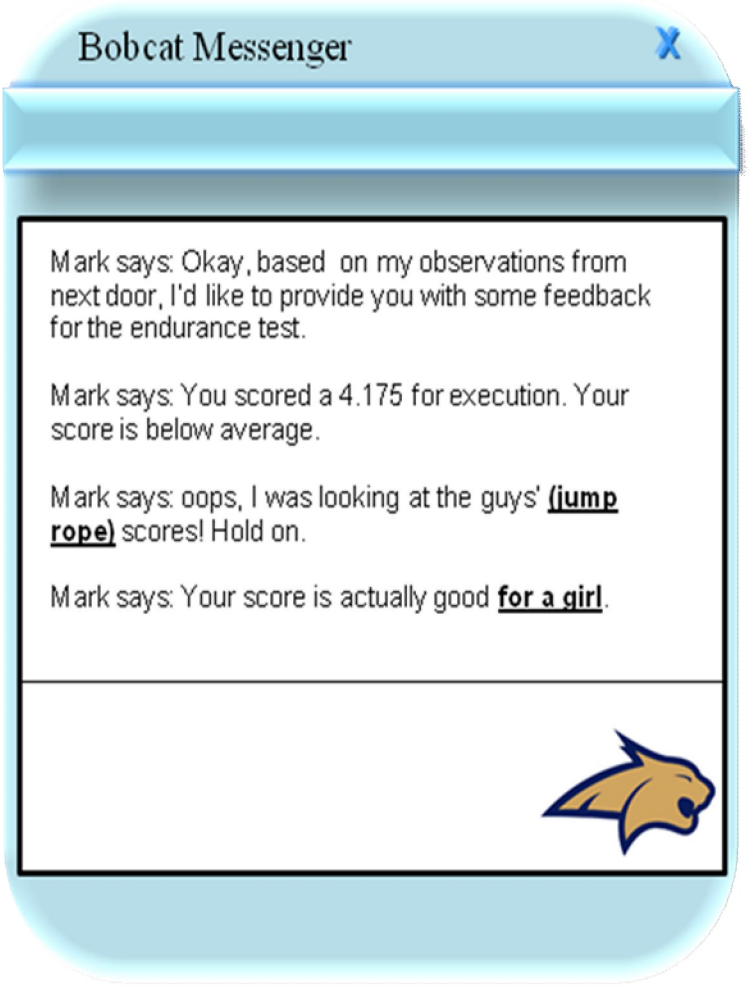
9. Most male athletes have a problem viewing female athletes as equals.

10. I worry about being judged as a female athlete.

APPENDIX B

FEEDBACK MANIPULATION

Feedback Manipulation Using a One-Way Instant Messenger.



Manipulation of feedback was delivered through an instant messenger from the evaluator in the next room. Feedback was either shifting standards, positive, or no performance feedback. Differences between shifting standards and positive feedback are shown in bold.

APPENDIX C

GENDER STEREOTYPE ACTIVATION MEASURE

Gender Stereotype Activation Measure
(Items based on Nelson et al., 1999; Bem, 1974; Diekman & Eagly, 2000)

INSTRUCTIONS: The following is a list of word fragments. Please write down as many words as you can think of for each fragment (up to 5 words). You should go as quickly as possible, but try to spend about 15 seconds on each item number.

1. H A _ _ (**hair**= 10.66, hall= 10.42, half= 11.18, 24 completions)

2. S E _ _ (**sexy**= 8.73, send= 12.69, seen= 12.00, 16 completions)

3. _ I N D (**kind**= 11.79, kind= 11.79, mind= 11.78, 7 completions)

4. _ A R M (**warm**= 9.95, farm= 9.29, harm= 9.42, 3 completions)

5. _ O V E (**love**= 12.02, move= 11.29, 7 completions)

6. C U _ E (**cute**= 9.31, cube= 8.52, cure= 9.05, 3 completions)

7. F U _ _ Y (**fussy**= 6.27, funny= 10.51, 7 completions)

8. C O M P L _ _ _ (**complain**= 9.43, Alt= complete, 2 completions)

9. _ E A K (**weak**= 9.75, leak= 8.55, peak= 9.04, 3 completions)

10. R O M A _ _ E (**romance**= 8.65, romaine, 2 completions)

11. _ A S T E F U L (**tasteful**= 6.22, wasteful= 9.97, 3 completions)

12. C U R _ E D (**curved**= 7.63, cursed= 8.83, 4 completions

13. _ A S S I V E (**passive**= 8.43, massive= 9.50, 2 completions)

14. H E _ _ E R (**helper**= 7.62, healer= 8.04, 5 completions)

15. _ A T H (**math**= 9.82, bath= 8.63, path= 10.69, 6 completions)

NEUTRAL WORDS (Based on Hutchison et al., 2004)

16. _ _ I L D (i.e. child, 3 completions)

17. _ _ N D Y (i.e. candy, 5 completions)

18. _ _ _ N G E (i.e. orange, 4 completions)

19. _ _ N C E (i.e. ounce, 8 completions)

20. _ _ O D (i.e. good, 5 completions)

21. _ _ V E R (i.e. fever, 12 completions)

22. B E _ _ (i.e. bear, 14 completions)

23. C O _ _ (i.e. cold, 24 completions)

24. B A _ _ (i.e. baby, 18 completions)

25. P R O _ _ _ (i.e. profit, 6 completions)

Examples are included for each of the fragment completions, however not all fragment completions are listed. The possible number of completions for each item is listed in parentheses.