Battling Doubt by Avoiding Practice: The Effects of Stereotype Threat on Self-Handicapping in White Athletes
Jeff Stone
DOI: 10.1177/014616702237648

The online version of this article can be found at:
http://psp.sagepub.com/content/28/12/1667
Battling Doubt by Avoiding Practice: The Effects of Stereotype Threat on Self-Handicapping in White Athletes

Jeff Stone  
University of Arizona

Two experiments examined the use of behavioral self-handicapping as a strategy for coping with stereotype threat. Using sports as the performance context, it was predicted that if a sports test was framed as a measure of “natural athletic ability,” White participants would feel threatened about confirming the negative stereotype about poor White athleticism and would practice less before the test as compared to control groups. The data from Experiment 1 supported the prediction and showed that the effect of stereotype threat on self-handicapping was moderated by participants’ level of psychological engagement in sports. Experiment 2 showed that engaged White participants practiced less than engaged Hispanic participants when their performance was linked to natural athletic ability. The discussion focuses on the processes by which the salience of a negative stereotype in a performance context induces proactive strategies for coping with the implications of a poor performance.

Gordon Allport (1954) observed that being the target of a negative stereotype engenders a threat to self-worth that is likely to create ego-defensive coping responses in target individuals. Indeed, social psychologists have found that rather than passively accepting the negative characterization, when confronted with a negative stereotype about an important social identity, target individuals actively defend themselves against the negative label (Crocker, Major, & Steele, 1998). However, recent research suggests that defending against being negatively stereotyped can have the ironic effect of leading to behavior that confirms the validity of the negative belief. The theory of stereotype threat (e.g., see Steele, 1997; Steele & Aronson, 1995) predicts that when a negative stereotype about a group becomes salient as the criterion for evaluating performance, individual group members may become concerned that their performance will confirm the validity of the negative stereotype. The increased concern imposed by stereotype threat adds an additional psychological burden to the task, which in turn reduces an individual’s ability to perform up to his or her potential. For example, Steele and Aronson (1995) showed that framing an academic test as diagnostic of innate intelligence caused poorer performance among Black participants compared to when the test was framed as measuring a non-stereotype-relevant dimension. Similarly, Spencer, Steele, and Quinn (1999) reported that when performance on a standardized math test was linked to gender differences in math, female participants performed more poorly than when the same test was framed as not showing gender differences in performance. Some evidence suggests that the poor performance in stereotype-threat situations is partially mediated by thoughts about race and self-doubt (Steele & Aronson, 1995), inefficient information processing (Ceryan & Bodenhausen, 2000), or feelings of anxiety (Spencer et al., 1999). The available evidence indicates that the salience of a negative stereotype in a performance context is capable of engendering a threat to self-worth that may consume the very psychological resources people need to overcome the negative characterization.

Less established, however, is the point at which targets in a performance context begin to experience the apprehension associated with stereotype threat. One possibil-

Author’s Note: I would like to thank Donald G. Hakes, Joseph M. Arguelles, Ross Parnell, and Matt Combs for their help in collecting the data. I also thank Toni Schmader and Mark Zanna for their comments on previous drafts of this article. Send inquiries to Jeff Stone, Department of Psychology, University of Arizona, Tucson, AZ, 85721; e-mail: jeffs@u.arizona.edu.

PSB, Vol. 28 No. 12, December 2002 1667-1678  
DOI: 10.1177/014616702237648  
© 2002 by the Society for Personality and Social Psychology, Inc.
Racial Stereotypes About Athletes

The use of self-handicapping in response to stereotype threat was examined in the current research in the context of sports, where there are negative stereotypes about White athletes (e.g., Sailes, 1996; Stone, Perry, & Darley, 1997). Racial stereotypes about athletes—particularly about Black and White athletes—represent the beliefs people generate to explain the perceived relationship between race and performance in sports. A major source for these beliefs is the historical debate over why Black athletes are overrepresented at the highest levels of some amateur and professional sports (see Wiggins, 1997, for a review). Whereas social and biological scientists have contributed to this debate, considerable theorizing about Black superiority in sports can be found in mass media presentations by sportswriters (e.g., Kane, 1971; Price, 1997), sports television commentators (e.g., Jimmy “The Greek” Snyder), contemporary authors (e.g., Entine, 2000; Hoberman, 1997; Schermer, 2001), and by filmmakers, such as in the 1989 movie “White Men Can’t Jump.” As this latter example illustrates, part of the debate over Black superiority in sports focuses on the causes for the perceived inferiority of White athletes (Price, 1997). The general sentiment seems to be that whereas White athletes are underrepresented in amateur and professional sports like basketball, boxing, or track and field because they have less “natural athletic ability” than Black athletes, Whites can succeed in sports when they use their superior intelligence to overcome their physical deficits.

The evidence that people hold racial stereotypes about Black and White athletes generally reflects the historical debate. For example, Devine and Baker (1991) found that the attributes assigned to the social category of “Black athlete” included “unintelligent” and “ostentatious,” and Krueger (1996) reported that Black men were perceived by both Black and White participants to be more “athletic” than White men (see also Biernat & Manis, 1994). Sailes (1996) reported that when asked to describe the attributes of Black and White college athletes, White participants rated Black athletes as significantly less intelligent, less academically prepared, and more temperamentally than White athletes, whereas Black participants rated White athletes as significantly less competitive and as exhibiting less “athletic style” compared to Black athletes. In a perceptual confirmation experiment, Stone et al. (1997) had White participants evaluate a basketball player while listening to a radio broadcast of a college basketball game. Half of the participants were led to believe the target player was White and half were led to believe the target was a Black athlete. The results showed that whereas the White target was perceived as exhibiting less “natural athletic ability” but more “court smarts” and “hustle,” the Black male target player was perceived as exhibiting less “court smarts” and “hustle” but more “natural athletic ability.” The Black target also was perceived to be a better basketball player, even though all participants heard the same performance in the basketball game (see also Johnson, Hallinan, & Westerfield, 1999). Thus, the empirical evidence indicates that people hold specific positive and negative racial stereotypes about athletes: Black athletes
are perceived to have high natural athletic ability but low sports intelligence, whereas White athletes are perceived to have low natural athletic ability but high sports intelligence.

In addition to influencing perceptions of an athlete’s performance, recent research indicates that the negative racial stereotypes can have a detrimental impact on performance in athletics. Stone, Lynch, Sjomeling, and Darley (1999) proposed that when made salient in a sports performance situation, the negative stereotypes about Black and White athletes might induce stereotype threat in those for whom athleticism represents an important aspect of their self-worth. In the first experiment, Stone et al. (1999) had Black and White college students complete a laboratory golf task that was described as a standardized test of either “natural athletic ability,” “sports intelligence,” or “sports psychology.” The results showed that both groups performed equally well on the golf task in the “sports psychology” control condition. However, Black participants performed significantly worse than control participants when performance on the golf task was framed as diagnostic of “sports intelligence,” whereas White participants performed worse than control participants when the golf task was framed as diagnostic of “natural athletic ability.” By showing that stereotype threat can impact the performance of White college students in sports, these studies support Steele’s (1997) contention that stereotype threat is a general psychological phenomenon that can have a negative impact on the performance of any group member, including traditionally nonstigmatized individuals (cf. Aronson et al., 1999; Croizet & Claire, 1998).

The Use of Self-Handicapping to Defeat Stereotype Threat

Steele and Aronson (1995) reported some evidence consistent with the hypothesis that the threat of confirming a negative stereotype through performance will engage a priori self-handicapping responses. In one experiment, after they were told that a standardized test was diagnostic of innate intelligence, Black participants reported that they had less sleep the night before, had been under more stress lately, felt less focused mentally, and that standardized tests were more biased against them, as compared to a non-stereotype-threat condition. Whereas these findings indicate that the targets of stereotype threat may resort to handicaps of the “claimed” variety (Arkin & Oleson, 1998), there is little published evidence that the targets of stereotype threat will purposefully “acquire” handicaps by erecting physical barriers to a successful performance (e.g., purposely stay up all night or fail to study prior to taking an important test). Claimed handicaps before a performance may represent dubious a priori excuses (e.g., Greenberg, Pyszczynski, & Paisley, 1985) or attempts to disengage self-worth from a potentially poor performance (e.g., Major & Schmader, 1998). The present research attempted to investigate whether stereotype threat would motivate targets to behaviorally acquire a handicap through the construction of barriers to success prior to a performance that held importance for their self-worth.

One behavioral self-handicapping strategy used by people in the face of a potentially threatening performance is to avoid preparing or practicing for the performance event. When individuals fail to adequately prepare or practice before an important performance, and their performance suffers, they can attribute their poor performance to the lack of effort and conclude that more practice would have achieved better results (Pyszczynski & Greenberg, 1983). For example, Tice and her colleagues (e.g., Tice, 1991; Tice & Baumeister, 1990) found that when participants anticipated that their self-esteem would be threatened by a poor performance on a personally important task, if afforded the opportunity to practice prior to the performance, threatened participants spent less time and effort practicing, as compared to control participants. The tendency to self-handicap by withholding practice effort also has been observed outside the laboratory among collegiate and professional athletes. Rhodewalt, Saltzman, and Wittmer (1984) found that high self-handicapping intercollegiate swimmers and professional golfers withheld practice effort relative to low self-handicappers before competitions that posed a threat to their self-worth. Thus, in a sports context, withholding practice effort appears to be a useful strategy for creating ambiguity about the meaning of a poor performance when perceptions of self-worth are on the line.

It was predicted in the current research that when the negative stereotype concerning poor White athleticism was linked to their performance, White sports participants would exhibit less practice effort on a laboratory golf task compared to participants for whom performance was linked to a non-stereotype-relevant attribute. Furthermore, it was predicted that the tendency to withhold practice effort would be moderated by the importance of performance in sports to perceptions of self-worth (Tice, 1991). Previous research indicates that stereotype threat has less impact on behavior when performance in a domain is relatively unimportant to a target’s self-worth (Steele, 1997). For example, Major, Spencer, Schmader, Wolfe, and Crocker (1998) reported that African American participants who were classified as “psychologically disengaged” from academics, based on responses to the Disengagement subscale of the Intellectual Orientation Inventory (IOI), did not
show decrements in self-esteem following failure feedback on a test of academic ability. Similarly, Stone et al. (1999, Experiment 2) modified the disengagement items on the IOI to measure disengagement from performance in sports and reported that when stereotype threat was salient, White participants whose self-worth was disengaged from performance in sports performed significantly better on a golf task as compared to participants for whom performance in sports represented an important measure of their self-worth. These findings suggest that only engaged White sports participants should feel threatened by being negatively stereotyped and, in response, be motivated to withdraw practice effort to self-handicap prior to their performance on a sports task.

The hypotheses were tested by replicating the laboratory golf task used by Stone et al. (1999) and allowing participants to practice on the first hole of the golf course ostensibly before they completed the entire course (Tice, 1991; Tice & Baumeister, 1990). A number of potential mediators (described below) of the relationship between stereotype threat and self-handicapping also were assessed. The design of Experiment 1 was a 2 (chronic athletic disengagement: engaged or disengaged) × 2 (test frame: high threat or low threat) factorial with the amount of practice effort on the first hole of the course as the primary measure of self-handicapping.

EXPERIMENT 1

Method

Participants. Participants were 36 female and 28 male undergraduates at the University of Arizona who participated in the study for partial course credit. All were recruited after they identified their ethnicity as White and met certain criteria on their self-ratings of athleticism and golf experience collected during a mass pretesting of the introductory psychology participant pool (Stone et al., 1999). On a short version of the Self-Attributes Questionnaire (Pelham & Swann, 1989), participants were asked to rate their perceived level of athleticism by comparing themselves to other college students their age on a scale that ranged from 1 (bottom 5%) to 11 (top 1%). Only those who rated their athletic ability above the midpoint of the scale (i.e., above the upper 50th percentile) were recruited to participate (sample \( M = 7.23, SD = 1.14 \)). Participants also responded to the question: “How many times a week do you play golf?” Only those who reported they played no more than one round of golf per week (scores ≤ 1.0, \( M = 0.9 \) days per week) were recruited to participate. In sum, participants were White college students who rated their athleticism as above average but who did not play enough golf to qualify as having strong skills in the game.

Also included in the pretest was a scale devised to measure the importance of performance in sports to their perceptions of self-worth (Stone et al., 1999). The Athletic Disengagement Scale (ADS) consisted of the items “No athletic test will ever change my opinion of how athletic I am,” “How I do athletically has little relation to who I really am,” and “I really don’t care what tests say about my athletic ability.” Each question was answered using a 1 (strongly disagree) to 7 (strongly agree) scale. The items on the ADS scale demonstrated an acceptable level of internal consistency in the sample (\( \alpha = .64 \)) and scores on the ADS were distributed normally among the selected sample despite limiting participation to only those who rated themselves as above average in athleticism (ADS \( M = 3.76, SD = 1.34, \) range 1-7).

As in our previous research (Stone et al., 1999), participants were classified as high or low in athletic disengagement based on a median split and were randomly assigned to the experimental conditions described below.

Procedure. Participants completed the procedures individually. When they arrived at the laboratory, the experimenter (a male who was blind to their level of psychological disengagement) explained that they would complete some brief questionnaires, perform a sports test based on the game of golf, and then answer questions about their performance after the test was completed.

The athletic test was based on the golf task described in Stone et al. (1999). Participants first read a handout that described the athletic task as a standardized measure of sports psychology called the “Michigan Athletic Aptitude Test” (MAAT). Participants were told that the MAAT was developed in 1988 by the exercise and sports psychology department at the University of Michigan. The handout noted that the test was based on the game of golf but had been normalized such that each successive level in the test represented a standard increase in performance difficulty. Ostensibly, performance on the test had been shown to correlate with actual performance on many of the physical and mental activities relevant to most college varsity sports, such as basketball, baseball, and hockey. At this point, the instructions changed course according to which test frame condition participants had been assigned to randomly.

Stereotype-threat manipulation. Participants in the high-threat condition read that the test was designed to measure “personal factors correlated with natural athletic ability.” Natural athletic ability was defined as “one’s natural ability to perform complex tasks that require hand-eye coordination, such as shooting, throwing, or hitting a ball or other moving object.” It was explained that as
test difficulty increased, so would the demand on their natural athletic ability or hand-eye coordination.

Participants in the low-threat condition read that the test was designed to measure “psychological factors correlated with general sports performance.” The handout explained that as test difficulty increased, so would the demand on the psychological factors that correlate with general sports performance.

After they read the handout, the experimenter reiterated the instructions and answered questions. Participants then completed a preperformance questionnaire packet that contained several potential mediators of the relationship between stereotype threat and self-handicapping.

Pre-performance measures of mediation. Before participants began the questionnaire packet, the experimenter stressed that their responses were confidential and would be kept anonymous. He showed them a large envelope that ostensibly held other questionnaires and told participants that when they were finished, they should place their questionnaire inside the envelope to maintain anonymity. He then handed them a packet and asked them to begin.

The first page of the packet was a short demographic questionnaire on which they indicated their gender, age, year in school, and ethnic identity. Next, to measure the accessibility of thoughts related to racial stereotypes and self-doubt about competence and ability in sports, participants performed a word-fragment completion task modeled after the one reported by Steele and Aronson (1995; see Gilbert & Hixon, 1991). Seven word fragments were designed to measure activation of thoughts associated with race: _ _ _ _ _ _ _ (WHITE), _ _ _ _ _ _ _ (BLACK), _ _ _ _ _ _ _ (INFERIOR), W __ _ _ _ _ _ (WEAK), and S H A __ _ _ _ _ (SHAME). Six word fragments were designed to measure activation of thoughts associated with self-doubt: H A _ _ _ _ _ _ _ (HARD), L O ___ ___ _ _ _ (LOSER), F L _ _ _ _ _ _ _ (FLUNK), _ _ _ _ _ _ _ (AWKWARD), _ _ _ _ _ _ _ (UNFIT). All of the target fragments were then randomly distributed among 12 filler words on one page. To complete the task, participants were instructed to fill in the blanks to form the first word that came to mind. They were told to work quickly, spending no more than 15 on each fragment, and that items could be left blank if no word came to mind.

After they completed the word fragment test, participants then completed a five-item self-report measure of situational anxiety (Mattsson, 1960). Participants were instructed to indicate their current level of anxiety by marking on 7-point scales with the endpoints (1) uneasy—at ease, (2) comfortable—uncomfortable, (3) upset—peaceful, (4) relaxed—tense, and (5) in control—not in control. Next, participants completed a state self-esteem measure based on the Rosenberg Self-Esteem Scale (1979). To specify their current state of self-esteem, the items on the Rosenberg were altered so that each began with “Right now . . .” Respondents indicated their response to each statement on a 1 (strongly disagree) to 4 (strongly agree) scale.

The final page of the packet held items designed to measure claimed handicaps and discounting processes. Specifically, for the claimed handicaps, participants indicated how many hours of sleep they had the night before and then indicated how “mentally focused” they felt and how “stressed out” they felt lately (e.g., Steele & Aronson, 1995). To measure discounting processes, participants rated their agreement with the statements “How biased or unfair are most standardized tests?” and “How biased or unfair are most standardized testing situations?” These items were rated using a 1 (very biased) to 7 (not biased) scale. Finally, participants indicated their agreement with the statement “I do not trust the results of most standardized tests,” which was also measured using a 7-point scale with the endpoints agree and disagree. Once participants completed their packet, the experimenter reminded them to place it in the large envelope. He then led participants into an adjoining room to complete the golf task.

The athletic performance test. Based on Stone et al. (1999), the task was designed to resemble a miniature golf course on which participants used a putter to hit a golf ball down a 34 ft × 10 ft stretch of carpet into a “hole.” Participants were told to start with their ball on a 4-in. × 6-in. computer-mouse pad. At the end of the carpet was the “hole” apparatus—an inclined felt mat with the three holes in the mat: a hole 5 in. in diameter, a hole 4 in. in diameter, and a hole 3 in. in diameter. To complete each phase of the test, participants were told that the ball had to roll up the incline and stop in one of the holes. Participants were told that they would complete eight different holes during the test. New holes would be created by placing 2 in. × 4 in. × 3 in. pieces of wood either on or under the carpet and by moving the hole apparatus up or down the carpet. Once the golf test began, the experimenter said he would change the putting surface according to a pretested pattern of increasing difficulty.
Participants were told that their goal on each course layout was to putt the ball into one of the three holes using the fewest strokes possible. In addition, they were told that the hole that received the ball would be recorded and that the small hole was worth one point, the middle hole was worth two points, and the largest hole was worth three points. At the end of the test, the points would be summed to yield an overall total performance score.

**Practice instructions.** At this point, the experimenter introduced the primary dependent measure. Participants were told that before they began the test, they could practice the first hole of the course as many times as they would like. The experimenter said that he was not going to keep their score but that he did need to keep track of how long they practiced because “we might need this later on to interpret your test score.” The experimenter then handed participants a putter and asked them to choose a golf ball. Finally, he reiterated that they could practice as much as they wanted to but that he needed to know when they were done so he could record the time.

**Practice effort measure.** Once participants placed their ball on the starting pad, the experimenter began timing their practice using a stopwatch. The experimenter stopped timing the practice when participants said verbally that they were finished practicing or that they were ready to begin the test. While participants practiced, the experimenter also kept track of the number of times participants hit the ball and the number of times participants placed the ball on the starting pad to practice the hole again. In combination, these three measures of practice effort served as the primary index of how much participants handicapped themselves prior to completing the athletic task.

**Performance expectancy measures.** After they had finished practicing, the experimenter told participants that they needed to complete a few other measures before they began the test. He then handed them a clipboard and asked participants to provide an estimate of how many strokes they expected to need and in which hole they expected the ball to stop on each hole of the course. To provide estimates of their performance, participants were instructed to review the diagrams of each hole on the course in the test booklet they had read during the introduction to the study. The pages of the test booklet had been photocopied and taped to one wall in the testing room. Participants were told their responses were anonymous and the experimenter left them alone in the test room while they completed their predictions. Participants recorded their estimates on a sheet of paper and, as with the pretest packet, were instructed to place their responses in an envelope that ostensibly held the predictions made by other participants. The experimenter returned when they indicated they were finished and collected the envelope. At this point, the experimenter announced that the study was finished and told participants that they would not have to complete the athletic task. He then led them back into the main lab office and conducted a thorough debriefing.

**Results**

One participant was excluded from the data analysis for indicating a non-White racial identity at the start of the experiment and one was excluded for failing to complete several of the pretest and posttest questionnaires. This left the data from 64 participants available for analysis. An initial analysis of the data using gender as a variable revealed that it did not significantly moderate any of the results reported below. It was thus dropped from the subsequent analyses.

**Practice effort.** The amount of time (in seconds) they practiced, the number of strokes, and the number of times they practiced the first hole were found to be highly correlated (range = .75 to .85, all ps < .0001). A composite was computed by standardizing each measure and summing them to create one overall measure of practice effort (α = .93). The practice effort composite was then analyzed using a 2 (chronic disengagement: high or low) × 2 (test frame: high or low threat) ANOVA, which revealed a significant main effect for the threat manipulation, $F(1, 60) = 13.42, p < .0001$. Participants practiced less on the first hole of the course when it was framed as a measure of natural athletic ability (high threat) compared to when the task was framed as a measure of general sports performance (low threat).

The main effect for the threat manipulation was qualified by the predicted Psychological Disengagement × Test Frame interaction, $F(1, 60) = 4.72, p < .03$. Table 1 presents the standardized practice effort composite scores for each experimental condition. As can be seen, a one versus three planned comparison revealed that as predicted, participants classified as psychologically engaged in their performance in sports practiced significantly less ($M = –0.60$) in the high-threat condition as compared to engaged participants in the low-threat condition ($M = 0.71$) and as compared to the disengaged participants in the high threat ($M = –0.21$) and low threat ($M = 0.11$) conditions, $F(1, 60) = 9.64, p < .003$. Further contrasts between the control groups revealed that engaged participants in the low-threat condition practiced more than both disengaged groups, $F(1, 60) = 7.60, p < .01$, but that there was no difference between the disengaged groups as a function of the level of threat, $F(1, 60) = 1.11, p < .30$. A planned comparison also showed that engaged participants in the high-threat condition practiced less compared to disengaged partic-
effects of test frame and psychological disengagement on the standardized measure of practice effort and activation of athleticism

participants completed the word hard, F(1, 60) = 7.71, p < .007, and somewhat more engaged participants completed the word awkward, F(1, 60) = 3.46, p < .067. As can be seen in Table 1, there was a tendency for engaged participants in the high-threat condition to complete more of these words compared to participants in the control conditions. For the self-doubt fragment set, a multivariate main effect for disengagement was observed, F(5, 56) = 2.95, p < .02. Univariate tests revealed that significantly more engaged participants completed the word inferior (19%) as compared to disengaged participants (0%), F(1, 60) = 7.71, p < .007, and somewhat more engaged participants completed the word hard (22%) as compared to disengaged participants (6%), F(1, 60) = 3.21, p < .08. A MANOVA on the target word fragments that were related to race was not significant, and none of the within-cell correlations between the completion of the words and the measure of practice effort approached significance.

Other dependent measures. There were no significant effects for the experimental variables on the prepractice self-report measures of anxiety, self-esteem, excuse making, discounting, or perceptions of trust. In addition, no significant effects emerged on the postpractice measures of performance expectancies.

Discussion

The results of Experiment 1 supported the hypothesis that linking performance to natural athletic ability would cause White sports participants to exhibit less practice effort before the test as compared to when the test was linked to a nonstereotyped dimension. Also as predicted, the effect of stereotype threat on self-handicapping was significantly moderated by how important performance in sports was to perceptions of self-worth. When the test was framed as diagnostic of natural athletic ability, participants whose self-worth was engaged in sports practiced less compared to engaged participants for whom the task was linked to a non-stereotyped domain. In addition, when stereotype threat was high, more engaged White participants tended to complete word fragments related to the negative stereotype about White athletes compared to the control conditions, suggesting that the natural athletic ability test frame activated more thoughts about poor athleticism compared to the control groups. These findings are consistent with the hypothesis that natural athletic ability represents a negative stereotype about White athletes, and when made salient in a sports performance context, induces the self-doubt theorized to drive both stereotype threat (Steele, 1997) and self-handicapping behavior (Arkin & Oleson, 1998).

The data also indicate that participants whose perceptions of self-worth were disengaged from performance in sports were unaffected by how the task was framed, suggesting that disengaged participants were less concerned overall with how they performed on the athletic test. However, the difference in practice effort observed between disengaged and engaged participants when the test was framed as diagnostic of the stereotype about poor White athleticism was not reliable. Whereas the difference was in the predicted direction, the weakness of the effect raises questions about the moderating role of psychological disengagement in self-handicapping as a reaction to stereotype threat.

In addition, whereas the effect for how the test was framed was consistent with the stereotype-threat hypothesis, there is at least one alternative interpretation of the observed differences in practice effort between the threat conditions. It is possible that participants prac-

<table>
<thead>
<tr>
<th>Practice effort[a]</th>
<th>Engaged</th>
<th>Disengaged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High Threat (n = 16)</td>
<td>Low Threat (n = 16)</td>
</tr>
<tr>
<td>M</td>
<td>-0.60</td>
<td>0.71</td>
</tr>
<tr>
<td>SD</td>
<td>0.40</td>
<td>1.28</td>
</tr>
</tbody>
</table>

Word completions (%)

<table>
<thead>
<tr>
<th></th>
<th>Engaged</th>
<th>Disengaged</th>
</tr>
</thead>
<tbody>
<tr>
<td>awkward</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>effort</td>
<td>31</td>
<td>13</td>
</tr>
</tbody>
</table>

TABLE 1: Effects of Test Frame and Psychological Disengagement on the Standardized Measure of Practice Effort and Activation of Athleticism

(a. Lower scores indicate less practice effort.)
ticed less in the high-threat condition because they perceived “natural athletic ability” to be a stable attribute or fixed entity that is unlikely to improve through practice. In contrast, when framed as a measure of general sports performance, participants may have perceived the test as tapping a more malleable dimension that might improve with practice (Hong, Chiu, Dweck, Lin, & Wan, 1999). Hence, participants practiced less when the test was framed as a measure of their natural athletic ability not because they were motivated to avoid being stereotyped negatively but because they rationally concluded that practice would not improve their performance on the test.

One way to examine the validity of the “entity versus incremental” alternative interpretation of the practice effects observed in Experiment 1 is to examine the practice effort of a group of participants for whom natural athletic ability is not perceived as a negative stereotype. A literature review of the psychological and sociological literature on stereotypes about athletes revealed that there may not be very clearly defined or firmly held stereotypes about Hispanic athletes. For example, Johnson et al. (1999) had 248 undergraduates indicate which of eight attributes contributed to the success of Black, White, and Hispanic basketball players. The results showed that whereas Black athletic success was attributed to physical characteristics such as jumping ability and White success was attributed to hard work and access to better athletic resources, Hispanic athletic success was not attributed to any of the measured physical or social factors. Other studies suggest that people do not readily generate or report specific beliefs about Hispanic athletes (e.g., Sabo, Jansen, Tate, Duncan, & Leggett, 1996). Thus, the available data implies that there may not be specific cultural stereotypes about Hispanic athletes.

According to the stereotype-threat hypothesis, if natural athletic ability represents a negative cultural stereotype about White but not Hispanic athletes, then when allowed to practice on the first hole of the course, White participants should exhibit less practice effort as compared to Hispanic participants. Moreover, this effect should be moderated by psychological disengagement, such that White participants for whom performance in sports represents an important measure of their self-worth should practice less compared to engaged Hispanic participants, when the test is framed as a measure of natural athletic ability. Disengaged White and Hispanic participants, in comparison, should practice longer than the engaged White participants and approximately the same as the engaged Hispanic participants. In contrast, if practice time in Experiment 1 was driven by an “entity” interpretation of the natural athletic ability test frame, then both White and Hispanic participants should practice for less time when the test is framed as a measure of natural athletic ability, and this effect should not be moderated by their level of psychological disengagement from sports. The design of Experiment 2 was a 2 (race: White or Hispanic) × 2 (chronic athletic disengagement: engaged or disengaged) factorial in which all participants were led to believe that the test measured natural athletic ability. As in Experiment 1, the amount of practice effort demonstrated on the first hole of the golf course was the primary measure of behavioral self-handicapping.

EXPERIMENT 2

Method

Participants. Participants were 38 Hispanic and 36 White male and female undergraduates at the University of Arizona who participated in the study for partial course credit. All participants were recruited after they identified their ethnicity as either White or Hispanic and met the same criteria on their self-ratings of athleticism and golf experience as reported in Experiment 1. Thus, all were college students who rated their athleticism as above average (M = 7.24, SD = 1.33) but who did not play more than one round of golf per week (M = 0.12, SD = 0.32). In addition, participants were classified as high or low in athletic disengagement based on a median split of the Athletic Disengagement Scale (ADS) collected during a mass pretest session of the psychology participant pool.

Procedure. The procedures for Experiment 2 replicated those described for the natural athletic ability test condition used in Experiment 1. Participants completed a similar pretest measure, were allowed to practice for as long as they wanted on the first hole of the golf course, and were then asked to predict their performance on each hole prior to taking the test. After they completed their predictions, the experimenter (a male who was blind to the hypothesis and to their level of psychological disengagement) announced that the study was complete and provided a thorough debriefing.

Results

An initial analysis of the data did not reveal any significant effects involving gender, so it was dropped from the analyses reported below.

Practice. As in Experiment 1, the amount of time they practiced, the number of strokes, and the number of times they practiced the first hole were highly correlated (range = .75 to .88, all ps <.0001) and were thus standardized and summed to create one overall composite measure of practice effort (α = 0.93). A race (White or Hispanic) × disengagement (engaged or disengaged) ANOVA on the amount participants practiced on the
first hole revealed a significant main effect for the race variable, $F(1, 70) = 14.78, p < .0003$, which showed that White participants practiced significantly less overall compared to Hispanic participants. The main effect was qualified by the predicted Race × Disengagement interaction, $F(1, 70) = 4.29, p < .04$. A one versus three planned comparison of the means shown in Table 2 revealed that as predicted by the stereotype-threat hypothesis, White participants classified as psychologically “engaged” in their performance in sports practiced significantly less ($M = –0.74$) as compared to chronically engaged Hispanic participants ($M = 0.50$) and as compared to chronically disengaged White ($M = 0.29$) and Hispanic ($M = 0.29$) participants, $F(1, 70) = 15.73, p < .0002$. A planned comparison also revealed that engaged White participants practiced significantly less than disengaged White participants, $F(1, 70) = 4.64, p < .03$. Further analyses between the control groups revealed no significant differences.

**Thought activation.** The word fragment completions were analyzed separately using Race × Psychological Disengagement Multivariate Analysis of Variance (MANOVA) procedures. A MANOVA on the athleticism words revealed a significant Race × Disengagement interaction, $F(4, 67) = 2.46, p < .03$. Univariate tests revealed that this was primarily attributed to a significant Race × Disengagement interaction on the word **awkward**, $F(1, 70) = 5.90, p < .02$. The data in Table 2 show that more engaged White participants completed the word **awkward** compared to the other three groups (one versus three contrast), $F(1, 70) = 9.91, p < .002$. No effects on the other words approached significance.

For the self-doubt fragment set, the MANOVA showed a significant multivariate Race × Disengagement interaction, $F(6, 65) = 2.19, p < .05$. Univariate tests revealed a significant Race × Disengagement interaction for the word fragment **hard**, $F(1, 70) = 10.24, p < .002$. The data in Table 2 indicate that more disengaged Whites and engaged Hispanic participants completed this word compared to engaged White and disengaged Hispanic participants. No effects approached significance for the MANOVA on the race word fragments.

**Pre-performance measures.** There was a significant Race × Disengagement interaction on the hours of sleep participants reported getting the night before the golf task, $F(1, 70) = 10.27, p < .002$. A post hoc comparison of the means in Table 2 indicated that White disengaged participants claimed to have slept fewer hours the night before compared to White engaged participants ($p < .05$), with the Hispanic groups falling in between. In addition, there was a significant main effect for Disengagement on how much stress participants reported they had been under lately, $F(1, 70) = 5.02, p < .02$, which showed that dis-engaged participants reported to have been under more stress lately ($M = 4.55$) compared to engaged participants ($M = 3.69$). Similarly, a Race × Disengagement ANOVA on the measure of self-reported anxiety showed a marginal main effect of the Disengagement variable, $F(1, 70) = 3.40, p < .06$. Disengaged participants reported higher levels of anxiety before the test ($M = 2.74$) compared to engaged participants ($M = 2.31$). An analysis of the pre-performance expectancy measure revealed a significant main effect for participant Race, $F(1, 69) = 11.42, p < .001$. White participants expected to complete the course in fewer strokes ($M = 25.39$) compared to Hispanic participants ($M = 37.95$). There were no significant differences between the groups on the other pre-performance measures.

**Mediational analyses.** Of the pre-performance measures that showed a significant Race × Disengagement interaction, only the accessibility of the poor athleticism word **awkward** correlated significantly with how much participants practiced on the first hole of the course ($r = –.25, p < .03$). Mediational analyses showed that when the accessibility of the word **awkward** was held constant, the interaction between Race and Disengagement on practice effort was no longer significant ($β = 0.31, p < .10$). However, a test of the influence of the mediator was not significant (Sobel Test = 0.96, $p < .33$). Thus, the data indicate that the accessibility of the word **awkward** only partially mediated the interactive effect between Race and Psychological Disengagement on how much participants practiced before the test.

**Discussion**

Several of the results in Experiment 2 converge to support the stereotype-threat interpretation of the low practice effort observed among engaged White participants in Experiment 1. Experiment 2 was designed such

| TABLE 2: Effects of Racial Identity and Psychological Disengagement on the Standardized Measure of Practice Effort and on the Activation of Athleticism and Self-Doubt Words Completed Before the Golf Task in Experiment 2 |
|----------------------------------|-----------------|-----------------|-----------------|-----------------|
|                                  | Engaged         | Hispanic        | Disengaged      | Hispanic        |
|                                  | ($n = 18$)       | ($n = 18$)       | ($n = 20$)       | ($n = 20$)       |
| Practice effort$^a$              |                 |                 |                 |                 |
| $M$                              | –0.74           | 0.50            | –0.09           | 0.29            |
| $SD$                             | 0.49            | 1.32            | 0.66            | 0.91            |
| Word completions (%)             |                 |                 |                 |                 |
| Awkward                          | 22              | 0               | 5               | 0               |
| Hard                             | 0               | 33              | 22              | 5               |
| Pre-performance                  |                 |                 |                 |                 |
| Hours of sleep                   | 7.47            | 6.19            | 5.89            | 6.90            |

$^a$ Lower scores indicate less practice effort.
that if natural athletic ability represented a dimension that could not be improved with practice, all of the groups would show low practice effort on the first hole of the course. Instead, engaged Hispanic and disengaged White participants practiced significantly more compared to engaged White participants, implying that only the engaged White group was motivated to ameliorate the meaning of a poor performance. In addition, significantly more engaged White participants completed the poor athleticism word awkward compared to the other groups, and the interaction between race and psychological disengagement on the amount of practice was partially mediated by the accessibility of the word awkward. This is consistent with the stereotype-threat assumption that the salience of a negative stereotype in a performance context engenders heightened concern in targets, which subsequently motivates coping strategies designed to mitigate the threat.

GENERAL DISCUSSION

The results of both experiments supported the hypothesis that linking performance in sports to natural athletic ability would induce an acquired or behavioral self-handicapping strategy among White sports participants. Poor natural athletic ability appears to be a widely held negative stereotype about White athletes that when made salient in a sports performance context raises the concern among White sports participants that they may confirm the veracity of the negative stereotype (Stone et al., 1999). To reduce the threat, White participants in the current studies used the opportunity to create ambiguity about the cause of a poor performance by reducing practice effort prior to completing a golf task. These findings extend previous research on stereotype threat and claimed self-handicaps (Steele & Aronson, 1995) to include acquired or behavioral self-handicapping. Taken together, it appears that the process of stereotype threat can include a stage prior to performance during which targets anticipate and begin to defend themselves against the possibility of confirming the stereotype through a poor performance. Once the negative stereotype is linked to an upcoming performance, targets may anticipate the potential for threat and begin to actively battle against the negative characterization, in this case, by withdrawing preparatory effort so that they could deflect a stable dispositional attribution for a poor performance.

In addition, the data suggest that the effect of stereotype threat on self-handicapping is moderated by how important performance is for a target's self-worth. In Experiment 2, White participants whose self-worth was engaged in performance in sports practiced less when the test was framed as diagnostic of natural athletic ability compared to White participants for whom performance in sports was not important to their self-worth. This conceptually replicates previous research on the effects of stereotype threat among White sports participants (Stone et al., 1999) and adds further support to the claim that stereotype-threat processes are most likely to have a detrimental impact on targets for whom performance in a domain holds significance for their self-worth (Steele, 1997).

The accessibility of thought related to poor athleticism also was affected by the manipulation of stereotype threat. Thoughts and feelings related to self-doubt have been documented during stereotype-threat processes (e.g., Steele & Aronson, 1995), and they are considered to be a central driving feature of self-handicapping behavior (see Arkin & Oleson, 1998). The current research draws a connection between these two phenomena by showing that when natural athletic ability was linked to performance, more engaged White participants completed the word fragment awkward as compared to control groups in each experiment. And although the influence was weak, the activation of the word awkward was found to partially mediate the amount of self-handicapping behavior exhibited by participants in Experiment 2. Thus, whereas the salience of the negative stereotype about natural athletic ability made at least one thought about poor athleticism more accessible, it played only a small role in the self-handicapping behavior that followed from the anticipation of stereotype threat for engaged White sports participants.

The observed differences between White and Hispanic participants in Experiment 2 on the measures of practice and the activation of thoughts about poor athleticism are consistent with the assumption that it was the threat of confirming a negative stereotype, and not the perception of natural ability as a fixed entity, that caused engaged White participants to withdraw practice effort before the task. Nevertheless, the data do not indicate that Hispanics were generally unfazed by the test of natural athletic ability. For example, the data show that thoughts about the difficulty of the test were accessible for Hispanic participants, and they also reported lower expectancies for their performance compared to the White participants, both of which are consistent with the higher practice effort observed for the Hispanic groups. Thus, it may be the case that Hispanic athletes were concerned about the prospect of displaying poor natural athletic ability but they chose to cope with the threat by increasing their practice effort to perform well on the test. The specific beliefs that represent stereotypes about Hispanic athletes, and the processes by which they cope with negative stereotypes in athletics, represent important directions for future research.
There are some questions and qualifications to the present findings. One question concerns the effect of chronic psychological disengagement on the use of claimed handicapping strategies. In Experiment 2, disengaged White participants claimed to have slept less the night before, and overall, disengaged participants claimed they were under more stress lately, as compared to engaged participants. The tendency for disengaged individuals to use claimed handicapping could suggest that disengaged participants were also threatened by the link between the stereotype and their performance but they coped with it through excuse making prior to their opportunity to practice. The reliability of this effect, however, is not clear; no effects for disengagement were observed on the same pretask measures collected in Experiment 1, and disengaged White sports participants did not show evidence of claimed handicaps in previous research using similar measures (Stone et al., 1999). It therefore seems prudent to interpret the current evidence for the use of claimed handicaps by disengaged participants with caution.

Another potential qualification to the conclusion that stereotype threat can lead to behavioral self-handicapping is the relatively public setting of the practice situation created in the present research. In each study, participants were told that the experimenter was monitoring their practice, which may have focused them on using practice effort to control the attributions of the experi- menter. Previous research on self-handicapping indicates that people are more likely to handicap themselves when in public (Arkin & Oleson, 1998), suggesting that if their practice session would have been private, perhaps engaged White sports participants would have coped with the threat by using a different strategy. However, it is important to note that sports at all levels are practiced and played in public, and as a result, many practices and performances are subject to the evaluation of important others such as coaches, managers, or fans (e.g., Schlenker, Phillips, Bonicki, & Schlenker, 1995). Thus, whereas behavioral self-handicapping may not present a viable coping response in some performance situations, it may be an especially prevalent response for people who play sports and often practice in the presence of an important audience.

Summary

The present research indicates that behavioral self-handicapping is a strategy by which the targets of a negative stereotype may proactively cope with their predica ment (e.g., Crocker & Major, 1989; Major & Schmader, 1998). In this case, it appeared that White sports participants reduced practice effort before an athletic task to create ambiguity about the meaning of a poor performance (e.g., Rhodewalt et al., 1984). However, the irony of this response is that as a strategy for defeating a negative stereotype, withdrawing preparatory effort to control attributions may be self-defeating in the long run. If not carefully managed, for example, a poor performance could provide perceivers with evidence that the negative stereotype is accurate. For the target individual, with repetition, poor preparation and the failure it incurs could begin to change perceptions of self-worth in the performance domain. In time, self-handicapping behavior may contribute to the process by which targets disidentify and ultimately withdraw from participation in domains such as education or sports.

REFERENCES
