Psychology of Sport and Exercise 15 (2014) 48–55

Contents lists available at ScienceDirect

Psychology of Sport and Exercise

journal homepage: www.elsevier.com/locate/psychsport

"You throw like a girl:" The effect of stereotype threat on women's athletic performance and gender stereotypes

Kimberly Hively [1], Amani El-Alayli*

Psychology Department, Eastern Washington University, 151 Martin Hall, Cheney, WA 99004, United States

ARTICLE INFO

Article history:
Received 30 November 2012
Received in revised form
24 August 2013
Accepted 5 September 2013
Available online 15 September 2013

Keywords:
Stereotype threat
Athletic performance
Sports
Athletes
Gender
Gender stereotypes

ABSTRACT

Objectives: "Stereotype threat" occurs when people perform worse at a task due to the pressure of a negative stereotype of their group's performance. We examined whether female athletes may underperform at an athletic task if prompted to think about gender stereotypes of athleticism. We also explored whether gender stereotypes regarding general athletic ability would be affected by a standard stereotype threat induction.
Design: We used a 2 (participant gender) × 2 (stereotype threat manipulation) factorial design with task performance and gender stereotypes of athleticism as dependent measures.
Method: Female and male tennis and basketball college student athletes performed two athletic tasks relevant to their sport: a difficult concentration task and an easier speed task. Participants were told beforehand that (1) there was a gender difference on the tasks (to induce stereotype threat) or (2) there was no gender difference (to remove any preexisting stereotype threat).
Results: On the difficult task, women performed worse than men only when stereotype threat was induced. Performance on the easier speed task was unaffected by the stereotype information. Interestingly, women's beliefs regarding women's and men's *general* athleticism were also affected by the manipulation.
Conclusions: We concluded that one minor comment regarding a very specific athletic task may sometimes impair task performance and alter gender stereotypes of athleticism among women. Some implications for preventing negative stereotype threat effects are discussed.

Introduction

In 2008, Gretchen Bleiler attempted to be the first snowboarder to achieve a second consecutive gold medal in the Winter X Games superpipe competition. As Bleiler attempted an impressive jump with a 900° turn, she took a big spill. Could the pressure for female athletes to prove their athleticism result in such negative performance outcomes? The commentator described Bleiler as the "most visible woman in snowboarding" at the time. If Bleiler had won this second consecutive gold medal, she would have set a record among both females and males in her sport. Thus, it is possible that gender was on her mind during the competition. We propose that when female athletes think about their gender, societal stereotypes regarding lower female athleticism could become salient, thus impacting their performance.

Stereotype threat

Researchers studying a phenomenon known as "stereotype threat" have demonstrated that people can underperform at tasks when thinking about the negative performance expectations for their group (Steele, 1997, 1998). For example, a woman's performance on a difficult math test may suffer if she is told that women tend to underperform in math or on that particular test. Thoughts about the negative gender stereotype may cause the woman to worry that her performance, if poor, would verify the negative stereotype of her group. Consequently, she may become particularly motivated to disprove the stereotype. Unfortunately, this excessive concern about performance can sometimes impair actual performance outcomes (e.g., O'Brien & Crandall, 2003; Spencer, Steele, & Quinn, 1999).

To support this notion, Spencer et al. (1999) found that female college students who were proficient in math performed worse on

* Corresponding author. Tel.: +1 509 359 7088; fax: +1 509 359 4366.
E-mail address: amani@ewu.edu (A. El-Alayli).
[1] Now practicing as a School Psychologist for the North Thurston Public School District in Lacey, WA, United States.

1469-0292/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.psychsport.2013.09.001

a difficult math test than their male counterparts when told that there was "a gender difference" on the test. However, the women and men performed equally well when told that there was no gender difference on the math test. It is noteworthy that a task must be quite difficult in order for stereotype threat cues to impair performance (O'Brien & Crandall, 2003). In fact, easy tasks sometimes result in improved performance under the pressure of a negative stereotype (O'Brien & Crandall, 2003).

Negative stereotype threat effects may occur even when no explicit stereotype regarding gender expectations is presented (Ben-Zeev, Fein, & Inzlicht, 2005). For example, Ben-Zeev et al. believe that stereotype threat is the "default" experience when women who care about math are taking a math test in a mixed-sex setting. If the same is true in the context of sports, then female athletic performance may regularly suffer due to the prevalent stereotypes about lower female athleticism. Thus, it may be important to remove already existing gender stereotypes by describing athletic tasks as gender-neutral.

Gender stereotypes in sport

Perceptions of lower female athleticism are pervasive (see Chalabaev, Sarrazin, Fontayne, Boiché, & Clément-Guillotin, 2013, for a recent review). Gender stereotypes are evident in media coverage of sports (e.g., Knight & Giuliano, 2001); referees' calls during games (Souchon, Coulomb-Cabagno, Traclet, & Rascle, 2004); and funding of athletic programs (Hardin, Simpson, Whiteside, & Garris, 2007). They are also evident in teachers' and parents' views and treatment of girls and boys. Although a little research finds that physical education teachers give more encouragement feedback to girls than to boys (Nicaise, Bois, Fairclough, Amorose, & Cogèrino, 2009), most studies have shown that physical education teachers interact more with boys (e.g., Duffy, Warren, & Walsh, 2001; MacDonald, 1990) and are more encouraging of boys' involvement in sport (see Cann, 1991, for a review). Physical education teachers also have gender-biased performance expectations in sports which are inconsistent with real group differences (Chalabaev, Sarrazin, Trouilloud, & Jussim, 2009). Parents display the same general patterns, favoring boys both in perceptions of athletic ability and in encouragement of their children's involvement in sport (Fredricks & Eccles, 2005).

It is not surprising, then, that boys have better perceptions of their athletic ability (e.g., Biddle, Atkin, Cavill, & Foster, 2011; Fredricks & Eccles, 2005; Hilland, Stratton, Vinson, & Fairclough, 2009) and greater motivation to participate in sports (e.g., Knisel, Opitz, Wossmann, & Keteihuf, 2009). Just thinking about the common expression, "You throw like a girl," conjures up an impression of women as unathletic. People seem to equate athleticism with masculinity rather than femininity (e.g., Fredricks & Eccles, 2005; Koivula, 1999).

Stereotype threat in sports

Negative beliefs regarding female athleticism may impede girls and females from performing to their true potential in sports contexts (Chalabaev et al., 2013). Interestingly, women need not endorse these gender stereotypes in order for their performance to suffer (Chalabaev, Sarrazin, Stone, & Cury, 2008). They may experience stereotype threat even if they disagree with the stereotype or believe that it does not apply to them personally (Steele, 1997) because their goal of disproving the stereotype may be present regardless. For example, a female athlete may have a coach who believes that men outperform women at her sport. Though the athlete may disagree, she may still fear that if she performs poorly in front of her coach, then that would (falsely) prove to the coach that the negative stereotype was true. Thus, the mere existence and reminder of gender stereotypes regarding athleticism may continuously harm female athletes' performance.

Stereotype threat in the context of sports has been investigated in only five prior studies (Beilock, Jellison, Rydell, McConnell, & Carr, 2006; Beilock & McConnell, 2004; Chalabaev et al., 2008; Stone, Lynch, Sjomeling, & Darley, 1999; Stone & McWhinnie, 2008), and all but one investigated the same athletic task, golf putting. These researchers have focused primarily on the effects of gender stereotypes on male athletes or the effect of racial stereotypes on the athletic performance of different racial groups. Although these issues are important, the prevalence of stereotypes regarding female athleticism warrants investigation on how gender stereotypes may impair *women's* athletic performance. Only two studies have examined this issue (Chalabaev et al., 2008; Stone & McWhinnie, 2008). The present study builds on this past work, which sets the foundation for an important new focus in the stereotype threat literature.

Stone and McWhinnie (2008) examined golf performance in a sample of White female college students who were novice golfers but were at least somewhat athletic. The participants' task was to putt a golf ball into one of three holes in each of eight different putting mat setups. Participants were to aim specifically for the smallest of the three different-sized holes, and the total number of strokes required to sink all eight balls was tallied. Participants took more strokes (performed worse) when the task was initially described as one with a gender difference in performance, as opposed to one with a racial difference or one in which no group differences were specified.

The same study was the first to discover that even a subtle stereotype threat cue may impair female athletes' performance. Stone and McWhinnie (2008) found that the presence of a male experimenter reduced their female participants' performance "accuracy," which was defined as the number of times (out of eight) that participants sunk the ball specifically into the smallest of the three holes. Thus a blatant cue, the mention of a gender difference in performance, affected only the total number of strokes, whereas a subtle cue, the presence of a male experimenter, affected only the accuracy of the final putt. The authors supported their proposal of a "dual process" model, which suggests that threat cues may operate independently and affect different types of performance outcomes.

The authors explained that blatant stereotype threat cues, such as the mention of a gender difference in performance, may disrupt performance on only tasks requiring fluid, continuous motions, whereas more subtle cues may disrupt performance on only tasks requiring careful concentration. Their perspective is that a blatant stereotype threat cue causes individuals to become prevention-focused, due to a fear of failure. This would then result in an overly conservative approach, such as taking smaller strokes in a putting task in order to get the ball closer and closer to the hole, which would disrupt overall performance on a continuous task. However, careful concentration tasks may be more disrupted by subtle stereotype threat cues, given that they are more ambiguous and thus more likely to consume some of the necessary working memory for the task at hand (Stone & McWhinnie, 2008).

Other researchers have shown that blatant (gender) stereotype threat *can* reduce athletic performance on careful concentration tasks that also involve golf putting (Beilock et al., 2006; Beilock & McConnell, 2004). One difference may be that Stone and McWhinnie tested novice golfers, whereas the participants in the other studies were expert athletes in the relevant sport. Perhaps expert athletes are more globally impacted by stereotype threat cues. The strongest stereotype threat effects tend to occur for individuals who identify strongly with the domain in which they are negatively stereotyped (Leyens, Désert, Croizet, & Darcis, 2000;

Steele, 1997, 1998; Stone, 2002; Stone et al., 1999). Such individuals may be particularly worried about how their performance, and that of their group, is being judged. Aside from the added pressure, experts' performance may be especially harmed by negative stereotypes because their well-practiced, automatic, fluid movements become disrupted when they over-think them in an attempt to perform well and disprove stereotypes (Beilock et al., 2006). It is possible that subtle cues may still have stronger effects than more blatant cues on careful concentration tasks, as Stone and McWhinnie (2008) suggest, but perhaps either could result in negative outcomes for athletes whose identity is linked with the relevant sport.

Thus far, only Chalabaev et al. (2008) have investigated how gender stereotypes may negatively affect expert female athletes' performance. In their study, competitive female soccer players in France used one foot to dribble a soccer ball through a difficult slalom course. The athletes performed the task less successfully if they were told that the task was an assessment of natural athletic ability pertaining to strength, speed, and power, than if told that the task measured psychological factors. The assumption is that the description of the task as assessing natural athletic ability would prompt female athletes to automatically think about gender stereotypes regarding women's athleticism. As such, their desire to prove the stereotypes wrong would backfire, causing them to underperform at the challenging task. Interestingly, describing the task as assessing technical soccer ability had no such effect, though there was a marginally significant trend in the same direction. This suggests that female athletes may be more threatened by the conceptualization of athletics as requiring natural athletic ability, as opposed to learned technical skills. One reason may be the lower perceived malleability of natural athletic prowess.

The present study

Thus, past work has begun to unravel some interesting complexities regarding when stereotype threat cues may have a negative impact on females performing athletic tasks. Our goal was to contribute to this line of research by comparing female athletes' performance against that of male athletes under two circumstances: when a stereotype reminder was present and when the threat of a default stereotype was specifically removed. Like the original work examining gender stereotype threat and math performance, we included male participants in our study in order to see whether between-gender comparisons in athletics resembled those observed on math tasks. We also build on prior work by examining basketball and tennis, two sports which had not yet been examined in stereotype threat research. Like Chalabaev et al. (2008), we included participants who were athletes within the specific sport in question, due to the potential for stronger and more global decrements in performance among the female athletes. It is also important to investigate this group because these athletes would have more at stake in the real world if their athletic performance was influenced by negative stereotypes.

We also examined whether stereotype threat manipulations influence athletes' general perceptions of men's and women's ability in sports. If athletes are told that there is a gender difference in performance on a specific athletic task, would they generalize that information and come to believe that men are generally better athletes than women? Similarly, if they are told that there is no gender difference on a particular task, do they come to believe that men and women perform equally well in sports in general? If so, then the pervasive subtle reminders of gender differences may have continuous and perhaps cumulative effects on female athletes' performance. A study by Chalabaev, Sarrazin, and Fontayne (2009) found that the mere belief that girls are poor soccer players caused junior high school girls to actually perform worse at a soccer task. Specifically, the general stereotype affected the girls' own perceived ability, which ultimately impaired their performance.

In the current study, we had male and female basketball and tennis athletes perform two athletic tasks within their sport. Participants were initially told either that there was a gender difference in performance or that there was no gender difference in performance on the tasks. The first sports task was a challenging task involving careful concentration. The second task was not as difficult, but was performed under time-pressure. This supplemental task was included to investigate whether time pressure could make an easier task challenging enough for stereotype threat effects to occur. We also examined whether our sample of athletes believed that men outperform women in athletics, and if such beliefs were influenced by the specific stereotype threat manipulation.

We hypothesized that female athletes would perform worse than similarly skilled athletes when a stereotype threat cue was provided than when stereotype threat pressure was lifted. This would provide a conceptual replication of the math performance findings of Spencer et al. (1999) and others, but in the context of athletic performance. It would also provide a conceptual replication of the athletic performance findings of Stone and McWhinnie (2008) and Chalabaev et al. (2008) but with a focus on comparing female and male athletes, and with the inclusion of a condition in which default gender stereotypes are specifically removed. Our use of two different athletic tasks addresses the potentially important issue of task difficulty, as raised by O'Brien and Crandall (2003). Our assessments of gender stereotypes in athletics were included on an exploratory basis to examine whether general attitudes may be influenced by a stereotype threat prime.

Method

Participants

Participants were 61 athletes and students (M_{age} = 21) at a northwestern university. They consisted of 17 basketball players (10 male and 7 female) and 13 tennis players (6 male and 7 female) from the National Collegiate Athletic Association (NCAA) divisions, as well as 33 other students (15 male and 16 female) who identified themselves as basketball or tennis athletes. The coaches of the NCAA teams assisted us in recruiting and scheduling their players for participation. Non-NCAA players were recruited from intramural teams, psychology courses, and open-gym. The psychology students were either pre-screened for their athletic status (via a written survey asking them if they identified as tennis/basketball athletes) or recruited via flyers that specifically requested basketball and tennis athletes. Open-gym participants were approached directly while they were playing or practicing their sport. Participants were primarily volunteers, but psychology students received course credit as compensation.

Design

A 2 (participant sex) × 2 (threat condition) between-subjects factorial design was used. Participants performed either basketball or tennis tasks and were either NCAA or non-NCAA athletes within that sport. Because our sample size was limited by our decision to include athletes only, we combined the data across the different types of athletes (basketball and tennis NCAA and non-NCAA athletes) in order to garner more statistical power. The threat and no threat conditions were alternated separately within each of eight groups that were created based on participant gender (female or male), athlete level (NCAA or non-NCAA) and athlete

type (basketball or tennis). Thus, the first participant in any of the eight groups (e.g., female non-NCAA basketball players) would be placed in one threat condition, and then the next participant within that same group would be placed in the other threat condition, and so on, so that each group had an even split between no-threat and threat-exposed participants. There were two dependent measures: performance on a difficult concentration task and performance on an easier speed task.

Materials and procedure

A female experimenter escorted one player at a time to an available basketball or tennis court, which served as the location for the study. The experimenter had one or two assistants to help her fetch balls during the study.

Threat manipulation

All participants were given a card that indicated that they would be performing two basketball/tennis tasks. In the threat condition, the card also read that the tasks have been shown to provide a good assessment of people's natural athletic ability, and that men and women perform at different levels on the tasks. In the no-threat condition, the card read the tasks have been shown to provide a good assessment of people's visual ability and that men and women perform equally well on them. Once participants finished reading the card, the experimenter (who was unaware of the experimental condition) provided the detailed instructions regarding how to perform the two sports tasks.

We were able to keep the experimenter blind to the threat condition by placing code numbers on the back of the instruction cards that contained the threat information. The experimenter handed the participant card "1" or "2" without knowing the content on the other side, and the participant proceeded to read the manipulation-relevant information silently. The experimenter wrote each participant's condition number on his or her score sheet for tracking purposes. The experimenter and her ball-retrieving assistants were able to remain unaware of which code number corresponded to which threat condition throughout the entirety of the study. An example card that was given to the participants read as follows:

> You will be asked to complete two difficult [basketball] shooting tasks. These two tasks have been shown to provide a good assessment of people's natural athletic ability. Research finds that there is a gender difference in performance on these tasks. Men and women perform at different levels on the specific basketball tasks you are about to complete. Task 1: Shooting baskets from different locations. Task 2: Shooting free throws within a time limit. Let the experimenter know when you are ready to begin.

Perceptions of gender differences in athleticism

Afterwards, participants completed a pre-task questionnaire, which consisted of items assessing participants' impression of gender differences in performance on each of the two sports tasks they were about to complete, as well as within that sport as a whole (e.g., "in tennis, in general"), and in sports, in general. Participants responded on a scale from 1, *Men perform much better than women*, to 7, *Women perform much better than men*, with the midpoint of the scale (4) labeled as *No gender difference in performance.*

Because perceptions of general basketball/tennis performance and general sports performance were highly correlated, $r(59) = .71$, $p < .001$, they were averaged to create a composite measure ($M = 2.46$, $SD = 1.01$). Perceptions of Task 1 performance gender differences ($M = 3.31$, $SD = 1.26$), and Task 2 performance gender differences ($M = 3.85$, $SD = 0.96$) were treated as separate measures. When one-sample *t*-tests were used to compare each of the three means against 4, the midpoint of the scale, only the first two comparisons yielded significant effects, $ts(60) = -11.97$ and -4.27, $ps < .001$. In other words, the sample viewed men as significantly better at sports and at performing Task 1 specifically, but did not view any gender difference in performance on Task 2. Participants made these estimates after the task descriptions and the threat manipulation, but before actually completing the athletic tasks. Thus, their estimates may have been influenced by what they were told, but not by how they performed.

Sports task and performance measures

There were two measures of task performance: Task 1 (difficult concentration task) and Task 2 (speed task) performance. For the basketball athletes, Task 1 involved shooting the basketball from specific locations around the 3-point line (see Fig. 1). Participants took two shots from each of four difficult locations and attempted to make as many shots as possible. For Task 2, participants retrieved basketballs from a basketball rack while trying to make as many free throws as they could within 30 s. The experimenter's assistants fetched and returned the balls during the task.

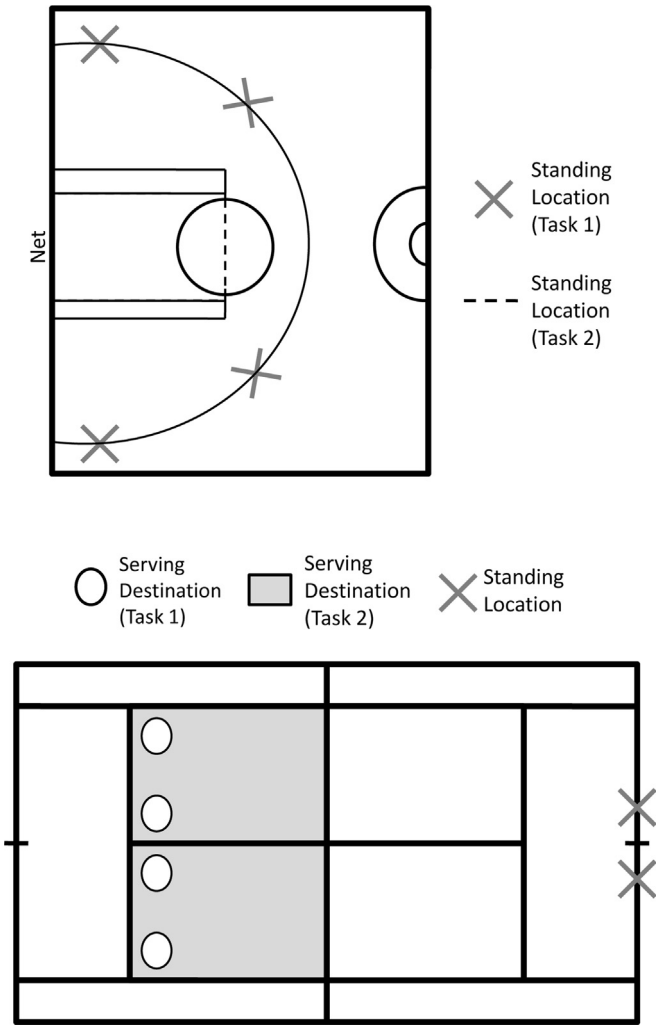

Fig. 1. Basketball and tennis task locations.

For the tennis athletes, Task 1 involved serving two tennis balls to four small targets (about 16 inches in diameter) on the other side of the court (see Fig. 1). For Task 2, participants served as many serves in bounds they could within 30 s. Participants pulled their own tennis balls from a bucket during the task. Participants were not allowed to practice either Task 1 or Task 2 before the testing began.

For Task 1, the total number of shots/serves made (out of 8) was summed for the basketball task ($M = 2.72$, $SD = 2.03$) and the tennis task ($M = 0.45$, $SD = 0.67$). For Task 2 (the timed task), the total number of shots/serves made in 30 s was recorded for the basketball task ($M = 5.95$, $SD = 3.60$) and the tennis task ($M = 6.36$, $SD = 2.28$). Because of the limited number of athletes within each sport, we standardized Task 1 and Task 2 scores within each sport and then combined the scores between the two sports. In other words, the basketball and tennis performance scores were changed to *z*-scores, so that the mean within each sport would be set to 0 and the standard deviation for each would be set to 1. This allowed us to equate basketball and tennis scores, making it appropriate to combine them into the same variable so that we would have more participants in each condition (*ns* = 14 to 17 per sex × threat condition).

Measures of sports investment, importance, and ability

To verify that our participants were athletes in the relevant sport, they completed a post-task questionnaire at the end of the study. It included estimates of their performance ability and playing frequency within the sport, how important it was for them to perform the sport well, and how much playing the sport was part of their life and their identity. The seven relevant items were averaged together ($M = 5.09$, $SD = 1.33$, $\alpha = .89$). There were an additional 10 items which assessed general athletic identity. These included items regarding perceptions of athletic ability, enjoyment of sports, frequency of playing sports, and investment in performing well ($M = 5.41$, $SD = 0.86$, $\alpha = .83$). Participants responded on a scale from 1, *not at all*, to 7, *extremely*. Both means were quite high (falling between "very" and "substantially" on the response scale), illustrating participants' strong athletic identification both within the relevant sport and in general.

Measure of suspicion

Suspicion was assessed by asking the participants to write what they thought was the purpose of the study. No responses reflected awareness of the hypotheses. Participants also responded to some written demographic questions as well as some supplemental questions not relevant to the present purpose. When the participants finished the study, they were thanked, debriefed, and asked not talk about the study's purpose or tasks with anyone who might still participate.

Results

Task performance

Task 1: difficult concentration task

A 2 (gender) × 2 (threat condition) Analysis of Variance (ANOVA) was performed on the Task 1 performance score. The main effects were not significant but the predicted interaction was, $F(1, 57) = 4.07$, $p = .048$ ($\eta^2 = .07$). As seen in Fig. 2, women performed significantly worse than men in the threat condition, $F(1, 57) = 4.71$, $p = .034$ ($\eta^2 = .14$), but not in the no-threat condition, $F(1, 57) = 0.42$, $p = .521$ ($\eta^2 = .01$).

Because performance on the first tennis task was significantly worse than performance on the first basketball task, $t(59) = 5.07$, $p < .001$, we also ran the ANOVA with sport entered in as a covariate, and found that the gender × threat interaction remained significant. We did not have enough data to split the analyses by sport type, but we do not believe that the effect was driven by one particular sport given that the pattern of means for the interaction was the same when examining performance scores within each sport separately.

Task 2: speed task

An ANOVA performed on Task 2 performance revealed only a gender main effect, $F(1,57) = 5.79$, $p = .019$ ($\eta^2 = .09$). Overall, men ($M = 0.29$, $SD = 0.78$) performed better than women ($M = -0.32$, $SD = 1.11$) on the task. Thus, participants' performance was affected by the stereotype threat manipulation on Task 1, but not on the easier speed task.

Perceptions of gender differences in athleticism

We also performed a 2 (gender) × 2 (threat condition) ANOVA on each of the three perceived gender differences measures. There were no significant effects on perceptions of gender differences in Task 1 performance. This is interesting considering that participants' actual performance was affected by what they were told regarding gender differences on the task.

The only significant effect on perceptions of gender differences in Task 2 performance was a main effect of participant gender, $F(1, 57) = 4.01$, $p = .05$ ($\eta^2 = .07$). Specifically, men perceived more of a gender difference in Task 2 performance ($M = 3.61$, $SD = 0.84$) than did women ($M = 4.10$, $SD = 1.03$), who viewed men and women as performing about equally.

When examining general stereotypes regarding gender differences in sports performance, there was a participant gender main effect, $F(1, 57) = 5.73$, $p = .020$ ($\eta^2 = .09$). Specifically, male participants perceived more of a gender difference favoring men's athletic performance ($M = 2.16$, $SD = 1.03$) than did female participants ($M = 2.77$, $SD = 0.90$). More importantly, there was a significant gender × threat condition interaction, $F(1, 57) = 5.04$, $p = .029$ ($\eta^2 = .08$). As seen in Fig. 3, men's and women's perceptions of gender differences in sport were similar when told there was a (non-specified) gender difference on the athletic tasks used in the present study and that the tasks assessed natural athletic ability. However, when told there was no gender difference and

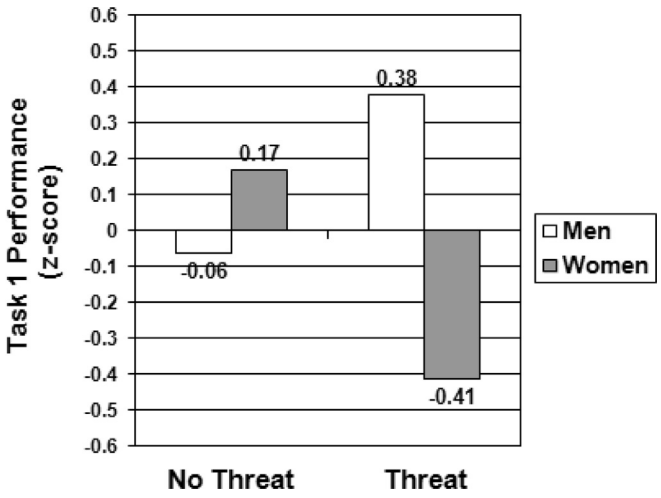

Fig. 2. Task 1 performance (standardized and combined across tennis and basketball tasks) as a function of stereotype threat and participant gender. A score of 0 reflects the average score in the group, with positive scores being above average and negative scores being below average.

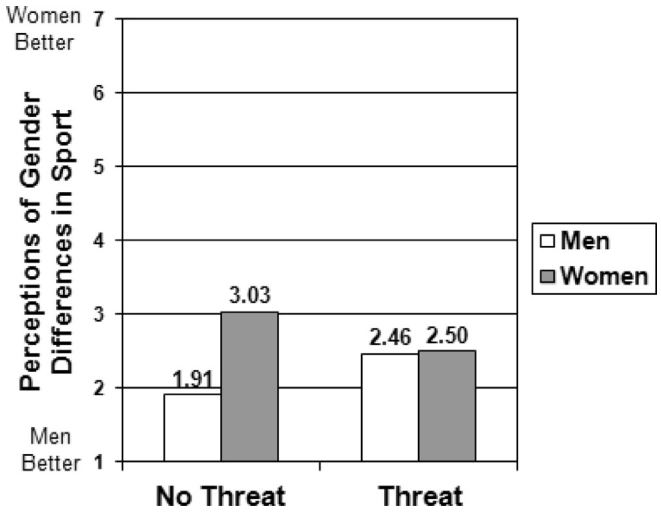

Fig. 3. Gender stereotypes regarding sports performance as a function of stereotype threat and participant gender.

that the tasks assessed visual ability, women had more egalitarian views than men, a difference which was found statistically significant using simple effects tests ($p = .001$). Thus, the female athletes may have been more responsive to generalizing from information regarding gender equivalence in athletic performance. It is important to note, however, that each of the four means was still significantly below the midpoint of the scale ($ps \leq .002$), so participants in all four groups perceived men to be better than women in sports.

As a supplemental analysis, we examined whether participants' general stereotypes regarding athletic ability may have influenced their task performance. Specifically, we performed a General Linear Model (GLM) analysis which simultaneously investigated the effects of participant gender, threat manipulation, and participants' gender stereotypes (measured on a continuous scale) on each task performance measure. There were no significant effects for Task 1 performance, but there were two significant effects for Task 2 performance. Specifically, there was a main effect of stereotypes, $F(1, 53) = 3.88$, $p = .054$ ($\eta^2 = .07$), with participants generally performing better on Task 2 to the extent that they believed that men outperform women in sports. Interestingly, this pattern was driven by the female participants, as reflected in the significant participant gender $\times$ stereotypes interaction, $F(1, 53) = 4.04$, $p = .050$ ($\eta^2 = .07$), shown in Fig. 4. As seen in the Figure, male athletes' performance was unrelated to their beliefs regarding gender differences in athletic ability. Among female athletes, however, those who had stronger stereotypes regarding men's superiority in sports actually performed better than those who had more egalitarian views regarding athletic performance.

Discussion

This was the first research to examine whether female athletes perform worse than male athletes in sports when faced with a gender stereotype, but not when the stereotype is removed. This pattern emerged on the more difficult basketball/tennis task, but not on the speed task. On the difficult concentration task, women performed worse than men when told that the task assessed natural athletic ability and that there was a gender difference on the task. However, they performed at the same level as men when the same task was described as assessing visual ability and as having no gender difference in performance. Thus, as in the previous two studies (Chalabaev et al., 2008; Stone & McWhinnie, 2008), we also identified boundary conditions with regard to when stereotype threat negatively impacts women's athletic performance.

The speed task, Task 2, was included to examine whether time pressure alone would sufficiently increase the difficulty level of an athletic task to produce stereotype threat effects. One might expect that this would especially be the case because adding time pressure inherently made the task less familiar to the athletes. Performance on this task was not, however, influenced by our stereotype threat manipulation. Serving a tennis ball into the service area and making free throws in basketball are very basic tasks that are well-practiced among basketball and tennis athletes. O'Brien and Crandall (2003) found that athletes' performance on easy tasks may even be enhanced in the face of stereotype threat information (O'Brien & Crandall, 2003). Our results suggest that adding time pressure to an easy task is not enough to reverse that tendency and impair performance under stereotype threat conditions.

However, the conclusions we can draw from our second task are limited. Because of our small sample size and our primary focus on the concentration task, we did not counterbalance the order of the two tasks. Although the tasks were quite different, it is still possible that practice effects improved performance and/or reduced anxiety among participants when they performed Task 2. In addition, Task 2 was more delayed from the presentation of the stereotype threat information. Although the study was quite short, perhaps that extra time and the completion of Task 1 caused participants to forget the stereotype information they were presented with.

Interestingly, our stereotype threat manipulation did not influence participants' views of gender differences in Task 1 or Task 2 performance, but did influence their more general gender stereotypes of athleticism, at least among women. The female athletes in our study had more egalitarian views regarding gender differences within the relevant sport and within sports in general if they were told that there was no gender difference on the specific athletic tasks they encountered. Perhaps female athletes' attitudes regarding gender differences are more easily shaped by what others believe. This may be due to conflicting information regarding women's athletic performance. They may perceive

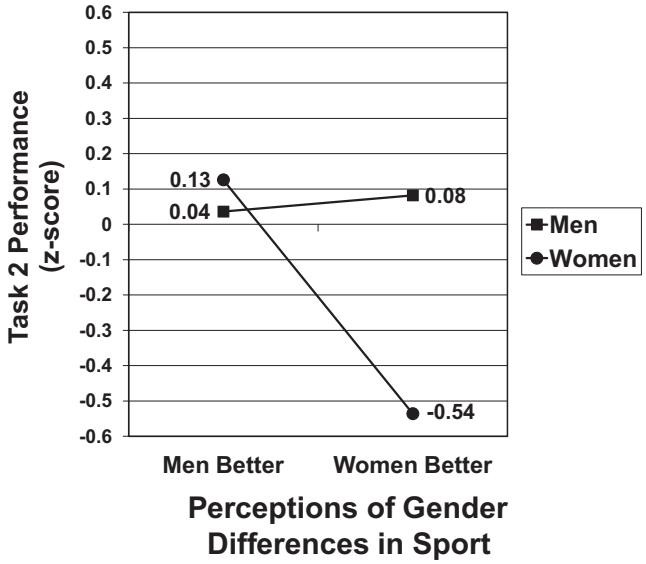

Fig. 4. Task 2 performance as a function of participant gender and gender stereotypes regarding sports performance. A score of 0 reflects the average score in the group, with positive scores being above average and negative scores being below average.

themselves and their female teammates as strong athletes, while simultaneously being bombarded with media images and interpersonal comments suggesting lower athleticism in women. As such, one mere context-specific comment regarding gender differences may be enough for women to change their perceptions regarding how women and men generally perform in sports. Thus, we believe it is important for future researchers to include assessments regarding general stereotypes when conducting this type of research.

Generally, it would seem that perceiving better athletic ability in one's own gender group would be most adaptive. However, our study found that women's opinions regarding men's superiority in sport were associated with better performance on Task 2. This is in contrast to Chalabaev, Sarrazin, and Fontayne's (2009) finding that female soccer players perform better when they believe girls generally perform soccer well. Our stereotype measure differed in that it directly compared perceptions of men versus women. A focus on female athletic ability irrespective of any comparison to male athletic ability may be more beneficial in impacting female athletes' performance. In Chalabaev et al.'s study, perceptions of boys' soccer abilities were unrelated to girls' soccer performance. It seems reasonable that perceptions of another group's ability level would not impact one's own task performance. Had we simply assessed perceptions of female athleticism individually, we may have observed a positive relationship with athletic performance among women. It would also be interesting to examine whether gender identification influences how much gender stereotypes affect athletic performance. Perhaps gender stereotypes have a larger effect on performance when female athletes have a strong gender identity.

The biggest limitation of this study was the limited sample size. We believed it was important to acquire a sample of participants who were experts within the relevant sport. However, doing so came at the cost of excluding many college students from participating in our study. Although we had sufficient statistical power to detect some significant stereotype threat effects, replication of the present findings in future research would help increase our confidence regarding the effects we observed.

Researchers are only scratching the surface in investigating the effects of gender stereotype threat on female athletic performance. In one line of future work, researchers could compare stereotype threat effects in sports that are more traditionally masculine (e.g., football) and those that are more traditionally feminine (e.g., ballet). Although the connection between athletics and masculinity is clearly documented, researchers are also noting that some sports are perceived as more masculine whereas others are perceived as feminine (Hardin & Greer, 2009). Perhaps males' athletic performance would be negatively affected by stereotype threat when the sport is deemed more feminine.

Stereotype threat research has been notoriously inconsistent, with the effects emerging in some circumstances but not others. The same seems true within the context of sports. More research is needed to elucidate the boundary conditions of the stereotype threat effect. If we can further identify when, how, and why female athletes experience negative consequences from stereotype threat, then we can take measures to prevent or minimize those consequences. Aside from telling women that there is no gender difference in performance or deemphasizing the relevance of natural athletic ability, we may also be able to reduce stereotype threat effects by educating female athletes on the phenomenon (Johns, Schmader, & Martens, 2005), training them to think about outstanding women who have succeeded in a male-dominated field (McIntyre, Paulson, & Lord, 2003), or having them think about themselves as athletes, rather than as women, before sports competitions (see Shih, Pittinsky, & Ambady, 1999; Yopyk & Prentice, 2005, for similar notions). For example, Shih et al. found that Asian-American women perform better at a difficult math test after thinking about the fact that they are Asian than after thinking about the fact that they are women. Although we can never know whether Gretchen Bleiler's crash during the Winter X Games was due to stereotype threat, we can ideally take measures to minimize the likelihood that stereotype threat could have such devastating effects during important athletic events.

References

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

Beilock, S. L., & McConnell, A. R. (2004). Stereotype threat and sport: can athletic performance be threatened? *Journal of Sport & Exercise Psychology, 26*, 597–609. Retrieved from http://journals.humankinetics.com/jsep.

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

Biddle, S. J. H., Atkin, A. J., Cavill, N., & Foster, C. (2011). Correlates of physical activity in youth: a review of quantitative systematic reviews. *International Review of Sport and Exercise Psychology, 4*, 25–49. http://dx.doi.org/10.1080/1750984X.2010.548528.

Cann, A. (1991). Gender expectations and sports participation. In L. Diamont (Ed.), *Psychology of sports, exercise, and fitness: Social and personal issues* (pp. 187–214). Washington, DC: Hemisphere Publishing Corporation.

Chalabaev, A., Sarrazin, P., & Fontayne, P. (2009). Stereotype endorsement and perceived ability as mediators of the girls' gender orientation-soccer performance relationship. *Psychology of Sport and Exercise, 10*, 297–299. http://dx.doi.org/10.1016/j.psychsport.2008.08.002.

Chalabaev, A., Sarrazin, P., Fontayne, P., Boiché, & Clément-Guillotin, C. (2013). The influence of sex stereotypes and gender roles on participation and performance in sport and exercise: review and future directions. *Psychology of Sport and Exercise, 14*, 136–144. http://dx.doi.org/10.1016/j.psychsport.2012.10.005.

Chalabaev, A., Sarrazin, P., Stone, J., & Cury, F. (2008). Do achievement goals mediate stereotype threat? An investigation on females' soccer performance. *Journal of Sport & Exercise Psychology, 30*, 143–158. Retrieved from http://journals.humankinetics.com/jsep.

Chalabaev, A., Sarrazin, P., Trouilloud, D., & Jussim, L. (2009). Can sex-undifferentiated teacher expectations mask an influence of sex stereotypes? Alternative forms of sex bias in teacher expectations. *Journal of Applied Social Psychology, 39*, 2469–2498. http://dx.doi.org/10.1111/j.1559-1816.2009.00534.x.

Duffy, J., Warren, K., & Walsh, M. (2001). Classroom interactions: gender of teacher, gender of student, and classroom subject. *Sex Roles, 45*, 579–593. http://dx.doi.org/10.1023/A:1014892408105.

Fredricks, J. A., & Eccles, J. S. (2005). Family socialization, gender, and sport motivation and involvement. *Journal of Sport & Exercise Psychology, 27*, 3–31. Retrieved from http://journals.humankinetics.com/jsep.

Hardin, M., & Greer, J. D. (2009). The influence of gender-role socialization, media use, and sports participation on perceptions of gender-appropriate sports. *Journal of Sport Behavior, 32*, 207–226. Retrieved from http://www.southalabama.edu/psychology/journal.html.

Hardin, M., Simpson, S., Whiteside, E., & Garris, K. (2007). The gender war in U.S. sport: winners and losers in news coverage of Title IX. *Mass Communication and Society, 10*, 211–233. Retrieved from http://www.tandf.co.uk/journals/HMCS.

Hilland, T. A., Stratton, G., Vinson, D., & Fairclough, S. (2009). The physical education predisposition scale: preliminary development and validation. *Journal of Sports Sciences, 27*, 1555–1563. http://dx.doi.org/10.1080/02640410903147513.

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: teaching stereotype threat as a means of improving women's math performance. *Psychological Science, 16*, 175–179. http://dx.doi.org/10.1111/j.0956-7976.2005.00799.x.

Knight, J. L., & Giuliano, T. A. (2001). He's a laker; she's a "looker": the consequences of gender-stereotypical portrayals of male and female athletes by the print media. *Sex Roles, 45*, 217–229. http://dx.doi.org/10.1023/A:1013553811620.

Knisel, E., Opitz, S., Wossmann, M., & Keteihuf, K. (2009). Sport motivation and physical activity of students in three European schools. *International Journal of Physical Education, 46*, 40–53. Retrieved from http://www.schulsport.m-m-sports.com/wissenschaft/front_content.php?idcat=120.

Koivula, N. (1999). Sport participation: differences in motivation and actual participation due to gender typing. *Journal of Sport Behavior, 22*, 360–380. Retrieved from http://www.southalabama.edu/psychology/journal.html.

Leyens, J., Désert, M., Croizet, J., & Darcis, C. (2000). Stereotype threat: are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin, 26*, 1189–1199. http://dx.doi.org/10.1177/0146167200262002.

MacDonald, D. (1990). The relationship between the sex composition of physical education classes and teacher-pupil verbal interaction. *Journal of Teaching in*

Physical Education, 9, 152–163. Retrieved from http://journals.humankinetics.com/jtpe-contents.

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

Nicaise, V., Bois, J. E., Fairclough, S. J., Amorose, A. J., & Cogèrino, G. (2009). Girls' and boys' perceptions of physical education teachers' feedback: effects on performance and psychological responses. *Journal of Sports Sciences, 25*, 915–926. http://dx.doi.org/10.1080/02640410600898095.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: effects on women's math performance. *Personality and Social Psychology Bulletin, 29*, 782–789. http://dx.doi.org/10.1177/0146167203029006010.

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

Souchon, N., Coulomb-Cabagno, G., Traclet, A., & Rascle, O. (2004). Referees' decision making in handball and transgressive behaviors: influence of stereotypes about gender of players? *Sex Roles, 51*, 445–453. http://dx.doi.org/10.1023/B:SERS.0000049233.28353.f0.

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. http://dx.doi.org/10.1006/jesp.1998.1373.

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629. http://dx.doi.org/10.1037/0003-066X.52.6.613.

Steele, C. M. (1998). Stereotyping and its threats are real. *The American Psychologist, 53*, 680–681. http://dx.doi.org/10.1037/0003-066X.53.6.680.

Stone, J. (2002). Battling doubt by avoiding practice: the effects of stereotype threat on self-handicapping white athletes. *Personality and Social Psychology Bulletin, 28*, 1667–1678. http://dx.doi.org/10.1177/014616702237648.

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

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology, 44*, 445–452. http://dx.doi.org/10.1016/j.jesp.2007.02.006.

Yopyk, D. J. A., & Prentice, D. A. (2005). Am I an athlete or a student? Identity salience and stereotype threat in student-athletes. *Basic and Applied Social Psychology, 27*, 329–336. http://dx.doi.org/10.1207/s15324834basp2704_5.