The Journal of Social Psychology, 155: 238–254, 2015

ISSN: 0022-4545 print / 1940-1183 online
DOI: 10.1080/00224545.2014.998979

Routledge
Taylor & Francis Group

Racial Athletic Stereotype Confirmation in College Football Recruiting

GRANT THOMAS
JESSICA J. GOOD
ALEXI R. GROSS
Davidson College

ABSTRACT. The present study tested real-world racial stereotype use in the context of college athletic recruiting. Stereotype confirmation suggests that observers use stereotypes as hypotheses and interpret relevant evidence in a biased way that confirms their stereotypes. Shifting standards suggest that the evaluative standard to which we hold a target changes as a function of their group membership. We examined whether stereotype confirmation and shifting standards effects would be seen in college football coaches during recruiting. College football coaches evaluated a Black or White player on several attributes and made both zero- and non-zero-sum allocations. Results suggested that coaches used the evidence presented to develop biased subjective evaluations of the players based on race while still maintaining equivalent objective evaluations. Coaches also allocated greater overall resources to the Black recruit than the White recruit.

Keywords: athletic, football, race, recruiting, shifting standards, stereotype, stereotype confirmation

IMAGINE THE FOLLOWING SITUATION: An excited college football coach notices a standout Black player while attending a high school football game. He raves about the young man's "natural athletic ability" and "God-given talents" and rushes to invite the player on an official campus visit. At the same game, the coach also notices the White running back on the opposing team. For some reason, the coach is not as awe-struck with this player as he was with the first, and it is not until mid-game that he realizes the two players actually have similar game statistics. In this hypothetical example, two players of differing racial backgrounds but with virtually the same performance—same number of yards, same number of carries, same number of touchdowns and big plays—are evaluated differently. Why might this happen? Two psychological theories help to account for this possible differential treatment based on player race: stereotype confirmation (see Nickerson, 1998) and shifting standards (Biernat, Manis, & Nelson, 1991). To our knowledge, researchers have not investigated how these theories might be evidenced in real-world contexts such as the one described above. In the present research, we seek to remedy

Address correspondence to Jessica J. Good, Davidson College, Department of Psychology, Box 7136, Davidson, NC 28035, USA. E-mail: jegood@davidson.edu

this gap in the literature by testing the use of race-based athletic stereotypes during college football recruitment.

Racial Stereotyping

The over-simplified recruiting situation presented above demonstrates how racial stereotypes can be used as deciding factors in the recruitment process. Racial stereotypes are overly generalized beliefs about a person that are based upon one's racial identity (Whitley & Kite, 2009). These broad, overarching generalizations come in both positive and negative forms (Whitley & Kite, 2009). For example, Asian Americans are positively stereotyped as very intelligent and academically gifted and, Asian men at least, are simultaneously negatively stereotyped as being sexually inadequate (Lin, Kwan, Cheung, & Fiske, 2005; Wong, Owen, Tran, Collins, & Higgins, 2012). Regardless of the level of positivity or negativity, the use of any form of stereotype is problematic because it takes away from the unique abilities and accomplishments of an individual person. Additionally, the evaluator is disadvantaged by having a misrepresentation of their target.

Blacks, as one of the largest minority groups in the United States (Census.gov, 2011), have many stereotypes associated with their group. From lazy, to dumb, to criminal, Blacks in America have dealt with broad judgments about their group for centuries (Feagin, 1991). Although many stereotypes about Blacks are negative, there is also a widely held positive stereotype of the group in which they are believed to be superior athletes (Hoberman, 1997). The current research uses this racially based athletic stereotype as the basis for which to test stereotype confirmation and shifting standards in the context of college football recruiting.

Stereotype Confirmation

Stereotype confirmation is simply the application of the confirmation bias to the use of stereotypes when evaluating others (Nickerson, 1998). In a classic social psychological experiment, Darley and Gross (1983) showed that when observers knew about a child's membership in a stigmatized group (low socioeconomic status) and had seemingly sufficient evidence on which to base evaluations of her academic performance (ambiguous video of the child completing a test), their evaluations tended to be consistent with the stereotype of lower academic performance. Without any evidence (no video), observers did not utilize stereotypes in their evaluations. When provided with evidence, however, participants used the stereotype as a hypothesis and paid greater attention to hypothesis-confirming evidence, thus showing stereotype consistent evaluations. Nickerson (1998) describes several mechanisms through which stereotype confirmation can occur, including preferential treatment of evidence supporting existing beliefs, use of a positive-case strategy, and overweighting positive confirmation instances.

In the present research, we examine a college football recruiting context. Generally, coaches are given fact sheets with game statistics about potential recruits and either watch the players' games in person or watch a video of highlighted plays. Just as in Darley and Gross's (1983) landmark study, coaches may use the positive athletic stereotype of Blacks as a hypothesis when evaluating the player evidence given to them. We predict that this will lead to a biased evaluation of recruits such that a Black recruit will be evaluated more favorably and given greater resources than an identical White recruit.

Shifting Standards

The Shifting Standards Model builds on stereotype confirmation research and proposes that standards of judgment are developed based on stereotypes, or expectations about a group (Biernat et al., 1991). An individual will be evaluated relative to the group to which they belong and the stereotypes associated with that group. For example, a "tall" man and a "tall" woman may be described in the same subjective terms, but the underlying objective meaning is understood to be different; he is tall *for a man* and she is tall *for a woman.* That is, the objective meaning ascribed to the subjective descriptor changes depending on the context in which it is used (Biernat, Kobrynowicz, & Weber, 2003). This difference in the objective meaning behind subjective evaluations is the basis of the Shifting Standards Model (Biernat, 2012). Therefore, subjective language may have a different meaning when used to describe people of different sexes, races, religious beliefs or economic statuses. Shifting standards have been shown in a variety of areas including physical attributes, emotional characteristics, verbal ability, and work-related behaviors (Biernat et al., 1991; Biernat & Kobrynowicz, 1997; Biernat & Manis, 1994). The world of athletics is also subject to shifting standards (Biernat & Vescio, 2002); the effect is usually observed such that a White individual is held to the lower standard because it is Blacks who are stereotyped as better in athletic situations (Biernat & Manis, 1994).

Shifting standards have also been shown to affect individuals' behaviors toward targets, and specifically their willingness to allocate resources to targets. Researchers have found that both zero-sum resource allocations and non-zero-sum resource allocations can be predicted by objective and subjective evaluations respectively (Biernat & Fuegen, 2001; Biernat & Vescio, 2002). In the literature, a zero-sum resource is conceptualized as anything that a participant has a fixed number of; in other words, giving some of this particular resource to one target means that the participant cannot give it to another target (Biernat & Vescio, 2002). A non-zero-sum resource has been conceptualized in two different ways: something that is unlimited in supply (Biernat & Fuegen, 2001) or something that the evaluator can allocate at no cost to him or herself (Biernat & Vescio, 2002). Both of these conceptualizations (unlimited quantity and no-cost resource) were shown to fit the shifting standards model, such that they were both predicted by the participant's subjective evaluations of the target.

The Shifting Standards Model also informs how people decide whether or not someone possesses a particular trait. Researchers have argued that evaluators will rely on the contrasting and assimilative effects of stereotypes to set the standard at which they judge targets (Biernat & Kobrynowicz, 1997). A contrasting effect is seen when a person stands out by behaving in a way that contradicts a stereotype, while an assimilative effect is one in which a person behaves in a way that is congruent with the stereotype of their particular group (Biernat, 2003). For example, since Blacks are stereotyped as athletic, a Black player will likely be held to a lower confirmatory standard of athletic competence; the player will not need to demonstrate as much athletic prowess in order for observers to pronounce him athletically competent (because he is assimilated to the stereotype). In comparison, since Whites are stereotyped as less athletic than Blacks, a White player will need to demonstrate greater athletic prowess in order for observers to pronounce him athletically competent; his athletic ability is in contrast with the racial stereotype. In two studies using a simulated job applicant evaluation situation and based on the stereotype that women have fewer job-related competencies than men, Biernat and Kobrynowicz (1997) found that participants tended to set higher confirmatory standards for women when compared

to men (stereotypically believed to be high on job-related competencies). In other words, people who are stereotyped to be high on a particular characteristic will have to show less evidence of actually possessing that trait when they are being evaluated in order to surpass the confirmatory competence threshold (Biernat, 2003).

In the hypothetical example mentioned at the beginning of the article, the Black player was believed to be an excellent athlete immediately; because of racial stereotypes, he had to show less evidence of being athletic than did the White player, who did not get a serious look until the game was practically over. In other words, the Black player was offered a zero-sum resource (campus visit) sooner than the White player because it did not require as much evidence for him to demonstrate athletic competence. The White player needed to show more evidence of actually being athletic in order to surpass the higher confirmatory athleticism standard that was set for him.

Athletic Recruiting

The college football recruiting situation presented above illustrates the basic process; a coach sees a player either in person or on film, makes a judgment about that player, and then decides whether to pursue him and offer resources (like official campus visits) to try to entice him to come to a particular school. Coaches do not base their recruiting decisions on precise facts or formulas. Instead, the college football mantra appears to claim that a coach just knows good talent when he sees it (Feldman, 2007). Although a coach possesses a great deal of football knowledge and experience, is he[1] really that skilled of an evaluator, or is it possible that he may overestimate his own abilities? Factors other than the players' statistics—personal beliefs, past experiences, stereotypes—may also inform his evaluations and subsequent decisions.

We do not claim that when making a recruiting decision, coaches simply rely on race-based athletic stereotypes. They receive quite a bit of supplementary information about a recruit, like résumés, highlight tapes, and statistical information, on which to base their decisions. Despite this information, it could be that coaches still rely on their feelings or instincts about recruits to make "high-stakes" decisions; that is, whether or not to offer them a spot on the team or a scholarship. Research has shown that people feel better about decisions when they have made them based on their gut feelings or intuition (Mikels, Maglio, Reed, & Kaplowitz, 2011); thus coaches may similarly be more confident in their gut decisions (or personal, subjective evaluations) in recruiting situations. Coaches invest a great deal of time and money into college athletes based on minimal interaction, and therefore they likely need to feel confident in their decision to recruit.

The Present Study

Using a naturalistic recruiting situation, the current study investigates how race-based athletic stereotypes are utilized by a sample of actual college football coaches. The use of domain experts in the current research (individuals who work in the area that is being investigated, doing the same or similar job as the evaluator in the experiment) is unique in its direct testing of stereotype use. Other studies, especially in the industrial/organizational field, have used domain experts like hiring managers, but have not directly had these individuals evaluate applicants in ecologically valid ways or using ecologically valid materials (Agerstrom, Bjorklund, Carlsson, & Rooth, 2012; Agerstrom & Rooth, 2011; Duehr & Bono, 2006; Horverak, Bye, Sandal, & Pallesen, 2013;

Schein, 1973). Only one other study that we are aware of has attempted to use domain experts in the investigation of shifting standards but used a much more limited evaluation procedure than the current research, and focused on the effect of gender and parental status on applicant ability (Fuegen & Endicott, 2010). Therefore, the current research represents a substantial extension of the literature, testing the application of shifting standards to a more "real-world" context.

In the present study, Division I FCS football coaches were asked to evaluate a recruit and, based on their evaluations of him, to allocate some amount of both zero- and non-zero-sum resources to the recruit. The sample was comprised exclusively of full-time coaches because they are most familiar with making recruiting decisions, thus making them not only experts in the domain being tested (football), but also in the task (recruiting).

The current study draws heavily on the work of Biernat and Vescio (2002). These researchers asked participants to imagine themselves filling the role of a coed softball team manager. As manager, the participants were shown pictures of the people on their team and were told to evaluate them both subjectively and objectively, and then to allocate both a zero- and non-zero-sum resource. In this particular study, the zero-sum resource was playing time; participants had to create a starting line-up with a limited number of spots to allocate. The non-zero-sum item was how much praise participants would give a male vs. a female player on their team when the player hit a single. We utilize a similar paradigm, although our participants are not simply imagining themselves as coaches; instead, as full-time employed coaches, they are enacting their normal job duties by evaluating a recruit and choosing to allocate resources to him. Thus, although our study still relies on a hypothetical situation (i.e., the target player is not a real recruit), the task being asked of participants is one very familiar to them.

Hypotheses

We predict that the signature shifting standards pattern (a difference in objective vs. subjective evaluations of target group members) will be observed in this sample. However, we predict that the difference in subjective versus objective evaluations will be in the opposite direction than that found by Biernat and Vescio (2002). Coaches will use the substantial evidence provided to them to evaluate the players, resulting in similar objective evaluations of a Black and White player. Yet, due to different confirmatory standards, coaches will rate a Black football player more positively than a White player on a subjective scale (*Hypothesis 1 [H1]*). We believe that the amount of specific competency information about the target (e.g., season statistics), coupled with coaches' experience observing players, will prevent them from giving different objective evaluations of the two players. However, we predict that racial stereotypes will surface in coaches' subjective evaluations of the recruits, where "gut feelings" and personal beliefs are more typically valued. This is consistent with a stereotype confirmation prediction, suggesting that coaches will use their stereotypes as hypotheses and selectively attend to evidence that confirms their preconceived expectations (Darley & Gross, 1983; Nickerson, 1998). Additionally, we predict that the Black player will receive greater zero-sum resources (scholarships, official visits, and roster spots) than the White player, however the two players will be allocated similar non-zero-sum resources (recruitment packages, visits and calls from coaches, and requests for more information; *H2*). Given two players with objectively identical performance, coaches will likely allocate similar amounts of non-zero-sum resources to each. Due to a lower standard for confirming athletic

competence for the Black recruit, we predict that coaches will be more likely to allocate costly, limited zero-sum resources to the Black player. In essence coaches will be more likely to give the Black player the benefit of the doubt regarding athletic competence. Finally, we hypothesize that the zero-sum and non-zero sum resource allocations will be differentially predicted by the coaches' objective and subjective evaluations, such that subjective evaluations will predict zero-sum resource allocation but objective evaluations will not (*H3*). In other words, since coaches will be given sufficient information about the recruit to render similar objective evaluations for both the White and Black player, we do not predict that objective evaluations will predict resource allocation. However, if coaches' subjective evaluations are influenced by player race, we predict those same evaluations to predict allocation of limited and highly important resources (i.e., zero-sum resources).

METHOD

Participants

The sample consisted of 47 male, full-time NCAA Division I FCS football coaches[2] with an average of 14.48 years of coaching experience (median = 10 years, min = 1 year, max = 40 years). With the exception of the West coast and Pacific Northwest, the sample was reasonably representative of every region in the country: 45.7% of participants were from the Northeast, 25.7% were from the Midwest, 14.3% were from the Mid-Atlantic, 11.4% were from the Southeast, and 2.9% were from the Southwest. The majority (85.1%) of participants self-reported as White or Caucasian and 14.9% self-reported as Black[3]. The sample was, on average, 38.06 years of age ($SD = 10.82$). Furthermore, the sample was made up of 3 head coaches and 44 assistant coaches who coached a wide variety of positions. Only one coach reported being a running backs coach; however, at the college level, all coaches know more than enough about each position to make a valid evaluation about any recruit that comes across their desk. Coaches were recruited via email and were offered the opportunity to enter a raffle for one of seven $50 gift cards in exchange for their participation.

Materials

Materials can be accessed at https://osf.io/5bdwk/. To make this study as ecologically valid as possible, it was important that coaches be given materials they would normally receive in an actual recruiting package. Therefore, we created an athletic résumé and highlight video for a potential running back recruit. All participants viewed the same recruitment materials, with the exception of the player race manipulation described below.

The running back position was chosen for this study because it is, arguably, one of the most stereotypically "Black" positions in the game of football today. In fact, over the past three years, fans and coaches have voted to send 21 running backs to the National Football League's (NFL) annual all-star game, The Pro Bowl (NFL Pro Bowl). Every one of those players was Black, with no White players even serving as alternates. The predominance of Black athletes at this position is also visible at the college level. At the end of every season, the Associated Press releases their

All-American Selections. Over the last 3 years, with 18 possible spots available, only one White player has received the honor (Stanford running back Toby Gerhart in 2009; CBSsports.com).

Athletic résumé. We created an athletic résumé or fact sheet for a high school senior player in the running back position. Such a sheet is typical and is expected by coaches as it gives them an easy and quick way to learn about the player they are about to evaluate. The résumé outlines very general academic information like the player's grade point average and standardized test scores. It also includes physical attributes like height and weight. The most important part of the résumé describes the player's athletic accomplishments, all-star selections, awards won, and statistics. For the present study, we created a résumé that portrayed a player who demonstrated average-level performance, due to research showing that people are more likely to use shifting standards when evaluating an average or run-of-the-mill target (Biernat, 2009; Biernat & Vescio, 2002).

Highlight tape. The most important part of a recruiting package is the highlight tape because it allows the coach to actually see the player in action. A highlight tape is a collection of plays from games throughout an athlete's career in which he has made a particularly great play. For the present study, we obtained the highlight tape of a former college running back who has a mixed racial heritage (Black and White). Given the little amount of skin exposed by the player's football uniform, the distance at which the footage was taken, and the medium skin tone of the player, the same highlight tape was able to be used regardless of the race manipulation. As an added precaution, we edited the raw highlight tape slightly to remove any plays where the race of the player was made obvious; that is, any plays where the camera zoomed in on him. We also removed any plays in which the player was hard to identify, any play where he got lost in the pile of other players, or view of him was blocked for an extended amount of time. Finally, we removed a select few plays that made the player seem extraordinary; this was done to mimic the statistics portion of the athletic résumé and similarly depict an average high school running back. The final 5-minute long highlight tape was posted privately on YouTube and was then embedded in the survey materials.

Experimental manipulation. In order to manipulate the race of the player, a headshot of the player was shown, and the Black and the White recruit were given names that clearly alluded to their ethnicity. Four names, two for each race, were pretested among an independent sample of undergraduate students ($N = 31$). The participants were shown each of the names and were then asked to independently generate the race of a person with that name. All four names strongly matched the intended race; the majority of participants generated the intended race for each target name. Out of the four, we chose "Montrael Jordan" for the Black recruit (93.11% of participants labeled the name as Black) and "Tanner Klein" for the White recruit (96.56% participants labeled the name as White).

Next, we chose a headshot for each player to accompany the recruiting materials. We selected three photos of high school aged men for each race (Black and White) from the Center for Vital Longevity's face database (Minear & Park, 2004); photos were selected based on medium skin tone so that the headshot could realistically match the highlight tape. Using the same independent sample as for the player names, the 6 photos were pretested for attractiveness, race, and athleticism. The men in the photos wore plain colored t-shirts, neutral facial expressions, and were shown only from the shoulders up. Participants rated each photo's attractiveness on a scale

of 1 (*very unattractive*) to 10 (*very attractive*). Participants rated each photo for athleticism on a scale of 1 (*not at all athletic*) to 5 (*very athletic*). For each photo participants also rated both to what extent they thought the person looked Black and to what extent they thought the person looked White on a scale of 1 (*not at all*) to 5 (*very much*). It was important to include this dimension of race in the photo pretest because the photos that were to be used had to correspond with the highlight tape of the medium skin-toned running back, yet still clearly communicate White versus Black racial group membership. Pretesting materials and data can be found at https://osf.io/5bdwk/.

We chose the two photos that did not differ on attractiveness, $t(28) = -.25$, $p = .81$, $d = -.05$, or athleticism, $t(28) = 1.22$, $p = .23$, $d = .23$. With regard to race, the White photo was rated as appearing significantly more White than the Black photo, $t(28) = 30.76$, $p < .001$, $d = 5.71$, and appearing significantly less Black than the Black photo, $t(26) = -39.66$, $p < .001$, $d = -7.63$. In addition, a one sample *t*-test comparing each photo to the midpoint of the scale showed that the White player appeared more White than the midpoint $t(29) = 21.78$, $p < .001$, $d = 3.98$, and the Black player appeared more Black than the midpoint $t(28) = 19.33$, $p < .001$, $d = 3.59$. Note that four participants neglected to answer all questions, and thus degrees of freedom differ for the various tests presented.

Subjective and objective evaluations. We asked participants to evaluate the players on the dimensions of speed, agility, strength, jumping ability, overall football playing ability, and "field-vision." Each dimension listed above included both a corresponding objective and subjective item. The subjective items asked the participants to rate the player along each dimension on a 6-point scale while the objective items asked the participants to give an estimation of that particular attribute. For example, for the characteristic of speed, the subjective item read: "Rate this player's speed" (1 = *very slow*, 6 = *very fast*). The corresponding objective item read: "Estimate this player's 40 yard dash time," and provided a space where the coach could write in his estimate (the 40 yard dash is the standard indicator of speed in the sport of football). Research on shifting standards has shown that the order in which objective and subjective questions are presented to participants does not affect whether they show the shifting standards pattern (Biernat & Vescio, 2002). Participants in the current study were shown the subjective items first.

Because all of the objective items were on different scales (seconds, starts, etc.), they were standardized before computing a mean score. Additionally, two of the items, agility and speed, were recoded to reflect the fact that a smaller response to these items indicated a more positive rating of the player. Once standardized, correlation analyses revealed that one of the objective items, the estimate of the "good cuts" the player made in the highlight tape, did not relate well to the other objective items, most likely because of the large range of responses (min = 5, max = 40) which may have occurred because the item was not explicitly defined for the participants. Looking at the raw data it appears that some participants estimated the number of *especially* good cuts the player made while others attempted to count *every single* cut the player made. After removing this item, scale reliability for the objective items was acceptable (Cronbach's $\alpha = .75$). An overall objective rating score was calculated by averaging the scores of the remaining 5 items for each participant.

So that the subjective scale appropriately mirrored the objective scale, the field-vision item (subjective equivalent to "good cuts") was removed from the subjective items. Scale reliability was then calculated for the subjective items and revealed good reliability (Cronbach's $\alpha = .89$).

An overall subjective evaluation score was computed as a mean score of all the subjective items. To directly compare objective and subjective scores within each participant, scores were standardized.

Zero- and non-zero-sum resource allocations. The final portion of the player evaluation was made up of conceptual replications of Biernat and Vescio's (2002) zero- and non-zero-sum items. Three zero-sum and three non-zero-sum items were included and participants rated these on scales of 1 (*not at all*) to 6 (*very much*/*very likely*). The zero-sum items were: "Assuming you had a limited number of running back spots available on your team, would you want this player on your team?," "Assuming you had a limited number of athletic scholarships at your disposal, how likely would you be to offer a scholarship to this player?" and, "How likely would you be to extend an invitation for an official visit to this player?" These items are conceptualized as zero-sum because coaches have a limited amount of roster spots, scholarships, and official visits at their disposal. The non-zero-sum items were: "How likely would you be to send this player a recruitment letter/package?" "How likely would you (or someone in your staff) be to go visit this player at his high school or home?" and, "Would you want to request additional tape or information from this player?" These items were conceptualized as non-zero-sum because a coach can send out as many recruiting letters as he wants. A coach can also go visit as many players and see as much film as time allows.

Scale reliability was calculated for both the zero-sum (Cronbach's $\alpha = .94$) and non-zero-sum (Cronbach's $\alpha = .80$) resource allocation items. Overall zero-sum resource allocation scores were computed by taking the average of all three zero-sum items; the same procedure was used to create an overall non-zero-sum resource allocation score.

Procedure

Research assistants gathered the email addresses of all Division I FCS football coaches from their respective university athletic websites. We then sent e-mails to each of the coaches, soliciting their participation in the study. The study, however, was presented as one investigating the effect of peripheral cues in the highlight tape, such as background music or spot-shadowing, on the coaches' perception and subsequent evaluation of the recruit. Coaches were not informed of the true purpose of the study until the end of the study.

All experimental materials were posted online using SNAP survey software and the link to the survey was included in the initial email to coaches. Upon opening the link, participants first saw an informed consent screen. If a coach consented to participate in the study, he was randomly routed to one of two versions of the online materials (Black recruit materials or White recruit materials). Coaches first saw a picture of the recruit they would be evaluating along with his name, position, and jersey number. Next, the coaches were shown the athletic résumé and could look over it for as long as they wanted. The coaches then viewed the highlight tape, and began the player evaluation. They could watch highlight tape as often as they wished, but were unable to return to the highlight tape or résumé while completing the evaluation. Finally, participants were asked to enter non-identifying demographic information. After doing so, a suspicion check was included that asked participants to describe what they thought the purpose of the study was. This was done to ensure that a participant did not realize the true purpose of the study

prior to debriefing. After answering this question participants were brought to a debriefing page explaining the true purpose of the study and thanking them for their participation.

RESULTS

Suspicion Check

On the open-ended suspicion check question, six coaches who evaluated the Black recruit and two coaches who evaluated the White recruit (17% of the sample) gave answers that suggested that they had some idea of the true purpose of the study. All eight of these participants indicated that they thought the study had to do with "biases" or "prejudice" in recruiting. Analyses were conducted both with and without these eight participants, and revealed similar patterns. In an effort to maintain a large enough sample to maximize statistical power, these participants were retained in the following analyses.

Descriptive Statistics

Data can be accessed at https://osf.io/5bdwk/. Table 1 includes descriptive statistics of the four dependent variables, presented by condition. Table 2 includes zero-order correlations between all dependent variables. The two resource allocation variables were positively correlated, however objective and subjective ratings were unrelated.

TABLE 1
Descriptive Statistics for Dependent Variables, by Recruit Race

	Overall		*Black*		*White*		*Black* v. *White comparison*		
	Mean	SD	*Mean*	SD	*Mean*	SD	t	p	d
Subjective rating score	0.00	1.00	0.31	0.96	−0.37	0.94	2.40	0.02	0.72
Objective rating score	0.00	1.00	0.02	1.07	−0.02	0.93	0.11	0.91	0.04
Zero-sum resource allocation score	3.72	1.27	4.09	1.3	3.27	1.09	2.30	0.03	0.69
Non-zero-sum resource allocation score	4.62	1.14	4.86	1.14	4.37	1.1	1.50	0.14	0.44

Note. Subjective and objective scores were standardized.

TABLE 2
Zero-Order Correlations Among All Dependent Variables

	1	*2*	*3*	*4*
1. Subjective rating score	—			
2. Objective rating score	0.13	—		
3. Zero-sum resource allocation	0.82**	0.19	—	
4. Non-zero-sum resource Allocation	0.63**	0.26	0.76**	—

**$p < .01$.

Hypothesis Testing

In order to test *H1*, that is, that the Black player would be rated more positively than the White player on the subjective scale and that the two would be rated equivalently on the objective scale, a 2 × 2 mixed analysis of variance (ANOVA) was performed with recruit race as the between-subjects factor and rating type as the within-subjects factor. Results of the ANOVA revealed a marginal interaction between recruit race and rating type, $F(1,43) = 3.54$, $p = .07$, $\eta_p^2 = .08$. Although not significant, the interaction was consistent with our a priori prediction. Given our small sample and resulting limited statistical power, we decided to proceed with follow-up testing to probe the interaction. We used a Bonferroni correction to account for the inflated error resulting from multiple comparisons; our corrected alpha level was .025. Consistent with predictions, follow-up independent-samples *t*-tests revealed that the Black recruit was rated significantly higher subjectively than the White recruit, $t(44) = 2.40$, $p = .02$, $d = .72$, but the two players did not differ on the objective scale, $t(44) = .11$, $p = .91$, $d = .04$ (see Figure 1).

Next, a 2 × 2 mixed ANOVA was conducted to test *H2*, that is, that the Black player would receive more zero-sum resources than the White player and the White player would receive more non-zero-sum resources than the Black player. Using resource allocation type as the within-subjects factor and recruit race as the between-subjects factor, the ANOVA revealed a main effect of resource allocation type, $F(1,45) = 56.73$, $p < .01$, $d = -1.10$, and a main effect of recruit race, $F(1,45) = 4.20$, $p = .05$, $d = .60$. The interaction was not significant, $F(1,45) = 1.73$, $p = .20$. These results show that, overall, coaches gave out more non-zero-sum resources than zero-sum resources, which is to be expected given the limited nature of zero-sum resources. Furthermore, contrary to predictions, these results show that the Black recruit received more of both types of resources than the White recruit (see Table 1 for means and standard deviations).

Finally, to investigate *H3*, regarding the relationship between rating type and resource type, two regression equations were calculated, with overall subjective rating score and overall objective rating score predicting each type of resource allocation. Results of the first regression revealed that together, both subjective and objective ratings accounted for a large amount of the

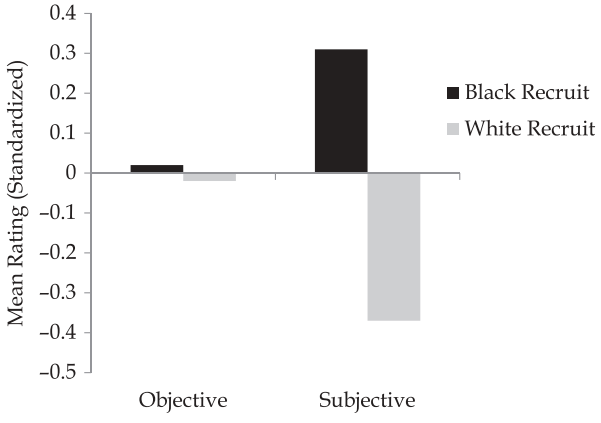

FIGURE 1 This figure illustrates the interaction between recruit race and rating type. Coaches rated the recruits the same objectively but rated the Black recruit significantly better subjectively than the White recruit.

TABLE 3
Regression Analysis of Objective and Subjective Rating Scores on Zero-Sum Resource Allocation

	Unstandardized coefficients		*Standardized coefficients*		
Model	*B*	*Std. error*	*β*	*t*	*Sig.*
(Constant)	3.75	0.11		34.82	0.00
Objective rating score	0.52	0.33	0.14	1.59	0.12
Subjective rating score	1.00	0.11	0.80	9.21	0.00

TABLE 4
Regression Analysis of Objective and Subjective Rating Scores on Non-Zero-Sum Resource Allocation

	Unstandardized coefficients		*Standardized coefficients*		
Model	*B*	*Std. Error*	*β*	*t*	*Sig.*
(Constant)	4.66	0.13		35.39	0.00
Objective rating score	0.75	0.40	0.22	1.89	0.07
Subjective rating score	0.68	0.13	0.60	5.15	0.00

variance in zero-sum resource allocation ($R^2 = .69$). The results showed that subjective ratings predicted zero-sum resource allocations, $\beta = .80$, $p < .01$, but objective ratings did not, $\beta = .14$, $p = .12$ (see Table 3).

For non-zero-sum resource allocations, the regression analysis revealed that both types of ratings accounted for less than half of the variance in the non-zero-sum resource allocations ($R^2 = .44$). Subjective ratings significantly predicted the allocation of non-zero-sum resources, $\beta = .60$, $p < .01$, while objective ratings marginally predicted non-zero-sum resource allocation, $\beta = .22$, $p = .07$ (see Table 4).

H1 stated that recruit race would affect subjective evaluations of the player and *H3* stated that subjective, but not objective, evaluations would predict zero-sum resource allocation. Together, these represent a mediational prediction. We therefore tested whether the relationship between recruit race and zero-sum resource allocation was mediated by subjective evaluations. Following the logic of Baron and Kenny (1986), the total effect of recruit race on zero-sum resource allocation ($\beta = -.36$, $p = .01$) was reduced to non-significance after accounting for subjective evaluations (direct effect $\beta = -.10$, $p = .31$; see Figure 2). The 95% confidence interval for the magnitude of the indirect effect ($-.80$, $-.12$) did not include zero, indicating significant mediation (Tofighi & MacKinnon, 2011). It is important to note, however, that given the correlational nature of this analysis, causal conclusions cannot and should not be made.

DISCUSSION

The present study was designed to expand upon the literature regarding stereotype confirmation and shifting standards, by testing target evaluations and resource allocation decisions within a

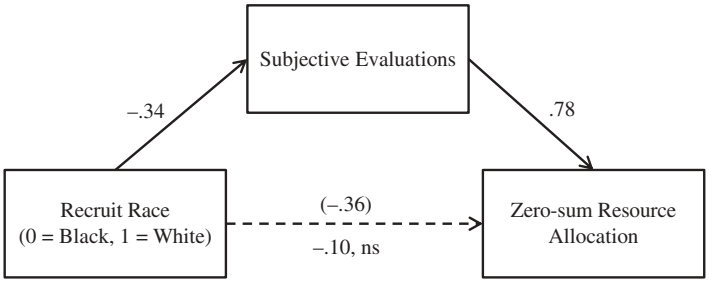

FIGURE 2 Significant mediation of the effect of recruit race on zero-sum resource allocation by subjective evaluations. Standardized regression coefficients are presented. Coefficients are significant at $p < .05$ unless otherwise noted. Dashed line represents mediated path.

football recruiting context. As expected, the Black recruit was rated more positively than the White recruit on a subjective scale and the two recruits were rated similarly on an objective scale. This finding is consistent with a stereotype confirmation effect similar to that found by Darley and Gross (1983). Coaches were given sufficient evidence on which to base their evaluations of the recruit (an athletic résumé and highlight tape); their racial stereotypes may have served as hypotheses, for which they then looked to the evidence for confirmation, unfortunately falling victim to the confirmation bias. On an objective rating scale, the coaches had no choice but to evaluate the recruits based on the concrete evidence given, and they therefore did not show a difference between the White and Black recruit. On the subjective rating scale however, coaches' stereotypes may have led them to set a lower standard of athletic competence for the Black recruit, such that the mediocre performance evidence was sufficient to rate him more positively than the White recruit. Of note, we did not specifically measure coaches' racial stereotypes; we were concerned that a self-report measure would give away the true purpose of the study and skew participants' responses, and an implicit measure (e.g., an implicit association task) would make the study too time intensive to obtain an adequate sample. However, the results of the present study suggest that stereotype confirmation may be the operating process, given that the two recruits were evaluated differently only on the subjective evaluations rather than the objective evaluations.

An informed reader will recognize that the present results differ from the classic shifting standards pattern in which targets are evaluated similarly on subjective measures, but differently on objective measures. However, by demonstrating a discrepancy in objective versus subjective evaluations based on target race, we believe the results still support a shifting standards explanation. One of the major differences between our study and past shifting standards research (e.g., Biernat & Vescio, 2002) is the amount of individuating or supplementary information that the coaches received. In an effort to make the materials and task as ecologically valid as possible, the coaches received all of the typical information about a recruit that they would normally receive: a résumé and a highlight tape. Biernat and Vescio (2002) however, only gave their participants photos to use in judging the athleticism of the targets. By giving our participants substantial information about the players, we equalized objective evaluations. As they tend to have a way of doing however, racial stereotypes may have still surfaced and influenced coaches' subjective evaluations of the players. Moreover, in laboratory simulations, the task is often framed such that the objective evaluation is more consequential than the subjective (e.g., Biernat & Kobrynowicz, 1997). However

in a college football recruiting situation, it is the subjective evaluation that may be more influential than the objective, given that there are many players who will possess similar performance statistics, but only a few who will be hailed as star recruits.

We also predicted (*H2*) that coaches would allocate more zero-sum resources to the Black recruit and more non-zero-sum resources to the White recruit. This hypothesis was partially supported. The results showed that despite rating both recruits the same objectively, coaches allocated significantly more of both zero-sum and non-zero-sum resources to the Black recruit. One possible explanation for this finding is people's belief that the stereotype of Blacks as superior athletes is a positive stereotype. Research has shown that people are less likely to hide their prejudices when they believe they are acting on a stereotype that is positive or complimentary to the target group (Czopp & Monteith, 2006; Kay, Day, Zanna, & Nussbaum, 2013). In this way, coaches may not consider their stereotypes to be problematic and thus may not make an effort to prevent those positive stereotypes from biasing their resource allocations. Furthermore, the results showed that resource allocation decisions were based primarily on coaches' subjective evaluations rather than objective evaluations. This is consistent with *H3*, and suggests that coaches not only may have allowed racial stereotypes to influence their subjective evaluation of concrete performance evidence, but then used those biased evaluations to allocate greater resources to the Black player as compared to the White player. Indeed, although causal inferences cannot be drawn, the results of our mediation analysis suggest that the Black recruit was allocated greater zero-sum resources due to increased subjective evaluations.

Ironically, in a recruiting situation a coach is given evidence about a player that should ideally limit his reliance on stereotypes or other irrelevant pieces of information about the recruit in question. Unfortunately due to a positive-test strategy, coaches may use the available evidence to seek out cases that confirm their initial expectations (racial stereotypes), thereby developing biased impressions of recruits. Although some shifting standards work has utilized supplementary information or evidence about targets prior to participants' evaluation (Biernat, Collins, Katzarska -Miller, & Thompson., 2009; Biernat & Fueguen, 2001; Biernat & Kobrynowicz, 1997), the way that it affects a person's decision making, evaluations, and resource allocation has never been systematically investigated in a shifting standards context. Future research should systematically vary whether supplementary information is given to participants prior to evaluation of a target. It may be that in the absence of evidence (simply a name or photo indicating group membership), participants rely on stereotypes because they have no other information on which to base their judgments. When given evidence however, participants may still make biased judgments, but in this case because they use a positive-case strategy (Nickerson, 1998) when evaluating the evidence. Testing the mediating processes producing these biased evaluations would be a fruitful next step in this line of research.

The present study is limited by its relatively small sample size, and the resulting inability to test whether coach race might influence the pattern observed here. Unfortunately, recruiting coaches to participate in the survey proved to be quite a difficult task. Future research could utilize a larger incentive in an effort to obtain a larger sample. Our sample did appear to be representative of teams across the United States however, leading us to have confidence in our findings. Future research should consider alternative scenarios that are also ecologically valid (e.g., such as employment situations), in order to test whether the results obtained in the present study replicate across various contexts. Additionally, we utilized a between-subjects design in the present study, in an effort to conceal the true purpose of the study. In a real-world recruiting

situation however, coaches receive information about many potential recruits, and compare one to another. Future research should utilize a within-subjects design in order to investigate the specific factors that predict coaches' allocation of resources (including, but not limited to player race) when directly comparing between multiple recruits. As with any study administered online, a limitation of the present study is its lack of control. Some coaches may have paid greater attention to the materials than others. However, given the target sample, we had no choice but to administer our materials via the internet. What the present study lacks in control, we believe it makes up in ecological validity by using real recruiting materials and actual college football coaches as evaluators.

Conclusions

Although the implications of the current research are not particularly encouraging, they are fairly straightforward; coaches differentially evaluated and allocated resources to a Black versus White recruit. Specifically, and of most importance, participants across the board allocated significantly more resources to the Black target. We showed a pattern of results in which coaches' subjective evaluations of recruits may have been influenced by their racial stereotypes, and these evaluations significantly predicted resource allocations that favored the Black recruit. Future research should examine the extent to which this pattern is generalizable to other sports. Would a White golfer receive more zero-sum resources than a Black golfer? Would a Hispanic baseball player receive more resources than a White baseball player? To the extent that racial group membership is stereotyped as contributing to athletic success in a given sport, we would predict that evaluators would allocate greater resources to the positively stereotyped player.

The current research, as well as the studies that inspired it, suggests to fans and coaches alike that they must make a concerted effort to fairly evaluate the teams and players that they view. In general, evaluators must attempt to evaluate others on their own individual merits and, more importantly, use these unbiased judgments to make fair resource allocations.

FUNDING

This research was supported by the Davidson College Psychology Department Fellowship and Davidson College George L. Abernethy Grant.

NOTES

1. For simplicity, we use masculine pronouns to refer to coaches and players throughout the article. Although we recognize that there are female football teams in existence and there is the possibility of a female coach, currently there are no female college football coaches or players. Football remains an incredibly male-dominated sport.
2. The sample was limited to full-time coaches because full-time coaches are the ones who are most familiar with making recruiting decisions. While many Football Champion Subdivision (FCS) programs have volunteer assistants, these coaches rarely, if ever, have a substantial say in recruit selection. This particular responsibility is reserved for those coaches who are part of the full-time staff.
3. Given the limited sample size and the proportionately small number of Black coaches, we were unable to test for racial differences in the way coaches evaluated the recruits.

AUTHOR NOTES

Grant Thomas graduated with a BS in Psychology from Davidson College in 2013. **Jessica J. Good**, PhD, is the L. Richardson King Assistant Professor of Psychology at Davidson College. **Alexi R. Gross** graduated with a BS in Psychology from Davidson College in 2014.

REFERENCES

Agerström, J., Björklund, F., Carlsson, R., & Rooth, D. (2012). Warm and competent Hassan = cold and incompetent Eric: A harsh equation of real-life hiring discrimination. *Basic and Applied Social Psychology*, *34*(4), 359–366. doi:10.1080/01973533.2012.693438

Agerström, J., & Rooth, D. (2011). The role of automatic obesity stereotypes in real hiring discrimination. *Journal of Applied Psychology*, *96*(4), 790–805. doi:10.1037/a0021594

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173

Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, *58*(12), 1019–1027. doi:10.1037/0003-066X.58.12.1019

Biernat, M. (2009). Stereotypes and shifting standards. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (1st ed., pp. 137–152). New York, NY: Psychology Press.

Biernat, M. (2012). Stereotypes and shifting standards: Forming, communicating, and translating person impressions. In P. Devine & E. A. Plant (Eds.), *Advances in experimental social psychology*, Vol. 45 (pp. 1–59). New York, NY: Elsevier.

Biernat, M., Collins, E. C., Katzarska-Miller, I., & Thompson, E. R. (2009). Race-based shifting standards and racial discrimination. *Personality and Social Psychology Bulletin*, *35*(1), 16–28. doi:10.1177/0146167208325195

Biernat, M., & Fuegen, K. (2001). Shifting standards and the evaluation of competence: Complexity in gender-based judgment and decision making. *Journal of Social Issues*, *57*(4), 707–724. doi:10.1111/0022-4537.00237

Biernat, M., & Kobrynowicz, D. (1997). Gender- and race-based standards of competence: Lower minimum standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, *72*(3), 544–557. doi:10.1037/0022-3514.72.3.544

Biernat, M., Kobrynowicz, D., & Weber, D. L. (2003). Stereotypes and shifting standards: Some paradoxical effects of cognitive load. *Journal of Applied Social Psychology*, *33*(10), 2060–2079. doi:10.1111/j.1559-1816.2003.tb01875.x

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*(1), 5–20. doi:10.1037/0022-3514.66.1.5

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, *60*(4), 485–499. doi:10.1037/0022-3514.60.4.485

Biernat, M., & Vescio, T. K. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin*, *28*(1), 66–77. doi:10.1177/0146167202281006

CBSsports.com. (2012). Retrieved from http://www.cbssports.com/collegefootball/story/12665429/-associated-press-allamerica-team

Census.gov. (2012). *State & County QuickFacts*. Retrieved from http://quickfacts.census.gov/qfd/states/00000.html

Czopp, A. M., & Monteith, M. J. (2006). Thinking well of African Americans: Measuring complimentary stereotypes and negative prejudice. *Basic and Applied Social Psychology*, *28*(3), 233–250. doi:10.1207/s15324834basp2803_3

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*(1), 20–33. doi:10.1037/0022-3514.44.1.20

Duehr, E. E., & Bono, J. E. (2006). Men, women, and managers: Are stereotypes finally changing? *Personnel Psychology*, *59*(4), 815–846. doi:10.1111/j.1744-6570.2006.00055.x

Feagin, J. R. (1991). The continuing significance of race: Antiblack discrimination in public places. *American Sociological Review*, *56*(1), 101–116. doi:10.2307/2095676

Feldman, B. (2007). *Meat market: Inside the smash-mouth world of college football recruiting*. New York, NY: ESPN Publishing.

Fuegen, K., & Endicott, N. F. (2010). Evidence of shifting standards in judgments of male and female parents' job-related ability. *Current Research in Social Psychology*, *15*, 52–61.

Hoberman, J. (1997). *Darwin's athletes: How sport has damaged Black America and preserved the myth of race*. New York, NY: Houghton Mifflin Company.

Horverak, J., Bye, H. H., Sandal, G. M., & Pallesen, S. (2013). Managers' evaluations of immigrant job applicants: The influence of acculturation strategy on perceived person-organization fit (P-O fit) and hiring outcome. *Journal of Cross-Cultural Psychology*, *44*(1), 46–60. doi:10.1177/0022022111430256

Kay, A. C., Day, M. V., Zanna, M. P., & Nussbaum, D. A. (2013). *The insidious (and ironic) effects of positive stereotypes*. *Journal of Experimental Social Psychology*, *49*, 287–291. doi:10.1016/j.jesp.2012.11.003

Lin, M. H., Kwan, V. S. Y., Cheung, A., & Fiske, S. T. (2005). Stereotype content model explains prejudice for an envied outgroup: Scale of anti-Asian American stereotypes. *Personality and Social Psychology Bulletin*, *31*(1), 34–47. doi:10.1177/0146167204271320

Mikels, J. A., Maglio, S. J., Reed, A. E., & Kaplowitz, L. J. (2011). Should I go with my gut? Investigating the benefits of emotion-focused decision making. *Emotion*, *11*(4), 743–753. doi:10.1037/a0023986

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*, 630–633.

NFL Pro Bowl. (2012). Retrieved 4/30, 2012, from http://www.nfl.com/probowl

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. doi:10.1037/1089-2680.2.2.175

Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, *57*(2), 95–100. doi:10.1037/h0037128

Tofighi, D., & MacKinnon, D. P. 2011. RMediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods*, *43*, 692–700. doi:10.3758/s13428-011-0076-x

Whitley, B. E., & Kite, M. E. (2009). *The psychology of prejudice and discrimination*. Belmont, CA: Wadsworth Publishing.

Wong, Y. J., Owen, J., Tran, K. K., Collins, D. L., & Higgins, C. E. (2012). Asian American male college students' perceptions of people's stereotypes about Asian American men. *Psychology of Men & Masculinity*, *13*(1), 75–88. doi:10.1037/a0022800

Received March 7, 2014
Accepted November 19, 2014