STEREOTYPE THREAT IMPACTS COLLEGE ATHLETES' ACADEMIC PERFORMANCE

Matthew Jameson
Wesleyan University

Robert Diehl
Wesleyan University

Henry Danso
Wesleyan University

ABSTRACT
This study examined the effects of negative stereotypes on college athletes’ academic performance. Seventy-two male college students on the university's football, basketball, or hockey teams took an intellectual test described as diagnostic or non-diagnostic of intelligence. Prior to testing, half the participants received a questionnaire designed to prime negative stereotypes about athletes, half received this questionnaire post-test. Participants underperformed on the test in the stereotype prime condition, even on the presumed non-diagnostic test. In addition, the more participants believed that athletic ability aided their admission into college the worse they performed on the test. These results indicate that negative stereotypes can affect athletes’ (a behaviorally-defined group) intellectual performance in college.
INTRODUCTION

Increasingly adolescents in North America are choosing to go to college (Steinberg, 1999). For many of these students, the increased academic demands and novel coursework may render the transition to university stressful (see, e.g., Baker & Siryk, 1984, Levitz & Noel, 1989). This may be especially true for student athletes who face the added pressure of succeeding in their respective sports (see Sellers, Kuperminc, & Damas, 1997). Not surprisingly, it has been reported that athletes tend to perform poorly academically compared to their non-athletes counterparts in college (e.g., Upthegrove, Roscigno, & Charles, 1999). Adler and Adler (1985), for instance, followed a group of incoming basketball players through their first four years of college and observed that the athletes' academic performance fell below that of non-athletes. The athletes also reported lower academic aspirations over the course of their four years in college. A recent NCAA report (2002) also revealed that male college football and basketball players tend to have lower graduation rates than do college athletes in other sports and students who are not athletes. Similarly, Purdy, Eitzen and Hufnagel (1982) found that students on athletic scholarship and those who play football and men's basketball tend to demonstrate the largest deficits in academic performance.

Some researchers have suggested that the academic performance deficits demonstrated by college athletes can be traced, in part, to their inadequate background preparation. Shulman and Bowen (2001) reported that athletes who play football, basketball, and hockey, often referred to as “revenue sports,” tend to have substantially lower incoming SAT scores and high school grades compared to their pears in college. In contrast, athletes in lower profile sports’ (e.g. tennis, swimming, track and field, etc) incoming SAT scores, high school grades, and collegiate academic performance more closely resemble that of the non-athlete population (see Richards & Aries, 1999). Papanikolaou, Nikolaidis, Patsiaouras and Alexopoulos (2003) argued that stress and poor coping strategies that lead to self-defeating behaviors disproportionately affect “revenue athletes” because of their high visibility and the pressures of their sports (see also Sellers, et al, 1997). Confirming this latter assertion, a study by Upthegrove, et al. (1999) found a significant disparity in academic performance between “revenue” and “non-revenue” college athletes. They posited that stresses due to institutional pressure, time constraints, and competitive intensity are significant factors that hinder revenue athletes' academic performance.

Killeya (2001) observed that the degree to which athletes viewed themselves as capable of performing both the athlete and student roles proficiently (role elaboration) predicted how well they performed academically and how they felt emotionally at school. This suggests that self-perceptions may play an important role in determining how athletes perform academically in the face of crippling stressors.

The Role of Negative Stereotypes

One stressor that has received relatively little attention is the threat of fulfilling negative stereotypes regarding athletes’ intellectual background and ability (see, Yopyk & Prentice, 2005). Stereotypes of athletes as “dumb jocks” or as criminals are fairly common in American society. Evidence of these negative stereotypes abounds in popular culture, such as in the film “Varsity Blues,” sports magazines, and websites. A limited number of empirical studies have
focused on attitudes about athletes or tendencies to spontaneously stereotype athletes. Some studies (e.g., Baucom & Lantz, 2001; Engstrom, Sedlacek & McEwen, 1995) have demonstrated that faculty and students tend to view athletes as less academically qualified than their non-athlete peers. One recent study demonstrated that students are also likely to spontaneously stereotype football players as being more narcissistic than non-athletes, suggesting that there may be a character component to the “dumb jock” stereotype (Elman & McElvie, 2003).

The present study examines the effect of such negative stereotypes on athletes’ academic performance. Research on stereotype threat (Steele, 1997) has found a link between negative stereotypes and poor performance across a broad range of domains. Steele and Aronson (1995) found that African American participants underperformed on a difficult intellectual ability test when negative stereotypes about their group’s intellectual ineptitude were made salient or when the test was characterized as diagnostic of ability. Similar performance decrements have been reported for various groups across different performance domains. For example, the negative effects of stereotype threat have been observed among women in the domain of quantitative ability (e.g. Schmader & Johns, 2003; Spencer, Steele and Quinn, 1999), white men, when compared to Asian men on quantitative tests (Aronson, Lusitana, Good, Keough, Steele, & Brown, 1999), and elderly people, when stereotypes regarding memory loss are salient (Hess, Auman, Colcombe, & Rahhal, 2003). A study by Stone, Lynch, Sjomeling, & Darley (1999) even demonstrated the effects of stereotype threat on an athletic task. Stone and his colleagues framed a miniature golf task as a “test of sports intelligence” or “a test of natural athletic ability.” In keeping with popular stereotypes, black subjects tended to outperform white subjects when “natural athletic ability” was primed, whereas white subjects dominated when the test was said to measure “sports intelligence.”

In keeping with Steele’s (1997; see also Spencer et al, 1999; Steele & Aronson’s 1995) theoretical model, priming negative stereotypes about one’s group creates performance pressure and subsequent underperformance in the relevant domain. However, reducing stereotype threat by rendering the stereotype less relevant to the performance situation (e.g., by characterizing the test as non-diagnostic of intelligence) improves test performance. Given the constellation of negative stereotypes identified by Baucom and Lantz (2001) and others (e.g., Elman & McKelvie, 2003) regarding athletes, it seems likely that intercollegiate athletes would suffer similar deficits in performance when negative stereotypes about their athlete identity are made salient. Upthegrove, et al. (1999), Shulman & Bowen (2001) and Richards & Aries, (1999) demonstrated that athletes in the “revenue sports” (football, men's basketball and men's hockey) tend to have relatively weak admissions credentials and subsequently poor academic performance in college. These factors are likely to prompt concerns about fulfilling negative stereotypes (e.g. Elman & McKelvie, 2003) and are very similar to the stereotype threat antecedents identified by Steele & Aronson (1995). Lower profile athletes are less likely to benefit from recruiting practices (i.e. lowered admissions standards) and generally more closely resemble non-athletes in terms of admissions credentials, socio-economic status and race (Richards & Aries, 1999, Shulman & Bowen, 2001). Engstrom, et al. (1995) also found that faculty members were more likely to have negative views of revenue athletes than non-athletes or non-revenue athletes. These lower-profile athletes are therefore less likely to be affected by negative stereotypes and subsequent performance deficits. The current study therefore focused on revenue athletes exclusively.
The principal goal of this study was to examine how negative stereotypes affect revenue athletes’ academic performance. One distinguishing feature of athletes (from other groups that have previously been studied), is that athletic identity is defined by behavior or by “choice” and members of the group might be considered to have an “escape” when confronted with the stereotypes. We proposed that since the stereotypes about athletes’ academic incompetence involve a character component, priming a college student’s identity as an athlete should lead to underperformance on an intellectual test, regardless of whether a test is thought to be diagnostic or non-diagnostic.

The Role of Self-Construal

Another important factor that might play a role in athletes’ academic performance is athletes’ perception of their academic ability. According to Bandura’s (1977) theory of self-efficacy, individuals’ understanding of their abilities serves as a guide that motivates them toward goal-attainment behaviors. Bandura conceptualized self-efficacy as a generative capability, which represents the ability to harness the cognitive, emotive, social, and behavioral skills that one has in a given situation to address a particular need. Consistent with this suggestion, we would expect that student athletes who feel efficacious about their academic competency at any particular time would be more likely to perform well on a given intellectual task. The second goal of this research was to determine how student athletes’ self-reported efficacy predicts their intellectual performance.

We anticipated that concerns about being judged by the negative stereotypes regarding athletes’ intellectual ability and character would cause them to perform poorly on an intellectual ability test when their athletic identity is primed. We expected this underperformance to occur even when the test is explicitly characterized as non-diagnostic of intellectual ability. We also expected that participants’ level of self-efficacy would predict their test performance, such that the more efficacious they feel, the better they would perform.

METHODS

Participants

Seventy-two male intercollegiate athletes who play football, basketball, or hockey at an elite liberal arts college in the Northeast were recruited to participate in the study. Participants (mean age of 19.29 years) were paid $5 compensation for their time.

Design & Procedure

The study was a 2 (Test characterization: Diagnostic vs. Non-diagnostic) x 2 (Identity prime: Salient vs. Non-salient) between-subjects design. The experimenter sent e-mail solicitation to a targeted list of athletes and scheduled a time for them to come to the lab. When participants arrived at the lab in groups of one to three, the experimenter introduced them to the study and exposed them to one of the experimental manipulations, depending on the randomly assigned condition.
**Test Characterization Manipulation**

All participants were presented with an ability test with either one or two slips of paper stapled to the front. For half of the participants (Non-diagnostic condition), the slip contained a benign characterization of the test that explicitly denied any relationship between performance on the test and intelligence. The other half of the participants (Diagnostic condition), on the other hand, read a rather direct message that suggested that their scores on the test would provide a reliable indicator of their intelligence.

**Identity Salience Manipulation**

In addition to the test characterization slip stapled to the front of the test, half of the participants received a second slip of paper attached to the ability test. Three questions asked them to 1. indicate whether they play any intercollegiate sport, 2. indicate which sport they play, if any, and 3. rate their likelihood of being accepted to the university without the aid of athletic recruiting on a seven-point Likert scale. Participants in the “Identity salient” condition responded to these questions prior to taking the ability test, whereas those in the “Non-salient” condition responded to the same set of questions after taking the test. In addition to serving as an identity prime, the third question (i.e., self-rated admissibility to college) also provided a useful gage of participants’ academic self-efficacy. In this respect, this measure tapped into some of the most psychologically troubling aspect of the “dumb jock” stereotype; the belief that athletes on some level do not deserve to attend selective colleges.

**Ability Test and Other Measures**

Following the manipulations, participants took a 20-question, relatively difficult, multiple-choice math test from the Graduate Records Examination (see Spencer, et al, 1997). Each participant was given 30 minutes to complete as many questions as possible. The dependent measures were the number of questions participants answered correctly and test accuracy, computed as a percentage of the number of correct responses out of number of questions answered. Finally, each participant responded to a set of questions regarding class year, age, ethnicity, and whether or not they were born the U.S.

**RESULTS**

**Number Correct & Accuracy**

We predicted that participants would score lower on the test when stereotypes about their identity as athletes was primed, regardless of how the test was characterized. A 2 (test characterization; diagnostic vs. non-diagnostic) x 2 (athletic identity; prime vs. non-prime) ANOVA revealed a main effect for identity prime, $F (1, 68) = 7.66, p <.01$. Consistent with our predictions, participants in the athlete stereotype condition answered less questions correctly ($M = 6.44, SD = 3.46$) than did those in the non-salient athlete identity condition ($M = 8.69, SD = 3.45$). The test characterization manipulation effect did not reach statistical significance, $F (1, 68) = 1.87, p >.10$, although the means were in the expected direction ($M = 7.00, SD = 3.14$ for the diagnostic and $M = 8.17, SD = 4.02$ for the non-diagnostic conditions, respectively). The
interaction effect was not significant but further analyses revealed that participants tended to perform equally in all conditions when the test was framed as diagnostic of intelligence. When the test was thought to be non-diagnostic, however, participants in the identity prime condition underperformed relative to those in the non-prime condition.

We also conducted analyses on accuracy (expressed as a percentage of the number of questions participants answered correctly out of those they completed). As with number of correct responses, we obtained a significant effect for identity prime, $F(1, 68) = 4.53, p < .05$. Participants in the athlete identity prime condition were less accurate on the test ($M = 43.86\%, SD = 22.87$) than were those in the non-prime condition ($M = 54.89\%, SD = 20.58$). The test characterization manipulation and the interaction between the two manipulations did not reach statistical significance, though, again, the means were in the expected direction for the test characterization effect.

Self-Perception

We examined the relations between participants’ ratings on the college admissibility measure and their test performance, as indexed by the number of questions answered correctly and accuracy. Participants’ responses on the college admissibility ratings were positively correlated with their test scores and accuracy on the test, $r(71) = .25$, $p < .05$ and $r(71) = .28$, $p < .05$, respectively. That is, the more participants believed that their athletic ability aided their college admission the worse they performed on the test. Because some of the participants completed the college admissibility measure before the test while others completed after the test, we examined the effects of the manipulations on the measure. A 2 (test characterization) x 2 (identity prime) ANOVA yielded no significant effects, suggesting that scores were equal across conditions.

DISCUSSION

This study tested the effects of negative stereotypes on college athletes’ intellectual test performance. We predicted that athletes would underperform on an intellectual test under high threat conditions. Consistent with this expectation, participants saw significant deficits in test performance when we primed negative stereotypes about athletes. Thus, athletes’ intellectual performance can be negatively affected by the “dumb jock” stereotype, especially if such stereotypes imply that athletes were given unmerited preferential treatment in college admissions. This finding suggests that athletes’ tendency to underperform academically may be fueled by salient negative stereotypes in higher education, as Shulman and Bowen (2001) noted (see also, Baucom & Lantz, 2001; Engstrom & Sedlacek, 1991; Elman & McElvie, 2003).

The role of the diagnosticity manipulation was less clear. We should note that average scores were lower among the participants who were led to believe that the test was diagnostic of intelligence compared to participants in the non-diagnostic condition, though differences were not statistically significant. This non-significant finding may actually reflect a different set of stereotype processes than those triggered by the stereotype salience manipulation. As Steele and Aronson demonstrated in 1995, the mere threat of intellectual evaluation may render stereotypes about race and other identity factors salient. The diagnosticity manipulation may therefore have activated racial stereotypes in participants who identified as people of color. Performance
deficits among this cohort may account for the differences between the diagnostic and non-diagnostic groups. Controlling for ethnicity did not render the differences significant, however, perhaps because of the small number of participants who identified as people of color.

It is quite plausible that the diagnosticity manipulation would not have automatically triggered negative stereotypes about athletes, because of dual “student-athlete” identity (Killeya, 2001). Most if not all of the athletes in the study—conducted at an elite university—would have experienced academic success through high school (most of the participants were freshman with little exposure to collegiate coursework). The institution in question was included in the Shulman & Bowen (2001) study, although only in the aggregate amongst comparable elite division III institutions. This study found that, although athletes at these institutions lag behind their non-athlete peers, they still possess exceptional SAT scores and high school grades. It is therefore possible that the diagnosticity manipulation did not prime “poor student” stereotypes as in the Steele and Aronson (1995) study.

We also predicted that performance on the intellectual ability test would vary as a function of self-efficacy, as measured by the self-reported likelihood of admission provided by each participant. This finding was confirmed, and, although somewhat intuitive, merits discussion. It suggests that academic performance may vary as a function of self-efficacy. Athletes who viewed themselves as more efficacious students—i.e., more likely to be accepted without the aid of athletic recruiting—performed at a higher level than participants who believed that they benefited from athletic recruiting and would not have otherwise been accepted. Given that stereotypes of athletes may reduce self-efficacy, at least with regard to specific domains, such stereotypes may directly hinder performance in those domains. More research is needed to verify this hypothesis, however. Clearly it is also possible that less academically suitable participants simply appraised their college admissibility as relatively low and then saw relatively poor performance on the test, commensurate with their intellectual abilities. According to this model, more apt participants would have rated themselves higher on the admissibility scale and seen higher test scores. Unfortunately, the design of the study did not include concrete indicators of past academic performance or potential. As noted earlier, athletes tend to have lower high school grades and standardized test scores compared to non-athletes (Shulman & Bowen, 2001). Therefore, their generally low admissibility ratings seem reasonably accurate. Future research would do well to include SAT scores and/or high school grades to determine the accuracy of the self-reported admissibility. Regardless, this finding is consistent with Killeya (2001), who found that role-elaboration—another proxy measure of intellectual self-efficacy—was related to academic performance in college athletes.

The findings obtained in the present study lend significant support to the notion that stereotype threat may affect athlete populations. College athletes fit the two criteria that Steele and Aronson identified as antecedents for stereotype threat. First, they are likely to underperform when compared to students who are not involved in athletics (NCAA, 2002, Purdy, Eitzen & Hufnagel, 1982). This is especially true of “revenue athletes” (male athletes involved in football, basketball and hockey) (Upthegrove, Roscigno & Charles, 1999, Purdy, Eitzen & Hufnagel, 1982). Second, athletes are likely to be exposed to a range of negative stereotypes in higher education (Baucom & Lantz, 2001 Engstrom & Sedlacek, 1991, Elman and McElvie, 2003, Shulman & Bowen 2001). These results, coupled with earlier findings regarding spontaneous negative stereotyping
of athletes underscore the need for interventions to attenuate the impact of stereotype threat. Recently, a number of studies have focused on creating mentoring relationships designed to buffer at-risk individuals from performance deficits due to stereotype threat. One such study by Good, Aronson, & Inzlicht, (2003) encouraged mentors (college students) to frame intelligence as inherently malleable. Mentors suggested to their 7th grade mentees that regardless of one's current state of knowledge and cognitive ability, infinite intellectual growth was possible in the future. Good et al. found that mentees who received this message were significantly more resistant to stereotype threat than mentees who received alternative messages.

Good et al. (2003) intervention method is particularly applicable to athletic venues, where young athletes are apt to admire older mentors, such as coaches and more experienced athletes. Interestingly, a study by Aronson, Fried & Good (2002) created a similar manipulation where college student “mentors” were led to believe that they were tape-recording a message for at-risk mentees. Some of the “mentors” were encouraged to express the same “malleability” message as in the Good, Aronson, & Inzlicht, (2003) study, while others expressed ostensibly neutral messages. The mentors were later subjected to an intellectual test under stereotype threat conditions. Aronson, Fried and Good found that those mentors who had recorded messages emphasizing malleability demonstrated greatest resistance to stereotype threat. This latter study is of particular interest because it suggests that a mentoring relationship may be mutually beneficial to the mentor and the mentee. College coaches and athletic directors would do well to incorporate such mentoring programs into athletes' activities. In this way, older athletes could provide younger athletes with the tools to resist the stigma of athletic stereotypes, as well as imparting such resistance to themselves.

REFERENCES


**APPENDIX A: DESCRIPTIVE STATISTICS**

**Dependent Variable: Number of Correct Responses**

<table>
<thead>
<tr>
<th>Diagnosticity</th>
<th>Stereotype</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-diagnostic No prime</td>
<td>9.6667</td>
<td>3.61370</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>prime</td>
<td>6.5882</td>
<td>3.90607</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>8.1714</td>
<td>4.01824</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Diagnostic no prime</td>
<td>7.7222</td>
<td>3.08327</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>prime</td>
<td>6.3158</td>
<td>3.11007</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>7.0000</td>
<td>3.13581</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Total no prime</td>
<td>8.6944</td>
<td>3.45435</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>prime</td>
<td>6.4444</td>
<td>3.45952</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>7.5694</td>
<td>3.61463</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>

**Dependent Variable: Accuracy (Correct Responses/Total Responses)**

<table>
<thead>
<tr>
<th>Diagnosticity</th>
<th>Stereotype</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-diagnostic No prime</td>
<td>57.8311</td>
<td>19.66074</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>prime</td>
<td>43.9759</td>
<td>22.39153</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>51.1014</td>
<td>21.87646</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Diagnostic no prime</td>
<td>51.9543</td>
<td>21.62207</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>prime</td>
<td>43.7539</td>
<td>23.90244</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>47.7435</td>
<td>22.88454</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Total no prime</td>
<td>54.8930</td>
<td>20.58417</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>prime</td>
<td>43.3758</td>
<td>22.87008</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>49.3758</td>
<td>22.30638</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B: CORRELATION TABLE

<table>
<thead>
<tr>
<th>Admissions Rating</th>
<th>Correct Responses</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admissions Rating</td>
<td>1</td>
<td>.249*</td>
</tr>
<tr>
<td>Correct Responses</td>
<td>.249*</td>
<td>1</td>
</tr>
<tr>
<td>Accuracy</td>
<td>.287*</td>
<td>.786**</td>
</tr>
</tbody>
</table>

*=P<.05  **=P<.005

APPENDIX C: STEREOTYPE PRIME

Participant Survey
For statistical purposes, please respond to the following questions before you take the test.
Thank you.
1. Do you play a varsity sport at Wesleyan?
2. If yes to 1, which sports? (List) _____________________________________________
3. If yes to 1, estimate the likelihood that you would have been accepted to the university if sports had not been a factor in admissions: (circle one number)
   1 2 3 4 5 6 7
   0% Chance 50% Chance 100% Chance

APPENDIX D: DIAGNOSTICITY MANIPULATIONS

Non-Diagnostic Condition
Stage 1: Early Trials
The test you are about to take is in the early stages of development. Your participation will help us to better understand which questions should be retained for the final version of the test and which should be omitted. Because the test is in such an early stage of completion, we are not concerned with your final score; it is not a reliable measure of your abilities. However, it is important that you take the test seriously and attempt to complete every question. In this way, we will have as much information as possible to determine which questions should be retained for the final version of the test.

If you are unable to solve a problem, don’t worry. Just continue on to the next question. You can spend as much time as you like on any one question, but be aware that we will be limiting you to 30 minutes for the entire test. Try to answer as many questions correctly as you can during the 30 minutes.
Good luck and do your best!

Non-Diagnostic Condition
Stage 5
The test you are about to take is in the final stages of preparation. The test provides a measure of your ability in mathematics: a good predictive measure of your future success in college. To get the best possible score, you should attempt to answer every question correctly. You will receive one point for each question you answer correctly. You will not receive a point for questions you answer incorrectly or fail to answer. You may not ask the researcher any questions during the test.
Good luck and do your best!
AUTHOR BIOGRAPHIES

Matthew Jameson is a recent graduate (2005) of Wesleyan University with a degree in Psychology and French Studies. He is currently employed at the Human Services Research Institute investigating mental health service delivery systems. He intends to pursue graduate study and a career in clinical psychology. He may be contacted at mjameson@wesleyan.edu or mjameson@hsri.org.

Robert Diehl is a recent graduate (2006) of Wesleyan University with a degree in Psychology. He is currently employed as a teacher and coach at Kent School in Kent, CT, and intends on pursuing a career in psychology and education. He may be contacted at rdiehl@wesleyan.edu.

Dr. Henry Danso, Ph. D is an Assistant Professor of Psychology at Wesleyan University. His research interests include attitudes toward immigrants and group-based influences on the academic performance of majority and minority groups, among other topics in social psychology. He may be contacted at hdanso@wesleyan.edu.