



Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders

Laurie A. Rudman ^{a,*}, Corinne A. Moss-Racusin ^b, Julie E. Phelan ^c, Sanne Nauts ^d

^a Rutgers University – New Brunswick, USA

^b Yale University, USA

^c Langer Research Associates, USA

^d Behavioural Science Institute, Radboud University Nijmegen, Netherlands

ARTICLE INFO

Article history:

Received 10 April 2011

Revised 30 August 2011

Available online 15 October 2011

Keywords:

Backlash effect

Sex discrimination

Gender stereotype

Gender prejudice

System justification theory

Impression management

ABSTRACT

Agentic female leaders risk social and economic penalties for behaving counter-stereotypically (i.e., *backlash*; Rudman, 1998), but what motivates prejudice against female leaders? The status incongruity hypothesis (SIH) proposes that agentic women are penalized for status violations because doing so defends the gender hierarchy. Consistent with this view, Study 1 found that women are proscribed from dominant, high status displays (which are reserved for leaders and men); Studies 2–3 revealed that prejudice against agentic female leaders was mediated by a dominance penalty; and in Study 3, participants' gender system-justifying beliefs moderated backlash effects. Study 4 found that backlash was exacerbated when perceivers were primed with a system threat. Study 5 showed that only female leaders who threatened the status quo suffered sabotage. In concert, support for the SIH suggests that backlash functions to preserve male dominance by reinforcing a double standard for power and control.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Professional women vying for leadership roles face a double bind. In order to be perceived as qualified, they must defeat gender stereotypes by presenting themselves as competent, confident, and assertive (i.e., *agentic*; Dodge, Gilroy, & Fenzel, 1995; Eagly & Karau, 2002; Glick, Zion, & Nelson, 1988). Thus, women's first hurdle to leadership is the "lack of fit" between feminine stereotypes and leadership qualities (Biernat & Fuegen, 2001; Foschi, 2000; Heilman, 1983). However, although women can clear this hurdle by demonstrating agency, they still face another roadblock. This second hurdle consists of backlash, whereby agentic women are perceived as highly capable, but risk prejudice and hiring discrimination for behaving counterstereotypically (Rudman, 1998). Specifically, backlash emerges when agentic women are judged as similarly competent, but less likable and hireable, compared with identically behaving men. This dilemma represents a Catch-22 for women, such that they must enact agency to be perceived as qualified for leadership, but risk penalties if they do so (e.g., Heilman & Okimoto, 2007; Heilman, Wallen, Fuchs, & Tamkins, 2004; Parks-Stamm, Heilman, & Hearn, 2008; Rudman, 1998; Rudman & Glick, 1999, 2001; for reviews, see Eagly & Karau, 2002; Rudman & Phelan, 2008). In essence, backlash forces women to choose between being

respected and being liked, which undermines their ability to obtain positions of status and power (e.g., Catalyst, 2010; Eagly & Karau, 2002; Rudman & Glick, 2008; Valian, 1999).

In their role congruity theory of prejudice toward female leaders (RCT), Eagly and Karau (2002) posit that the two hurdles for professional women stem from a perceived conflict between leadership roles and women's gender roles. However, RCT is limited in scope for three reasons. First, RCT does not account for evidence that agentic women suffer backlash even when they are not in leadership positions (e.g., Rudman, 1998; Rudman & Fairchild, 2004), or for the fact that atypical men are also at risk (e.g., Moss-Racusin, Phelan, & Rudman, 2010; Rudman & Fairchild, 2004). Second, Eagly and Karau broadly defined gender roles as "people's consensual beliefs about the attributes of women and men" (p. 574), without stipulating which aspects of gender roles are culpable in backlash. Third, RCT does not specify perceivers' motives for engaging in backlash. According to Rudman and Fairchild (2004), penalizing atypical targets is not engaged in arbitrarily; there must be justification for it. That is, people are viewed as deviant when they violate gender norms (Kobrynowicz & Biernat, 1998; Rudman & Fairchild, 2004), but backlash only results if perceivers feel justified.

The present research was designed to specify the primary motive for backlash, and which gender norm violations are culpable. First, we describe *gender rules*: gender stereotypes that stipulate how women and men should be (prescriptions), and how they should not be (proscriptions). Second, we consider the relationship between gender rules and status characteristics by presenting the results of a survey (Study 1). Third, we introduce the status incongruity hypothesis, which proposes

* Corresponding author at: Department of Psychology, Tillett Hall, Rutgers, the State University of New Jersey, 53 Avenue E, Piscataway, NJ 08854-8040, USA.

E-mail address: rudman@rci.rutgers.edu (L.A. Rudman).

that women are penalized specifically for status violations because defending the gender hierarchy is a key motive for backlash. Finally, we present four experiments that provide support for the status incongruity hypothesis.

Gender stereotypes as gender rules

Gender stereotypes are not merely descriptive but also prescriptive, consisting of rules concerning how men and women *should* behave (Eagly & Karau, 2002; Fiske, 1998; Fiske & Stevens, 1993). Agentic attributes (e.g., assertive, competitive, and independent) are prescriptive for men, whereas communal attributes (e.g., warm, kind, and supportive) are prescriptive for women – plausibly because agency is required for leadership and career success, whereas communality is required for caring for the welfare of others (Burgess & Borgida, 1999; Eagly, 1987; Prentice & Carranza, 2002; Williams & Best, 1990).

Gender rules also consist of *proscribed* characteristics, rules concerning how men and women *should not* behave. Proscriptions are relatively negative qualities that are prohibited for only one gender. For example, dominant masculine traits (e.g., controlling and arrogant) are proscribed for women but tolerated for men, and weak feminine traits (e.g., weak and naive) are proscribed for men but tolerated for women (Prentice & Carranza, 2002). From these examples, it seems likely that proscriptions serve to reinforce the gender hierarchy (e.g., that men ought not to be weak because it is low in status and women ought not to be dominant because it is high in status). Therefore, a consideration of how gender rules align with status characteristics should help to explain why agentic female leaders are perceived as highly competent but socially unattractive (i.e., why they are respected but not liked; e.g., Heilman et al., 2004; Rudman & Phelan, 2008). Given that women encounter backlash when they have merely demonstrated agency (and not negative traits), what is it about competent, ambitious women that puts them at risk for social rejection? Because prejudice against agentic women fully accounts for why they are less likely to be hired as leaders than agentic men (Rudman & Glick, 1999; 2001), it is imperative to answer this question to illuminate barriers to gender equality.

Gender stereotypes and status

The first step was to examine the extent to which gender rules (prescriptions and proscriptions) are aligned with status characteristics. To do so, Study 1 consisted of a large, online survey demonstrating that male agency prescriptions were aligned with high status, and male weakness proscriptions were aligned with low status. Thus, what men should be is high in status and what they should not be is low in status, and both rules reinforce the gender hierarchy. Of more importance, female dominance proscriptions were linked to high status, whereas female communality proscriptions consisted of traits that were, on average, status neutral. Thus, what women *should not be* is high in status, whereas what they should be does not necessarily reinforce the gender hierarchy. Although the finding that women's prescriptions are status neutral might seem surprising, past research has shown that agency is linked to high status more so than communality is linked to low status (Conway, Pizzamiglio, & Mount, 1996). Moreover, agency and communality tend to be positively associated in perceptions of individuals (e.g., Henik & Tzelgov, 1985; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005; Rosenberg, Nelson, & Vivekananthan, 1968), and in perceptions of ingroup members (Fiske, Cuddy, Glick, & Xu, 2002), suggesting that they are not bipolar qualities, or likely to be oppositional vis-à-vis status.

The results of Study 1 are described in detail below, but here we emphasize that they confirmed that agency is aligned with high status, but that communality is not necessarily low in status (Conway et al., 1996; Hoffman & Hurst, 1990). Instead, women are proscribed against high status, dominance displays (e.g., dominant, controlling,

and arrogant) – traits reserved for leaders and men. These findings provide a key to the puzzle of which stereotypes are likely to be culpable in backlash against agentic women, and why. That is, they begin to point toward status incongruities as a critical component of backlash.

The status incongruity hypothesis

Researchers distinguish between status based on achievement (earning your way to the top) and ascribed status based on personal characteristics (e.g., sex, race, and age; Berger, Webster, Ridgeway, & Rosenholtz, 1986; Ridgeway, 2001; Ridgeway & Bourg, 2004). Because of their gender, women are automatically linked to low status (Rudman & Kilianski, 2000) – so much so that when they enter an occupation in large numbers, its prestige can drop significantly (Nieva & Gutek, 1981; Touhey, 1974). The status incongruity hypothesis (SIH) proposes that women who possess or pursue power are de facto status incongruent, but particularly when their behavior violates status expectations. Given that agency is high in status, female agency is discrepant with women's low ascribed status, and this status incongruity elicits backlash. By exhibiting masculine competencies, agentic women undermine the presumed differences between the genders, and discredit the system in which men have more access to power and resources for ostensibly legitimate reasons. That is, agentic women should incur penalties because they threaten the gender hierarchy. As a result, women's perceived status violations (as opposed to any type of gender role violation) should account for backlash effects.

The dominance penalty

According to Study 1, female dominance proscriptions are the gender rules for women that are most strongly aligned with status and thus, reinforce the gender hierarchy. Therefore, agentic female leaders should be viewed as extreme on dominant traits (e.g., dominant, controlling, and arrogant), compared with agentic male leaders. Termed the *dominance penalty*, we expected these ratings to fully account for why agentic women are liked less than agentic men (i.e., prejudice). If so, then defending the gender status quo will be revealed as a key motive for backlash. Interestingly, Eagly, Makhijani, and Klonsky (1992) invoked a dominance penalty in their meta-analysis investigating evaluations of male and female leaders (p. 5):

In addition to a tendency for women in leadership roles to be evaluated negatively, their behavior may be regarded as more extreme than that of their male counterparts – that is, as more dominant and controlling, and, in general, as embodying a higher level of prototypical leadership qualities. This perception of female leaders as more extreme than their male counterparts would be likely to occur to the extent that female leadership behaviors are quite discrepant from people's stereotypes about women and are therefore perceptually contrasted from these stereotypes (see Manis, Nelson, & Shedler, 1988).

This analysis foreshadows a key prediction of the SIH, with two important caveats. First, leadership alone is not sufficient to provoke backlash because female leaders who lower their status (e.g., by adopting inclusive styles) are evaluated favorably (Eagly et al., 1992; Rudman & Glick, 2001). By contrast, female agency in the absence of leadership can elicit backlash (for a review, see Rudman, 1998), suggesting that status incongruity drives backlash more so than occupational role incongruity. Second, the SIH proposes a motivational, rather than a cognitive, explanation for the dominance penalty (protecting the gender hierarchy), whereas Eagly et al. (1992) proposed that perceptual contrast effects would yield extremely high ratings on prototypical leadership qualities such as dominance and agency. Specifically, behaviors that are “sufficiently discrepant with the stereotype should be perceptually contrasted

[away from the stereotype]" (Manis et al., 1988, p. 28). Indeed, if agentic female leaders are viewed as highly discrepant women, contrast effects should also yield extremely low ratings on communality and weakness (male proscriptions). However, the SIH predicts that ratings of agency, communality, and weakness will not account for prejudice against female leaders, thereby ruling out a cognitive explanation for backlash. Instead, it predicts that only the dominance penalty will account for (i.e., mediate) prejudice and thus, is key to understanding why leadership roles are more challenging for women than men.

System justification motives for backlash

But how does the dominance penalty reveal the motive for backlash? According to the SIH, status incongruent women jeopardize extant norms for power, forcing a reconsideration of the gender status quo (Ridgeway, 2001; Ridgeway & Bourg, 2004). Because motives to legitimize existing social structures are pervasive and often nonconscious (Jost & Banaji, 1994; Jost, Banaji, & Nosek, 2004; Rudman, Feinberg, & Fairchild, 2002), people may instinctively resist powerful women (Rudman & Kilianski, 2000). Thus, motives to defend the gender hierarchy may result in perceptions that agentic female leaders are "too powerful," as reflected by the dominance penalty, which is then used to justify prejudice. Consistent with this view, epithets for powerful women often cast them as destroyers of male virility (e.g., "ball-breaker" and "castrating bitch"), signaling the extent to which women are expected to yield to men — economically, politically, and sexually (Kanter, 1977). Women's ascribed status (as women) is incongruent with their achieved status (as leaders), as well as with the requisite agency demanded of leaders. By contrast, agentic male leaders are not at risk for status violations because their ascribed status (as men), achieved status (as leaders), and the requisite agency demanded of leaders are congruent. In essence, agentic male leaders support the gender hierarchy, whereas agentic female leaders subvert it. Judging only the latter as "too powerful" (i.e., dominant and controlling) is expected to signify a system-justifying motive for rejecting women who disrupt the status quo. Moreover, pinpointing the dominance penalty as the sole mediator of prejudice (ruling out extreme ratings on agency, communality, and weakness) will support our claim that it is used to justify backlash and thereby, to defend male hegemony. The present research sought to test this key motive for backlash in multiple ways, as outlined below.

Overview of the research and hypotheses

Study 1 examined the overlap between gender stereotypes and status to lay the foundation for testing predictions derived from the SIH. Studies 2–4 used a hiring paradigm, in which agentic applicants vied for leadership roles that varied with respect to gender typicality (English professor, pretested as gender neutral, and marketing manager, pretested as male dominated; see Footnotes 1 and 3). We expected agentic male and female applicants to be viewed as similarly qualified for the job, because strong individuating information defeats women's first hurdle (Eagly & Karau, 2002). Nonetheless, agentic women should encounter prejudice, which should fully account for hiring discrimination (Rudman & Glick, 1999, 2001), reflecting female leaders' second hurdle (i.e., backlash; Eagly & Karau, 2002).

To test the SIH, applicants in Studies 2–4 were rated on the four sets of gender rules (i.e., male and female prescriptions and proscriptions). We expected only female dominance proscriptions would (1) distinguish between agentic men and women, and (2) account for prejudice against agentic women. These results would show, remarkably, that high-powered traits that are viewed as more typical of men (indeed, are reserved for them) can in fact be attributed *more to women* when women are status incongruent (i.e., agentic leaders). As the only set of rules consisting of high status traits prohibited for women, female

dominance proscriptions should be used to justify backlash and thus, to defend the gender hierarchy.

Studies 3–5 sought additional evidence for the system-justifying motive underlying backlash. In Study 3, we expected that perceivers who endorsed the gender hierarchy as just and fair (i.e., gender system-justifiers; Jost & Kay, 2005) would be most likely to administer the dominance penalty. By contrast, gender system-justifiers should not penalize women as insufficiently communal because this rule is not aligned with status. In Study 4, we manipulated system threat to provide a causal test of the SIH's claim that backlash is used to protect the gender hierarchy (Kay et al., 2009). In Study 5, we "randomly assigned" male and female confederates to a leadership position while manipulating their agency. Because agency is necessary for leadership success, we expected participants (assigned to be subordinates) to sabotage only female leaders who were high (not low) on agency. This is because only women who effectively challenge the gender hierarchy (i.e., who exhibit masculine competencies) should be at risk for backlash. In concert, these results would bolster the SIH's claim that defending the gender status quo is a key motive for backlash.

Study 1

Study 1 examined the extent to which prescriptive and proscriptive gender stereotypes are aligned with status. To do so, we compiled 64 traits and administered one of six surveys to participants in order to determine (a) how stereotypical the traits were for men or women, (b) the extent to which each trait fell on a prescriptive-proscriptive continuum (i.e., reflected gender rules) for men or women, and (c) the extent to which each trait was characteristic of high status or low status people. By isolating each type of judgment, we sought to reduce context effects and demand in order to obtain relatively pure estimates of the degree to which gender rules overlap with status.

Participants

Volunteers ($N = 832$; 406 men, 415 women, and 11 who did not indicate their gender) participated in exchange for partial fulfillment of an Introductory Psychology course requirement. Of these, 366 (44%) were White, 259 (31%) were Asian, 77 (9%) were Latino, 63 (8%) were Black, and 67 (8%) reported another ethnicity.

Materials and procedure

We began with the traits shown by Prentice and Carranza (2002) to reflect gender rules. After adding traits often used in backlash research (e.g., "performs well under pressure"; "is sensitive to others' needs") and eliminating synonyms, we obtained 64 traits used to create six separate surveys. The two *gender rules* surveys asked participants to "indicate how desirable it is in American society for a woman [man] to possess each of the following characteristics" on a scale from 1 (*not at all desirable*) to 9 (*very desirable*). The two *status typicality* surveys asked participants to "indicate how common or typical you think each of the following characteristics is in someone who has high [low] status on a scale from 1 (*not at all typical*) to 9 (*very typical*)". Finally, the two *gender typicality* surveys asked participants to "indicate how common or typical you think each of the following characteristics is in women [men] in American society" on a scale from 1 (*not at all typical*) to 9 (*very typical*).

All measures were administered online. After reading the consent form, participants were randomly assigned to complete one of the six surveys. After completing the survey, participants were debriefed and credited.

Results and discussion

Prescriptive stereotypes

We defined male prescriptions as traits that were rated above 6 on the 1 (*not at all desirable*) to 9 (*very desirable*) scale for men, and that when compared to the desirability rating for women, also had a gender difference effect size greater or equal to $d = .40$. The top half of Table 1 shows these 16 traits, ranked by prescriptive d scores. As can be seen, men's prescriptions strongly reflect agency ($M d = .75$, range = .40 to 1.12). We defined female prescriptions as traits that were rated above 6 for women, and that also had a gender difference effect size less than or equal to $d = -.40$. These 16 traits are shown in the bottom half of Table 1, ranked by prescriptive d scores. As can be seen, these traits primarily reflect communality ($M d = -.76$, range: -1.12 to $-.42$).

Table 1's fourth column shows the effect size for typicality (positive d s indicate traits rated as more stereotypical for men; negative d s indicate traits rated as more stereotypical for women). As can be seen, men's agency prescriptions were generally rated as stereotypically male (M typicality $d = .53$), and women's communality prescriptions were all rated as stereotypically female (M typicality $d = -.82$).

The last column of Table 1 provides effect sizes for the difference between high status and low status ratings. Positive d scores reflect traits linked to high more than low status people; negative d scores reflect the reverse. Table 1 reveals that all of men's prescriptions were high status traits (M status $d = 1.26$, range = .43 to 1.60). By contrast, women's prescriptions reflected a mix (M status $d = .08$, range = $-.80$ to 1.20). Four of their prescriptions were low status traits (emotional, warm,

interested in children, and humble). However, four traits were rated high in status (cheerful, enthusiastic, excitable, and attends to appearance). The remaining eight traits were neutral in status (sensitive to others, good listener, cooperative, friendly, supportive, polite, helpful, and likeable). These results suggest that men should enact high status, agentic traits whereas women should enact communal traits that, on average, are status neutral.

Proscriptive stereotypes

We defined male proscriptives as traits that were rated below 4 on desirability for men, and that also had a gender difference effect size greater than or equal to $d = .40$. The top half of Table 2 shows these 10 traits, ranked by proscriptive d scores. As can be seen, men's proscriptives reflect vulnerable traits associated with weakness ($M d = .87$, range = .56 to 1.12). These attributes are perceived not only as feminine (M typicality $d = -1.05$), but also as low in status (M status $d = -.76$, range = -1.32 to .05); only two were status neutral (melodramatic and moody), and none were high in status.

Similarly, we defined female proscriptives as traits that were rated below 4 for women, and that also had a gender difference effect size less than or equal to $d = -.40$. The bottom half of Table 2 shows these 13 traits, ranked by proscriptive d scores. As can be seen, these traits primarily consist of dominance-related attributes that are prohibited for women ($M d = -.68$, range = -1.03 to $-.41$), and perceived as more typical of men (M typicality $d = .53$), with the exception of demanding and self-centered, which were gender neutral. Female dominance proscriptives were primarily high in status ($M d = .73$, range = $-.47$ to 1.42). Only two were low status traits (rebellious and angry), and only one was status neutral (cynical). These findings suggest that women cannot enact status-enhancing traits that are reserved for men and leaders (e.g., dominant, controlling, and arrogant), whereas men cannot enact status-attenuating traits that are acceptable for women (e.g., emotional, naive, and weak).

In summary, Study 1 revealed correspondence between dominance proscriptives (what women should not be) and high status

Table 1
Prescriptive traits for men and women (Study 1).

Trait	Prescriptive d	Male M	Female M	Typicality d	Status d
<i>Men's prescriptions</i>					
Career oriented	1.12	7.74	5.74	.49	1.57
Leadership ability	1.09	7.86	5.89	.79	1.45
Aggressive	1.03	6.16	3.91	1.36	.43
Assertive	1.01	7.26	5.20	.78	1.39
Independent	.98	7.67	5.57	.65	1.23
Business sense	.97	7.39	5.76	.86	1.60
Ambitious	.95	7.95	6.28	.35	1.37
Hard working	.80	8.08	6.78	.13	.91
Works well under pressure	.74	7.39	6.05	.65	1.26
Self-starter	.58	6.26	5.13	.53	1.18
Intelligent	.55	7.67	6.78	-.04	1.08
Analytical	.48	6.71	5.83	.02	1.29
High self-esteem	.48	7.29	6.56	.74	1.59
Persuasive	.41	6.44	5.74	-.04	1.30
Competitive	.40	7.67	5.28	1.08	1.43
Competent	.40	7.17	6.48	.14	1.12
<i>Women's prescriptions</i>					
Emotional	-1.12	3.87	6.51	-1.49	-.63
Warm	-1.03	6.07	7.99	-1.11	-.47
Interested in children	-1.00	5.92	7.82	-1.29	-.46
Sensitive to others	-1.00	5.48	7.52	-1.02	-.23
Good listener	-.89	6.14	7.82	-1.18	.07
Cheerful	-.87	6.06	7.58	-.83	.57
Enthusiastic	-.83	5.71	7.06	-.87	.91
Excitable	-.76	5.69	6.95	-.91	.42
Cooperative	-.75	6.29	7.50	-.45	.11
Friendly	-.71	6.77	7.76	-.70	.09
Supportive	-.69	6.69	7.75	-.73	.06
Polite	-.57	6.90	7.74	-.62	.04
Humble	-.56	5.77	6.80	-.43	-.80
Attends to appearance	-.46	6.68	7.51	-.34	1.20
Helpful	-.45	6.95	7.58	-.67	.25
Likeable	-.42	7.00	7.74	-.47	.15

Note. Positive d scores reflect stronger prescriptions or typicality for men than women, or stronger typicality for high than low status people. Negative d scores reflect the reverse. Conventional small, medium, and large effect sizes for d are .20, .50, and .80, respectively (Cohen, 1988).

Table 2
Proscriptive traits for men and women (Study 1).

Trait	Proscriptive d	Male M	Female M	Typicality d	Status d
<i>Men's proscriptives</i>					
Emotional	1.12	3.87	6.51	-1.49	-.63
Naive	1.03	2.35	4.55	-.88	-.78
Weak	.97	1.85	3.96	-1.02	-1.32
Insecure	.91	2.29	4.08	-1.06	-.96
Gullible	.89	2.80	4.67	-1.02	-1.07
Melodramatic	.88	2.87	4.80	-1.22	-.01
Uncertain	.80	2.78	4.13	-1.22	-1.22
Moody	.78	2.67	4.44	-.71	.05
Indecisive	.74	2.81	4.31	-1.18	-1.06
Superstitious	.56	3.16	4.15	-.74	-.64
<i>Women's proscriptives</i>					
Aggressive	-1.03	6.16	3.91	.43	1.36
Intimidating	-.98	5.37	3.29	.89	1.21
Dominating	-.94	5.74	3.54	.97	1.42
Arrogant	-.76	4.61	2.93	1.11	1.08
Rebellious	-.69	5.02	3.64	.66	-.40
Demanding	-.65	5.16	2.96	-.15	1.24
Ruthless	-.65	4.41	2.96	.64	.59
Angry	-.65	4.14	3.82	.71	-.47
Controlling	-.61	5.20	3.88	.42	1.33
Stubborn	-.55	4.74	3.63	.42	.65
Cold toward others	-.51	3.43	2.49	.38	.35
Self-centered	-.41	4.20	3.21	.14	1.05
Cynical	-.41	3.94	3.27	.28	.12

Note. Positive d scores reflect stronger proscriptives or typicality for men than women, or stronger typicality for high than low status people. Negative d scores reflect the reverse. Conventional small, medium, and large effect sizes for d are .20, .50, and .80, respectively (Cohen, 1988).

traits, whereas communality prescriptions do not necessarily reinforce the status quo. These findings lay the groundwork for the SIH's prediction that the dominance penalty (i.e., extreme ratings on status-enhancing traits that are typical of leaders and men, but prohibited for women) will be used to justify prejudice against agentic female leaders. The goal of Study 2 was to test this prediction.

Study 2

Participants read recommendation letters for candidates eligible for promotion to English professor at Yale University. The job was pretested to be gender neutral.¹ Candidates were lauded as internationally recognized authors and literary critics who were highly intelligent (e.g., winners of the MacArthur Genius Award). However, the author further described the candidate's style as a literary critic as "somewhat controversial" because it was either agentic (brutally honest) or communal (overly polite). In both cases, the author defended the candidate's style, claiming that agentic targets were highly critical "in order to maintain the high standards of the field," whereas communal targets were excessively diplomatic "in order to protect authors' fragile egos." Consistent with backlash effects, the agentic woman should be rated as similarly competent, but lower on liking and hiring dimensions than the agentic man, and prejudice should mediate hiring discrimination (Rudman & Glick, 1999, 2001). To support the SIH, we expected the agentic woman to be rated higher on female dominance proscriptions than a comparable man, and that the dominance penalty would mediate target gender differences in liking (i.e., prejudice against female agency). To rule out perceptual contrast effects, no other gender rules should significantly differ for agentic women and men.

Method

Participants

Volunteers ($N = 178$; 114 women) participated in exchange for credit toward their Introductory Psychology research participation requirement. Of these, 103 (58%) were White, 45 (25%) were Asian, 12 (7%) were Hispanic, 8 (4%) were Black, and 10 (6%) reported another ethnic identity.

Materials

Candidate style manipulation. To manipulate the candidates' critiquing style, four recommendation letters were created, for either Edward or Emily Mullen. The letters were ostensibly written by the (male) Chair of the English Department at Cornell University who strongly recommended the candidate be promoted to full professor at Yale. In all four letters, the candidate was lauded as supremely competent, with a prestigious academic background (e.g., a PhD from Harvard) and numerous accomplishments (e.g., over 40 publications, five well received novels, and recipient of the MacArthur Genius Award). The candidate was also described as a world-renowned literary critic. Appendix A contains the only section of the letter that differed by condition. In the communal condition, the candidate was described as a tactful reviewer who protected writers' egos, no matter how poor the work. In the agentic condition, the candidate was described as brutally honest when the work fell short. In each case, the author stated that the applicant's reviewing style was considered controversial by some, but that in his opinion it was necessary (either to maintain high standards or to protect fragile egos). Each letter concluded that the candidate was a genius who was highly deserving of the promotion, and that "if Yale is looking to promote the best literary minds of the era, then you need look no further."

¹ Students at Radboud University Nijmegen responded to two items, "When considering English professors at Yale University, what percentage of them are likely to be women [men]?" On average, percentages were 51% and 49%, respectively. The difference between these was not reliably different from zero, $t(42) < 1.00$, *ns*.

Candidate ratings. Participants responded to all dependent measures on scales ranging from 1 (*not at all*) to 6 (*very much*). Two items ("Did the applicant strike you as competent?" and "How likely is it that the applicant has the necessary skills for this job?") were averaged to form the competence index, $r(274) = .31$, $p < .001$. Three items were averaged to form the liking index: "How much did you like the applicant?"; "Is this person someone you want to get to know better?"; and "Would the applicant be popular with colleagues?" ($\alpha = .80$). Three items were averaged to form the hireability index: "Would you choose to interview the candidate?"; "Would you personally promote the candidate?"; and "How likely is it that the candidate will be promoted?" ($\alpha = .89$).

Gender rules. Four gender stereotype indexes were derived from Study 1. Male agency prescriptions were: *career-oriented, leadership ability, assertive, independent, business sense, ambitious, self-starter, intelligent, high self-esteem, and competitive* ($\alpha = .77$; M status $d = 1.38$). Female communality prescriptions were: *warm, sensitive to the needs of others, interested in children, cheerful, enthusiastic, cooperative, friendly, polite, and humble* ($\alpha = .90$; M status $d = .03$). Male weakness proscriptions were: *weak, emotional, naive, gullible, uncertain, and indecisive* ($\alpha = .76$; M status $d = -1.01$). Female dominance proscriptions were: *dominating, intimidating, arrogant, ruthless, controlling, cold toward others, and cynical* ($\alpha = .90$; M status $d = .87$).

Procedure

Participants were told that we were interested in their responses to a candidate being evaluated for promotion. After signing a consent form, participants were escorted to separate cubicles and randomly assigned to one of the four letters, which were written on Cornell stationery and addressed to the Dean of Faculty of Arts and Sciences at Yale. After reading the letter, the experimenter started a program to administer the measures in the order described above. The program randomly presented items within each measure. Finally, participants completed demographic items (age, gender, and race). They were subsequently debriefed and accredited.

Results and discussion

Candidate ratings

Table 3 shows the candidates' ratings as a function of target gender and critiquing style. Results of 2 (target gender) \times 2 (style: agentic, communal) \times 2 (participant gender) ANOVAs on each measure revealed no significant participant gender effects (consistent with previous research; Rudman & Phelan, 2008), all $F_s(1, 170) < 1.00$, *ns*. Because there were no significant effects for competence, all $F_s(1, 170) < 2.46$, $ps > .11$, the four letters were similarly successful in conveying the candidates' intelligence.

The liking and hireability indexes each showed main effects for target gender and style, qualified by Target Gender \times Style interactions; both $F_s(1, 170) > 3.92$, $ps < .05$. Table 3 shows no significant target gender differences for communal targets. By contrast, the agentic woman was judged as less likable, $t(82) = 4.47$, $p < .0001$, $d = .88$, and less hireable, $t(82) = 3.41$, $p < .01$, $d = .72$, compared with the agentic man.

Gender rules

Planned comparisons were conducted for the stereotype indexes.² Table 3 (bottom half) shows no target gender differences in the communal condition for any of the stereotyping indexes, all $t_s(92) < 1.00$,

² ANOVAs revealed large main effects for target style on dominance and communalism, both $F_s(1, 170) > 100.00$, such that agentic targets were rated as more dominant, but less communal, compared with communal targets. They were also rated as more agentic than communal targets, $F(1, 170) = 5.30$, $p < .05$. For the weak index, there was a main effect of participant gender, $F(1, 170) = 4.41$, $p < .05$, such that men perceived targets to be weaker than did women. Because these effects were not of theoretical interest, we report details only for the planned comparisons.

Table 3
Evaluations of candidates for full promotion to English Professor (Study 2).

	Communal targets				Agentic targets					
	Male		Female		d	Male		Female		d
	M	SD	M	SD		M	SD	M	SD	
Competence	5.35	.78	5.26	.74	.12	5.34	.64	5.17	.84	.23
Liking	4.18	1.45	4.16	1.06	.01	3.78 _a	1.34	2.53 _b	1.16	.88
Hireability	5.03	1.11	4.91	.96	.11	4.90 _a	.89	4.11 _b	1.22	.72
<i>Stereotypes</i>										
Agentic	5.74	.66	5.75	.79	-.01	5.91	.63	6.03	.58	-.18
Communal	4.62	.97	4.54	1.13	.07	3.52	.92	3.62	.70	-.09
Weak	2.27	.60	2.32	.74	-.06	2.13	.77	2.09	.90	.05
Dominant	3.15	1.00	3.34	1.16	-.13	4.94 _a	1.24	5.48 _b	.88	-.37

Note. Means with different subscripts differ by target gender within each condition ($p < .05$). Effect sizes (Cohen's d) represent target gender sex differences. Positive effect sizes favor male candidates; negative effect sizes favor female candidates. Conventional small, medium, and large effect sizes for d are .20, .50, and .80, respectively (Cohen, 1988).

ns. As predicted, the agentic woman was rated as higher on female dominance proscriptions than the agentic man, $t(82) = 2.31, p < .05, d = -.37$. No other target gender differences emerged in the agentic condition, all $t(87) < 1.00, ns$. These results support the SIH's prediction that agentic women suffer a dominance penalty, not a communality deficit, and further, they are not rated as extremely high on agency or low on weakness, ruling out contrast effects.

Accounting for backlash against female leaders

Prior backlash research has found that hiring discrimination was fully mediated by greater liking for agentic men than women (i.e., prejudice; Rudman & Glick, 1999, 2001). To test this prediction, we used the PRODCLIN program to compute confidence intervals based on an asymmetrical distribution of the mediated (indirect) effect (MacKinnon, Fritz, Williams, & Lockwood, 2007).³ Table 4's Model 1 shows the predicted results. The effect of target gender on hireability was reduced to nonsignificance after accounting for liking, and the 95% confidence intervals for the mediated effect did not include zero, resulting in full and reliable mediation.

New to the present research, Table 4's Model 2 tests the dominance penalty as a mediator of target liking differences. The effect of target gender on liking remained reliable, but was significantly reduced after controlling for the dominance index (i.e., the 95% confidence intervals for the mediated effect did not include zero). These results support partial but significant mediation for Model 2. Finally, Model 3 shows that communality ratings did not have a similar mediation effect; the effect of target gender was not reduced (in fact, it slightly increased), and the 95% confidence intervals for the mediated effect included zero.

In sum, Study 2 supports the SIH's claim that backlash depends on status violations (i.e., a conflict between women's ascribed and achieved status that is exacerbated by high status displays). Having an outstanding list of achievements did not protect a woman from backlash or the dominance penalty when she displayed high status behaviors reserved for men, even when they are necessary for effective leadership (i.e., frank assessment to maintain high standards). However, it did protect a communal woman from both hurdles, suggesting that women can be exceptionally intelligent and accomplished without risking backlash, provided they attenuate their status with extreme diplomacy. A contrast effect was not responsible

³ This approach avoids relying on Sobel's Z , which is an approximate test for the significance of the $a \times b$ (mediated) effect (Baron & Kenny, 1986) that assumes the sampling distribution of $a \times b$ in the population is normal (Shrout & Bolger, 2002). It is also superior to bootstrapping, which can result in inflated Type I error rates (up to 10%; MacKinnon et al., 2007), whereas computing asymmetrical confidence intervals offers the best balance of low Type I errors with increased power (see also MacKinnon, Lockwood, & Williams, 2004).

Table 4
Mediation analyses for agentic condition (Study 2).

Path/effect	B	SE	95% Confidence intervals
<i>Model 1</i>			
c (target gender \rightarrow hire)	-.36**	.11	
a (target gender \rightarrow like)	-.44***	.10	
b (like \rightarrow hire)	.33**	.12	
c'	-.21	.11	
$a \times b$ (mediation effect)	-.15*	.06	-.268, -.040
<i>Model 2</i>			
c (target gender \rightarrow like)	-.44***	.10	
a (target gender \rightarrow dominant)	.19*	.08	
b (dominant \rightarrow like)	-.38***	.06	
c'	-.37***	.10	
$a \times b$ (mediation effect)	-.07*	.04	-.161, -.010
<i>Model 3</i>			
c (target gender \rightarrow like)	-.44***	.10	
a (target gender \rightarrow communal)	.05	.08	
b (communal \rightarrow like)	.55***	.11	
c'	-.46***	.09	
$a \times b$ (mediation effect)	.03	.05	-.058, .120

Note. Target gender was coded 0 (male) 1 (female). Estimates are unstandardized. Confidence intervals for $a \times b$ are based on an asymmetrical distribution. Intervals that do not include zero support rejecting the null hypothesis that $a \times b = 0$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

for our findings because communality, agency, and weakness ratings were similar for women and men. Instead, the dominance penalty's emergence as a key factor in backlash supports the SIH's claim that agentic women are "demonized" to subordinate them, thus protecting the gender hierarchy. Study 3 further tested the SIH in a job interview context more comparable to past research, using live confederates. In addition, we investigated gender system-justification beliefs as a moderator of the dominance penalty.

Study 3

In Study 3, confederates responded to scripted questions while being interviewed for a marketing manager job pre-tested to be male-dominated.⁴ Participants conducted phone interviews with agentic (i.e., self-promoting and ambitious) targets trained to respond similarly using scripted answers. We expected to replicate Study 2's backlash pattern, whereby targets were viewed as similarly competent, but agentic women encountered prejudice (mediated by the dominance penalty) and hiring discrimination (mediated by prejudice).

The SIH posits that a key motive for backlash is to defend the gender status quo, which causes people to exaggerate the dominance of agentic women, rather than deny them communality (because only the dominance rule is aligned with high status). If that argument is correct, then people who endorse the gender status quo (i.e., who are gender system-justifiers) should be particularly likely to administer the dominance penalty, and to show prejudice and hiring discrimination. By contrast, gender system-justification beliefs should not moderate communality ratings because women's communal prescriptions are status neutral.

Method

Participants

Volunteers ($N = 74$; 38 women) participated in exchange for credit toward their Introductory Psychology research participation

⁴ Students at Radboud University Nijmegen responded to two items, "When considering marketing managers, what percentage of them are likely to be women [men]?" On average, percentages were 37% and 63%, respectively. The difference between these was reliably different from zero, $t(40) = 8.61, p < .001$.

requirement. Of these, 29 (39%) were White, 20 (27%) were Asian, 8 (11%) were Hispanic, 5 (7%) were Black, and 12 (16%) reported another ethnic identity.

Materials

Job description and confederates. The marketing manager position was described in a company advertisement as follows: “You are responsible for the formulation and execution of a marketing strategy, together with your team of eight experienced marketers. You coordinate market analyses aimed at identifying consumer needs, and introduce new products and services to strengthen our position in the market.” The qualifications included, “You have a masters degree in marketing; you like to take the initiative; you have excellent analytic and communication skills; and you can manage and inspire a team.” Confederates (two men, three women) were solicited through advertisements and paid \$10.00 an hour. Because the interviews were conducted by phone, confederates did not have to memorize their scripts. However, they were trained to present themselves similarly.

Applicant script. The confederates' script contained agentic responses to nine questions, during which applicants stressed their prior managerial success and their ability to initiate projects, lead a team, and work well under pressure. Appendix B provides more detail.

Applicant ratings. We used Study 2's liking index ($\alpha = .81$). Because the job was described as requiring both agentic and communal qualities, we designed the competence index to reflect these demands (see also Rudman & Glick, 2001). Participants responded to four items on a scale ranging from 1 (*not at all*) to 6 (*very much*) that asked whether the applicant had strong analytic skills, was a self-starter, a good listener, and was likely to get ahead in their career. Two items asked for estimates of the percentage of marketing problems the applicant would be able to solve independently, and the percentage of subordinates who would feel comfortable seeking help from the applicant (also scaled on 6-point scales). These six items were averaged to form the competence index ($\alpha = .72$). Using the same 6-point scale, hireability was assessed with three items in which participants indicated their likelihood of choosing to interview the candidate for the actual job, that they would personally hire the applicant, and that the applicant would be hired ($\alpha = .89$).

Gender rules. We adopted Study 2's stereotypic trait indexes. Reliability estimates ranged from $\alpha = .70$ (male agency prescriptions) to $\alpha = .82$ (female dominance prescriptions).

Gender-related system justification. Support for the gender hierarchy was assessed using the Gender System Justification Beliefs scale (GSJB; Jost & Kay, 2005). Participants indicated their agreement with eight items (e.g., “In general, relations between men and women are just and fair,” and “Society is set up so that men and women usually get what they deserve”), on a scale ranging from 1 (*strongly disagree*) to 9 (*strongly agree*). Responses were averaged so that high scores reflect endorsing the gender hierarchy ($\alpha = .75$).

Procedure

Participants were told that they would be helping a recent recipient of a Master's degree in Marketing rehearse for an upcoming phone interview. They read the job description and received their questions on index cards (in random order). After briefly meeting the confederate, participants were escorted to separate cubicles where they conducted the phone interview. Confederates answered using scripted responses. After the interview, participants completed their evaluations (competence, liking, hireability, and stereotypic traits) using a computer program. Gender system-justification beliefs were assessed weeks in advance as part of a departmental pretest.

After completing demographic items (gender and race), participants were subsequently debriefed and accredited.

Results and discussion

Applicant ratings

Preliminary analyses using the five individual confederates as a moderator revealed negligible effects on all dependent variables, all $F_s(4, 64) < 1.29$, *ns*. We therefore collapsed across this variable and submitted the competence, liking, and hireability indexes to separate 2 (target gender) \times 2 (participant gender) ANOVAs. As in Study 2, no significant effects involving participant gender emerged, all $F_s(1, 70) < 1.26$, *ns*.

Table 5 depicts the candidates' evaluations by target gender. Results revealed that agentic men and women were viewed as similarly competent, all $F_s(1, 70) < 1.55$, $ps > .21$. The expected backlash effects emerged: men were liked more than women, $F(1, 70) = 4.05$, $p < .05$ ($d = .43$), and were more likely to be hired than women, $F(1, 70) = 5.92$, $p < .05$ ($d = .51$). Table 5 also depicts the stereotypic trait ratings, by target gender. The last row reveals the dominance penalty for women, $t(71) = 2.03$, $p < .05$ ($d = -.46$). As in Study 2, results for the agency, communal, and weak indexes showed no target gender differences, all $t_s < 1.00$, *ns*.

Accounting for backlash toward female leaders

Table 6 provides a conceptual replication of Study 2. Models 1 and 2 reveal that the effects of target gender on hireability and liking were reduced to nonsignificance after accounting for liking (Model 1) or dominance (Model 2), and the 95% confidence intervals for the mediated effects did not include zero, resulting in full and reliable mediation in each case. Further, Model 3 shows that communality ratings did not mediate the effect of target gender on liking; although this effect was slightly reduced, the 95% confidence intervals for the mediated effect included zero.

Gender system justification

The SIH claims that defending the gender hierarchy is a key motive for backlash. If so, then gender system-justifiers should be especially likely to administer the dominance penalty (but not a communality deficit) when rating agentic women. To test these predictions, we standardized all variables and then hierarchically regressed female dominance prescriptions on target gender, gender system-justification beliefs (GSJB), and their interaction. We controlled for participant gender to be cautious, although results were virtually identical whether or not we did so, and no sex differences emerged on GSJB scores, $t(58) = 1.58$, *ns* ($M_s = 5.72$ and 5.40 for men and women). Beyond the known effect for target gender, the expected Target Gender \times GSJB interaction was significant, $\beta = -.33$, $p < .01$. Simple effects showed that for female targets, GSJB was positively correlated with perceived dominance, $r(29) = .46$, $p = .01$. For

Table 5
Evaluations of agentic male and female marketing manager applicants (Study 3).

	Male targets		Female targets		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Competence	4.88	.45	4.75	.61	.23
Liking	4.89 _a	.65	4.53 _b	.84	.43
Hireability	5.43 _a	.60	5.04 _b	.82	.51
Stereotypes					
Agentic	5.25	.54	5.27	.65	-.03
Communal	4.16	.60	4.15	.71	.02
Weak	1.78	.81	1.78	.64	.00
Dominant	3.03 _a	.76	3.44 _b	.98	-.46

Note. Means with different subscripts differ by target gender within each condition ($ps < .05$). Effect sizes (Cohen's *d*) represent target gender differences. Positive effect sizes favor male applicants; negative effect sizes favor female applicants. Conventional small, medium, and large effect sizes for *d* are .20, .50, and .80, respectively (Cohen, 1988).

Table 6
Mediation analyses for marketing manager applicants (Study 3).

Path/effect	B	SE	95% Confidence intervals
<i>Model 1</i>			
c (target gender → hire)	-.29*	.12	
a (target gender → like)	-.24*	.11	
b (like → hire)	.58***	.10	
c'	-.15	.10	
a × b (mediation effect)	-.14*	.07	-.283, -.015
<i>Model 2</i>			
c (target gender → like)	-.24*	.11	
a (target gender → dominant)	.20*	.10	
b (dominant → like)	-.35***	.12	
c'	-.18	.11	
a × b (mediation effect)	-.07*	.04	-.167, -.002
<i>Model 3</i>			
c (target gender → like)	-.24*	.11	
a (target gender → communal)	-.01	.10	
b (communal → like)	.80***	.09	
c'	-.23*	.08	
a × b (mediation effect)	-.01	.08	-.166, .149

Note. Target gender was coded 0 (male) 1 (female). Estimates are unstandardized. Confidence intervals for $a \times b$ are based on an asymmetrical distribution. Intervals that do not include zero support rejecting the null hypothesis that $a \times b = 0$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

male targets, this relationship was unreliably negative, $r(27) = -.15$, *ns*. Further, people who endorsed the gender status quo (i.e., who scored above the median on the GSJB) rated agentic women as significantly more dominant than agentic men, $t(28) = 2.12$, $p < .05$ ($d = .79$). For low gender system-justifiers (i.e., who scored at or below the median on the GSJB), there was no significant target gender difference, $t(28) = 1.62$, $p = .12$ ($d = .52$).

Communality ratings were then submitted to the identical analysis. As expected, results were negligible, all $\beta s < .13$, $ps > .34$. These findings bolster our claim that women's prescriptions do not reinforce the gender hierarchy; if they did, gender system-justifiers should have penalized agentic women with a communality deficit.

Results for the hiring index showed the known target gender effect and a reliable Target Gender \times GSJB interaction, $\beta = -.30$, $p < .05$. The correlation between GSJB and hireability was negative for agentic women, $r(31) = -.38$, $p < .05$, but weakly positive for agentic men, $r(29) = .14$, *ns*. For the liking index, the Target Gender \times GSJB interaction was marginally significant, $\beta = -.23$, $p < .08$. Simple slopes revealed a negative but unreliable relationship for agentic women, $r(31) = -.18$, $p = .32$, but a marginally positive relationship for agentic men, $r(29) = .35$, $p = .06$. Finally, people who endorsed the gender status quo rated agentic men as more hireable and likable than agentic women, both $ts(28) > 2.46$, $ps < .05$ ($ds = 1.13$ and 1.04 , respectively). For low gender system-justifiers, these target gender differences were negligible, both $ts(28) < 1.00$, *ns* ($ds = .01$ and $.06$ for hiring and liking).

Replicating Study 2, Study 3 showed that agentic female leaders suffered hiring discrimination that was fully mediated by prejudice, and prejudice that was fully explained by the dominance penalty. New to Study 3, gender-system justifiers were especially likely to administer the dominance penalty (but not a communality deficit). Further, gender-system justification beliefs moderated perceptions of hiring and liking. In each case, participants who defended the gender status quo were more likely to penalize agentic women, compared with agentic men.

Study 4

Study 3's results support the SIH's claim that gender system justification motives play a role in backlash, but they do not provide a

causal test. Prior research has provided direct support for system-justifying motives by exposing Canadians to information that their nation is on the rise or in decline (Kay et al., 2009). When under system threat, even women responded by defending gender stereotypes (e.g., that men are better at business than women); they were also likely to devalue a female business student (Kay et al., 2009). In Study 4, we manipulated system threat by exposing participants to news articles reporting that America was either in decline, on the rise, or nothing was said. To induce backlash, we used Study 2's recommendation letters for agentic male and female candidates.

Study 4's design is a 3 (system threat: high, low, control) \times 2 (target gender) \times 2 (participant gender) between subjects factorial. As in Study 2, we expected the agentic woman to be rated as more dominant, but less likable and hireable, compared with the agentic man. However, if backlash serves a system-justifying function, then all three penalties for female agency should be strongest in the high system threat condition, compared with the low threat condition and with controls.

Method

Participants

Volunteers ($N = 142$, 90 women) participated in exchange for partial fulfillment of their Introductory Psychology research requirement. Of these, 44% were White, 25% were Asian, 11% were Black, 6% were Hispanic, and 14% reported another ethnic identity.

Materials and procedure

System threat manipulation. Under the guise of piloting materials for an upcoming project, some participants read a news article describing America's status. In the high threat condition, they read an article titled, "America in Decline" that contained the following:

These days, many people in the United States feel disappointed with the nation's condition. Whether it stems from the economic meltdown and persistent high rates of unemployment, fatigue from fighting protracted wars in the Middle East that have cost America dearly in blood and treasure, or general anxieties regarding global and technological changes that the government seems unable to leverage to their advantage, Americans are deeply dissatisfied. Many citizens feel that the country has reached a low point in terms of social, economic, and political factors. It seems that many countries in the world are enjoying better economic and political conditions than the U.S. In recent nationwide polls, more Americans than ever before expressed a willingness to leave the United States and emigrate to other nations.

In the low threat condition, they read an article titled, "America on the Rise" that contained the following:

These days, despite the difficulties the nation is facing, many people in the United States feel safer and more secure than in the past. Whether it stems from the nation's relatively fast recovery from a global economic crisis, or the foreseeable end to protracted wars in the Middle East, or a general faith that the government will be able to leverage global and technological changes to their advantage, Americans are, on the whole, satisfied. Many citizens feel that the country is relatively stable in terms of social, economic, and political factors. It seems that compared with many countries in the world the social, economic, and political conditions in the U.S. are relatively good. In recent nationwide polls, very few Americans expressed a willingness to leave the United States and emigrate to other nations.

Following the threat manipulation, participants were encouraged to spend 3 min elaborating on what they had read and why they

thought the article's opinion was justified. Control participants (who did not read a news article) spent 3 min writing about their experiences the day before. To bolster the cover story, participants who read news articles reported the extent to which they found the news article clearly written, understandable, interesting, and compelling. Because no differences were found between high and low threat conditions, all $t_s(89) < 1.33, p_s > .19$, these results are not further discussed.

Candidate ratings. We used Study 2's competence, liking, and hireability indexes (all $\alpha_s > .81$). Because the results for competence did not differ as a function of target gender or threat condition, all $F_s < 2.12, p_s > .15$, we report only the findings for liking and hireability. We also employed the four sets of gender rules (all $\alpha_s > .75$). Because the results for agency and weakness did not differ as a function of target gender or threat condition, all $F_s < 1.83, p_s > .17$, we report only the findings for dominance and communality.

Procedure

Participants believed they would be completing two studies, one concerning piloting news articles for use in an upcoming project, and the other concerning their responses to a candidate being evaluated for promotion (to heighten the cover story, participants consented separately to the two studies). Participants were randomly assigned to threat condition by the computer program. After reading the news article (or not), they completed the writing task. They were then assigned by the computer program to read one of the two recommendation letters, ostensibly written by a Cornell English professor, for either a male or female candidate up for promotion (as in Study 2, except that all targets were agentic). After participants read the letter, the program administered the measures exactly as in Study 2 (items within each measure were randomly presented). Finally, participants completed demographic items (age, gender, and race). They were subsequently debriefed and accredited.

Results and discussion

Dependent measures were submitted to separate 3 (system threat: high, low, control) \times 2 (target gender) \times 2 (participant gender) ANOVAs.

Candidate evaluations

Replicating backlash effects, results for liking and hireability revealed main effects for target gender, both $F_s(1, 130) > 33.06, p_s < .001$, such that the male target was liked more than the female target ($M_s = 4.16$ vs. $3.11, d = .87$), and viewed as more hireable than the female target ($M_s = 5.55$ vs. $4.30, d = .95$). Of more importance, results for both measures showed the expected System Threat \times Target Gender interaction, both $F_s(2, 130) > 3.23, p_s < .05$. Fig. 1 shows the results.

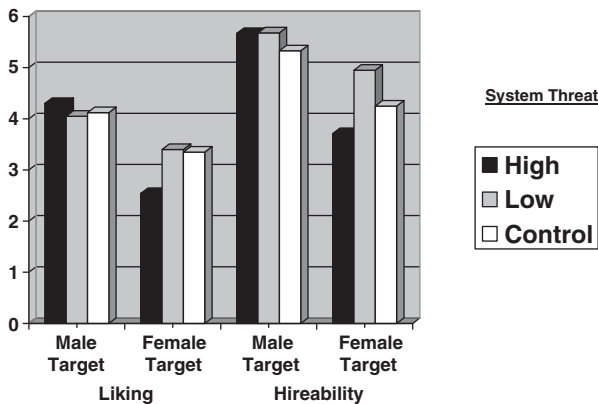


Fig. 1. Liking and hireability as a function of target gender and system threat (Study 4).

As predicted, the female target was liked less in the high threat condition compared with the low threat condition, $t(42) = 2.90, p < .01, d = .70$, and compared with controls, $t(44) = 2.53, p < .05, d = .66$. Similarly, the female target was viewed as significantly less hireable in the high threat condition compared with the low threat condition, $t(42) = 4.39, p < .001, d = .95$, and marginally, compared with controls, $t(44) = 1.80, p = .08, d = .41$. No significant differences for liking emerged between the low threat and control conditions, $t(44) < 1.00, ns$. However, low threat participants were more willing to hire the female target, compared with controls, $t(44) = 2.16, p < .05, d = .53$. For the male target, liking and hireability did not differ as a function of threat, all $t_s < 1.04, p_s > .38$.

Gender rules

Results for the dominance index revealed a target gender main effect, $F(1, 130) = 9.30, p < .01$, such that the agentic woman was viewed as more dominant than the agentic man ($M_s = 5.50$ vs. $4.36, d = .68$). In addition, the expected System Threat \times Target Gender interaction was significant, $F(2, 130) = 3.20, p < .05$. Fig. 2 shows the results. As predicted, the female target was judged to be more dominant in the high threat condition compared with the low threat condition, $t(42) = 3.03, p < .01, d = .79$, and marginally, compared with controls, $t(44) = 1.68, p < .10, d = .50$. No significant differences emerged between the low threat and control conditions, $t(44) < 1.00, ns$. The male target was viewed as similarly dominant in all threat conditions, all $t_s < 1.00, ns$.

By contrast, we found no significant target gender effect for the communality index, $F(1, 130) < 1.00, ns$. On average, male and female targets were viewed as similarly low on communality ($M_s = 2.98$ vs. 2.97). This replicates the null effects found in Studies 2 and 3. Unexpectedly, a marginally significant System Threat \times Target Gender interaction emerged, $F(2, 130) = 2.93, p < .06$. In no condition did we find significant target gender differences, all $t_s < 1.41, p_s > .15$. Instead, Fig. 2 shows that the male target was viewed as marginally more communal in the high, compared with the low, threat condition, $t(45) = 1.88, p < .07, d = .54$. No other reliable differences emerged for the male target, both $t_s < 1.64, p_s > .10$. Finally, the female target's communality did not differ across threat conditions, all $t_s < 1.28, p_s > .21$.

In summary, Study 4 bolsters the SIH's claim that backlash serves a system-justifying purpose. As expected, the female target was judged as particularly low on liking and hireability, and especially high on dominance, when participants were threatened with America's decline. By contrast, the male target was viewed similarly regardless of the threat manipulation (with the exception of a marginally

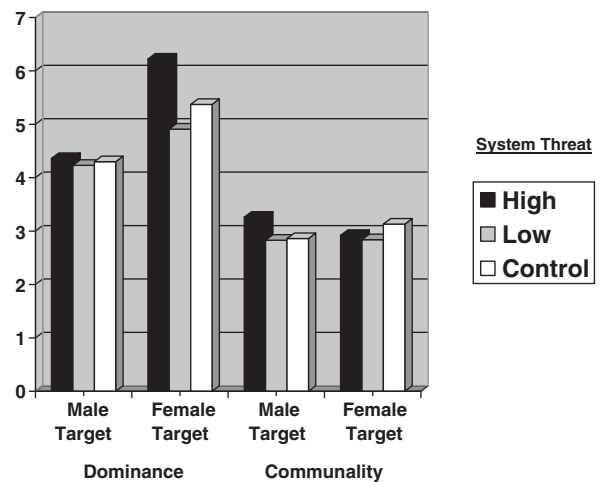


Fig. 2. Dominance and communality as a function of target gender and system threat (Study 4).

significant difference in his communality ratings). As in prior studies, the female target was viewed as higher on dominance than the male target, but not lower on communality, suggesting that backlash stems from perceived status violations. Given that people under system threat tend to defend their worldviews, which include gender status differences (Kay et al., 2009), and because female agency was especially rejected by people under system threat, Study 4 provides direct evidence that backlash functions to preserve the gender hierarchy.

Study 5

Thus far, we have tested the SIH using agentic targets competing for leadership roles — a situation that evokes the dominance penalty for women. In Study 5, we sought to remove the dominance penalty by making confederates “accidental” leaders. We also manipulated agency to instantiate female targets’ status incongruity. We relied on high or low leadership ability to operationalize agency because it is aligned with high status ($d = 1.45$, see Table 1). The SIH argues that agentic women elicit backlash because they are viewed as status-incongruent (their gender’s low status conflicts with their high status behavior). Agentic female leaders also directly challenge the gender hierarchy: women who have the traits required for successful leadership are precisely those who can bring about social change. By contrast, low agency female leaders are less likely to succeed, representing less of a threat to the status quo (indeed, if their anticipated failure occurred it would reinforce rather than challenge gender stereotypes that women make poor leaders, ultimately supporting the status quo). As a result, even in the absence of the dominance penalty, perceivers should exhibit backlash against agentic female leaders to defend the gender hierarchy.

After a computer program “randomly assigned” male and female confederates to a leadership role (and all participants to a subordinate role), participants were afforded the opportunity to sabotage confederates (i.e., undermine their ability to succeed). Because sabotage reflects active harm, it is a compelling form of backlash (Phelan & Rudman, 2010; Rudman & Fairchild, 2004). The SIH predicts that only high agency female leaders will be sabotaged because they effectively threaten the gender hierarchy. By contrast, low agency female leaders should be judged as less qualified than comparable male leaders. As a rule, women need to be exceptionally agentic before being viewed as men’s equals (e.g., Biernat & Fuegen, 2001; Eagly & Karau, 2002; Foschi, 2000; Glick et al., 1988). The problem for women is that being “men’s equals” yields backlash.

Method

Participants

Volunteers ($N = 237$, 123 women) participated in exchange for partial fulfillment of their Introductory Psychology research requirement. Of these, 52% were White, 26% were Asian, 7% were Black, 5% were Hispanic, and 10% reported another ethnic identity.

Agency manipulation

To manipulate confederates’ agency, participants completed a computerized “leadership aptitude test” in the belief that confederates did so also. The test consisted of 14 items with high face validity. Sample items included, “When I am in charge of a group, things always go smoothly,” and “I usually do not need deadlines and timetables to be productive.” Seven indeterminate items were included so that participants would find their own leadership aptitude score credible (e.g., “Given the choice, I would rather do something physical than read a book”). All participants were informed by the program that they scored in the 77th percentile, to heighten respect for agentic confederate leaders, who ostensibly scored at the 97th percentile. In the low agency condition, confederates scored at the 67th percentile.

Participants also completed a brief “participant profile” that the experimenter exchanged so that “everyone will know something about their partner” before they started working together. The profile included gender, college major, and leadership aptitude score. The main purpose of the profile was to deliver the agency manipulation, but it also underscored that leaders would have control over subordinates’ outcomes. For that purpose, two items asked, “If you are assigned to be the boss, how comfortable would you feel grading others’ work?” and “If you are assigned to be the boss, how comfortable would you feel criticizing others?” on scales ranging from 1 (*not at all*) to 6 (*very much so*). All confederates had circled 6 and 5 for the two items, respectively. At the top of the profile, the confederate’s test score (high or low) appeared prominently.

Dependent measures

Competence. Participants indicated whether they thought the other participant was competent, intelligent, and would make a good leader, on scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). These items were averaged to form the competence index ($\alpha = .80$).

Liking and dominance. Participants responded to three items on the same scales averaged to form the liking index ($\alpha = .84$). The items were, “I do not like my partner very much” (reverse coded), “I would like to get to know my partner better,” and “I look forward to working closely with my partner.” They also indicated whether they thought their partner was dominant, controlling, and arrogant, averaged to form the dominance index ($\alpha = .80$).

Sabotage task. Sabotage was operationalized as inhibiting targets’ ability to successfully perform a task (Rudman & Fairchild, 2004; Tesser & Smith, 1980). Participants were asked to program an anagram task for the leader while the experimenter ostensibly gave the leader further instructions. Participants received an anagram and its solution and were told that the confederate had to type in the correct answer within 30 s in order to score points. Participants were instructed to choose only one clue from a list of three possible clues to present to the confederate for each of 10 anagrams. The clues were pretested to vary in their helpfulness from low to high. A sample anagram was “CPESNRAA” (the answer is “PANCREAS”). The following clues were provided: “It starts with the letter ‘P’” (unhelpful), “It’s an organ in your body” (medium), and “It’s the organ in your body that starts with ‘P’” (helpful). The clues (which were unlabeled and presented in random order) were subsequently scored on a scale from 1 (*helpful*) to 3 (*unhelpful*) and summed to form the sabotage index (possible range 10–30, $\alpha = .94$).⁵

Procedure

Up to four participants were recruited for a “Work Roles” study and told as a group that they would be randomly partnered and assigned to either a leadership or subordinate role for an upcoming task. At this point, the confederate (male or female)⁶ expressed a preference for the leadership role with one statement (“I hope I get to be the leader”), in response to which the experimenter emphasized that the computer would randomly determine their role. If they were

⁵ To ensure that saboteurs knew they were harming confederates, we included three items on scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Sample items were “I do not expect the other participant will perform well on the anagram task,” and “I tried to be helpful as possible when programming the task” (reverse coded). We averaged the awareness of sabotage index ($\alpha = .69$) and found that it correlated robustly with the sabotage task, $r(235) = .70, p < .001$. Thus, participants were aware that the anagram task could be used to sabotage targets (see also Rudman & Fairchild, 2004). Because the findings for this index mirrored those for the sabotage task, we do not discuss it further.

⁶ We used two male and two female confederates. Preliminary analyses suggested we could collapse across individual targets for our focal analyses.

assigned to the leadership role, the experimenter would set them up with the materials for the study. Meanwhile, subordinates (in fact, all participants) would program a task for the leader designed to be challenging (the sabotage task).

Participants and confederates were then escorted to separate cubicles, where participants filled out the profile and handed it to the experimenter, who then started a computer program to administer the measures. Participants first completed the leadership aptitude test and received their score. At this point, the experimenter entered with the confederate's profile, stating that the confederate had been assigned as the participant's partner and pointing out that she or he had scored at the 97th percentile on the leadership aptitude test (high agency condition), or at the 67th percentile (low agency condition). Participants were then "randomly assigned" to the subordinate role by the computer program (with the experimenter noting this information). The experimenter then left participants alone to complete the sabotage task. Afterwards, they completed the sabotage awareness index, and competence, liking, and dominance ratings. Upon completion, they were fully debriefed and compensated.

Results and discussion

Dependent measures were submitted to separate 2 (target gender) \times 2 (target agency: high, low) \times 2 (participant gender) ANOVAs.

Competence

The competence index showed the expected main effect for target agency, $F(229) = 27.13$, $p < .001$, such that high agency targets were rated as more competent and qualified leaders ($M = 4.94$, $SD = .93$) than low agency targets ($M = 4.30$, $SD = 1.06$), $d = .62$. There was also a main effect for target gender, such that men were rated as more competent than women, $F(229) = 5.31$, $p < .05$, $d = .30$. However, this effect was qualified by the expected Target Gender \times Agency interaction, $F(229) = 6.05$, $p < .05$. Table 7 reveals that high agency men and women were viewed similarly ($d = .11$), but low agency men received higher scores than comparable women ($d = .46$). This demonstrates the need for women to overcome the "lack of fit" between their gender and leadership by being exceptionally qualified – women's first hurdle (Eagly & Karau, 2002).

Liking

Results for the liking index revealed a main effect for target gender, $F(229) = 5.26$, $p < .05$. On average, male leaders were liked more than female leaders ($M_s = 4.03$ vs. 3.60, $SD_s = 1.42$ and 1.44), $d = .30$. No other effects emerged, all $F_s < 1.76$, $ps > .18$.

Dominance

We predicted that women would be spared the dominance penalty because they were not actively competing for leadership. As expected, the only significant effect for dominance was that men were rated higher than women, $F(229) = 5.26$, $p < .05$ ($M_s = 4.32$ vs. 3.83, $SD_s = 1.55$ and 1.60), $d = .31$. Thus, participants likely relied on

their gender stereotypes when rating confederates' dominance (Snodgrass & Rosenthal, 1984; see Table 2).

Sabotage

Results revealed a main effect for target gender, $F(229) = 5.00$, $p < .05$, such that women were sabotaged more than men ($d = .30$). However, this effect was qualified by the expected Target Gender \times Agency interaction, $F(229) = 4.59$, $p < .05$. Table 7 reveals that high agency women were sabotaged more than comparable men, $t(115) = 2.74$, $p < .01$ ($d = .51$). By contrast, no differences emerged among low agency targets, $t(118) < 1.00$, *ns*. Further analyses showed that high agency female leaders were sabotaged more than low agency leaders, whether male ($d = .51$) or female ($d = .41$), both $ts > 2.08$, $ps < .05$.

Finally, although women avoided the dominance penalty, perceived dominance corresponded with sabotage, but only for high agency female leaders, $r(43) = .32$, $p < .05$. By comparison, this relationship was weak for all other targets, with r_s ranging from $-.03$ (for high agency men) to $.12$ (for low agency women), all $ps > .42$.

In summary, Study 5's results support the SIH because women who challenged the status quo by being highly qualified leaders suffered the most sabotage, even though they were viewed as similarly competent, compared with highly qualified men. Further, although we successfully removed the dominance penalty, perceived dominance correlated with sabotage only for agentic women. Of importance, low agency female leaders were rated as less qualified than comparable male leaders, but they were not at risk for sabotage, likely because they posed no threat to the gender hierarchy. Thus, women occupying a leadership role is not sufficient to evoke backlash, but instead must be coupled with agency (the high status qualities expected of a leader) because women who are likely to succeed as leaders are precisely those who challenge the status quo.

General discussion

The present research underscores women's Catch-22: because their gender status is lower than men's, they must enact agency to be viewed as "fit" for leadership – yet if they do so, they risk backlash. Past demonstrations of backlash have left open the question of what motivates it (Rudman & Fairchild, 2004; Rudman & Phelan, 2008), and which set of gender rules are used to justify it. The status incongruity hypothesis (SIH) provides a significant advance by specifying that defending the gender hierarchy is a primary motive for backlash and, as a result, prejudice against female leaders stems from perceived status violations.

To test the SIH, it was necessary to establish the status implications of gender rules. After Study 1 determined that female dominance proscriptions are more aligned with status than female communality prescriptions, the results of four experiments showed strong support for the SIH. Although agentic female leaders were viewed as men's equals vis-à-vis competence, they suffered hiring discrimination that was mediated by prejudice, and prejudice was mediated by the dominance penalty (extreme ratings on high status traits reserved for leaders and men) in Studies 2–4. In addition, Study 3 showed that gender system-justifiers were particularly likely to engage in backlash, including the dominance penalty, and Study 4 showed that prejudice, hiring discrimination, and the dominance penalty were exacerbated when perceivers were primed with a system threat. These findings support our claim that a key function of backlash is to defend the gender status quo. In Study 5, female leaders who were high (but not low) on agency were sabotaged more so than all other targets, and although they avoided the dominance penalty, perceived dominance was significantly related to sabotage only in this condition. When agentic female leaders are rejected or sabotaged for being "too powerful," the logical inference is that backlash serves to defend the gender hierarchy. Beyond Study 4's direct evidence for this motive, further evidence concerns the fact that when women

Table 7
Reactions to confederate leaders by gender and agency (Study 5).

	High agency				Low agency			
	Male leaders		Female leaders		Male leaders		Female leaders	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Competence	4.98	.96	4.87	.86	4.50 _a	1.09	4.02 _b	.97
Sabotage	13.00 _a	4.36	15.47 _b	5.29	12.99	4.41	13.48	5.42

Note. Means with different subscripts differ between male and female leaders ($ps < .05$). Competence was scored on a scale ranging from 1 to 7. Sabotage scores could range from 10 to 30.

leveled their status by means of extreme diplomacy (Study 2), or low agency (Study 5), they avoided backlash, even though they held leadership positions. Thus, the SIH argues that female leadership on its own is not sufficient to provoke backlash; what counts is female agency, which can lead to the perception that female leaders are “uppity” rather than entitled to wield power.

The SIH sets a high bar by predicting that agentic women will be viewed as more dominant than men, despite robust gender stereotypes claiming the reverse. The present research suggests that contrast effects are not culpable in backlash, given that no other gender rules distinguished between target men and women. Perceptual contrast effects would result in rating agentic women as exceptionally high on agency, and low on weakness and communality, which we did not find. Even gender system-justifiers in Study 3 eschewed rating agentic women as lower on “niceness” than agentic men, as did participants primed with a system threat in Study 4, which supports Study 1’s finding that communality prescriptions do not strongly reinforce the gender hierarchy. In tandem with the dominance penalty, these null effects contradict the notion that perceivers will use any means available to them to justify backlash and instead, support defending the status quo as a key motive for rejecting agentic women. As a result, female leaders may have to walk a fine line between presenting themselves as qualified for leadership while side-stepping perceived violations of female dominance proscriptions.

As noted, female leaders are viewed favorably when they avoid male leaders’ privileges (e.g., taking charge and giving directions; Eagly et al., 1992; Rudman & Glick, 2001). In essence, when women lead by “not leading,” they are spared backlash. Although unfair, this dilemma may encourage women to adopt participatory styles of leadership that are more effective than being directive (Eagly & Carli, 2007), which may be cause for optimism. However, our concern is that status and competence ratings are so highly related (Eckes, 2002; Fiske et al., 2002) that female leaders who lower their status may jeopardize their perceived competence. Moreover, in order to become leaders, women must first defeat the stereotype that they are “not fit” to lead, and agency is necessary to thwart that hurdle (Rudman & Phelan, 2008; cf. the present Study 5). That is, women must combat sexism based on perceived incompetence, but their hands are tied when female agency is reacted to negatively – as “too powerful” (for a woman). Gender status rules virtually guarantee men’s greater access to power and resources, resulting in a system that rewards men for leadership abilities while punishing comparable women, thereby reinforcing the perceived conflict between a woman’s gender and power. Because this conflict is the root of the prejudice women face, the problem is self-perpetuating. Attenuating one’s status to avoid backlash might spare individual women the dominance penalty, but it does little to lift women’s status as a group – which is the ideal remedy for backlash, and one that should be kept in sight.

Clarifying the underpinnings of backlash helps to illuminate residual barriers to women’s advancement (Catalyst, 2010; Valian, 1999). Until women’s cultural status is equal to men’s, perceivers are likely to reject powerful women, and unfairly resort to the dominance penalty to justify doing so. For example, the SIH may explain why women who are merely described as successful managers are viewed as interpersonally abrasive (e.g., hostile and manipulative; Heilman, Block, & Martell, 1995; Heilman et al., 2004; Parks-Stamm et al., 2008). In the absence of status-attenuating displays, perceivers may assume successful women are extreme on dominance. The SIH also accounts for why people who showed automatic prejudice against female authority figures did so to the extent they showed implicit gender-status associations, rather than implicit gender-stereotype or gender-role associations (Rudman & Kilianski, 2000). By specifying which aspects of gender roles are culpable in backlash, as well as a key motive for backlash, the SIH reaches beyond past theorizing (Eagly & Karau, 2002) to explain why agentic women suffer backlash

even when they are not leaders; for example, when they are merely self-confident or present themselves as having expertise (for a review, see Rudman, 1998), and why atypical men risk backlash (e.g., Heilman & Wallen, 2010), covered in more detail below.

Finally, the SIH helps to explain why women, who are less likely than men to endorse gender stereotypes (e.g., Twenge, 1997), tend to engage in backlash to the same degree as men (for a review, see Rudman & Phelan, 2008). Given that men and women are intimately interdependent, both genders have a stake in the status quo (Glick & Fiske, 1996, 2001) and therefore, may react negatively to status-incongruent women. Further, when women display agency in pursuit of power, they dispute the presumed division between the sexes, reflected in the common vernacular, “the opposite sex.” Both genders are steeped in this division from childhood on, which is reflected in the tradition that men should be leaders in the public sphere, whereas women are valued for devoting themselves to family (Eagly, 1987; Eagly, Wood, & Diekmann, 2000). Women who challenge this view subvert powerful cultural norms dictating what constitutes an appropriate and fulfilling life, as well as deeply ingrained beliefs about gender status hierarchies. As a result, we should perhaps not be surprised that women and men engage in backlash similarly, and for the same primary motive.

Limitations and future directions

Gender system-justification beliefs moderated the dominance penalty in Study 3, but other variables should be investigated. For example, the role of essentialism might be explored. Because agentic women challenge presumed gender differences, people who believe that women are “by nature” inherently different from men may resist powerful women more so than social constructionists. Essentialists are more likely to endorse gender stereotypes (e.g., Brescoll & LaFrance, 2004; Prentice & Miller, 2007), but whether this would influence reactions to female agency is a question worthy of future research.

The present research focused on prejudice toward female leaders because of its significance for gender equality. However, the SIH should also apply to men. In Study 1, male prescriptions and proscriptions alike were strongly aligned with status ($M ds = 1.27$ and $-.76$, respectively). Thus, to avoid status violations, men should be required to enact both the positive traits associated with leaders and to avoid any signs of weakness associated with low status. Consistent with this view, modest male job applicants were judged to be less likable than modest female applicants, and this effect was mediated by both male agency prescriptions and male weakness proscriptions (Moss-Racusin et al., 2010). No other stereotypic rules distinguished between modest men and women (including communality – it was not that modest men were “too nice”), and competence ratings were similar. Instead, because modesty is a status-attenuating behavior (Jones, 1964), and both sets of men’s rules are aligned with status, modest men were likely penalized for failing to uphold the gender hierarchy. Similarly, men who were described as highly successful in a female-dominated profession were viewed as wimpy, weak, and spineless, which corresponded to rejecting them as potential supervisors (Heilman & Wallen, 2010). Finally, male workers who requested a family leave to care for an ailing child were judged as high on weakness and low on agency, and these perceptions led to economic penalties (e.g., recommendations for demotion and termination; Rudman & Mescher, in press). Although further research is needed to test the SIH for men, these initial results provide promising support.

A general limitation of backlash research is that it has largely focused on reactions to atypical White women and men. Nonetheless, the SIH should have broader applications. Like gender, racial stereotypes inarguably serve to justify social hierarchies, with Whites being viewed as more competent and qualified for leadership, relative to racial minorities (for a review, see Pratto & Pitpitpan, 2008). As a

result, the strongest “rule” for Whites may be that they conform to status expectancies, rather than stereotypic expectancies (Phelan & Rudman, 2010). By contrast, reactions to minority members may be informed by dominance proscriptions, much as they are for White women. Although there is evidence suggesting this is true for Black men (Livingston & Richardson, submitted for review; Dovidio & Gaertner, 1981; Livingston & Pearce, 2009), to date, we know of only one study in which prescriptive stereotypes for Black men and women were compared and then used to inform a backlash investigation (Richardson et al., submitted for publication). The authors found that (a) proscriptions against dominance were stronger for Black men than for Black women, and (b) a highly competent but dominant Black man suffered more backlash than an identically described Black woman. These findings suggest that gender rules may differ (or even be reversed) for Blacks, but may similarly inform backlash. However, this arena remains wide open for further research. As a first step, researchers should undergo systematic investigations of prescriptive and proscriptive stereotypes and their status implications for the groups they examine.

Conclusion

Because gender rules are organizing principles that legitimize men's greater access to power and resources (Ridgeway, 2001), and individuals are motivated to justify the existing social structure (Jost et al., 2004; Rudman et al., 2002), behaviors that challenge the gender hierarchy are censured. The SIH proposes that defending the gender hierarchy is a primary motive for backlash and thus, women are penalized for perceived status violations. To the extent that agentic women are judged as brazen (rather than bold), they are held hostage to a set of rules that perpetuates inequality by reserving power and its privileges for men. Because exhibiting masculine traits is required for advancement in most high status careers, rejecting agentic women reinforces male hegemony. Until women are afforded the same latitude as men to strive for leadership, unequal opportunity and disparate treatment in the workplace will persist. As a result, a pressing challenge facing gender equality is the perceived conflict between a woman's gender, public positions of influence, and the agentic qualities necessary to be judged worthy of wielding power and control.

Acknowledgments

This research was partially supported by grant BCS-0443342 from National Science Foundation to the first author and an NSF Graduate Research Fellowship to the second author. We thank Roos Vonk and Susan Fiske for their invaluable assistance with Study 3, which was conducted as part of Sanne Nauts' internship, supervised by Vonk, Rudman, and Fiske. We also thank Nalini Ambady, Alice Eagly, Susan Fiske, and Peter Glick for their helpful comments on earlier versions of this paper. Correspondence concerning this article may be addressed to Laurie A. Rudman, Department of Psychology, Tillett Hall, Rutgers University, 53 Avenue E, Piscataway, NJ, 08854-8040. Electronic mail may be addressed to rudman@rci.rutgers.edu.

Appendix A

Excerpt from recommendation letter for Full English Professor (female candidate version).

All conditions

Although Professor Mullen has established herself as an accomplished novelist, she may be even better known for her reputation as a literary critic. Professor Mullen has made a name for herself as the most respected opinion on contemporary fiction, and her critiques have appeared in the most prestigious newspapers and

journals throughout the world. Her reviews have simply become the final word on what people ought (and ought not) to be reading.

Agentic condition

Professor Mullen has very high standards for contemporary fiction, and is not afraid to share her brutally honest opinion when her colleagues' work falls short. Although she can be merciless in her damning critiques, the ability to provide unvarnished, critical feedback is essential to moving our profession forward. Further, it should be noted that Professor Mullen is not discriminatory in her reviewing, treating newcomers and established writers alike to the same vicious knife. Her negative review of Margaret Crosley's dreadful last book as “the most lazily-written book I have ever read... having all the grace and subtlety of a television drama of the week” is just one example of her willingness to cut to the bone without regard for the famously fragile artistic ego. Moreover, I witnessed her public debate with Philip Baldwin at the 2007 ALA convention, in which she commented that his (terribly boring) recent book's commercial success was remarkable, given that it was “terrible, bloated, boring, gratuitous, and shamelessly uncontrolled” and that it “could not survive without the goodwill of its readers.” Although some thought she went “for the jugular,” others viewed it as a testimony to her commitment to maintaining high standards.

Communal condition

Professor Mullen has very high standards for contemporary fiction, and her reviews are particularly artful when her colleague's work falls short. Although she is at times viewed as overly polite and friendly because of her diplomatic critiques, the ability to provide tactful, critical feedback is essential to moving our profession forward. Further, it should be noted that Professor Mullen is not discriminatory in her reviewing, treating newcomers and established writers alike to the same sensitivity. Her review of Margaret Crosley's dreadful last book as “wordy, predictable, and conventional... although, to be fair, some readers may appreciate the familiarity and accessibility of the plot” is just one example of her willingness to cut to the chase while protecting the famously fragile artistic ego. Moreover, I witnessed her public debate with Philip Baldwin at the 2007 ALA convention, in which she commented that his (terribly boring) recent book's commercial success was remarkable, given that it was so “jam packed with verbiage, frustratingly complex character descriptions, and intricate landscape details” and that it was “truly a testament to his readers' willpower that they were able to appreciate the underlying story hidden beneath all those words.” Although some thought she went “too easy” on him, others viewed it as a testimony to her commitment to protecting authors' egos.

Appendix B

Excerpts from applicant scripts.

Q1. Can you give an example of a project you did in your former position at DWG, and what your role in this project was?

At DWG, I developed a new way to conduct market analysis that improved our response by 20%. Instead of conducting market analysis by phone, I initiated a project to switch to modern technologies, such as e-mail and using Facebook. We reached a much younger group of customers and dramatically improved our marketing plan.

Q2. Can you name your two most important qualities and a point for improvement?

One of my most important qualities is that I am good at analyzing complex situations. At DWG, I was often confronted with difficult situations in which I had to incorporate perspectives of different people

within the company, and determine what was important and needed to be addressed immediately and what could wait. I also know what I want and I like to make quick decisions, not debate endlessly about all the options that are on the table. It is inefficient to keep repeating arguments people already heard a couple of times; if you have all the necessary information, at a certain point you just need to stop talking and decide what to do.

A point of improvement? Let me think. Uh...I think a point of improvement is that I can be a little impatient now and then. Making decisions quickly is important, but you should also wait for others if they need a little more time to reach a certain solution.

Q3. Under which conditions do you work best?

I tend to perform best in pressure situations. If I have to do many things in a short period of time, I just try to plan things really well and I always get a lot of things done. If I have a deadline, I can work very efficiently.

References

- 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.
- Berger, J., Webster, M., Jr., Ridgeway, C. L., & Rosenholtz, S. J. (1986). Status cues, expectations, and behaviors. In E. Lawler (Ed.), *Advances in group processes*, Vol. 3. (pp. 1–22) Greenwich, CT: JAI Press.
- Biernat, M., & Fuegen, K. (2001). Shifting standards and the evaluation of competence: Complexity in gender-based judgment and decision making. *Journal of Social Issues*, 57, 707–724.
- Brescoll, V., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science*, 15(8), 515–520.
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive stereotyping in sex discrimination. *Psychology, Public Policy, and Law*, 5, 665–692.
- Catalyst (2010). Statistical overview of women in the workplace. Available online at <http://catalyst.org/publication/219/statistical-overview-of-women-in-the-workplace>.
- Cohen, J. (1988). *Statistical power for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Conway, M., Pizzamiglio, M. T., & Mount, L. (1996). Status, communality, and agency: Implications for stereotypes of gender and other groups. *Journal of Personality and Social Psychology*, 71, 25–38.
- Dodge, K. A., Gilroy, F. D., & Fenzel, L. M. (1995). Requisite management characteristics revisited: Two decades later. *Journal of Social Behavior and Personality*, 10, 253–264.
- Dovidio, J. F., & Gaertner, S. L. (1981). The effects of race, status, and ability on helping behavior. *Social Psychology Quarterly*, 44, 192–203.
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ.
- Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders*. Boston: Harvard Business School Press.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109, 573–598.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111, 3–22.
- Eagly, A. H., Wood, W., & Diekmann, A. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes, & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Erlbaum.
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles*, 47(3–4), 99–114.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), (4th ed.). *Handbook of social psychology*, Vol. 2. (pp. 357–411) New York: McGraw-Hill.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878–902.
- Fiske, S. T., & Stevens, L. E. (1993). What's so special about sex? Gender stereotyping and discrimination. In S. Oskamp, & M. Costanzo (Eds.), *Gender issues in contemporary society* (pp. 173–196). Thousand Oaks, CA: Sage.
- Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology*, 26, 21–42.
- Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491–512.
- Glick, P., & Fiske, S. T. (2001). Ambivalent sexism. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, 33. (pp. 115–188) Thousand Oaks, CA: Academic Press.
- Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions? *Journal of Personality and Social Psychology*, 55, 178–186.
- Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5, 269–298.
- Heilman, M. E., Block, C. J., & Martell, R. F. (1995). Sex stereotypes: Do they influence perceptions of managers? *Journal of Social Behavior and Personality*, 10, 237–252.
- Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. *Journal of Applied Psychology*, 92, 81–92.
- Heilman, M. E., & Wallen, A. S. (2010). Wimpy and undeserving of respect: Penalties for men's gender-inconsistent success. *Journal of Experimental Social Psychology*, 46, 664–667.
- Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, 89, 416–427.
- Henik, A., & Tzelgov, J. (1985). Control of halo error: A multiple regression approach. *Journal of Applied Psychology*, 70, 577–921.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology*, 58, 197–208.
- Jones, E. E. (1964). *Ingratiation: A social-psychological analysis*. New York: Appleton-Century-Crofts.
- 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.
- Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology*, 88, 498–509.
- Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, 89, 899–913.
- Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.
- Kay, A. C., Gaucher, D., Peach, J. M., Laurin, K., Friesen, J., Zanna, M., et al. (2009). Inequality, discrimination, and the power of the status quo: Direct evidence for a motivation to see the way things are as the way they should be. *Journal of Personality and Social Psychology*, 97, 421–434.
- 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.
- Livingston, R. W., & Pearce, N. A. (2009). The teddy bear effect: Does having a baby face benefit black chief executive officers? *Psychological Science*, 20, 1229–1236.
- Livingston, R. W., & Richardson, E. V. (submitted for review). The hubris penalty: Biased responses to “celebration” displays of Black NFL players.
- MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods*, 39, 384–389.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the products and resampling methods. *Multivariate Behavioral Research*, 39, 99–128.
- Manis, M., Nelson, T. E., & Shedler, J. (1988). Stereotypes and social judgment: Extremity, assimilation, and contrast. *Journal of Personality and Social Psychology*, 55, 28–36.
- Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: status incongruity and backlash toward modest men. *Psychology of Men and Masculinity*, 11, 140–151.
- Nieva, V. E., & Gutek, B. A. (1981). *Women and work: A psychological perspective*. New York: Praeger.
- 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.
- Phelan, J. E., & Rudman, L. A. (2010). Actors' reactions to ethnic deviance: The role of backlash in racial stereotype preservation. *Journal of Personality and Social Psychology*, 99, 265–281.
- Pratto, F., & Pittipati, 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. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science*, 16, 202–206.
- Richardson, E., Phillips, K. W., Rudman, L. A., & Glick, P. (submitted for publication). Double jeopardy or greater latitude: Do Black women escape backlash for dominance displays?
- Ridgeway, C. L. (2001). Gender, status, and leadership. *Journal of Social Issues*, 57, 627–655.
- Ridgeway, C., & Bourg, C. (2004). Gender as status: An expectation states theory approach. In A. H. Eagly, A. E. Beall, & R. J. Sternberg (Eds.), *The psychology of gender* (pp. 217–241). (2nd ed.). New York: Guilford Press.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 9, 283–294.
- 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., & 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., Feinberg, J. M., & Fairchild, K. (2002). Minority members' implicit attitudes: Ingroup bias as a function of group status. *Social Cognition*, 20, 294–320.
- 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: Guilford.
- Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*, 1315–1328.
- Rudman, L. A., & Mescher, K. (in press). Penalizing men who request a family leave: Is flexibility stigma a femininity stigma? *Journal of Social Issues*.
- Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. In A. P. Brief, & B. M. Staw (Eds.), *Research in organizational behavior, Vol. 4*. (pp. 61–79) New York: Elsevier.
- Shrout, P. T., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445.
- Snodgrass, S. E., & Rosenthal, R. (1984). Females in charge: Effects of sex of subordinate and romantic attachment status upon self-ratings of dominance. *Journal of Personality, 52*, 355–371.
- Tesser, A., & Smith, J. (1980). Some effects of task relevance and helping: You don't always help the one you like. *Journal of Experimental Social Psychology, 16*, 582–590.
- Touhey, J. C. (1974). Effects of additional women professionals on ratings of occupational prestige and desirability. *Journal of Personality and Social Psychology, 29*, 86–89.
- Twenge, J. M. (1997). Attitudes toward women, 1970–1995: A meta-analysis. *Sex Roles, 31*, 35–52.
- Valian, V. (1999). *Why so slow? The advancement of women*. Cambridge, MA: The MIT Press.
- Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multinational study*. Newbury Park, CA: Sage.