

# Reactions to Ethnic Deviance: The Role of Backlash in Racial Stereotype Maintenance

Julie E. Phelan and Laurie A. Rudman  
Rutgers University

Backlash effects are social and economic penalties for counterstereotypical behavior (Rudman & Phelan, 2008). Five experiments support a model of the role of backlash in racial stereotype maintenance from the standpoint of perceivers and actors (Rudman & Fairchild, 2004). In Experiment 1, perceivers sabotaged Asians and Whites for succeeding in counterstereotypical domains, thereby preventing their future success. In Experiment 2, a White rapper suffered prejudice and economic discrimination, relative to a Black rapper, and prejudice mediated discrimination. Further, actors threatened by backlash for achievement in cross-racial domains responded to success in ways that bolster ethnic stereotypes. For example, Black men and women who feared backlash for academic skill (Experiment 3), and non-Black (Experiment 4) and non-White (Experiment 5) men who experienced backlash for cross-racial achievement, resorted to defensive strategies that preserve racial stereotypes (e.g., refusing to publicize and pursue counterstereotypical talents). Implications for cultural stereotype maintenance are discussed.

*Keywords:* social cognition, stereotypes, stereotype maintenance, impression formation, backlash

When stereotypes are violated, people take notice. The woman smoking a cigar is singled out from the crowd. The man sobbing in the grocery store becomes the disquieting center of attention. People who deviate from expectations pique our interest, but increased attention is not the only consequence. A growing body of research has demonstrated that expectancy violators suffer social and economic penalties (termed *backlash effects*; Rudman, 1998; Rudman & Fairchild, 2004), particularly when gender stereotypes are disconfirmed. For example, female agency and male communality are viewed negatively in job applicants and can lead to discrimination in hiring and promotion decisions, negative performance evaluations, and even sabotage (for reviews, see Eagly & Karau, 2002; Rudman & Phelan, 2008). But do backlash effects extend to other groups? In the present research, we examined the consequences of violating racial stereotypes as a critical step toward answering this question. Specifically, we investigated (a) perceivers' reactions to ethnic deviants and (b) actors' defensive reactions to learning they had violated a racial stereotype.

Because stereotypes are consensually held beliefs about social groups, they function as normative expectancies for group members' attributes and behaviors. As with the violation of injunctive social

norms (Cialdini & Trost, 1998), when group members deviate from stereotypical expectations, they risk negative reactions from perceivers. Although prior backlash research has focused on penalties for stepping out of gender bounds, ethnic lines may be similarly policed. Derogatory labels such as "Uncle Tom" and "Oreo" (black on the outside, white on the inside) for Blacks and "Wigga" for Whites stigmatize people who cross racial bounds. Moreover, research suggests that disconfirming racial stereotypes can lead to social costs. Fryer and Torelli (2006) found that for Black and Hispanic high school students, academic achievement (viewed as "acting White") was associated with decreased popularity, whereas for White students it was associated with increased popularity. In addition, researchers have identified own-group conformity pressure (i.e., jeer pressure) as a significant stressor for ethnic minority college students (Contrada et al., 2001). Thus, deviating from racial stereotypes may result in backlash, and as a consequence, members of ethnic minorities may behave defensively to avoid it. By investigating how perceivers react to targets who violate racial stereotypes as well as how actors react to their own counterstereotypical success, we tested the claim that backlash not only is harmful to targets who violate expectancies but also functions as a mechanism to maintain racial stereotypes in the culture at large.

## Backlash and Cultural Stereotype Maintenance: The Role of the Perceiver

Rudman and Fairchild (2004) proposed a model that outlines the role of backlash in stereotype maintenance from the perspective of both perceivers and actors (see Figure 1). As outlined in the top row of Figure 1, cultural stereotypes evoke a normative standard to which the target is compared; when individuals fail to fit the standard, a contrast effect may cause them to be viewed as deviant (Biernat, Ma, & Nario-Redmond, 2008; Kobrynowicz & Biernat, 1998). This can result in backlash effects, provided perceivers feel justified (e.g., when they feel threatened or desire revenge). Ac-

---

Julie E. Phelan and Laurie A. Rudman, Department of Psychology, Rutgers University.

This research was partially supported by Grant BCS-0109997 from the National Science Foundation. We thank Doris Huang, Patrick Lee, John Norton, Marsha Pederson, Chris Rochford, Jessica Sasko, Jennifer Su, and James Yu for their assistance with Experiment 1; Justin Schleifer for his assistance with Experiment 3; and Marco DiBonaventura and Julian Walcott for their assistance with Experiment 4. We also thank Monica Biernat, Peter Glick, Phil Atiba Goff, Stephanie Goodwin, Corinne Moss-Racusin, and Dave Wilder for their comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Julie E. Phelan, Department of Psychology, Tillett Hall, Rutgers University, 53 Avenue E, Piscataway, NJ 08854-8040. E-mail: jephelan@eden.rutgers.edu

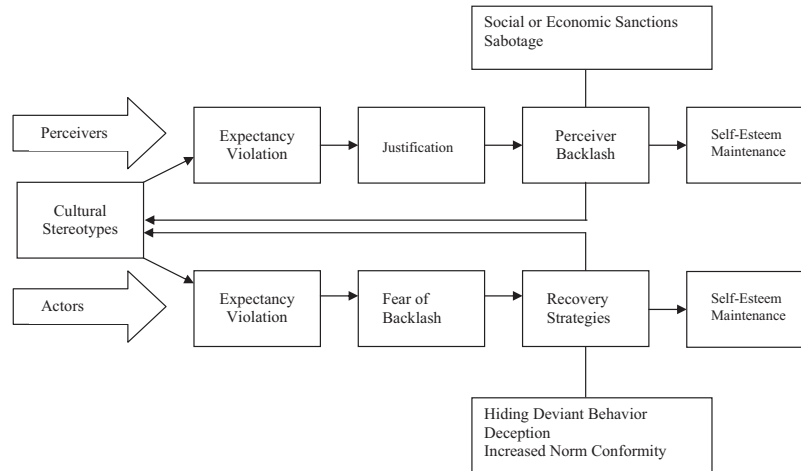


Figure 1. Model of the role of backlash in stereotype maintenance processes (adopted from Rudman & Fairchild, 2004). The top and bottom rows show the hypothesized sequence of events from the standpoint of perceivers and actors, respectively.

According to the model, undermining deviants may afford psychological benefits for perceivers via self-esteem maintenance (see also Branscombe & Wann, 1994; Spencer, Fein, Wolfe, Fong, & Dunn, 1998). This may be especially true when perceivers engage in backlash in response to threats to self-worth (e.g., when deviants outperform participants), because opposing deviants becomes a means of protecting personal self-esteem (Parks-Stamm, Heilman, & Hearn, 2008; Tesser, 1988, 2000). Finally, backlash ultimately reinforces stereotypes by curbing the success and visibility of people who challenge cultural beliefs. Atypical individuals effectively undermine stereotypes (Brewer, 1988; Fiske & Neuberg, 1990), yet backlash disrupts their ability to do so.

In support of the top half of the model, Rudman and Fairchild (2004) found that people who lost a competition to a gender-deviant target were more likely to sabotage the target's future success than were participants who lost to a normative target. Further, sabotaging normative (but not deviant) targets was associated with decreased self-esteem, suggesting that perceivers felt less compunction for undermining deviants. Because sabotaging deviants meant that the deviants would not receive public acclaim, future perceivers' stereotypes would not be challenged. Consequently, people who sabotaged deviants also estimated greater gender stereotyping on the part of future perceivers, suggesting they were cognizant of the crucial role that successful deviants play in dismantling stereotypes (Brewer, 1988; Fiske & Neuberg, 1990).

### Extending the Model to Backlash for Ethnic Deviance

How might backlash toward gender deviants translate to racial deviants? On the one hand, men and women's complementary gender roles and interdependence may cause gender stereotypes to be more fiercely defended than racial stereotypes would be (Rudman & Glick, 2008). Because gender stereotypes are prescriptive (Prentice & Carranza, 2002), people who violate them are viewed as transgressors and thus are at risk for penalties. However, if racial stereotypes are descriptive rather than prescriptive, we

should not expect penalties for crossing racial bounds. Indeed, when minorities disconfirm negative stereotypes, they can be preferred over typical counterparts (e.g., a Black conservative was preferred over a Black panhandler; Rosenfield, Greenberg, Folger, & Borys, 1982; see also Jussim, Coleman, & Lerch, 1987). Nonetheless, this pattern can be reversed when there is justification for backlash, suggesting that racial stereotypes can be prescriptive. For example, threatening perceivers' worldview caused them to prefer a stereotype-consistent Black student over a "bookish" Black student and an effeminate gay man over a masculine counterpart (Schimmel et al., 1999). Further, a surly handicapped target was preferred over a cheerful handicapped target (Katz, 1981), suggesting that minorities can be penalized merely for violating stereotypical expectations. This should not be surprising, given that stereotypes serve numerous cognitive, motivational, and social functions (e.g., Ashmore & Del Boca, 1981; Jost & Banaji, 1994; Snyder & Miene, 1994). That is, negative reactions to atypical targets may signify the extent to which people rely on stereotypes.

Because these findings pertain to minority targets, it is unclear whether stereotypes of dominants (i.e., Whites) are similarly confining and prescriptive. Some evidence suggests that high status affords people greater immunity from the costs of violating group expectancies (Hollander, 1958; Magee & Galinsky, 2008). Compared with low-status groups, high-status groups are perceived as more heterogeneous (Guinote, Judd, & Brauer, 2002; Pratto & Pitpitan, 2008), and their members behave with more variability across situations (Guinote, 2008). Therefore, the relative elasticity and heterogeneity of perceptions of Whites might afford them greater latitude for crossing racial bounds. However, stereotypes based on race and gender inarguably serve to justify social hierarchies (Jost, Banaji, & Nosek, 2004; Rudman, Phelan, Moss-Racusin, & Nauts, 2009), with Whites and men being viewed as more competent and qualified for leadership, relative to racial minorities and women (Foschi, 2000; Pratto & Pitpitan, 2008; Ridgeway, 2001). As a result, defending the social hierarchy may motivate backlash toward dominant group members. If so, the

strongest “rule” for Whites may be that they conform to status expectancies rather than stereotypes. By contrast, minority group members may have less latitude for disconfirming stereotypes than do Whites simply because stereotypes about their group are more narrow and confining (Guinote et al., 2002; Pratto & Pitpitan, 2008).

Experiments 1 and 2 examined reactions to ethnic deviants in the expectation that minorities would be penalized for stereotype violation, whereas Whites would be penalized for status violations. In addition, Experiment 1 provided a test of Figure 1’s hypothesis that backlash toward ethnic deviants functions to maintain perceivers’ self-esteem (Rudman & Fairchild, 2004). Taken together, results of the two experiments should reveal that targets poised to challenge cultural beliefs are prevented from doing so when perceivers undermine their ability to succeed. Therefore, backlash is costly not only to targets (whose aspirations are thwarted) but also to society (because stereotypes are preserved), even though it may psychologically benefit perceivers.

### **Backlash and Cultural Stereotype Maintenance: The Role of the Actor**

The bottom row of Figure 1 outlines how actors who fear backlash may contribute to the preservation of cultural stereotypes. Individuals are likely aware of stereotypic expectancies, and if they deviate from them, they may (rightly) fear backlash from others. From the model we predict that deviants who fear backlash will be motivated to avoid it by engaging in recovery strategies, such as hiding or lying about their counterstereotypical behavior. Rudman and Fairchild (2004) found that, in support of the model, individuals who feared backlash for cross-gendered success cloaked their success, lied about it, and conformed to gender norms on a subsequent task. The result is the maintenance of cultural stereotypes because behaviors that would challenge stereotypic expectations are hidden and/or overcompensated for in a manner that reinforces cultural beliefs.

Finally, because social rejection impacts negatively on self-esteem (Leary & Baumeister, 2000; Williams, 2007), the model predicts that deviants who fear backlash may use recovery strategies to protect their self-regard. That is, actors may resort to defensive behaviors that allow them to avoid backlash from others and thus a loss of self-esteem. However, Rudman and Fairchild (2004) found that gender deviants’ self-esteem was not protected by the use of recovery strategies. Instead, they reported low state self-esteem to the extent they feared backlash, suggesting the power of (even imagined) social rejection to decrease self-worth (Leary & Baumeister, 2000). In Experiment 5, we reexamined this hypothesis with a measure of implicit self-esteem, which should be less reactive than are self-reports (Rudman, Dohn, & Fairchild, 2007).

### **Extending the Model to Backlash for Ethnic Deviance**

Experiments 3, 4, and 5 examined how actors’ defensive reactions to counterstereotypical success can perpetuate cultural stereotypes (see the bottom row of Figure 1). Our goal was to provide participants with information that they had talent in a cross-racial domain in a manner that would mirror real-world situations—an approach that was meant to improve upon past research (Rudman

& Fairchild, 2004). We used cross-racial achievement to instantiate deviance because only successful role models are likely to undermine stereotypes.

Experiment 3 provided an initial test of Figure 1’s predicted links between stereotypes, fear of backlash, and defensive behaviors on the part of atypical actors. We decided, on the basis of prior research suggesting that academically gifted Blacks suffer jeer pressure (Fryer & Torelli, 2006), to inform Black participants that they had excelled on a test of leadership ability and to provide them with the opportunity to publicize their success. We expected that, in accord with Figure 1, participants who perceived Black leaders to be atypical would report more fear of backlash and that fear of backlash would predict their hiding their achievement. These findings would extend Rudman and Fairchild’s (2004) results beyond gender, but they would not afford testing causality.

In Experiments 4 and 5, we corrected this limitation by experimentally manipulating backlash (rather than measuring fear of it). To do so, we led non-Black (Experiment 4) and non-White (Experiment 5) men to believe they had performed well on a counterstereotypical task (i.e., rapping or country singing, respectively). Pretesting determined that each talent was associated more with Blacks or Whites, respectively, than with other ethnic groups.<sup>1</sup> Confederates then administered backlash (or social support) in the expectation that suffering backlash would increase deviant actors’ likelihood of resorting to recovery strategies, including hiding their success, reluctance to pursue their talent, and decreased identification with the atypical domain. These defensive strategies would protect actors from future backlash, but they would also preclude their ability to challenge future perceivers’ racial stereotypes.

Counterstereotypical achievement should serve as a powerful means of reducing cultural stereotypes, but only if actors can be encouraged to take pride in their success. By including social support, in Experiments 4 and 5 we also uniquely tested an intervention strategy designed to remove the threat of backlash. Because social rejection (or the mere threat of it) often results in conforming to group norms to avoid exclusion (Asch, 1955; Leary & Baumeister, 2000; Rudman & Fairchild, 2004), social support should be an effective antidote by assuaging actors’ fear of backlash. Such assurances should then prevent actors’ capitulation to recovery strategies and resistance to becoming an atypical role model. If social support is effective in reducing the threat of backlash, then actors should be less likely to engage in recovery strategies when they are supported for counterstereotypical achievement compared with when they suffer jeer pressure for transgressing racial bounds.

<sup>1</sup> Participants in a pilot test ( $N = 145$ ) indicated how surprised they would be if a Black man were successful at rapping on a scale ranging from 1 (*not at all*) to 5 (*extremely*). They responded to the same item for other ethnicities, including a White man. The difference between the Black and White ratings yielded a large effect size favoring Blacks ( $d = 0.95$ ). We subtracted similar ratings for White versus Black male success at country singing and found a robust effect size favoring Whites ( $d = 1.10$ ). This suggests that rapping is viewed as a typically Black talent (see Experiment 4) and country singing a White talent (see Experiment 5). Effect sizes comparing Latinos and Asians to Blacks (for rapping) and to Whites (for country singing) ranged from 0.73 to 1.12.

## Experiment 1

In Experiment 1 we examined backlash effects for ethnic deviance with Rudman and Fairchild's (2004) sabotage paradigm, because intentionally harming deviants signals the expectancies that are prescriptive and are therefore policed. Participants were defeated by an Asian or White (same-gendered) confederate on a computer task that pretesting indicated was either (a) more typical for Whites than Asians (knowledge of domestic vs. imported beers) or (b) more typical for Asians than Whites (knowledge of Japanese vs. Chinese culture). A third contest condition that was included was atypical for both groups (knowledge of Black jazz vs. hip-hop musicians). A separate test of the status implications of these tasks (described in the Method section) showed that knowledge of Asian culture was status enhancing for Whites, whereas knowledge of beer or Black musicians was status attenuating. These findings afforded the means by which to test our predictions.

After losing the contest, participants were given the chance to sabotage the confederate. If stereotype violation is costly for minorities, Asians should be sabotaged more in the White and Black contest conditions, compared with the Asian domain. Further, perceivers' stereotypes should predict sabotage. However, if Whites suffer backlash for status violation, then they should be sabotaged more in the White and Black contest conditions, compared with the Asian domain, because the latter is high in status despite being atypical. Moreover, perceivers' stereotypes should not predict sabotage.

Finally, we assessed perceivers' self-esteem before and after sabotage. If defending stereotypes justifies backlash toward minorities, then sabotaging normative Asians should be associated with low Time 2 self-esteem (adjusting for Time 1), whereas sabotaging deviant Asians should protect perceivers from this outcome (Rudman & Fairchild, 2004). By contrast, if defending the status hierarchy justifies backlash toward dominants, then sabotaging Whites should be linked to low self-esteem only when their atypical behavior is status enhancing, not when it is status attenuating. In other words, people's self-esteem should be maintained provided backlash is justified; otherwise, undermining others is likely to result in compunction.

## Method

**Testing the status of the tasks.** To pretest the status implications of the contest domains, we asked participants ( $N = 66$ ) if Whites knowledgeable about Asian culture, beer, or Black musicians would raise or lower the status of their group, on a scale ranging from 1 (*lower the status*) to 7 (*raise the status*). Results showed that Asian culture expertise would elevate the status of Whites ( $M = 5.21$ ,  $SD = 0.88$ ) more so than would knowledge of beer ( $M = 4.43$ ,  $SD = 1.04$ ),  $t(65) = 5.07$ ,  $p < .001$ ,  $d = 0.62$ , or Black musicians ( $M = 3.85$ ,  $SD = 1.37$ ),  $t(65) = 7.21$ ,  $p < .001$ ,  $d = 0.89$ . Knowledge of Black musicians was rated as lower in status than knowledge of beer,  $t(65) = 2.95$ ,  $p < .01$ , but the effect size was small ( $d = -0.36$ ). These results afforded the opportunity to test our hypothesis that Whites will suffer backlash for status violations rather than stereotype violations.

We asked the same questions, replacing Whites with Asians, and found similar results when comparing knowledge of Asian culture ( $M = 4.85$ ,  $SD = 1.14$ ) with knowledge of beer ( $M = 3.94$ ,

$SD = 0.94$ ),  $t(65) = 4.95$ ,  $p < .001$ ,  $d = 0.59$ , and Black musicians ( $M = 3.86$ ,  $SD = 1.30$ ),  $t(65) = 4.56$ ,  $p < .001$ ,  $d = 0.54$ . No differences emerged between Asians familiar with beer versus Black musicians,  $t(65) < 1.00$ , *ns*. Although the status results are similar for both groups, we expected them to predict sabotaging only Whites; we expected perceivers' stereotypes to predict sabotaging Asians.

**Participants.** Volunteers ( $N = 204$ ; 141 women) participated in exchange for partial credit toward their introductory psychology research participation requirement. Of these, 87 (43%) were White, 61 (30%) were Asian, 20 (10%) were Black, and the remaining 36 (17%) reported another ethnicity.

### Materials and procedure.

**Precontest measures.** To measure stereotypic expectancies, we asked participants to separately indicate the percentage of Whites and Asians expected to perform well on the elimination round task. As filler, they also reported the extent to which various other groups would perform well. We measured self-esteem as well: once at Time 1, prior to the elimination round, and again at Time 2. To reduce consistency effects, we used the State Self-Esteem Scale (SSES; Heatherton & Polivy, 1991) at Time 1 and the Self-Esteem Scale (SES; Rosenberg, 1965) at Time 2. The SSES was administered on scales ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Sample items include "I feel confident about my abilities" and "I feel inferior to others at this moment." After appropriate recoding, we combined these items to form the Time 1 self-esteem index ( $\alpha = .88$ ).

**Computerized knowledge tests.** The beer contest (White domain) required people to correctly categorize brands of beer as either imported (e.g., Bass Ale) or domestic (e.g., Chili's). The culture contest (Asian domain) required categorizing people, places, and objects (e.g., clothing and vases) as either Japanese or Chinese. The music contest required correctly recognizing photos of Black musicians as either jazz or rap artists.

**Sabotage measure.** Following their loss, participants were asked to program the qualification round, which involved selecting clues for the confederate's "upcoming task" (see Rudman & Fairchild, 2004, for details about the task). Participants were instructed to choose one clue from a list of three possible clues to present to the confederate for each of 12 questions. The clues were selected on the basis of pretesting to vary in their helpfulness from 1 (*helpful*) to 3 (*unhelpful*). Responses were summed to form the sabotage index (range = 12–36;  $\alpha = .69$ ).

**Postcontest self-esteem.** Following the contest, participants completed the SES ( $\alpha = .88$ ). The relationship between Time 1's and Time 2's explicit self-esteem measures was robust,  $r(202) = .64$ ,  $p < .001$ , suggesting that they assess similar constructs.

**Procedure.** Participants were admitted to the laboratory ostensibly with another participant (in fact, an Asian or White same-gendered confederate) for an experiment described as "examining competition under time pressure." The experimenter explained that, in Phase 1, the elimination round, the participants and confederates would compete against each other on a computerized knowledge test that required both speed and accuracy. In Phase 2, the winners would go on to compete in the qualification round. If they performed well on this task, they would be entered in a \$100 cash prize drawing. In order to fulfill their experimental obligation, the losers of the elimination round would help the experimenter set up Phase 2 (in fact, the sabotage task) and would complete a

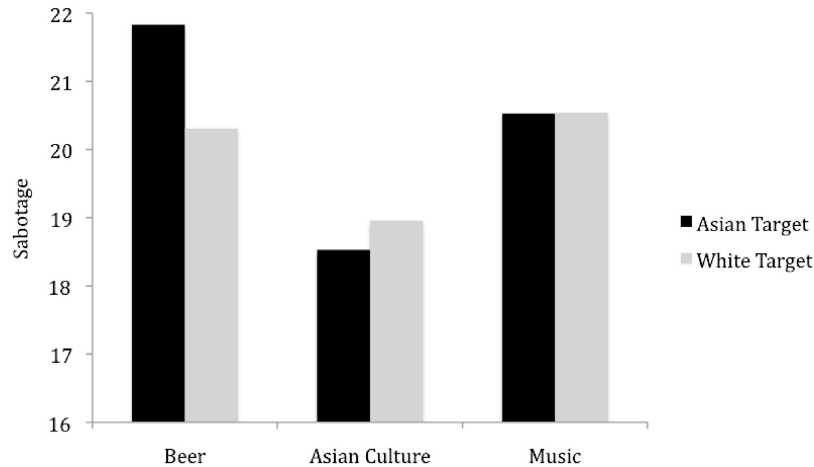


Figure 2. Sabotage as a function of contest domain and target race.

survey. The participants then selected a contest topic by choosing a slip of paper from a box.

Participants were then escorted to separate cubicles, where they completed the precontest measures (SSES and stereotypic expectancies) on a standard desktop PC. The program randomly presented items within each measure. The participant and confederate were then brought together and given instructions for completing the computerized knowledge test. To underscore the confederate's victory, the confederate recorded their final scores and speed and gave the information to the experimenter, who complimented the winner. Following this, the participant and confederate were separated, ostensibly so that each could prepare for the second phase of the experiment.

At this point, participants (as the eliminated "loser") selected clues for the confederate's qualification round, used to measure sabotage (see Rudman & Fairchild, 2004, for details). Participants were then escorted to a private booth, where they completed the SES. Following this, participants were probed for suspicion (no participant indicated that the confederate was the researcher's accomplice) and received a full debriefing.

## Results and Discussion

**Backlash for ethnic deviance.** The main objective was to test whether backlash extends to social categories beyond gender. To do so, we submitted sabotage scores to a 3 (contest: beer, Asian culture, music)  $\times$  2 (target race) between-subjects analysis of variance (ANOVA).<sup>2</sup> Results yielded the expected main effect for contest,  $F(2, 198) = 10.04, p < .001$ . Simple effects showed higher sabotage scores in the beer and music contests, compared with the Asian culture contest (both  $t_s > 2.19, p_s < .05; M_s = 21.05, 20.53, \text{ and } 18.74$  for the beer, music, and Asian culture contests, respectively). No other effects approached significance (all  $F_s < 2.00, p_s > .14$ ).

As can be seen in Figure 2, results confirmed that Asians suffered backlash for stereotype violation. They were sabotaged more in the beer contest than in the Asian culture contest,  $t(61) = 5.44, p < .001, d = 1.07$ , and more in the music contest than in the Asian culture contest,  $t(64) = 2.66, p = .01, d = 0.65$ . They were sabotaged similarly in the beer and music contests,  $t(71) = 1.80,$

$p > .08, d = 0.42$ . Finally, Asians were sabotaged more so than Whites, but only in the beer contest,  $t(69) = 2.52, p = .01, d = 0.50$ . Thus, Asians were more likely to be sabotaged when they performed well at counterstereotypical tasks than when they succeeded in a stereotypical domain.

Figure 2 also confirms expectations for White targets, who were sabotaged more in the beer contest compared with the Asian culture contest,  $t(60) = 2.06, p < .05, d = 0.51$ , and marginally more in the music contest compared with the Asian culture contest,  $t(65) = 1.91, p = .06, d = 0.43$ , but similarly in the beer and music contests,  $t(75) < 1.00, ns$ . Thus, there was no evidence that Whites were sabotaged for violating stereotypes. In fact, when Whites performed well at the Asian culture contest, they were less likely to be sabotaged than if they performed well in the beer or music contests. Because the status of the beer and music contests was lower than the status of the Asian contest, this suggests that Whites risk backlash when their behavior is status attenuating but not when it is status enhancing.

Finally, we expected only Asians to be sabotaged as a function of perceivers' stereotypes. The correlation between stereotypic expectations and sabotage was reliably negative for Asian targets,  $r(99) = -.23, p < .04$ , further supporting the hypothesis that minorities are penalized for violating stereotypes. For Whites, this relationship was weakly negative,  $r(101) = -.15, ns$ . Thus, Whites were not punished for violating stereotypes. Instead, they suffered backlash when their behavior violated status expectancies and therefore undermined the racial hierarchy.

**Psychological consequences of sabotage.** In Rudman and Fairchild's (2004) research, perceivers who sabotaged normative targets reported lower self-esteem than did those who sabotaged deviants. By measuring self-esteem before and after sabotage in Experiment 1, we could control for a generalized self-esteem

<sup>2</sup> There were eight confederates, crossed on ethnicity and gender. Preliminary analyses revealed a null effect for this variable on the sabotage index,  $F(7, 170) < 1.00, ns$ . In addition, there were no main or interaction effects for participant gender (all  $F_s < 2.63, p_s > .08$ ) or participant race (all  $F_s < 2.27, p_s > .11$ ). We therefore collapsed across these variables to increase statistical power.

effect. To do so, we standardized all variables and submitted Time 2 self-esteem to a univariate ANOVA with contest, target race, and sabotage entered as predictors, while adjusting for Time 1 self-esteem. We predicted that sabotaging typical (but not atypical) Asians would be linked to low self-esteem, whereas sabotaging status-enhancing (but not status-attenuating) Whites would be linked to low self-esteem. For both target groups, this would result in more negative correlations when they were sabotaged in the Asian culture contest, compared with the other two contests. Results confirmed the expected Sabotage  $\times$  Contest interaction,  $F(2, 189) = 3.15, p < .05$ . Collapsed across target race and adjusted for Time 1 self-esteem, sabotage was linked to low self-esteem at Time 2 in the Asian culture contest,  $r(48) = -.32, p < .05$ . By contrast, sabotage had no effect on Time 2 self-esteem in the beer contest,  $r(65) = .09, ns$ , or the music contest,  $r(76) = -.11, ns$ .<sup>3</sup>

Table 1 shows the expected first-order correlations among sabotage and Time 2 self-esteem scores within each contest condition, separately for participants who lost to Whites (above the diagonal) and Asians (below the diagonal). Time 1 self-esteem scores are included for the sake of completion; they were not a reliable predictor of sabotage in any condition. As predicted, participants who sabotaged Whites in the Asian culture contest showed low self-esteem at Time 2,  $r(23) = -.40, p < .05$ , suggesting that undermining atypical, high-status Whites evoked compunction. By contrast, self-esteem was unrelated to sabotage in the beer and music contests, indicating that undermining low-status Whites afforded protection from compunction, whether or not their behavior was typical. In addition, sabotaging Asians in the Asian culture contest was linked to low self-esteem at Time 2,  $r(24) = -.41, p < .05$ , whereas this relationship was negligible in the beer and music contests. The pattern supports our prediction that undermining atypical minorities would reap a benefit for saboteurs in the form of self-esteem maintenance.

In sum, Experiment 1 extends backlash effects to penalties for ethnic deviance and supports our prediction that different motives

justify backlash toward minorities and dominants. Asian targets were sabotaged when they transgressed racial bounds, and sabotaging Asians (but not Whites) positively correlated with perceivers' stereotypic expectancies. In addition, perceivers suffered low self-esteem when they sabotaged typical (but not atypical) Asians. Taken together, these findings suggest that backlash toward minorities is justified by defending stereotypes, which, for minorities, may be relatively prescriptive (Katz, 1981; Schimel et al., 1999). By contrast, Whites were sabotaged more so when their behavior—whether typical or atypical—was status attenuating than when their atypical behavior was status enhancing. Further, perceivers suffered low self-esteem when they sabotaged status-enhancing Whites but not when their skills were low in status. Thus, when Whites upheld status expectancies, they were relatively immune to backlash, whereas perceivers not only sabotaged status-attenuating Whites but they did so without compunction. In tandem, these results suggest that backlash toward dominants has a system-justification motive and that, for Whites, deviance is defined more in terms of status violation than stereotype violation. However, regardless of the motives underlying backlash toward ethnic deviants, Experiment 1 provides initial support for the process shown in the top row of Figure 1. When backlash is justified, perceivers engage in behavior that harms ethnic deviants and therefore undermines their ability to challenge racial stereotypes in the culture-at-large.

## Experiment 2

Experiment 1 showed that reactions to Whites varied as a function of the status implications of their behavior, such that low status skills yielded more backlash than did high status skills. This suggests that the strongest prescriptive rule for dominants is that they legitimize the social hierarchy. However, this hypothesis should be tested in a context free of revenge to determine whether revenge is necessary to elicit penalties for deviant Whites. In Experiment 2, conducted under the guise of a market research project, we investigated reactions to aspiring male rap artists (pretests showed rap to be counterstereotypical and status attenuating for Whites). Evidence of backlash would be indicated by prejudicial evaluations of, and withholding economic support from, a White artist compared with a Black artist. Further, economic discrimination should be fully accounted for by greater liking for the Black artist. These results would extend backlash effects commonly found in investigations of gender deviance (Rudman & Phelan, 2008). These findings would further support our hypothesis that (a) Whites are prescribed to behave in a manner that reflects their higher status and (b) motives for backlash toward deviant Whites stem from legitimizing the status quo even when vengeance is absent from the situation.

## Method

**Testing the status implications of rapping.** As noted in Footnote 1, pretesting confirmed that rapping is counterstereotypical for Whites, but Experiment 1 suggested that atypical Whites risk backlash only when their behavior lowers the status of their

Table 1  
Correlations as a Function of Target Race and Contest Condition (Experiment 1)

Target race and contest	1	2	3
Asian culture			
1. Sabotage	—	-.16	-.40*
2. Self-esteem (T1)	-.24	—	.53**
3. Self-esteem (T2)	-.41*	.76**	—
Beer contest			
1. Sabotage	—	-.28	.16
2. Self-esteem (T1)	.12	—	.76**
3. Self-esteem (T2)	-.07	.70**	—
Music contest			
1. Sabotage	—	-.26	-.14
2. Self-esteem (T1)	-.01	—	.59**
3. Self-esteem (T2)	-.06	.51**	—

Note. Correlations for participants who lost to White targets are shown above the diagonal. Correlations for participants who lost to Asian targets are shown below the diagonal. Correlations for T2 self-esteem are partial, controlling for T1 self-esteem. Zero-order correlations were similar. T1 = Time 1; T2 = Time 2.

\*  $p < .05$ . \*\*  $p < .01$ .

<sup>3</sup> There was also a robust main effect for Time 1 self-esteem,  $F(1, 189) > 100.00, p < .001$ , but no other effects were significant (all  $F$ s  $< 2.62, ps > .08$ ), including the three-way interaction,  $F(2, 189) < 1.00, ns$ .

group. In a separate test ( $N = 69$ ), participants were asked, "Does a White man who pursues a career as a rapper raise or lower the status of his ethnic group?" on a 1 (*lowers the status*) to 7 (*raises the status*) scale. They responded to the same item for a Black man. Results showed that pursuing a rap career was perceived as lowering the status of Whites ( $M = 3.33$ ,  $SD = 1.17$ ) more so than the status of Blacks ( $M = 4.30$ ,  $SD = 1.12$ ),  $t(68) = 6.07$ ,  $p < .001$ ,  $d = -0.73$ . Consequently, a White rapper should be at risk for social and economic penalties.

**Participants.** Volunteers ( $N = 167$ ; 108 women) participated in exchange for introductory psychology research credit. Of these, 105 (63%) were White, 36 (22%) were Asian, 10 (6%) were Latino, 10 (6%) were Black, and the remaining 6 (4%) reported another ethnicity.

**Cover story.** Participants believed they were providing feedback about "up and coming" musical artists for a local market research firm. They were asked to listen to two songs that would be "randomly selected" from a database of unsigned artists and to provide their opinions of the music and the artists. Participants were informed that their feedback would be given to a major record label to inform decisions about which new artists to sign. All participants listened to the same rap song and rock song (used to bolster the cover story). The rap song was paired with a photo of either a White rap artist or a Black rap artist (on the basis of random assignment). Two rap songs and four photos (two of a White artist, two of a Black artist) were included to enhance generalizability (a procedural variable that did not influence results). The photo that accompanied the rock song consisted of a band that had both White and Black members.

#### Materials.

**Rap songs.** Two rap songs were downloaded from a website that allows users to upload their self-recorded music (garageband.com). Our selection was guided by (equal) length, likelihood of novelty, and whether artists were unknown and unsigned by a record company. Two filler rock songs were chosen by the same procedure.

**Likability.** Participants indicated how much they liked the song, how much they would like to hear other songs by the artist, whether they were interested in meeting the artist, and how much they thought they would like the artist if they met him on scales ranging from 1 (*not at all*) to 6 (*very much*). Responses were averaged to create the liking index ( $\alpha = .93$ ).

**Musical competence.** Participants also indicated whether they thought the artist was talented and how much they respected the artist on a 1 (*not at all*) to 6 (*very much*) scale. Responses to these items were averaged to form the musical competence index,  $r(166) = .61$ ,  $p < .001$ .

**Economic support.** Participants were asked if they would buy the artist's CD or the song single or go to a concert on a 1 (*definitely*) to 6 (*definitely not*) scale. In addition, two open-ended questions inquired how much money they would be willing to spend to buy the artist's CD or to go to a concert by the artist. Responses were restricted to a whole number from \$0 to \$100. After appropriate recoding and standardizing, these five items were combined to form the economic support index ( $\alpha = .93$ ), on which higher scores indicate greater support.

**Procedure.** Participants completed the study in a private cubicle with a standard desktop PC with speakers. After providing instructions, the computer program randomly selected one of the

Table 2

*Likability, Musical Competence, and Economic Support as a Function of the Race of the Rap Artist (Experiment 2)*

Variable	Black		White		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Likability	2.55 <sub>a</sub>	1.20	1.97 <sub>b</sub>	1.10	0.50
Musical competence	2.84 <sub>a</sub>	1.19	2.47 <sub>b</sub>	1.18	0.31
Economic support	0.14 <sub>a</sub>	0.90	-0.14 <sub>b</sub>	0.85	0.33

*Note.* Economic support is standardized. Means not sharing a subscript differ for Black and White targets ( $ps < .01$ ). By convention, 0.20, 0.50, and 0.80 correspond to small, medium, and large effects, respectively (Cohen, 1988).

two rock songs to play. Participants then completed the measures listed above in random order. The program then randomly selected one of the two rap songs to play, as well as one of the four artist photos to display. After listening to the rap song, participants completed the measures again. Afterward, they reported their gender and race. They were subsequently thanked and debriefed.

## Results and Discussion

**Backlash for ethnic deviance.** Because pursuing a rap career attenuates the status of Whites, we expected participants to like the White rap artist less, perceive him as less musically competent, and offer him less economic support, relative to the (identically talented) Black rap artist. If so, backlash toward Whites who pursue cross-racial talents would be observed, and their ability to function as atypical role models capable of thwarting racial stereotypes would be curbed. Table 2 shows support for each hypothesis. Despite listening to the identical song, participants who believed the rap artist was Black liked the artist more,  $t(164) = 0.24$ ,  $p < .001$ , thought he was more musically competent,  $t(164) = 1.98$ ,  $p < .05$ , and gave him more economic support,  $t(164) = 2.10$ ,  $p < .05$ , compared with those who believed the rap artist was White.<sup>4</sup> Not surprisingly, all three indexes were highly correlated, all  $rs(164) > .58$ ,  $ps < .001$ .

**Mediation analyses.** Gender research consistently demonstrates that prejudice toward agentic women fully accounts for hiring discrimination (i.e., more willingness to hire agentic men than women is mediated by greater liking for agentic men; e.g., Rudman & Glick, 1999, 2001; Rudman et al., 2009). To extend this finding to atypical Whites, we hierarchically regressed economic support on target race (coded 0 = *White*, 1 = *Black*) and likability to test for mediation (Baron & Kenny, 1986). The significant link between target race and economic support in Step 1 ( $\beta = .16$ ,  $p < .05$ ) was reduced to nonsignificance ( $\beta = -.05$ ,  $p = .31$ ) when likability ( $\beta = .85$ ,  $p < .001$ ) was entered as a predictor

<sup>4</sup> The data did not afford testing for gender and race interactions, so we tested these potential moderators separately. A 2 (target race)  $\times$  2 (participant gender) ANOVA revealed no significant main effects or interactions involving participant gender (all  $F_s < 1.19$ ,  $ps > .28$ ). Similarly, a 2 (target race)  $\times$  6 (participant race) ANOVA revealed no significant main effects or interactions involving participant race (all  $F_s < 1.91$ ,  $ps > .10$ ). Therefore, we collapsed across these variables for the focal analyses.

in Step 2. A Sobel's test confirmed successful mediation ( $z = 3.20$ ,  $p < .01$ ). These results suggest that atypical White artists suffer economic discrimination because they are not liked as much as typical Black artists.

In sum, the White artist suffered social and economic backlash for pursuing a low-status, counterstereotypical talent, even though revenge was removed as a possible catalyst. Moreover, prejudice (i.e., relative disliking) fully explained why the White rapper received less economic support than did the Black rapper. In tandem with Experiment 1, the findings in Experiment 2 suggest that people who violate racial stereotypes suffer penalties that reinforce cultural beliefs. Whether sabotaging Asians and Whites who succeeded in atypical domains (e.g., knowledge of Black musicians) or denying the White rapper social acceptance and economic success, perceivers acted as guardians of racial boundaries by hindering atypical role models' ability to challenge racial stereotypes.

Taken together, the results of Experiments 1 and 2 extend the processes described in the top row of Figure 1 beyond gender by demonstrating that ethnic deviants risk backlash in the form of sabotage, prejudice, and discrimination. In Experiments 3, 4, and 5 we examined the processes described in the bottom row of Figure 1 by testing whether fear of backlash causes deviant actors to react defensively to their own counterstereotypical achievements (e.g., by hiding their success and self-selecting away from cross-racial pursuits), thus allowing racial stereotypes to persist in the culture-at-large.

### Experiment 3

In Experiment 3 we made a first examination of Figure 1's links between deviant actors' stereotypes, fear of backlash, and defensive behaviors. Given prior research suggesting that Blacks suffer backlash for academic achievement (Contrada et al., 2001; Fryer & Torelli, 2006), we informed Black participants that they had performed extremely well on a logic and reasoning test (using items from the Law School Admission Test [LSAT]) that ostensibly predicted "leadership in politics and the law." We then offered them the choice to publicize their success on a national website. In addition, we measured participants' stereotypes about Blacks' achievement in politics and law. Support for extending the processes described in the bottom row of Figure 1 to race would be demonstrated if Blacks feared backlash to the extent that they perceived leadership in politics and law to be counterstereotypical and if those who feared backlash were more likely to hide their success. Finally, by measuring participants' actual scores on the test (taken from the LSAT), we were able to assess whether academic achievement for Blacks positively predicts their fear of backlash, suggesting that a history of academic success heightens the salience of its social costs (Contrada et al., 2001; Fryer & Torelli, 2006).

### Method

**Participants.** Black men and women ( $N = 64$ ; 50 female,  $M$  age = 19) were recruited through university and community flyers to participate in exchange for \$20.

**Cover story.** Participants were informed that they were taking part in a nationwide pilot test of "the Leadership and Reasoning Aptitude Task," which was created to identify potential leaders.

The instructions explained that the task assessed non-math-based logic and reasoning skills, because "aptitude in these domains is associated with success in law and politics." Further, participants were told that their scores would be uploaded to a national database and that they would be given feedback on their performance relative to other test takers.

#### Materials.

**Leadership and reasoning aptitude task.** Thirteen questions from the critical reasoning section of the LSAT were chosen to be challenging enough that participants would have difficulty discerning their score. Although all participants received feedback indicating they had scored exceptionally well on the task (in the 98th percentile), we also computed the number of items that participants answered correctly (range = 1.00–9.00,  $M = 5.00$ ,  $SD = 2.10$ ).

**Publicity index.** After receiving feedback on the test, participants were invited to advertise their success on a national website that publicized the top scorers. Specifically, participants decided whether to post their name and/or digital photo to the website by selecting 1 (*yes*) or 0 (*no*). In two separate questions, participants were asked whether they felt uncomfortable posting their name and picture on the website, and also whether they felt obligated to post their name and picture, on scales ranging from 1 (*not at all*) to 7 (*a great deal*). After appropriate recoding, these six items were standardized and averaged to form the publicity index ( $\alpha = .80$ ), on which high scores indicated more willingness to publicize success.

**Fear of backlash.** Regardless of whether they publicized their success, participants were asked to imagine that others knew how well they had performed. They then responded to 13 questions to assess fear of backlash on a scale from 1 (*not at all*) to 6 (*very much so*). Sample items included "Would you be concerned that others would think you were odd?" and "Would your friends be likely to tease you?" (cf. Rudman & Fairchild, 2004). The items were averaged to form the fear of backlash index ( $\alpha = .76$ ), on which high scores indicate more fear of backlash.

**Racial stereotyping.** Participants were asked to estimate what percentage of successful U.S. political leaders and lawyers are Black on a scale from 1 (0–10%) to 10 (90–100%). Responses to these two items were averaged to form the racial stereotyping index,  $r(64) = .75$ ,  $p < .001$ .

**Procedure.** Participants were escorted to individual cubicles, where they were given instructions before the experimenter started the computer program. The program administered the leadership and reasoning aptitude task, followed by the remaining measures (in the order described; items within each measure were randomly presented). Participants then indicated their gender, race, and age. After completion of the test, participants were debriefed, thanked, and compensated for their participation.

### Results and Discussion

Preliminary analyses revealed that men ( $M = 1.79$ ,  $SD = 0.43$ ) marginally feared backlash more than did women ( $M = 1.51$ ,  $SD = 0.48$ ),  $t(62) = 1.99$ ,  $p = .05$ . However, controlling for participant gender did not alter our results, so we present the zero-order correlations in Table 3. As can be seen, results supported the process seen in the bottom row of Figure 1. First, fear of backlash was negatively correlated with stereotyping,  $r(62) = -.31$ ,  $p < .05$ , such that participants feared backlash for their

Table 3  
Correlations Among Dependent Measures for Black Participants  
(Experiment 3)

Variable	Fear of backlash	Stereotyping	Publicity index
Stereotyping	-.31*	—	
Publicity index	-.27*	-.07	—
Test score	.27*	-.09	-.13

Note. A high score on stereotyping reflects greater estimation of the number of Black leaders and politicians. A high score on the publicity index reflects willingness to publicize a high score on the Law School Admission Test (LSAT) items used as a measure of leadership aptitude. Test score reflects actual score on the LSAT items.

\*  $p < .05$ .

success to the extent that they estimated low numbers of successful Black lawyers and political leaders. This indicates that fear of backlash stems from perceived expectancy violations, as predicted in Figure 1. Second, fear of backlash was negatively correlated with participants' willingness to publicize their success on the national website,  $r(62) = -.27, p < .05$ . Thus, in accord with Figure 1, fear of backlash predicted defensive behaviors that prevent atypical actors from challenging cultural stereotypes.

Finally, fear of backlash was positively correlated with actual performance on the leadership and reasoning aptitude task,  $r(62) = .27, p < .05$ . That is, Blacks who scored high on the test were also likely to fear backlash for their success. Because we used actual LSAT items, this suggests that Blacks who perform well academically have a history of suffering backlash and therefore were aware of the threat. Although this interpretation is speculative, it does suggest that fear of backlash was not driven by skepticism of the feedback (leading to the threat of future failure) but is more plausibly derived from past experience with jeer pressure for success in "White" domains (Contrada et al., 2001; Fryer & Torelli, 2006).

In sum, Experiment 3 provides initial support for the process outlined in the bottom row of Figure 1. Specifically, the fact that lower estimates of the number of Black lawyers and politicians positively predicted fear of backlash suggests that cultural stereotypes guide social expectations, and as a result, people who violate them fear backlash. In addition, the association between fear of backlash and the publicity index indicates that fear of backlash for racial deviance operates in much the same way as does fear of backlash for gender deviance (Rudman & Fairchild, 2004), with both resulting in defensive behaviors that are designed to avoid social rejection but that ultimately reinforce cultural stereotypes. Finally, the positive link between actual test performance and fear of backlash suggests that Blacks who excel in stereotypically White domains may have more past experience with backlash and therefore more reason to avoid it in the future. This interpretation is consistent with Fryer and Torelli's (2006) observation that academic achievement was perceived as "acting White" (p. 3) and was associated with decreased popularity for Blacks and Hispanics.

#### Experiment 4

The results of Experiment 2 indicated that White rappers suffer backlash for a stereotypically Black talent, so we returned to this domain in Experiment 4 to experimentally test Figure 1's path

between fear of backlash and defensive strategies. We also tested social support as a potential antidote to the stereotype-reinforcing sequence of events outlined in the model. To do so, we examined behaviors likely to be used by actors eager to avoid backlash. Specifically, we predicted that non-Blacks (i.e., primarily Whites, but also Asians and Latinos) who succeeded at rapping would be less likely to publicize and pursue their atypical talent, and less likely to identify with it, if they received backlash, compared with receiving social support, from a White confederate. Thus, participants who suffered backlash for deviant success were expected to engage in defensive behaviors that would prevent them from challenging stereotypes. These results would provide further support for the process in the bottom row of Figure 1 as it applies to backlash for ethnic deviance. They would also confirm social support as an antidote for conforming to stereotypic norms.

#### Method

**Participants.** Male volunteers ( $N = 138$ ) participated in exchange for credit toward their introductory psychology research participation requirement. Of these, 75 (54%) were White, 34 (25%) were Asian, 14 (10%) were Latino, and the remaining 15 (11%) reported another (non-Black) ethnicity.

**Experimental design and cover story.** The design of the study was a 2 (music video: White or Black rap artist)  $\times$  2 (backlash condition: backlash, social support)  $\times$  2 (confederate gender) between-subjects factorial. We manipulated the race of the rap artist in the event that Whites would fear backlash for "acting Black" primarily when the artist they imitated was Black.<sup>5</sup> Confederates were always White, but we included three men and three women to check on the generalizability of the predicted effects. Participants believed they were auditioning for a future project designed to study "rhythmic ability." It was stressed that no experience was necessary because we were interested in finding people with "raw talent." They watched a music video of a rap song and, following a brief training session, performed a verse from that song while being videotaped.

#### Materials.

**Music videos.** The White music video showed a live performance of "Without Me," by Eminem, and the Black music video showed a performance of "More Money, More Cash, More Hoes," by Jay-Z. After watching the video (ostensibly for inspiration), participants watched a training video. For each condition, we prepared a looped DVD that presented the words of the verse along with the video. The verses were chosen to contain relatively inoffensive lyrics.

**Publicity index.** After being told they had passed the audition—and following confederate backlash or social support—participants were asked whether they would be willing to release their videotape to be shown on a website advertising the project. They were then asked whether they would provide contact information in order to be recruited for the upcoming project that would help them develop their rapping abilities. The experimenter coded responses as 1 (*no*) or 2 (*yes*). Subsequently, participants were asked privately (by a computer program) to respond to two

<sup>5</sup> Manipulation checks confirmed that the two rap videos were rated as similarly likable and helpful and that the difficulty of the audition was similar in both conditions, all  $t_s(136) < 1.61, ns$ .

items—"Were you willing to release your audition videotape?" and "Were you willing to be contacted for the upcoming project?"—on a scale ranging from 1 (*not at all*) to 7 (*a great deal*). These four items were standardized and averaged to form the publicity index ( $\alpha = .74$ ), on which high scores indicated more willingness to publicize success and develop atypical talent.

**Domain identification.** To measure rap identification, we asked participants to respond to five items on a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Sample items were "I think I would be a good rap artist" and "I would be happy learning how to be a rap artist." Items were averaged to form the rap identity index ( $\alpha = .81$ ; overall  $M = 3.75$ ,  $SD = 1.33$ ).

**Procedure.** Participants were admitted to the laboratory with another volunteer (in fact, a male or female confederate) for the "Rhythmic Potential Project." After hearing a description of the audition, participants and confederates were escorted to separate cubicles where they viewed either the White or Black music video, followed by the training video. Participants were then brought to an audition room, where they performed karaoke style, with an instrumental CD (with the lyrics displayed on a computer monitor). When the participants had finished, the experimenter complimented them (by saying "Good job!") regardless of actual ability. Participants were then escorted to a private cubicle, where they waited while the confederate ostensibly auditioned.

After "auditioning," confederates were seated with participants. The experimenter announced that both were eligible for the upcoming project but told the participant, "You were especially good." After the experimenter left to "get the release forms," confederates administered social support or backlash, allowing the experimenter to remain blind to condition. Supportive confederates expressed enthusiasm for the upcoming study and complimented the participant ("That was fun! Should be a fun study to participate in. Good for you for doing so well—the experimenter said you were especially good!"). In contrast, backlash-administering confederates mocked the participant for his success ("That was dumb! The experimenter said you were especially good. What's wrong with you? Do you think you're Black?").

The experimenter returned to ask participants whether they were willing to release their audition tape and provide contact information for the upcoming project (part of the publicity index). After escorting the confederate to a separate cubicle, the experimenter started a computer program that administered the remaining measures in the following order: rap identity index, the remaining publicity index items, manipulation checks, and demographics (ethnicity and age). Items within each measure were randomly presented.

Participants were then probed for suspicion (nobody was eliminated on this basis) and received a full debriefing, including an apology from the confederate for participants in the backlash condition. Enthusiasm for the project was remarkably high, and no participants expressed lingering concerns, suggesting that debriefing was successful.

## Results and Discussion

**Consequences of backlash for stereotype maintenance.** An advantage of the present research was that the threat of backlash was manipulated rather than merely assessed as in Experiment 3. Therefore, if participants refused to publicize or pursue their

success in the wake of backlash, the causal claims of the model in Figure 1 would be supported.

All measures were submitted to 2 (music video: White or Black rap artist)  $\times$  2 (backlash condition: backlash, social support) between-subjects ANOVAs.<sup>6</sup>

**Publicity index.** Results for the publicity index yielded the expected main effect for backlash,  $F(1, 134) = 15.02$ ,  $p < .001$ . No other effects were reliable, all  $F$ s ( $1, 134$ )  $< 1.00$ , *ns*. Men in the social support condition ( $M = 3.43$ ,  $SD = 0.88$ ) were more likely to publicize their success and to pursue their talent in the upcoming study, compared with men in the backlash condition ( $M = 2.74$ ,  $SD = 1.22$ ), resulting in a moderate effect size for this difference ( $d = 0.64$ ).

**Rap identification.** Another potential means by which backlash promotes cultural stereotypes is by discouraging actors from identifying with the skill in question. However, as with the publicity index, we predicted that social support would reduce this tendency. Results showed the expected main effect for backlash on rap identification,  $F(1, 134) = 8.23$ ,  $p < .001$ . No other effects were reliable, all  $F$ s ( $1, 134$ )  $< 1.18$ , *ns*. Men in the social support condition ( $M = 4.06$ ,  $SD = 1.24$ ) were more likely to identify with rap, compared with men in the backlash condition ( $M = 3.42$ ,  $SD = 1.34$ ), resulting in a reasonable effect size for this difference ( $d = 0.48$ ).

**Including a control condition.** To determine whether actors must suffer overt backlash before engaging in recovery strategies, we conducted a poststudy in which White men ( $N = 31$ ) experienced the identical procedure but without confederates (i.e., they received neither backlash or social support). On the publicity index ( $\alpha = .70$ ), their mean score was 2.99 ( $SD = 0.85$ )—a result that was in between Experiment 4 participants who suffered backlash ( $M = 2.74$ ,  $SD = 1.23$ ) and those who received social support ( $M = 3.43$ ,  $SD = 0.82$ ). Planned comparisons revealed that control men scored lower than did supported men,  $t(100) = 2.91$ ,  $p < .01$ , resulting in an adequate effect size ( $d = 0.43$ ). However, the difference between control men and those who suffered backlash was nonsignificant,  $t(96) = 1.13$ ,  $p = .26$ , resulting in a small effect size ( $d = 0.25$ ). Results were similar for the rap identity index for the control men ( $M = 3.35$ ,  $SD = 1.05$ ;  $\alpha = .75$ ). Controls scored lower than did men who received social support,  $t(100) = 2.76$ ,  $p < .01$ ,  $d = 0.54$ , but results were similar for controls and those who suffered backlash,  $t(96) < 1.00$ , *ns*. Thus, without assurances that cross-racial achievement will not be penalized, actors may hide their success, refuse to pursue their talent, and dissociate from their talent. In tandem with the results of Experiment 3, this suggests that actors are aware of backlash for ethnic deviance even without directly experiencing it, just as they are for gender deviance (Rudman & Fairchild, 2004).

In sum, the results of Experiment 4 extended the generalizability of the process in the bottom row of Figure 1 to actors' reactions to ethnic deviance. We also successfully expanded the consequences

<sup>6</sup> Because participant race was not equally distributed across all conditions, it could not be fully analyzed. However, there were no main effects for this variable, all  $F$ s ( $4, 133$ )  $< 1.97$ , *ns*, and including it as a covariate in our analyses did not influence our findings. Further, examining only White participants (and eliminating Asians and Latinos) did not alter our results, so we chose to be inclusive.

of backlash for maintaining cultural stereotypes. Participants who suffered backlash were less likely to publicize their success, pursue their atypical talent, and identify with rapping, compared with participants who received social support. Thus, backlash was shown to have a causal influence on atypical actors' recovery strategies (in support of Figure 1), but social support effectively thwarted the threat of being rejected for counterstereotypic skills. That is, atypical actors who received support were more likely to identify with rapping (in addition to publicizing and pursuing their success) and were therefore poised to disconfirm the stereotype that only Blacks can rap.

In Experiment 5, we sought to experimentally test Figure 1 with backlash toward minority men for acting White. In addition, although the use of several confederates (of both genders) in Experiment 4 increased the external validity of the research, in Experiment 5 we sought more control over the delivery of backlash and social support by using taped recordings.

### Experiment 5

In Experiment 5, non-White men (i.e., primarily Blacks, but also Asians and Latinos) received feedback that they were successful at singing country music. As in Experiment 4, backlash or social support was manipulated, but we used audio recordings (of male confederates) to standardize feedback delivery. We also varied the confederate's race and assessed actors' racial identity to test these as potential moderators.

People derive their self-esteem partly from their social identity (Tajfel & Turner, 1979, 1986), but they differ in the extent to which they identify with their social groups (e.g., Crocker & Luhtanen, 1990). For most White Americans, ethnicity is not an important part of their identity and they can choose what role, if any, it will have in their life (Waters, 1990). But for ethnic minorities, race can be a salient aspect of their identity (Deaux, 1992; Phinney & Alipuria, 1990). Thus, for minority men, reactions to ethnic deviance may depend on the strength of their racial identity and the race of the person administering feedback.

We expected a buffering effect for people highly identified with their race, such that they would be relatively impervious to backlash. This is because such people may be so secure in their racial identity as to not be affected by anyone who questions their ethnicity (e.g., by accusing them of acting White; see Cross, Parham, & Helms, 1998). In support of this view, high minority group identification has been associated with low rates of depression (Arroyo & Zigler, 1995; Munford, 1994), high self-esteem (e.g., Branscombe, Schmitt, & Harvey, 1999;), positive psychological adjustment (e.g., Sellers, Caldwell, Schmeelk-Cone, & Zimmerman, 2003), and academic achievement (Wong, Eccles, & Sameroff, 2003). In sum, minority men high on racial identity might be immune to jeer pressure because racial identification may act as a global buffer against threats to ethnic identity.

By contrast, we expected minority-group members who were low on racial identity to be particularly vulnerable to backlash from majority group members (i.e., a White male). Not only do they lack the buffer provided by racial identification but, compared with high identifiers, low identifiers are less committed to their group and more likely to desire social mobility (Ellemers, Spears, & Doosje, 1997). In other words, they strive for acceptance among higher status groups. Thus, for Black low identifiers, backlash

from a White confederate should be more threatening than backlash from a Black confederate, and social support from a White confederate should be more reassuring than support from a Black confederate.

Finally, we assessed implicit self-esteem by measuring participants' preference for their birthday date and month (e.g., DeHart, Pelham, & Tennen, 2006; Koole, Dijksterhuis, & van Knippenberg, 2001). According to Figure 1, people threatened with backlash risk lowered self-esteem. However, those who engage in recovery strategies should maintain their self-esteem because they eliminate the risk of future backlash by hiding their deviant behavior. Threat-congruency effects have been shown for a similar implicit self-esteem measure based on participants' preference for their own initials (the name-letter effect, or NLE; Nuttin, 1985, 1987). For example, Koole et al.'s (2001) NLE test revealed that failure feedback reduced self-esteem. Because the NLE is well correlated with birthday number preference (Kitayama & Karasawa, 1997; Koole et al., 2001), we expected it to perform similarly. Although racial identity should moderate the use of recovery strategies, we suspected that implicit self-esteem effects would generalize across respondents who were either high or low on racial identity. One reason for predicting this pattern is that people under various types of threat generally respond defensively to protect their implicit self-esteem (Rudman et al., 2007). Moreover, implicit self-esteem measures are often dissociated from their explicit counterparts, indicating that they tap different processes (e.g., Bosson, Swann, & Pennebaker, 2000). Thus, although Rudman and Fairchild (2004) found evidence for reduced self-esteem in deviants who feared backlash (no matter their recovery strategies), an implicit measure might be more sensitive to the defensive process outlined in Figure 1.

### Method

**Participants.** Participants ( $n = 83$ ) included 49 Black male students who were recruited through various strategies (e.g., advertising flyers) and compensated with \$20 and 34 other minority male students who volunteered for credit toward their introductory psychology research requirement (14 were Asian, 12 were Latino, and the remaining 8 participants reported another [non-White] ethnicity).<sup>7</sup>

**Experimental design and cover story.** The design of the study was a 2 (confederate race: Black, White)  $\times$  2 (backlash condition: backlash, social support) between-subjects factorial. Confederates were phantom men whose ethnicity was manipulated through the use of White and Black names and dialects. Participants were told they were auditioning for a project designed to study "stage presence." It was stressed that past experience was not necessary because we were interested in people who had raw talent. Although stage presence is not counterstereotypical for

<sup>7</sup> The initial sample size was  $N = 103$ . Data from 20 participants were dropped either because problems with the intercoms resulted in suspicion ( $n = 15$ ) or because the experimenter made an error. Suspicion did not vary as a function of backlash condition,  $\chi^2(1) = 0.45, ns$ .

minority men, we relied on our instantiation of the domain (country music) to transmit cross-racial deviance (see Footnote 2).

#### Materials.

**Music video.** Because Experiment 4 ruled out artist race as a significant factor, we did not vary the race of the artist in Experiment 5. The music video was "Last of a Dying Breed," by Neal McCoy. After watching the complete music video, participants then watched a training video (as in Experiment 4). This video was chosen on the basis of its potential to heighten deviance for minority men (e.g., the song extols the virtues of "rednecks").

**Publicity index.** As in Experiment 4, the publicity index averaged public behaviors (releasing the audition tape for the website and providing contact information for the upcoming project to develop the participants' talent) with private measures of their willingness to do both ( $\alpha = .71$ ). High scores indicated more willingness to publicize and develop counterstereotypical talent.

**Domain identification.** Three items assessed country music identification on a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The items were "I think I would be a good country singer," "I have imagined myself as a country singer in the past," and "I would be happy learning how to be a country singer" ( $\alpha = .69$ ; overall  $M = 3.80$ ,  $SD = 1.00$ ).

**Racial identity.** We used the four-item identity subscale from the Collective Self-Esteem Scale to measure racial identity (Luhtanen & Crocker, 1992). The items include "The racial/ethnic group I belong to is an important reflection of who I am" and "In general, belonging to my race/ethnicity is an important part of my self-image." Responses were indicated on a 1 (*strongly disagree*) to 7 (*strongly agree*) scale and were averaged ( $\alpha = .74$ ;  $M = 3.80$ ,  $SD = 1.00$ ).

**Implicit self-esteem.** Participants evaluated their birth numbers (day and month) as a measure of implicit self-esteem on a scale ranging from 1 (*dislike very much*) to 9 (*like very much*). Following past research, we separately computed the difference between participants' rating of their birth day (and month) and the mean liking for this number provided by people whose birthday did not fall on that day (or in that month); we then averaged the two difference scores. On average, participants showed implicit self-esteem (i.e., scores were greater than zero),  $t(81) = 4.18$ ,  $p < .001$  (overall  $M = 0.77$ ,  $SD = 1.68$ ,  $d = 0.45$ ).

**Procedure.** The lab was set up to suggest that another volunteer had already arrived (e.g., a coat and books were displayed in the waiting area). After each participant entered the laboratory, the experimenter said, "The other participant, [confederate's name: Donald in the White condition, Jamal in the Black condition], stepped out to take a phone call, but we can start without him." The procedure was similar to that in Experiment 3, with the exception that the participant's cubicle was equipped with an intercom, which he was told to use to communicate with the experimenter.

After each participant's audition, the experimenter escorted him back to his cubicle and then pretended to go get the confederate to bring him into the audition room. To make it appear as though the confederate had tampered with the intercom, the experimenter said over the intercom, "Hey! Stop that! Don't play with those buttons!" The experimenter then returned to the participant to announce that his audition was "especially good" and that he qualified for the upcoming project. The experimenter then left "to audition the other participant." Shortly thereafter, the experimenter played a recording over the intercom so that the participant would

overhear either backlash or social support (from either a White or Black confederate). The recording went as follows:

*Experimenter:* [static] . . . the headphones have been acting up lately. Did they work for you?

*Confederate:* No problem.

*Experimenter:* Most people don't ask to watch somebody else's audition before they do their own, so I'm not sure I should have let you see the guy who auditioned before you. What'd you think?

In the backlash condition, the confederate laughed derisively and said: "Oh man, what's wrong with him? Does he think he's White?" In the social support condition, the confederate said: "That was pretty good! I hope I can do it half as well!" The experimenter's reply was inaudible due to static, and the intercom was then cut off.

After a few minutes, the experimenter returned to the cubicle, offering the participant a tape release form (for the website publicizing the project) and a contact sheet if he agreed to be recruited for the upcoming project. The experimenter then started a computer program to administer the remaining measures (the country music identification index, the racial identity index, willingness to publicize and pursue atypical talent, the implicit self-esteem measure, and demographic items, in that order). Items within each measure were randomly presented. Participants were then probed for suspicion and received a full debriefing that included an apology to those in the backlash condition.

## Results and Discussion

Because participant race was not equally distributed across all conditions, it could not be fully analyzed. However, there were no main effects for this variable, all  $F_s(4, 78) < 2.09$ , *ns*. Further, controlling for participant race in each analysis did not alter our findings, so we collapsed across this variable in the results reported below. In addition, examining only Black participants did not change our results, so we chose to be inclusive.

**Consequences of backlash for stereotype maintenance.** After standardizing, we regressed the publicity index on confederate race, backlash condition, racial identity, and all interaction terms. Only the expected three-way interaction was significant ( $\beta = .25$ ,  $p < .01$ ; overall  $R^2 = .10$ ). The remaining effects were weak (all  $\beta_s < .15$ ,  $p_s > .16$ ). To decompose the interaction, we dichotomized participants on their racial identity (by using the mean). As expected, participants high on racial identity showed no reliable effects (all  $\beta_s < .22$ ,  $p_s > .19$ ;  $R^2 = .06$ ). That is, high identifiers were not significantly influenced by backlash or social support from either White or Black confederates (see Table 4). In contrast, participants low on racial identity showed a main effect for backlash, qualified by the predicted Confederate Race  $\times$  Backlash Condition interaction ( $\beta = -.48$ ,  $p < .01$ ;  $R^2 = .29$ ).

Table 4 shows the simple effects for the two-way interaction. As predicted, low identifiers were less likely to publicize their success and pursue atypical talent following backlash from the White confederate, compared with following social support ( $d = -1.45$ ). This contrast was not significant in the Black confederate condition, suggesting that feedback from Black confederates was ineffective for low identifiers ( $d = 0.22$ ). Comparing across confederate race, low identifiers were less likely to publicize and pursue

Table 4  
*Publicity Index as a Function of Racial Identity, Backlash Condition, and Confederate Race (Experiment 4)*

Publicity index	Low racial identity				High racial identity			
	Backlash		Social support		Backlash		Social support	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
White confederate	<b>2.07<sub>a</sub></b>	0.99	<b>3.78<sub>b</sub></b>	0.61	3.46 <sub>b</sub>	0.34	3.04 <sub>b</sub>	1.61
Black confederate	<b>3.13<sub>a</sub></b>	1.10	<b>2.86<sub>a</sub></b>	1.06	3.03 <sub>a</sub>	1.06	3.71 <sub>a</sub>	1.59

Note. Means not sharing a subscript differ within rows ( $p < .05$ ). Means in bold face type differ between White and Black confederate conditions ( $p < .05$ ).

deviant success following backlash from White, compared with Black, confederates ( $d = -0.89$ ) but were more likely to do so following support from White, compared with Black, confederates ( $d = 0.76$ ). In sum, low identifiers were sensitive to feedback from the White (but not Black) confederate.

**Country music identification.** Results of the same regression analysis for identification with country music showed a main effect for only backlash ( $\beta = .25, p < .05$ ). The three-way interaction was weak ( $\beta = .12, p = .31$ ; overall  $R^2 = .09$ ). On average, participants identified with country music more in the social support condition ( $M = 4.04, SD = 1.18$ ) than in the backlash condition ( $M = 3.56, SD = 0.72$ ), resulting in an adequate effect size ( $d = 0.42$ ). Thus, as in Experiment 4, backlash decreased identification with a counterstereotypical domain—an effect that did not depend on confederate race or participants’ racial identity.

**Consequences of backlash and recovery strategies for self-esteem.** According to Figure 1, counterstereotypical actors who suffer backlash and engage in defensive behaviors (hiding and refusing to pursue their talent) should show higher self-esteem, compared with counterparts who refrain from recovery strategies. Because preliminary analyses revealed the expected null effects for racial identity, we collapsed across this variable.

After standardizing all measures, we regressed implicit self-esteem on backlash, confederate race, the publicity index, and their interaction terms. Results revealed the expected Backlash  $\times$  Publicity Index interaction ( $\beta = .23, p < .05$ ). As shown in Table 5, participants in the backlash condition showed higher implicit self-

esteem if they used recovery strategies, compared with when they did not ( $d = 0.88$ ). Participants in the social support condition did not show this difference ( $d = 0.11$ ). These results support the prediction in Figure 1 that recovery strategies protect deviant actors’ self-esteem when they are threatened by backlash.

In sum, minority men low on racial identity responded to feedback from a White confederate in a way similar to that of the (primarily) White participants in Experiment 4. When threatened by backlash, they were more likely to closet their success and refuse to develop atypical talents, compared with counterparts in the social support condition. In line with Figure 1, these results again suggest that backlash perpetuates cultural stereotypes by depressing atypical abilities and achievement. In contrast, minority men high on racial identity were impervious to feedback from either source. Secure in their racial identity, they were less threatened by intimations that they were acting White. However, identification with a White domain decreased solely as a function of backlash. Moreover, recovery strategies served to protect threatened deviants’ implicit self-esteem irrespective of confederate’s race or participant’s racial identity. Thus, when extending Rudman and Fairchild’s (2004) model to reactions to ethnic deviance on the part of minority men, public recovery strategies (hiding success and pursuing atypical talent) may be a function of confederate’s race and participant’s racial identity, whereas private consequences (domain identification and implicit self-esteem) may reflect the extant model. We discuss the implications of these results in greater detail in the following section.

Table 5  
*Implicit Self-Esteem as a Function of Backlash Condition and Recovery Strategies (Experiment 4)*

Backlash condition	Publicity index				Group difference	
	Low		High		<i>t</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Backlash	1.60	1.03	0.12	1.73	3.00**	0.88
Social support	0.94	1.58	0.83	1.80	0.20	0.11

Note. Small, medium, and large effect sizes correspond to 0.20, 0.50, and 0.80, respectively (Cohen, 1988). Positive effect sizes reflect greater self-esteem for participants who scored low on the publicity index (i.e., who engaged in recovery strategies by hiding their success and refusing to pursue atypical talent).

\*\* $p < .01$ .

## General Discussion

In five experiments we investigated the role of backlash in racial stereotype maintenance, with the goal of extending Rudman and Fairchild's (2004) model beyond gender. In concert, the findings support the model's claim that backlash functions as a mechanism that reinforces cultural stereotypes and social hierarchies by undermining counterstereotypical individuals' ability to weaken them. To our knowledge, the present research provides the first evidence that justification for backlash may differ for minorities and dominants, such that minorities are penalized for stereotype violation, whereas Whites are penalized for status-attenuating behaviors (whether atypical or not). In Experiment 1, atypical Asians were sabotaged for succeeding in cross-racial domains, and sabotage was predicted by perceivers' racial stereotypes. By contrast, Whites were sabotaged when they showed skill in low-status domains but not when their atypical skill was status enhancing. These results suggest that Whites are permitted to succeed in cross-racial domains provided their behavior upholds the racial status quo.

In Experiment 2, we investigated reactions to a White rapper after determining that pursuing a rap career was both atypical and status attenuating for Whites. Of importance, we found that backlash emerged even when revenge motives are absent, as is typically the case for backlash toward gender deviants (Rudman & Phelan, 2008). Although perceivers listened to the identical song, they penalized a White rapper by rating him as less likable, talented, and worthy of economic support, relative to a Black rapper. Moreover, likability fully mediated target race differences in perceived talent and economic support, suggesting that the White rapper was discriminated against because perceivers were prejudiced against him. Taken together, results from Experiments 1 and 2 support the process in the top row of Figure 1 by showing that individuals who step outside racial bounds are at risk for penalties that prevent them from becoming successful role models who challenge cultural beliefs. Although the motives for backlash toward minority and majority group members may differ, the stereotype-preserving consequences of backlash are identical.

Figure 1 posits that, beyond perceiver backlash, actors' fear of backlash is another means by which racial stereotypes are reinforced. In Experiment 3, Black participants who feared backlash for excelling on a leadership and reasoning aptitude task were unwilling to publicize their success on a national website. Moreover, fear of backlash was predicted by stereotypic expectancies and actual test performance, suggesting that Blacks who are either aware of racial stereotypes or academically talented feel more threatened by backlash for academic success. In Experiment 4, non-Black men who were mocked for acting Black were more likely to hide their atypical success and refuse to pursue their talent, compared with counterparts who received social support. In Experiment 5, we found comparable results for minority men (through country music singing), provided they were low on racial identity and received feedback from a White confederate. In both Experiments 4 and 5, actors who suffered jeer pressure for crossing racial lines showed low identification with their atypical talents, compared with those who received social support. Because identifying with atypical talents and taking pride in one's success are important steps toward defying stereotypes, these experiments

further underscore the role of backlash in maintaining racial stereotypes.

In concert, our results suggest that people who violate racial stereotypes and/or status expectancies are at risk for negative reactions, which presents an obstacle to personal aspirations and ambitions as well as the defeat of cultural stereotypes. But we also found evidence for the self-esteem effects outlined in Figure 1. First, in Experiment 1, perceivers who sabotaged typical Asians showed low self-esteem, whereas sabotaging deviant Asians did not inflict this cost. Moreover, only perceivers who sabotaged status-enhancing Whites showed low self-esteem; self-esteem was protected when status-attenuating Whites were sabotaged. These findings further supported our hypothesis that stereotype (status) violation justifies backlash toward Asians (Whites). Second, minority men in Experiment 5 showed decreased self-esteem if they suffered backlash for acting White and did not engage in recovery strategies. Thus, although backlash from perceivers and steps to avoid it from actors may harm individuals (by thwarting talents) and society (by preserving stereotypes), they may also serve a protective function for self-esteem.

Finally, in Experiment 5 we found that the path between fear of backlash and recovery strategies in Figure 1 was moderated by the public versus private nature of the behavior. Minority men low on racial identity who suffered backlash from White men reacted defensively when behaviors were public (i.e., by refusing to publicize and pursue their atypical talent), whereas high identifiers did not. By contrast, when responses were private (i.e., identification with country music and implicit self-esteem), results showed straightforward support for the model. Thus, our results suggest that the protective effects of racial identification (e.g., Phinney & Alipuria, 1990) may extend to some, but not all, of the damaging consequences of backlash. That is, identification with one's race may act as a buffer in terms of public behaviors, but private or less controlled beliefs may still be affected.

## Limitations and Future Directions

We have suggested that, when justified, minorities suffer backlash for stereotype violation, whereas Whites are penalized for failing to uphold the racial status quo. Thus, justification for backlash toward minorities may involve defending stereotypes, suggesting that (some) racial stereotypes, like gender stereotypes, can be prescriptive (Prentice & Carranza, 2002) as well as restrictive (Guinote et al., 2002; Pratto & Pitpitan, 2008). However, there may be other reasons why stereotype disconfirmation can jeopardize minorities. For example, because stereotype confirmation is self-validating for perceivers (Clark, Wegener, Brinol, & Petty, 2009), they may react negatively to stereotype violators because they invalidate their beliefs. Alternatively, or in addition, stereotype violators might cast perceivers into a state of uncertainty, and people respond to uncertainty by defending their worldviews (McGregor, Zanna, Holmes, & Spencer, 2001), which include racial stereotypes (Jost & Banaji, 1994; Pratto & Pitpitan, 2008). In contrast, Experiments 1 and 2 suggest that justification for backlash toward dominant group members may involve defending the status hierarchy rather than stereotypes per se. However, future research should fully cross status and stereotypicality to determine whether they interact. For example, Whites who excel in stereotypic, high-status tasks (e.g., knowledge of European culture) may

be spared backlash even more so than those who excel at counterstereotypic, high-status tasks (e.g., knowledge of Asian culture).

Rudman and Fairchild (2004) argued that the backlash paradigm is useful for determining whether stereotypic norms are merely descriptive or injunctive (Cialdini & Trost, 1998), and our results promisingly support that suggestion. Although the present research extends Figure 1 beyond gender, there are many other groups that could be tested, including other ethnic groups and groups based on religious, sexual, and political orientation. In addition, there may be other ways that actors threatened by backlash unwittingly perpetuate stereotypes. For example, people may avoid even trying to succeed in atypical domains if they are given a choice (Prentice & Miller, 2006). Self-selection biases likely play a prominent role in bolstering stereotypes, because people shy away from domains that appear to disadvantage their ingroup (e.g., adolescent girls and math; Eccles, 1984; Jacobs & Eccles, 1992). When jeer pressure is added to the mix, a powerful recipe for maintaining cultural stereotypes is obtained.

On that note, we are not claiming unique effects of social threat and social support for actors' reactions to their own deviance in the backlash paradigm. If people fear ostracism for any reason, they will behave defensively (Williams, 2007). The importance of threat in the process shown in the bottom row of Figure 1 is that defensive behaviors designed to avoid backlash are also likely to reinforce cultural stereotypes, contributing to their self-sustaining prophecy. Poignantly, Experiments 1 and 2 demonstrate that people who transgress racial bounds are likely to suffer backlash, and the results of Experiments 3, 4, and 5 indicate that this threat results in defensive strategies that protect the actor while maintaining the very beliefs that promote rejecting deviants.

Finally, further exploration of the moderating role of racial identification among minorities is needed. As noted, Experiment 5 found that racial identity (and confederate race) moderated public reactions to backlash (publicizing and pursuing success) but not private reactions (identification with a White domain and self-esteem). However, the public and private distinction is tentative and requires additional evidence. In addition, because ethnicity is a more salient self-component for racial minorities, compared with Whites (e.g., Deaux, 1992; Phinney & Alipuria, 1990), we did not examine racial identity as a moderator in Experiment 4. However, future research might do so in the event that we overlooked its significance.

## Conclusion

Disconfirming racial stereotypes is a powerful means of combating them in the culture-at-large, yet doing so risks backlash—a mechanism that limits equal opportunity for individuals and curbs their ability to stand out as successful, atypical role models. The present research suggests that perceivers play a role in perpetuating racial stereotypes by penalizing counterstereotypical actors in ways that can prevent them from challenging cultural beliefs. Investigating backlash from counterstereotypical actors' perspective suggests that they also play a role in preserving stereotypes when the threat of social rejection causes them to refrain from proudly exhibiting, developing, and identifying with atypical talents. Taken together, our findings suggest that backlash is an effective mechanism for maintaining racial stereotypes because people not only police others for crossing ethnic bounds but also

police themselves in order to avoid jeer pressure. As a result, stereotypes are allowed to persist, unchallenged, in the culture-at-large.

## References

- Arroyo, C. G., & Zigler, E. (1995). Racial identity, academic achievement, and the psychological well-being of economically disadvantaged adolescents. *Journal of Personality and Social Psychology, 69*, 903–914.
- Asch, S. E. (1955). Opinions and social pressure. *Scientific American, 193*, 31–35.
- Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1–36). Hillsdale, NJ: Erlbaum.
- 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.
- Biernat, M., Ma, J. E., & Nario-Redmond, M. R. (2008). Standards to suspect and diagnose stereotypical traits. *Social Cognition, 26*, 288–313.
- Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*, 631–643.
- Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology, 77*, 135–149.
- Branscombe, N. R., & Wann, D. L. (1994). Collective self-esteem consequences of outgroup derogation when a value social identity is on trial. *European Journal of Social Psychology, 24*, 641–657.
- Brewer, M. B. (1988). A dual-process model of impression formation. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.
- Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 151–192). New York, NY: McGraw-Hill.
- Clark, J. K., Wegener, D. T., Brinol, P., & Petty, R. E. (2009). Discovering that the shoe fits: The self-validating role of stereotypes. *Psychological Science, 20*, 846–852.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Contrada, R. J., Ashmore, R. D., Gary, M. L., Coups, E., Egeth, J. D., Sewell, A., . . . Chasse, V. (2001). Measures of ethnicity-related stress: Psychometric properties, ethnic group differences, and associations with well-being. *Journal of Applied Social Psychology, 31*, 1775–1820.
- Crocker, J., & Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality and Social Psychology, 58*, 60–67.
- Cross, W. E., Parham, T. A., & Helms, J. E. (1998). Nigrescence revisited: Theory and research. In R. L. Jones (Ed.), *African American identity development: Theory, research, and intervention* (pp. 3–71). Hampton VA: Cobb & Henry.
- Deaux, K. (1992). Personalizing identity and socializing self. In G. Breakwell (Ed.), *Social psychology of identity and the self concept* (pp. 9–33). San Diego, CA: Academic Press.
- DeHart, T., Pelham, B. W., & Tennen, H. (2006). What lies beneath: Parenting style and implicit self-esteem. *Journal of Experimental Social Psychology, 42*, 1–17.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573–598.
- Eccles, J. S. (1984). Sex differences in achievement patterns. In T. Sonderegger (Ed.), *Nebraska Symposium on Motivation* (Vol. 32, pp. 98–132). Lincoln: Nebraska University Press.

- Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology, 72*, 617–626.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York, NY: Academic Press.
- Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology, 26*, 21–42.
- Fryer, R. G., & Torelli, P. (2006). *An empirical analysis of "Acting White."* Unpublished manuscript, Harvard University, Cambridge, MA.
- Guinote, A. (2008). Power and affordances: When the situation has more power over powerful than powerless individuals. *Journal of Personality and Social Psychology, 95*, 237–252.
- Guinote, A., Judd, C. M., & Brauer, M. (2002). Effects of power on perceived and objective group variability: Evidence that more powerful groups are more variable. *Journal of Personality and Social Psychology, 82*, 708–721.
- Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*, 895–910.
- Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological Review, 65*, 117–127.
- Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology, 63*, 932–944.
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system justification and the production of false-consciousness. *British Journal of Social Psychology, 33*, 1–27.
- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25*, 881–919.
- Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology, 52*, 536–546.
- Katz, I. (1981). *Stigma: A social psychological analysis*. Hillsdale, NJ: Erlbaum.
- Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin, 23*, 736–742.
- Kobrynowicz, D., & Biernat, M. (1998). Considering correctness, contrast, and categorization in stereotyping phenomena. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 11, pp. 109–126). Mahwah, NJ: Erlbaum.
- Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80*, 669–685.
- Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 1–62). New York, NY: Academic Press.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin, 18*, 302–318.
- Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals, 2*, 351–398.
- McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology, 80*, 472–488.
- Munford, M. B. (1994). Relationship of gender, self-esteem, social class, and racial identity to depression in Blacks. *Journal of Black Psychology, 20*, 157–174.
- Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology, 15*, 353–361.
- Nuttin, J. M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. *European Journal of Social Psychology, 17*, 381–402.
- Parks-Stamm, E. J., Heilman, M. E., & Hearn, K. A. (2008). Motivated to penalize: Women's strategic rejection of successful women. *Personality and Social Psychology Bulletin, 34*, 237–247.
- Phinney, J. S., & Alipuria, L. L. (1990). Ethnic identity in college students from four ethnic groups. *Journal of Adolescence, 13*, 171–183.
- Pratto, F., & Pitpitan, E. V. (2008). Ethnocentrism and sexism: How stereotypes legitimize six types of power. *Social and Personality Psychology Compass, 2*, 2159–2176.
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*, 269–281.
- Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science, 17*, 129–135.
- Ridgeway, C. L. (2001). Gender, status, and leadership. *Journal of Social Issues, 57*, 627–655.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosenfield, D., Greenberg, J., Folger, R., & Borys, R. (1982). Effect of an encounter with a Black panhandler on subsequent helping for Blacks: Tokenism or confirming a negative stereotype? *Personality and Social Psychology Bulletin, 8*, 664–678.
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology, 74*, 629–645.
- Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology, 93*, 798–813.
- Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology, 87*, 157–176.
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology, 77*, 1004–1010.
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 732–762.
- Rudman, L. A., & Glick, P. (2008). *The social psychology of gender: How power and intimacy shape gender relations*. New York, NY: Guilford Press.
- Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for counterstereotypical behavior in organizations. In A. Brief & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 28, pp. 61–79). New York, NY: Elsevier.
- Rudman, L. A., Phelan, J. E., Moss-Racusin, C. A., & Nauts, S. (2009). *Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice toward female leaders*. Manuscript submitted for publication.
- Schimel, J., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., Waxmonsky, J., & Arndt, J. (1999). Stereotypes and terror management: Evidence that mortality salience enhances stereotypic thinking and preferences. *Journal of Personality and Social Psychology, 77*, 905–926.
- Sellers, R. M., Caldwell, C. H., Schmeelk-Cone, K. H., & Zimmerman, M. A. (2003). Racial identity, racial discrimination, perceived stress, and psychological distress among African American young adults. *Journal of Health and Social Behavior, 44*, 302–317.
- Snyder, M., & Miene, P. (1994). On the functions of stereotypes and

- prejudice. In M. P. Zanna & J. M. Olson (Eds.), *The Ontario Symposium: Vol. 7. The psychology of prejudice* (pp. 33–54). Hillsdale, NJ: Erlbaum.
- Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin*, *24*, 1139–1152.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–48). Monterey, CA: Brooks–Cole.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup relations* (7–24). Chicago, IL: Nelson–Hall.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 181–227). San Diego, CA: Academic Press.
- Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, *4*, 290–299.
- Waters, M. (1990). *Ethnic options: Choosing identities in America*. Berkeley: University of California Press.
- Williams, K. D. (2007). Ostracism. *Annual review of psychology* (Vol. 58, pp. 425–452). New York, NY: Annual Reviews.
- Wong, C. A., Eccles, J. S., & Sameroff, A. (2003). The influence of ethnic discrimination and ethnic identification on African American adolescents' school and socioemotional adjustment. *Journal of Personality*, *71*, 1197–1232.

Received December 6, 2008

Revision received November 2, 2009

Accepted November 3, 2009 ■