THIS IS A PREPRINT THAT IS CURRENTLY UNDER PEER-REVIEW. THEREFORE,
THE DOCUMENT IS SUBJECT TO CHANGE. KEEP THIS IN MIND WHEN CITING IT.

REFERENCE:

Born, N.M., & Erle, T.M. Are we there yet? Shifting towards a more positive perception of men and women countering gender stereotypes in the workplace. Preprint at: <https://osf.io/wpx7m/> (2020).

Are we there yet? Shifting towards a more positive perception of men and women countering
gender stereotypes in the workplace

Nadja M. Born*

Massachusetts Institute of Technology

Thorsten M. Erle*

Tilburg University

Please address correspondence to: Dr. Thorsten Michael Erle, Tilburg School of Social and Behavioral Sciences, Department of Social Psychology, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, Phone: (+31) 1334664165, E-mail: t.m.erle@tilburguniversity.edu

This research was funded by a Prolific Junior Researcher Grant awarded by Prolific Academic (<https://prolific.co>). The authors would like to thank Madeline Heilman and Tyler Okimoto for providing their stimulus materials, and Victoria Striewe for valuable feedback on the results of Experiment 2.

* shared first authorship

Abstract

Individuals who violate gender stereotypes such as agentic female leaders or communal male subordinates often suffer social or economic penalties, so-called backlash. Two high-powered experiments (total $N = 975$) aimed to answer the question whether backlash occurs due to the perception of gender-specific proscription or prescription violations. A proscription-penalty implies that a person exhibits an excess of behaviors or traits that are stereotypically associated with the other gender. A prescription-deficit implies that a person lacks the qualities that are stereotypically expected of his or her gender. To test this, participants rated gender-congruent and incongruent job applicants on indicators of gender-proscriptions, -prescriptions, and backlash. In Experiment 1, the job applicants applied for a desirable non-leadership position. The results strongly supported the idea that gender-incongruent individuals are perceived as having a prescription-deficit. But surprisingly no backlash for gender-incongruent individuals was found. Therefore, Experiment 2 was a pre-registered conceptual replication of the first experiment in a leadership context where backlash is more likely. Although Experiment 2 replicated the absence of backlash at the leadership level, the data did neither support a prescription deficit, nor a proscription penalty. Cross-experimental comparisons unearthed potential explanations for these findings, based on which implications for human resources and future research are outlined.

Keywords: gender stereotypes, prescriptions, proscriptions, communality, agency

Introduction

“A father and his son are in a car accident. The father dies at the scene and the son, badly injured, is rushed to the hospital. In the operating room, the surgeon looks at the boy and says, ‘I can’t operate on this boy. He is my son.’ How can this be?” (Banaji & Greenwald, 2013, p. 71)

This riddle elegantly demonstrates an automatic association (i.e., surgeon = male) that leads to the fallacious conclusion that the riddle is unsolvable. As obvious as the solution is, only few think of it straight away: the surgeon is the boy’s mother. Our daily lives are filled with gender stereotypes - many occupations and personality traits are automatically associated with one gender (Banaji & Hardin, 1996; Blair & Banaji, 1996). Even when people do not consciously agree with gender stereotypes, they are often influenced by them.

Typical women are thought to be caring, sensitive, friendly, and understanding – traits summarized under *communality* while typical men are described as ambitious, assertive, independent, and strong – traits associated with *agency* (Eckes, 2010; Heilman, 2012; Rudman & Glick, 2010). As Rudman and Glick (2010) put it, men are stereotyped as “bad but bold” and women as “wonderful but weak” (p. 82).

According to Rudman and Glick (2010), stereotypes have a *prescriptive* function, defining how men and women should be or behave (see also Burgess & Borgida, 1999; Heilman, 2001; Prentice & Carranza, 2002). While women are prescribed to behave in a communal way, men are prescribed to behave in an agentic way (Prentice & Carranza, 2002). In addition to prescriptions there are gender-specific *proscriptions*, that is, “generally undesirable qualities that [are] prohibited for one gender, but tolerated for the other” (Rudman, Moss-Racusin, Glick, & Phelan, 2012a, p. 181). Dominance-related traits such as being controlling and arrogant are proscribed for women; in contrast, vulnerability-related traits such as weakness and naiveté are proscribed for men (Prentice & Carranza, 2002).

Gender stereotypes can also function as gender rules. When people violate them, they risk suffering *backlash*. Backlash manifests itself for instance as social rejection, hiring discrimination, or sabotage (Rudman & Phelan, 2008). There has been extensive research regarding such backlash against gender-incongruent women showing that when women are perceived as competent or agentic, they often are also seen as less likable and more hostile (e.g., Carli, 2001; Cuddy, Glick, & Beninger, 2011; Heilman, Wallen, Fuchs & Tamkins, 2004; Rudman, 1998; Rudman & Glick, 1999; 2001; Rudman & Phelan, 2008), indicative of hostile sexism (Cikara & Fiske, 2009; Gaunt, 2013; Glick & Fiske, 1996; Wade & Brewer, 2006). Defying stereotypes in the workplace can also lead to backlash against men (Rudman et al., 2012a), although comparatively less evidence for this exists and more research on this underexamined inequality is needed (Croft, Schmader, & Block, 2015). So far, gender backlash has been observed against male elementary school teachers (Moss-Racusin & Johnson, 2016), men who request a family leave (Rudman & Mescher, 2013), and modest men (Moss-Racusin, Phelan, & Rudman, 2010). Importantly, while gender-incongruent women are disliked and interpersonally derogated, gender-incongruent men are perceived as ineffectual and less respectable than gender-congruent women or men. Both gender-incongruent men and women were seen as less desirable as a boss than their gender-congruent counterparts (Heilman & Wallen, 2010).

Two potential underlying mechanisms for backlash have been discussed: First, studies suggest that backlash results from a perceived deficit in communality for women (Heilman & Okimoto, 2007) and a perceived deficit in agency for men (Heilman & Wallen, 2010), respectively (i.e., from the violation of prescriptions). Second, the *Status Incongruity Hypothesis* (SIH; Rudman, Moss-Racusin, Phelan, & Nauts, 2012b) argues that backlash is motivated by a need to preserve the status hierarchy and consequently for women, results only from status-relevant proscription violations (i.e., a dominance penalty), but not from status-neutral prescription violations (i.e., a communality deficit). For men however, the SIH would

predict that prescription (i.e., an agency deficit) as well as proscription violations (i.e., a weakness penalty) would lead to backlash, since both are status relevant.

Early accounts like Heilman and Okimoto (2007) tested the idea that backlash for women's success in traditionally male tasks occurs due to a perceived deficit in communality, that is implied by their success. The authors expected that providing information of communality would prevent backlash and their results indeed suggested that the backlash directed toward successful female managers was mitigated when there was additional information about their communality. However, Heilman and Okimoto (2007) only measured backlash (i.e. likeability, perceived interpersonal hostility and boss desirability) and prescriptions (i.e. agency and communality), but not proscriptions (i.e. dominance and weakness) and according to Rudman et al. (2012b, p. 190), they themselves "repeatedly failed to support the communality deficit explanation using self-reports: agentic women simply are not rated as less communal than agentic men (Rudman et al., 2012). Instead, support for the communality deficit explanation was inferred from research showing that communal displays appear to inoculate female leaders against backlash".

In order to explicitly pit the prescription and proscription explanation against each other Rudman et al. (2012a) compared agentic men and women and measured female and male gender rules (i.e. prescriptions for agency and communion, proscriptions for dominance and weakness). Their results supported the predictions of a proscription explanation by the SIH, as agentic women were seen as more dominant than agentic men but not as less communal than agentic men. However, in this previous research supporting a dominance penalty rather than a communality deficit, agentic women were only compared to agentic men. Similarly in accounts about atypical men supporting the SIH, modest men were compared with modest women (Moss-Racusin et al., 2010) and received both higher weakness ratings (i.e. proscription violation) and lower agency ratings (i.e. prescription violation) than equally depicted female targets.

Nevertheless, in contrast to the original article about an implied communality deficit for women who succeed at male tasks by Heilman and Okimoto (2007), there were no neutral control groups (i.e., individuals framed as neither agentic nor communal) in these later studies. In our opinion, in order to examine a deficit explanation for backlash and whether gender-incongruent individuals actually lose communality or agency, gender-incongruent individuals should most importantly be compared to a neutral control group of their own gender, rather than a comparable group of the opposite gender. For example, the observed difference in perceived dominance between agentic women and men (Rudman et al., 2012a) could be attributed to a dominance penalty for women. However, it could similarly be caused by agentic men being perceived as less dominant than regular men (potentially because their dominance is misinterpreted as agency). Without neutral female and male control groups, any explanation based on increases or decreases on any dimension is inherently confounded with an inverse decrease or increase for the other group.

We therefore want to answer the important open research question, for both men and women, whether backlash is the consequence of not fulfilling gender prescriptions (as argued by Heilman & Okimoto, 2007; Heilman & Wallen, 2010) or of fulfilling gender proscriptions (as argued by the SIH; e.g., Rudman et al., 2012a). For instance, are female leaders punished for not being communal enough or for seeming too dominant? And are men penalized for not being agentic enough or for seeming too weak? To answer the question whether the violation of proscriptions or prescriptions is responsible for the backlash experienced by gender-incongruent individuals, it is necessary to directly assess both proscriptions and prescriptions at the same time and to compare both gender-congruent and incongruent individuals to a control group. This was done in two high-powered experiments.

Open Practices, Power Analysis, and Data Treatment

Both experiments were designed to be sensitive for effect sizes of $\eta_p^2 = .02$ with a power of $(1-\beta) = .95$. The target sample size was computed as $N = 468$ using G*Power 3 (Faul,

Erdfelder, Lang, & Buchner, 2007). In a sequential analysis (Lakens, 2014) with interim analyses after every 160 participants, 480 participants were collected. Using a linear spending function, alpha boundaries of .009 and .011 for the interim analyses and .014 for the final analysis resulted. There were no exclusion criteria. All measures and manipulations of both studies are reported, and all data and materials, as well as the pre-registration of Experiment 2 are available at <https://osf.io/k4zjq/>.

Experiment 1

In Experiment 1, participants were asked to evaluate job applicants on gender prescriptions (communality, agency), gender proscriptions (dominance, weakness), and indicators of backlash (respect, likeability, hostility, and hireability). All participants evaluated both female and male applicants that applied for the same desirable position in a large company. In addition to information that made the applicants' competence unambiguously clear, participants received either additional communal or agentic information about the applicant, or no additional information (control group) to establish the gender-(in)congruency of the applicants.

Experiment 1 had three main goals. First, to replicate the finding that gender-incongruent applicants (compared to gender-congruent and control applicants) suffer backlash. For women, this would mean that they are rated as more hostile, less likeable and thus as less hireable. For men, backlash would mean that they are rated as less respectable and thus as less hireable. The second goal of the present project was to investigate whether gender-incongruent job applicants are primarily perceived as lacking gender-prescriptive attributes or as exhibiting gender-proscriptive attributes. Finally, we aimed to investigate whether these mechanisms (lack of prescriptive attributes or excess of proscriptive attributes) mediate the effects of our independent variables on gender backlash.

Methods

Materials and Procedure. Participants were told that they had to review two applicants, randomly selected from a larger group, applying for the same position in a large company, which was then described in detail. Then they saw a letter of recommendation for either a female or a male applicant adapted from Heilman and Okimoto (2007) that included the manipulations of applicant gender and attribute type (see below). All applicants were described as belonging to the top 5% of all employees at their level, as “stellar performer[s]” or “rising stars” (Heilman et al., 2004, p. 418). After the letter of recommendation, participants completed all dependent measures for the current applicant. Then, a similar letter was displayed for the second applicant, followed again by the same dependent measures. Presentation order of the applicants and dependent measures was randomized.

Experimental Manipulations. The study had a 2 (Gender: male vs. female applicant; within) $\times$ 2 (Job Type: female vs. male-gendered job; between) $\times$ 3 (Attribute Type: agentic vs. communal vs. control; between) design. All letters and job descriptions were closely adapted from previous research and are available at <https://osf.io/k4zjq/>.

Applicant gender. All participants reviewed one female and one male applicant. Applicant gender was manipulated by altering their names (Andrea vs. James) and gender pronouns in the letters of recommendation (Heilman & Okimoto, 2007).

Attribute type. At the end of the letter of recommendation, information indicating that the applicant was understanding, concerned about others, and caring (communal condition) or ambitious, rational, and decisive (agentic condition) or no additional information (control condition) was included (Heilman et al., 2004).

Job type. The applicants either applied for a male-gendered position of a Financial Advisor or a female-gendered position of an Employee Relations Counselor, adapted from previous research (Heilman & Okimoto, 2007). The job descriptions indicated that 88% of Financial Advisors/Employee Relations Counselors are men/women and that employees in both jobs typically exhibit stereotypically male/female characteristics. The salary was the same for

both jobs. Heilman and Wallen (2010) successfully used this manipulation and showed that the jobs are equal in status and prestige. We implemented this manipulation to safeguard us against the case that the descriptions of the applicants alone were not sufficient to create impressions of gender-(in-)congruency.

Dependent Measures. All scales were adopted from previous research (Heilman & Okimoto, 2007; Heilman et al., 2004; Heilmann & Wallen, 2010; Moss-Racusin et al., 2010; Rudman & Mescher, 2013; Rudman et al, 2012a; 2012b). Items for competence, communality, dominance, agency, weakness, hostility, respect, liking, and hireability had 9-point bipolar scales. The applicants' status was assessed using two MacArthur Scales of Subjective Social Status (Hu, Adler, Goldman, Weinstein, & Seeman, 2005), representing the applicant's relative position within the company and within all applicants, respectively. As a manipulation check, participants were asked what the typical gender of an employee in the advertised job was.

Sample. $N = 483$ (228 female, 252 male, 3 no answer; age: $M = 36.09$, $SD = 11.53$; work experience in years: $M = 14.83$, $SD = 11.20$) participants were recruited via Prolific Academic (Damer & Bradley, 2014) for a 10-minute study and received a compensation of £1.

Results

All dependent variables were subjected to 2x2x3 mixed ANOVAs with an alpha-level of $\alpha = .014$ in line with the sequential analysis protocol.

Manipulation Checks. For competence, although the descriptions between conditions were identical, there was a main effect of Gender with women ($M = 7.75$, $SD = 1.15$) being rated as more competent than men ($M = 7.58$, $SD = 1.27$), $F(1,477) = 9.88$, $p = .002$, $\eta_p^2 = .02$. There was also a main effect of Job Type with applicants for the male-gendered job were rated as more competent ($M = 7.87$, $SD = 0.89$) than applicants for the female-gendered job ($M = 7.45$, $SD = 1.17$), $F(1,477) = 19.18$, $p < .001$, $\eta_p^2 = .04$ (all other F s < 3.87 , all other p s $\geq .022$). For status, there were no significant effects (all F s < 5.40 , all p s $\geq .021$). Despite these small and unexpected differences, overall, the applicants were rated as highly competent.

The Job Type manipulation was also mostly successful as 73.77% described the female-gendered job as a female position, and 82.75% described the male-gendered job as a male position. Analyzing the data without participants who answered the manipulation check incorrectly did not affect the reported results.

Female Gender Rules: Communality and Dominance. A prescription-deficit would be supported by a Gender x Attribute Type interaction for the Communality ratings. Specifically, gender-incongruent (i.e., agentic) women should be perceived as less communal than the female control group, while there would be no difference between agentic men and the male control group. A proscription-penalty, on the other hand, would be supported by a Gender x Attribute Type interaction for the Dominance ratings with gender-incongruent (i.e., agentic) women scoring higher on dominance than the female control condition, while agentic men would not be rated differently than the male control group. The results of these two ANOVAs are plotted in Figure 1.

Communality. This analysis yielded main effects of Gender, $F(1,477) = 92.80, p < .001$, $\eta_p^2 = .16$ (with women receiving higher communality ratings than men) and Attribute Type, $F(2,477) = 94.99, p < .001$, $\eta_p^2 = .29$. Communal targets were rated more communal than both control, $t(315) = 11.78, p < .001, d = 1.32$, and agentic applicants, $t(327) = 12.40, p < .001, d = 1.36$. Most importantly, there was a significant Gender $\times$ Attribute Type interaction, $F(2,477) = 5.64, p = .004$, $\eta_p^2 = .02$ (all other F s < 1.16 , all other p s $\geq .315$). Supporting a prescription-deficit, agentic women were rated significantly less communal than control women, $t(307) = 2.51, p = .013, d = 0.28$, whereas agentic men were not rated significantly less communal than control men, $t(318) = 1.35, p = .264$.

Dominance. The analysis yielded significant main effects of Gender, $F(1,477) = 54.84, p < .001$, $\eta_p^2 = .10$ (with women receiving lower dominance ratings than men) and Attribute Type, $F(2,477) = 68.35, p < .001$, $\eta_p^2 = .22$ (all other F s < 2.57 , all other p s $\geq .110$). Communal

applicants received lower dominance ratings than control, $t(315) = 9.25, p < .001, d = 1.04$, and agentic applicants, $t(315) = 10.43, p < .001, d = 1.15$. The non-significant Gender x Attribute Type interaction, $F(2,477) = 1.75, p < .001, \eta_p^2 = .22$, does not support a specific proscription penalty for agentic female job applicants.

Figure 1. Mean Communality and Dominance ratings. Error bars indicate +/-1 SEM.

Male Gender Rules: Agency and Weakness. A prescription-deficit would be supported by a Gender x Attribute Type interaction for the agency ratings. Specifically, gender-incongruent (i.e., communal) men should be perceived as less agentic than the male control group, while there would be no difference between communal women and the female control group. A proscription-penalty, on the other hand, would be supported by a Gender x Attribute Type interaction for the weakness ratings with gender-incongruent (i.e., communal) men scoring higher on weakness than the male control condition, while communal women would not be rated differently than the female control group. The results of these two ANOVAs are plotted in Figure 2.

Agency. The analysis yielded significant main effects of Gender, $F(1,477) = 17.04, p < .001, \eta_p^2 = .03$, and Attribute Type, $F(2,477) = 12.91, p < .001, \eta_p^2 = .05$. Most importantly,

these main effects were qualified by a significant Gender $\times$ Attribute type interaction, $F(2,477) = 5.07, p = .007, \eta_p^2 = .02$ (all other F s < 0.75 , all other p s $\geq .388$). In line with the prescription-deficit hypothesis, communal men were perceived as less agentic than control men, $t(315) = 3.68, p < .001, d = 0.41$, whereas communal women were not rated significantly less agentic than control women, $t(315) = 0.43, p = .667$.

Weakness. This ANOVA yielded only a significant main effect of Attribute Type, $F(2,477) = 9.07, p < .001, \eta_p^2 = .04$ (all other F s < 1.67 , all other p s $\geq .190$). Agentic applicants were rated as less weak than both control, $t(318) = 4.30, p < .001, d = 0.48$, and communal applicants, $t(327) = 2.67, p = .008, d = 0.30$. Thus, the gender-incongruency did not affect weakness perception, speaking against a proscription penalty.

Figure 2. Mean Agency and Weakness ratings. Error bars indicate ± 1 SEM.

Backlash: Interpersonal Hostility, Respect, Likeability and Hireability. Across all analyses, there were no signs of backlash against gender-incongruent applicants. Gender-incongruent women were not rated as less hireable or likeable or more hostile than other women. Gender-incongruent men were also not rated as less respectable or hireable than other men. The full backlash analyses are documented at <https://osf.io/k4zjq/>.

Discussion

Experiment 1 showed that counter-stereotypical job applicants (i.e., agentic women and communal men) are rated as lacking gender-prescriptive attributes (i.e., communality for women and agency for men), but not as displaying higher levels of gender-proscriptive attributes (i.e., dominance for women and weakness for men). Thus, Experiment 1 supported the idea of a prescription-deficit more than that of a proscription penalty.

Furthermore, and to our surprise, no signs of backlash were observed. This was even the case for gender-incongruent job applicants applying for a gender-incongruent job. In general, the job type manipulation had little effect on the observed results. The absence of gender backlash could be due to multiple reasons. For example, it could be a function of our sample, which exhibited a lot of job experience (while many previous studies used undergraduates; e.g. Heilman & Okimoto, 2007; Heilman & Wallen, 2010; Moss-Racusin et al., 2010; Rudman et al. 2012b), due to changing societal norms that are more open towards gender-incongruent employees (Eagly, Nater, Miller, Kaufmann, & Sczesny, 2019; Rudman & Glick, 1999, 2001; Steffens, Schult, & Ebert, 2009), or due to the specific materials employed. For example, although the jobs were described as clearly prestigious and desirable, backlash for gender-incongruent applicants might only emerge at the top of a company's hierarchy (e.g., top level management). Although this seems unlikely given that the job materials were adapted from previous research conducted at lower hierarchical levels (Heilman & Wallen, 2010), but to conclusively rule this out, a second experiment was conducted.

Experiment 2

Experiment 2 was a pre-registered conceptual replication of Experiment 1 with only one major change: While the jobs that participants applied for in Experiment 1 were clearly rated as high in competence and status they were not part of the management within the company. This might have prevented backlash against gender-incongruent job applicants. To address this concern, in Experiment 2, participants applied for a leadership position.

Methods

The methodology was identical to Experiment 1, except for the following changes: First, the job descriptions were changed to clearly denoted leadership positions, adapted from previous research (Heilman et al., 2004). The applicants either applied for a position as Assistant Vice President of Human Resources in the elderly care industry (a classically female-dominated area) or as Assistant Vice President of Sales in the automotive industry (a classically male-dominated area; Eagly, Nater, Miller, Kaufmann, & Sczesny, 2019; Joshi, Son, & Roh, 2015; Koch, D'Mello, & Sackett, 2015). To make it clear that the advertised positions were in leadership, the job responsibilities included supervising staff and representing the company and salary and experience level were increased compared to the jobs in Experiment 1.

Second, instead of rating hireability, participants rated the applicants' boss desirability on a scale from 1 = "not at all" to 9 = "very much" (Heilman & Okimoto, 2007; Heilman & Wallen, 2010).

Sample. $N = 492$ (253 female, 237 male, 2 no answer; age: $M = 35.50$, $SD = 12.91$; work experience in years: $M = 14.07$, $SD = 11.76$) participants were recruited via Prolific Academic and received a compensation of £1.

Results

We carried out the same analyses as in Experiment 1, in line with the study's pre-registration (see <https://osf.io/k4zjq/>).

Manipulation Checks. For competence, we observed the same effects as in Experiment 1. There were main effects of Gender, with women ($M = 7.47$, $SD = 1.21$) being rated as more competent than men ($M = 7.31$, $SD = 1.19$), $F(1,486) = 9.47$, $p = .002$, $\eta_p^2 = .02$, and Job Type, $F(1,486) = 9.02$, $p = .003$, $\eta_p^2 = .02$, with applicants for male-gendered jobs receiving higher competence ratings ($M = 7.53$, $SD = 1.15$) than applicants for female-gendered jobs ($M = 7.25$, $SD = 1.23$; all other F s < 2.53 , all other p s $\geq .081$). For status, there were again

no significant effects (all F s < 2.90, all p s $\geq$.089). Overall, all applicants again were rated as highly competent.

The Job Type manipulation was also mostly successful, as 72.02% of the participants described the female-gendered job as a female position, and 89.16% described the male-gendered job as a male position. Analyzing the data without participants who answered the manipulation check incorrectly did not affect the reported results.

Female Gender Rules: Communality and Dominance. The results of these two ANOVAs are plotted in Figure 3.

Communality. Contrary to Experiment 1, this ANOVA indicated only significant main effects of Gender, $F(1,486) = 27.99$, $p < .001$, $\eta_p^2 = .05$, and Attribute Type, $F(2,486) = 69.55$, $p < .001$, $\eta_p^2 = .22$ (all other F s < 5.09, all other p s $\geq$.025). Women were rated as more communal than men and communal applicants were rated more communal than both control, $t(330) = 9.27$, $p < .001$, $d = 1.02$, and agentic applicants, $t(318) = 10.64$, $p < .001$, $d = 1.21$.

Crucially, the expected Gender $\times$ Attribute Type interaction did not replicate, $F(2,486) = 0.08$, $p = .920$. Speaking against a communality-deficit, both agentic women, $t(330) = 2.39$, $p = .017$, and agentic men, $t(330) = 2.31$, $p = .021$, were not rated significantly less communal than the control group. Although the effects in both conditions were significant using an unadjusted significance criterion, the stronger argument against a communality-deficit was that the effect was virtually identical for both genders.

Dominance. For Dominance the results were similar to Experiment 1 and there were only main effects of Applicant Gender, $F(1,486) = 10.89$, $p = .001$, $\eta_p^2 = .02$, and Attribute Type, $F(2,486) = 54.27$, $p < .001$, $\eta_p^2 = .18$ (all other F s < 1.74, all other p s $\geq$.177). Women were rated as less dominant than men and communal applicants received lower dominance ratings than both control, $t(330) = 6.37$, $p < .001$, $d = 0.70$, and agentic applicants, $t(318) = 9.94$, $p < .001$, $d = 1.11$.

Figure 3. Mean Communality and Dominance ratings. Error bars indicate ± 1 SEM.

Male Gender Rules: Agency and Weakness. The results of these two ANOVAs are plotted in Figure 4.

Agency. Contrary to Experiment 1, there was only a significant main effect of Attribute Type, $F(2,486) = 19.99, p < .001, \eta_p^2 = .08$ (all other F s < 2.97 , all other p s $\geq .053$). Agentic applicants received higher agency ratings than both control, $t(330) = 2.65, p = .008, d = 0.29$, and communal applicants, $t(327) = 6.03, p < .001, d = 0.67$. Interestingly, in contrast to Experiment 1 we did not find a significant main effect for Gender, men were not rated as more agentic than women, indicating that in leadership positions, agency is also implied for women.

Crucially, the expected Gender $\times$ Attribute Type interaction did not replicate, $F(2,486) = 2.91, p = .056$. Contrary to the prescription-deficit hypothesis, communal men were not perceived as less agentic than control men, $t(330) = 2.35, p = .019$, whereas communal women were unexpectedly rated as significantly less agentic than control women, $t(330) = 4.14, p < .001, d = 0.45$.

Weakness. For Weakness, results were overall similar to Experiment 1 and no significant effect emerged, all F s < 3.13 , all p s $\geq .045$.

Figure 4. Mean Agency and Weakness ratings. Error bars indicate ± 1 SEM.

Backlash: Interpersonal Hostility, Respect, Likeability and Boss Desirability.

Across all analyses, there were again no signs of backlash against gender-incongruent applicants. Gender-incongruent women were not rated as less desirable as a boss or likeable or more hostile than other women. Gender-incongruent men were also not rated as less respectable or desirable as a boss than other men. The full backlash analyses are documented at <https://osf.io/k4zjq/>.

Discussion

Experiment 2 demonstrated that the absence of backlash in Experiment 1 cannot be attributed to the specific materials. Even our efforts to manipulate gender-incongruency via attributes of the job and the applicants, as well as moving to a leadership context could not produce backlash – although all materials were adapted from previous research. Although initially surprising, these results are in line with a recent meta-analysis that also found no support for backlash-effects (Koch et al., 2015), no difference in ratings of men and women for male-dominated jobs, and that men were even rewarded for being highly competent in female-dominated jobs. Koch et al. (2015) interpreted this in line with a contrast effect (Kunda &

Thagard, 1996), in which unambiguous information about a counter-stereotypical trait leads to extreme ratings on this trait.

However, Experiment 2 differed regarding the perception of gender-incongruent applicants. Specifically, both agentic women and men were numerically seen as less communal than control women and men, but this effect was non-significant ($ps > .019$) and more importantly equally strong for both genders. Thus, for leadership positions both male and female agentic applicants are seen as slightly less communal. For communal job applicants, the pattern contradicted the prescription-deficit hypothesis even more: Whereas communal men were not seen as less agentic than control men, this difference was significant for female applicants. We will address this issue in the following.

Cross-Experimental Comparisons

In both studies, we observed no backlash against gender-incongruent job applicants. However, the perception of applicants regarding gender-prescriptions and proscriptions differed between studies, which could be considered indicative of stereotypes about leaders. To explore these differences further, we conducted 2 (Gender: male vs. female; within) $\times 2$ (Position: non-leadership (Experiment 1) vs. leadership (Experiment 2); between) $\times 3$ (Attribute Type: agentic vs. communal vs. control; between) ANOVAs on all gender-proscriptive and gender-prescriptive attributes.

For both attributes related to female gender rules, we observed Gender $\times$ Position interactions, (communality: $F(1,969) = 10.20, p < .001, \eta_p^2 = .01$; dominance: $F(1,969) = 10.31, p < .001, \eta_p^2 = .01$). While men were not rated as differently communal or dominant as a function of position (both $ts < 0.55$, both $ps \geq .587$), female superiors were generally rated as less communal, $t(973) = 3.07, p = .002, d = 0.20$, and more dominant, $t(973) = 3.99, p < .001, d = 0.26$. Although female superiors were rated as less dominant, $t(491) = 3.29, p = .001, d_z = 0.15$, and more communal, $t(491) = 5.38, p < .001, d_z = 0.26$, than male superiors, they

were still perceived as violating gender prescriptions and proscriptions compared to female applicants for a non-leadership position, while this was not the case for men.

For both attributes related to male gender rules, we observed Gender x Position x Attribute Type interactions, (agency: $F(2,969) = 8.09, p < .001, \eta_p^2 = .02$; weakness: $F(2,969) = 4.52, p = .011, \eta_p^2 = .01$). While there were generally no differences in the communal and agentic conditions (all t s < 1.69 , all p s $\geq .093$), in the control condition the pattern of Experiment 1 (male colleagues were rated as more agentic and less weak than female colleagues) was (by trend) reversed in Experiment 2 with male superiors receiving lower agency, $t(171) = 1.68, p = .096$, and higher weakness, $t(171) = 2.59, p = .010, d_z = 0.20$, ratings than female superiors.

Hence, even though the Job Type (i.e. female- vs. male-gendered job) manipulation failed to produce effects in both experiments, we could still see a cross-experimental effect of position level. While in Experiment 1 (for non-leadership positions) the judgment of gender-incongruent job applicants fell in line with the expectations of prescription violations (i.e., female controls were seen as more communal than gender-incongruent agentic women, and male controls were seen as more agentic than gender-incongruent communal men), in Experiment 2 the control conditions moved away from these prescriptions (i.e., female control leaders were seen as both less communal and more dominant than female control non-leaders, and male control leaders were seen as less agentic than male control non-leaders).

These effects for control leaders could stem from inherent associations of leadership positions. For female leaders, there might be an underlying assumption that women who have reached a leadership position must have shown less communal and more dominant traits to do so, even if there is no additional information about them such as in the control group. The fact that in Experiment 2 the significant main effect for gender on agency perception from Experiment 1 disappeared (i.e., on a leadership level, men were not rated as more agentic than women anymore) points into a similar direction. This is in line with previous research and could again stem from a contrast effect (Kunda & Thagard, 1996), indicating that the unambiguous

information about a woman's success in a leadership position lead to extreme ratings on the traits that are thought to be necessary for a woman to reach a leadership position.

However, it is unclear why male control leaders were rated as less agentic and weaker than female control leaders. A possible explanation might be, that as soon as there are managerial responsibilities involved, as it is the case in leadership positions, interpersonal (e.g. communal) skills become more important for men and are presumed to be necessary to reach the leadership position. While women had to be agentic, less communal and more dominant to reach a leadership position and might be perceived extreme on this traits due to a contrast effect, men who are inherently agentic do not have to prove their agency anymore but rather their interpersonal skills for a leadership position. We return to the implications of this pattern in the general discussion.

General Discussion

Regarding our initial research question whether gender-incongruent individuals are perceived as violating gender prescriptions or proscriptions, we found mixed results. While our first experiment clearly spoke for a prescription (i.e., communality and agency) deficit explanation, for both men and women, our second experiment could not fully replicate those findings for leadership positions. Furthermore, based on cross-experimental comparisons it seems that women are still penalized to an extent when applying for leadership positions – especially when no individuating information was provided. In the absence of individuating information, female leaders were perceived as both lacking gender-prescriptive and as exhibiting gender-proscriptive traits compared to female non-leaders, while male leaders were surprisingly perceived as lacking gender-prescriptive traits compared to male non-leaders.

Fortunately, in no condition of the experiments did we observe any backlash against gender-incongruent job applicants – regardless of their gender and the position they applied for. Finally, we found that gendered job descriptions did not affect the perception of applicants in

meaningful ways. Instead, framing a job as being at the top of a company's hierarchy had implications for the perception of applicants, especially female applicants.

In contrast to many previous studies (for a review, see Rudman & Phelan, 2008) females were actually perceived as more competent than their male counterparts, even though male-gendered jobs were still perceived as more competent than female-gendered jobs. Even though surprising, the higher competence ratings are not without base, considering that for example in the U.S women have been the majority of college-educated adults since the 1980s and that in 2019 they earned about 57% of bachelor's degrees (U.S. Bureau of Labor Statistics, 2020). This development in higher education seems to finally be reflected in the perception of men and women, as recent meta-analytical results of nationally representative U.S. public opinion polls on gender stereotypes from 1946 to 2018 in fact also indicate increased gender equality in perceived competence (Eagly et al., 2019). This meta-analysis further indicated an increase in women's relative advantage in perceived communality and a concurrent decrease in men's relative advantage in perceived agency over the years (Eagly et al., 2019).

Similarly, there were no penalties for gender-incongruency in any of the present studies, which ran counter to previous work using similar manipulations and measures (e.g., Heilman & Okimoto, 2007; Heilman & Wallen, 2010; Rudman, 1998; Rudman & Fairchild, 2004; Rudman & Glick, 1999; 2001), but in line with a more recent meta-analysis of experimental simulations of employment decisions, that could also not find a backlash-effect (Koch, D'Mello, & Sackett, 2015).

In regards to gender-incongruent women we did not only find no backlash effect, but females and communal targets were actually perceived as more likeable, more respectable, less interpersonally hostile and more hireable. As nowadays, a more inclusive and democratic leadership style is more common and valued (Eagly, Johannesen-Schmidt, & van Engen, 2003; Judge & Piccolo, 2004) and the importance of interpersonal skills seems to be growing, communal targets and women, who are automatically perceived as more communal, might be

more respected, more likeable, and therefore more hireable. These findings are in line with other recent reports showing a preference for women in academic STEM positions, typically thought of as male-gendered fields (Williams & Ceci, 2015) and a “feminization of management” (Rudman & Glick, 1999, 2001; Steffens, Schult, & Ebert, 2009) as well as earlier work on the “women are wonderful” effect (Eagly & Mladinic, 1989; Eagly, Mladinic, & Otto, 1991), which might also affect evaluations in favor for female targets. The relative preference for women and communal targets could also stem from a “stand-out effect” from expectations based on descriptive stereotypes. A following subtyping (Wade & Brewer, 2006; Weber & Crocker, 1983) of these remarkable individuals who are both highly competent and communal might lead to these positive evaluations. But further research is needed to assess this directly.

Regarding gender-incongruent men we did not only find no backlash effect, but gender-incongruent male leaders were also not perceived as less agentic or weaker than control male leaders, even when they applied for gender-incongruent jobs. This might be due to the still highly competent and prestigious positions they applied for. Even though some of the advocated jobs were in a female domain, they still were on a leadership level, which might still have made the applicants gender-congruent, as leader stereotypes are masculine (Koenig, Eagly, Mitchell, & Ristikari, 2011). Previous research found different results for men applying for female-gendered jobs that are perceived lower in status such as elementary school teachers (Moss-Racusin & Johnson, 2016) and so did we in the first experiment with non-leadership positions, where gender-incongruent men were perceived as violating gender prescriptions. Nonetheless, this could also be a positive sign for gender-incongruent men, indicating that they do not have to fear backlash for communal behavior. This would be of great practical importance, as men’s involvement in communal roles still has not increased as much as women’s involvement in agentic roles, even though it would be a crucial step to true gender equality and beneficial for society, men themselves as well as their families (Croft et al., 2015).

An advantage of the present study compared to previous research and another possible explanation for the absence of backlash was that participants were no college students, but mostly working professionals with an average work experience of 15 years. As samples of many previous studies being composed mostly of students from introductory psychology classes might have limited external validity (Henry, 2008), previous research clearly called for such representative workplace samples (Heilman & Okimoto, 2007). In addition to being more representative than the samples of previous studies, the sample size for the present experiments was also increased (often twelvefold compared to similar previous studies). The nature of the sample might partially account for the missing backlash as working professionals are expected to have more relevant experiences with the value of interpersonal skills in job contexts and thus might rate communal targets and women more positively than college students.

However, even though in the present experiments women and communal targets were rated more positively, in real work settings women still often seem to struggle in male-dominated areas, are still on average earning less than men (OECD, 2020a), are still under-represented in boardrooms across all industries (OECD, 2020b), hold less leadership positions in general and even if they shatter the glass ceiling, more often than men land in precarious leadership positions – on the so called glass cliff (Ryan & Haslam, 2007). Hence, future research should keep investigating the processes of bias towards gender-incongruent individuals in order to support women as well as men who want to succeed in gender-incongruent ways.

Also, it is important to note that the present studies did not entail real hiring decisions. Even though our approach allowed us to test the role of prescriptions and proscriptions in a controlled, experimental manner, the degree to which conclusions can be drawn about how professional recruiters and colleagues react to gender-incongruent individuals in actual work situations is limited. Future research should therefore further explore gender prescriptions and proscriptions in field settings considering the complexity of real organizational settings.

Another limitation of the current studies is the failing of the job type manipulation, as in contrast to previous research using this manipulation (e.g., Heilman & Wallen, 2010) most dependent measures did not yield differences for the different job types, even though the manipulation checks clearly revealed that the positions were perceived as gendered. On these grounds, additional research is encouraged to examine in which cases positions are gendered and under which circumstances they trigger backlash. Moreover, since participants were not asked what they suspected the study to be about, it cannot completely be ruled out that the participants were aware of the study's intention and overcorrected their answers in a socially desirable way. Future research should include an awareness check to investigate this.

In conclusion, for non-leadership positions violating gender-rules seems to affect only the perception of gender prescriptions, but not proscriptions. For women in leadership positions, however, gender prescriptions as well as proscriptions seem to be affected. Provided that high competence is unambiguous, agency and communality both are perceived as positive traits, albeit communality more strongly so. Being aware of these biases is important for recruiters as well as gender-incongruent job applicants. Understanding these biases might help to prevent potential backlash and biased decision making. Eventually, female surgeons and male nurses might be nothing perplexing anymore and the riddle that was posed at the beginning of this article will no longer be a brainteaser.

References

- Banaji, M. R., & Greenwald, A. G. (2013). *Blind spot: Hidden biases of good people*. New York: Delacorte Press.
- Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, 7(3), 136–141. doi:10.1111/j.1467-9280.1996.tb00346.x
- Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, 70(6), 1142–1163. doi:10.1037/0022-3514.70.6.1142
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law*, 5(3), 665–692. doi:10.1037/1076-8971.5.3.665
- Carli, L. (2001). Gender and social influence. *Journal of Social Issues*, 57(4), 725–741. doi:10.1111/0022-4537.00238
- Cikara, M., & Fiske, S. T. (2009). Warmth, competence, and ambivalent sexism: Vertical assault and collateral damage. In M. Barreto, M. K. Ryan, & M. T. Schmitt (Eds.), *The glass ceiling in the 21st century: Understanding barriers to gender equality* (pp. 73–96). Washington, DC: American Psychological Association. doi:10.1037/11863-004
- Croft, A., Schmader, T., & Block, K. (2015). An Underexamined Inequality: Cultural and Psychological Barriers to Men's Engagement With Communal Roles. *Personality and Social Psychology Review*, 19(4), 343–370. doi:10.1177/1088868314564789
- Cuddy, A. J. C., Glick, P., & Beninger, A. (2011). The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior*, 31, 73–98. doi:10.1016/j.riob.2011.10.004
- Damer, E., & Bradley, P. (2014). Prolific Academic [Computer Software]. Available at: <https://www.prolific.ac/>.
- 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., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, 129(4), 569–591. <https://doi.org/10.1037/0033-2909.129.4.569>
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. doi:10.1037/0033-295X.109.3.573
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111(1), 3–22.
- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, 15(4), 543–558.
- Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men? An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly*, 15(2), 203–216. doi:10.1111/j.1471-6402.1991.tb00792.x
- Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2019, July 18). Gender Stereotypes Have Changed: A Cross-Temporal Meta-Analysis of U.S. Public Opinion Polls From 1946 to 2018. *American Psychologist*. Advance online publication. <http://dx.doi.org/10.1037/amp0000494>
- Eckes, T. (2010). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung: Theorie, Methoden, Empirie* (pp. 178–189). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92041-2_21
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- 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(6), 878–902. doi:10.1037/0022-3514.82.6.878
- Gaunt, R. (2013). Ambivalent sexism and perceptions of men and women who violate gendered family roles. *Community, Work & Family*, 16(4), 401–416. doi:10.1080/13668803.2013.779231
- Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70(3), 491–512. doi:10.1037/0022-3514.70.3.491
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.
- Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, 50(1), 1–22. doi:10.1080/00273171.2014.962683
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57(4), 657–674. doi:10.1111/0022-4537.00234
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113–135. doi:10.1016/j.riob.2012.11.003
- Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *The Journal of Applied Psychology*, 92(1), 81–92. doi:10.1037/0021-9010.92.1.81
- 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(4), 664–667. doi:10.1016/j.jesp.2010.01.008
- 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(3), 416–427. doi:10.1037/0021-9010.89.3.416
- Henry, P. (2008). College sophomores in the laboratory redux: Influences of a narrow data base on social psychology's view of the nature of prejudice. *Psychological Inquiry*, 19, 49–71. <http://dx.doi.org/10.1080/10478400802049936>
- Hu, P., Adler, N. E., Goldman, N., Weinstein, M., & Seeman, T. E. (2005). Relationship between subjective social status and measures of health in older Taiwanese persons. *Journal of the American Geriatrics Society*, 53(3), 483–488.
- Joshi, A., Son, J., & Roh, H. (2015). When can women close the gap? A meta-analytic test of sex differences and rewards. *Academy of Management Journal*, 58(5), 1516–1545. <http://dx.doi.org/10.5465/amj.2013.0721>
- Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89, 755–768. doi:10.1037/0021-9010.89.5.755
- Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, 100(1), 128–161. <http://dx.doi.org/10.1037/a0036734>
- Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, 137(4), 616–642. <https://doi.org/10.1037/a0023557>
- Kunda, Z., & Thagard, P., (1996). Forming impressions from stereotypes, traits, and behaviors : A parallel-constraint-satisfaction theory. *Psychological Review*, 103(2), 284–308.
- Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*, 44(7), 701–710. doi:10.1002/ejsp.2023
- Moss-Racusin, C. A., & Johnson, E. R. (2016). Backlash against male elementary educators. *Journal of Applied Social Psychology*, 46(7), 379–393. doi:10.1111/jasp.12366
- Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules:

- Status incongruity and backlash against modest men. *Psychology of Men & Masculinity*, 11(2), 140–151. doi:10.1037/a0018093
- OECD (2020a). *Gender wage gap (indicator)*. doi: 10.1787/7cee77aa-en (Accessed on 29 June 2020)
- OECD (2020b). *Gender Equality – What big data can tell us about women on boards*. <https://www.oecd.org/gender/data/what-big-data-can-tell-us-about-women-on-boards.htm>
- Prentice, D., & 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(4), 269–281. doi:10.1111/1471-6402.t01-1-00066
- 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(3), 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(2), 157–176.
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, 77(5), 1004–1010. doi:10.1037/0022-3514.77.5.1004
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *The Journal of Social Issues*, 57(4), 743–762. doi:10.1111/0022-4537.00239
- Rudman, L. A., & Glick, P. (2010). *The social psychology of gender: How power and intimacy shape gender relations*. New York: The Guilford Press.
- Rudman, L. A., & Mescher, K. (2013). Penalizing men who request a family leave: Is flexibility stigma a femininity stigma? *Journal of Social Issues*, 69(2), 322–340.

doi:10.1111/josi.12017

- Rudman, L. A., Moss-Racusin, C. A., Glick, P., & Phelan, J. E. (2012a). Reactions to vanguards. Advances in backlash theory. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology*, 45, (pp. 167–227). Burlington: Academic Press. doi:10.1016/B978-0-12-394286-9.00004-4
- Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012b). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48(1), 165–179. doi:10.1016/j.jesp.2011.10.008
- Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, 28, 61–79. doi:10.1016/j.riob.2008.04.003
- Ryan, M. K., & Haslam, S. A. (2007). The Glass Cliff: Exploring the Dynamics Surrounding the Appointment of Women to Precarious Leadership Positions. *Academy of Management Review*, 32(2), 549–572. <https://doi.org/10.5465/amr.2007.24351856>
- Steffens, M. C., Schult, J. C., & Ebert, I. D. (2009). Feminization of management leads to backlash against agentic applicants: Lack of social skills, not gender, determines low hireability judgments in a student sample. *Psychology Science Quarterly*, 51(1), 16–46.
- U.S. Bureau of Labor Statistics (2020). *Labor Force Statistics from the Current Population Surve*. Retrieved from <https://data.bls.gov/PDQWeb/ln>
- Wade, M. L., & Brewer, M. B. (2006). The structure of female subgroups: An exploration of ambivalent stereotypes. *Sex Roles*, 54(11–12), 753–765. doi:10.1007/s11199-006-9043-x
- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45(5), 961–977. doi:10.1037/0022-3514.45.5.961
- Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty

preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, 112(17),5360-5365. doi:10.1073/pnas.1418878112