



**University of
Zurich^{UZH}**

Department of Business Administration

UZH Business Working Paper Series

Working Paper No. 375

**Fix the Game–Not the Dame: A Team Intervention to Restore
Gender Equity in Leadership Evaluations**

Jamie L. Gloor, Manuela C. Morf, Samantha Paustian-Underdahl and
Uschi Backes-Gellner

January 2017

Published as: "Fix the Game, Not the Dame: Restoring Equity in Leadership Evaluations." *Journal of Business Ethics* (2018). By Jamie L. Gloor, Manuela Morf, Samantha Paustian-Underdahl and Uschi Backes-Gellner.
DOI: <https://doi.org/10.1007/s10551-018-3861-y>

University of Zurich, Plattenstrasse 14, CH-8053 Zurich,
<http://www.business.uzh.ch/forschung/wps.html>

UZH Business Working Paper Series

Contact Details

Jamie L. Gloor

Technical University of Munich

School of Management

Manuela Morf

University of Lucerne

Center for Human Resource Management

Samantha Paustian-Underdahl

Florida International University

College of Business

Uschi Backes-Gellner

University of Zurich

Department of Business Administration

Fix the Game—Not the Dame:**A Team Intervention to Restore Gender Equity in Leadership Evaluations**

Jamie L. Gloor, PhD¹
Technical University of Munich

Manuela C. Morf, PhD
Erasmus University Rotterdam

Samantha Paustian-Underdahl, PhD
Florida International University

Uschi Backes-Gellner, PhD
University of Zurich

The leadership literature shows consistent, sizeable, and persistent effects indicating that female leaders face significant biases in the workplace compared with male leaders. However, the social identity leadership literature suggests these biases might be overcome at the team level by adjusting the number of women in the team. Building on this work, we conducted 2 multiple source, multiple wave, multi-level randomized field experiments to test if the gender composition of teams helps to restore equity in leadership evaluations of men and women. Across two samples of university students engaged in a team-building exercise, we find that male leaders are rated as more prototypical leaders than female leaders despite no differences in leaders' self-reported prototypicality; however, this male leadership advantage is eliminated in gender-balanced teams. In Study 2, we extend this finding by supporting a moderated mediation model showing that leader gender and the team's gender composition interact to relate to perceived trust in the leader, through the mediating mechanism of leader prototypicality. Findings support the social identity model of organizational leadership and indicate a boundary condition of role congruity theory, bolstering our need for a more social relational, context-based approach to leadership.

Keywords: gender; leadership; teams; social identity; prototypicality

¹Corresponding author: Jamie.Gloor@gmail.com; Working Paper Version January 2017.

**Team Design with (Female) Leaders in Mind:
Restoring Equity in Leadership Evaluations**

Despite women's considerable progress entering the managerial ranks in recent decades and making up at least half of the workforce and higher education degree earners (Perry, 2013), women remain a small minority at the highest levels (i.e., the "glass ceiling;" Catalyst, 2015; Morrison et al., 1994). One explanation for this discrepancy focuses on the widespread societal stereotypes and prototypes of leaders, which largely associate men with leadership more than women (Eagly & Karau, 2002; Lord & Hall, 2003). Indeed, decades of research show that men are typically seen as more suited for and more effective in leadership roles than women (e.g., Eagly, Makhijani, & Klonsky, 1992; Schein, 1973, 2007).

While gender stereotypes—widely-held persistent and pervasive oversimplified images or ideas of a particular type of person or thing (Schein, Mueller, Lituchy, & Liu, 1996)—can be extremely difficult to change (Moss-Racusin et al., 2014), leader prototypes are more malleable and are largely influenced by contextual norms. Prototypicality is a "set of characteristics possessed by most category members" (Cronshaw & Lord, 1987: 97) and may be benchmarked according to the leader (i.e., attributes that characterize "leaders," such as gender) or the group (i.e., attributes that characterize the follower group). Since leaders' demographic characteristics are generally immutable, yet group prototypes are easily alterable with timely feedback (e.g., Giessner et al., 2009; van Knippenberg, 2011), group prototypes serve as an appropriate mechanism to potentially increase gender equity amongst leaders. Indeed, leader prototypicality is a key determinant of leadership effectiveness (Hogg, 2001; van Knippenberg, 2011; van Knippenberg & Hogg, 2003).

Guided by role congruity theory (Eagly & Karau, 2002), and the social identity model of organizational leadership (van Knippenberg & Hogg, 2003), we argue that the gender composition of the group (i.e., the percentage of women in the team) may weaken or even

override leader prototypes to improve followers' responses to female leaders. According to the social identity theory of leadership, as a leader is perceived to be group prototypical – that is, to embody the group (team, organizational) identity or “who we are” – the leader builds influence and legitimacy from group members who believe the leader represents what is group-normative (Hogg, 2001; van Knippenberg & Hogg, 2003). Role congruity theory (Eagly & Karau, 2002) suggests that in general, prejudice toward female leaders follows from the perceived incongruity between the characteristics of women and the requirements of leader roles. Eagly and Karau (2002) also propose that prejudice toward female leaders can vary depending on features of the leadership context as well as characteristics of leaders' evaluators. We integrate the social identity model of organizational leadership with role congruity theory to examine how team gender composition may serve as a contextual moderator affecting how leader gender relates to team perceptions of leader prototypicality and leader trustworthiness. We also conduct several robustness checks in an attempt to rule out possible alternative explanations including leader self-perceptions of prototypicality, as well as follower gender-match with the leader (see Tsui & O'Reilly, 1989).

A strength of our intervention is that we aim to improve gender equity in leadership evaluations by adjusting team composition, as opposed to asking leaders to change their behaviors in the workplace. Indeed, research shows that women leaders often face backlash, or penalties at work, for adopting masculine behaviors often seen as congruent with leadership prototypes, but incongruent with their gender role (Brescoll, 2012; Brescoll & Uhlmann, 2008; Rudman, 1998; Rudman, Moss-Racusin, Phelan, & Nauts, 2012). We propose that female leaders can be seen as more prototypical—or rather as prototypical as male leaders—without cost, training, or backlash, through a contextual change at the team-level. We test this proposition with two randomized, multiple wave, and multiple-source field experiments with teams of new business, economics, and informatics students.

In the following, we present the relevant literature pertaining to gender and leadership with a focus on the context within which leadership is enacted: teams. We define and describe leader and group prototypicality, summarizing the relevant empirical literature and supporting the key proposition we aim to test. Finally, we outline several theoretical and practical implications, including the team gender composition's relevance for managers, organizational design, and female leaders, as well as ideas for future research.

HYPOTHESIS DEVELOPMENT

Leadership as a Group Process

Leadership is defined by its context because a leader cannot exist without followers, and teams represent a contextual element of leadership with potent implications for followers' responses to leadership via group identity. Indeed, leadership is a process enacted within group settings, and as such, a key factor not to be overlooked is that "leaders not only lead groups of people, but are also themselves members of these groups" (van Knippenberg & Hogg, 2003: 244). Thus, shared group membership has implications for leaders as well as for followers and groups.

People are quick to categorize the self and others into groups, which are cognitively represented by prototypes (Turner, 1985; Turner et al., 1987). Derived from cognitive psychology (Rosch, 1978), prototypes are "fuzzy sets of attributes that define and prescribe attitudes, feelings, and behaviors that characterize one group and distinguish it from other groups" (Hogg, 2001: 187). Comparable to stereotypes, prototypes serve as mental heuristics that are retrieved in relevant situations to guide perception, self-conception, and eventual action (Cronshaw & Lord, 1987; Hogg, 2001). However, prototypes also comprise a contextual element, which allows them to be responsive to specific social contexts (Hogg, Fielding, Johnson, Masser, Russell, & Svensson, 2006; Hogg, Hains, & Mason, 1998). For example, a liberal, democratic leader may be viewed as prototypical of her team in

metropolitan New York, but less prototypical in conservative, rural South Carolina. Indeed, group prototypes are inherently context-based according to theory (Hogg, 2001; van Knippenberg & Hogg, 2003), so that as group composition changes, group and leader prototypes also evolve accordingly. In other words, the context determines what kind of prototype is used for benchmarking as well as shaping how this prototype is characterized (i.e., what constitutes a typical “leader”), with direct implications for leadership effectiveness (see van Knippenberg, 2011).

Relating this concept to gender, Hogg and colleagues (2006) propose that group prototypes may be gendered and thus inform responses to male and female leaders. In an experiment, they compared perceptions of male and female leaders as a function of whether group norms emphasized stereotypically masculine or stereotypically feminine qualities. They argued that the match between the gendered group norm and leader gender would influence the degree to which male or female leaders were seen as more group prototypical, thus influencing their perceived effectiveness. Consistent with this notion, they found that gendered group norms render leader group prototypicality contingent on leader gender.

Leader Gender, Team Gender Composition, & Leadership Evaluations

We build on this work to examine the effects of the gender composition of teams on perceptions of leadership prototypicality for male and female leaders. Role congruity theory is based in social role theory (Eagly, 1987), which explains that historical distributions of men and women into breadwinner and homemaker roles (respectively) have produced societal gender norms as well as actual differences in behavior. Women and men are expected to have attitudes and skills congruent with their traditional roles, which create stereotypes that foster gendered responses to leadership and leadership selection (Eagly & Karau, 2002). Meta-analytic results bolster this assertion, indicating men are perceived as more prototypical leaders and are evaluated more favorably than women (Eagly et al., 1992).

Yet, more recently, a meta-analysis found that gendered responses to leadership vary according to certain contextual moderators including the percent of male raters evaluating the leader (Paustian-Underdahl, Walker, & Woehr, 2014).

This fits with the social identity model of organizational leadership (van Knippenberg & Hogg, 2003), which argues that the group context within which leadership is enacted influences followers' responses to leadership beyond individual leader's characteristics (e.g., physical or objective characteristics such as sex as well as subjective characteristics such as attitudes; Giessner et al., 2013, Study 1; Hains et al., 1997; Monzani et al., 2014; van Knippenberg & van Knippenberg, 2005). In other words, leaders are more effective in mobilizing and influencing followers as the leader's identity more closely reflects that of the team or group (Hogg, 2001). This theory is based on social identity and social influence theories (e.g., Ashforth & Mael, 1989; Hogg & Abrams, 1988; Tajfel & Turner, 1986), which posit that groups are a critical source of social influence and information used for prototype benchmarking. A wealth of evidence supports this proposition, as research examining leaders' group prototypicality has reported consistent positive effects on leadership effectiveness since its earliest tests dating back to the 1990s (see van Knippenberg, 2011).

In summary, our key theoretical proposition is derived from dovetailing propositions from role congruity (Eagly & Karau, 2002) and social identity theories (Hogg, 2001; van Knippenberg & Hogg, 2003): the negative consequences for women's leadership ratings are more pronounced when leadership occurs in numerically male-dominated teams. This proposition is also supported by multiple meta-analyses (e.g., Eagly et al., 1995; Eagly et al., 1992; Paustian-Underdahl et al., 2014), and research on group composition and tokenism.

According to tokenism theory (Kanter, 1977), as the percentage of males increases within a particular setting, women leaders' feminine or female-stereotypical qualities become increasingly salient, which also increase their perceived lack of fit as leaders. Through a case

study of 20 saleswomen in a 300-person sales force at a multinational, Fortune 500 corporation, Kanter (1977) found that tokenism has several consequences for minority group members within the workplace including higher visibility (and increased scrutiny), exaggeration of differences from majority group members, exclusion from informal workplace interactions, and assimilation (i.e., tokens are forced into stereotypical categories defined by the majority group members). In other words, tokens may not be viewed for who they are, and instead, are often simplified as symbols of their category with caricatured roles. The negative consequences of token status for women have been replicated across a variety of professional settings, whereas token status often results in null or positive effects for men (Budig, 2002; Williams, 1992).

Given the negative consequences of tokenism within groups that have not yet been examined across levels, and because leadership roles are currently and historically male-skewed (e.g., Catalyst, 2015; Eagly & Karau, 2002; State Secretary for Economic Affairs, 2012), leader gender and its effects on leadership outcomes are especially pronounced in male-dominated settings, making it a prime point for potential change. Thus, we propose that the immediate team gender context may reduce or even override the more general, societal leader prototypes to improve followers' responses to female leaders as the team gender composition transforms from male majority to gender-balanced. In other words, a woman leader should be perceived as more representative of the group with increasing shares of women in the team, and hence will be seen as a more prototypical leader.

Hypothesis 1: Leader gender will interact with team gender composition such that the team will rate female leaders as less prototypical than male leaders in male majority teams, and this effect will be weaker in gender-balanced teams.

Leader Prototypicality and Trust

A core element within leader group prototypicality research is that leader prototypicality is an important driver of leadership and group effectiveness, because group

prototypical leaders are trusted to pursue the group's best interest (Giessner & van Knippenberg, 2008; Giessner, van Knippenberg, & Sleebos, 2009; van Knippenberg & Hogg, 2003; van Knippenberg, 2011). This also means that leaders are granted more leeway in their actions. For example, leadership ratings and effectiveness become less contingent on leaders' behaviors and more contingent on their perceived prototypicality (Giessner, van Knippenberg, & Sleebos, 2009). Indeed, van Knippenberg and van Knippenberg (2005) found that leader's prototypicality of the collective moderates the effects of leader self-sacrifice on leadership effectiveness. They argued that this would occur because leader prototypicality raises trust in the leader and should therefore render leadership effectiveness less contingent on the display of group-oriented behavior like leader self-sacrifice. Yet, the authors did not actually measure trust in the leader as an outcome of leader prototypicality. Thus, we make a similar proposition while also testing the extent to which team perceptions of leader prototypicality relate to team perceptions of leader trustworthiness.

Hypothesis 2: Team ratings of leader prototypicality are positively associated with team ratings of trust in the leader.

Altogether, we propose that team gender composition will interact with leader gender to effect ratings of leader trust through the explanatory mechanism of leader prototypicality. Indeed, teams made up of more men are likely to see a male leader as more prototypical of the team, and thus, are more trustworthy. However, as teams become more gender-balanced, this male advantage should dissipate, bringing about more gender equity in leadership evaluations (see Figure 1).

Hypothesis 3: Leader prototypicality ratings mediate the interactive effect of leader gender and team gender composition on leader trust, such that the indirect effect of leader gender on leader trust via leader prototypicality will be negative for female leaders in male-dominated teams, an effect that should be weaker in more gender balanced teams.

Insert Figure 1 about here

Alternative Explanations and Robustness Checks

Although there is ample evidence to guide our predictions for followers' responses to leaders per leader gender and group gender composition, it is unclear if only the followers' perceptions of leaders change, or if the leaders' conceptions of themselves also change (i.e., leaders' self-perceived similarity to the leader prototype). Leaders conceptions could change based on the gender composition of the teams that they lead, because this could alter leaders' objective prototypicality of the group. In turn, leaders' group prototypicality is strongly and positively associated with leader prototypicality and leadership effectiveness. This argumentation builds on principles from the social identity theory of leadership (Hogg, 2011; van Knippenberg & Hogg, 2003), such that leaders' self-perceived prototypicality and degree of team identification also predicts leaders' team-oriented attitudes and behavior.

However, the vast majority of research on leader prototypicality has examined followers' perceptions of leaders' prototypicality (see van Knippenberg, 2011). To our knowledge, only one paper (Giessner, van Knippenberg, van Ginkel, & Sleebos, 2013) and two conference presentations (Giessner & van Knippenberg, 2007; 2009) provide preliminary evidence that more group-prototypical leaders exhibit more group-oriented behaviors (i.e., behaviors that serve the interests of the team and its members). Importantly, when leaders successfully achieve group-oriented behaviors, it further increases and reinforces followers' ratings of their leaders' prototypicality (Giessner et al., 2009).

Given this paucity of research paired with the only recently emerging research on leader's self-definitions of leadership or its role in shaping leader actions (van Knippenberg, 2011), we include a rather exploratory assessment of leaders' self-reported leader prototypicality in addition to assessing followers' ratings of leaders' prototypicality. To assess leaders' baseline perceptions of their leadership prototypicality which could be tied to their gender (Eagly & Karau, 2002), as well as their post-leadership perceptions which could

be affected by their group prototypicality (i.e., team gender compositions) as described above, we measure leaders' ratings of their prototypicality both before and after they lead their teams. In addition to informing the paucity of literature on the topic, this exploratory analysis provides insight as to whether our effects might be partially attributable to leaders' own conceptions of themselves as leaders, which could alter their leadership behavior.

However, of note, this analysis chiefly serves as a robustness check given that we hypothesize that the changes in followers' responses to leaders are primarily driven by group-based prototype processes within the team, as we previously argued.

Finally, given our knowledge of relational demography theory and findings (e.g., Tsui & O'Reilly, 1989), it could also be the dyadic gender match that improves followers' ratings of leaders in more gender balanced teams rather than the leader being more representative of the group gender composition more broadly. That is, with more women in the team, there are more gender matched pairs, which could provide an alternative explanation for our findings. Thus, we include additional analyses of the dyadic gender match as an additional robustness check to rule out these potential alternative explanations. As with the previous research question, this analysis chiefly serves as a robustness check given that we hypothesize changes in followers' responses to leaders are driven by group-based prototype processes within the team, not between individual followers and leaders.

STUDY 1

METHODS

Sample and Procedure

We conducted a randomized field experiment among teams of students from economics, business, and informatics at a large university in Western Europe. Followers were incoming first year students and leaders were more experienced senior students. Followers

were students of a freshmen orientation event starting on the first day of the semester and lasting for the duration of the semester.

Leaders applied for our course and received academic credit upon completion. A total of 32 leaders were selected using systematic criteria (e.g., previous academic performance) and then trained for two days. Leaders were not trained in a specific leadership style; instead, they received a general theoretical overview of leadership and group organization. This training was meant to help leaders manage their own teams more effectively during the orientation event. Each leader had up to 30 students in the freshmen group they lead.

Data were collected via pencil and paper surveys provided in person. Surveys were administered in the participants' native language of German with items forward- and back-translated from English. Surveys were completed after team members spent approximately 6 hours with their teams and leaders during the orientation event. Leaders were responsible for designing activities for their groups including study tips and strategies for academic success, as well as physical and social orientation to the university campus. Leaders also organized subsequent events for their teams with an academic and a social focus. Thus, this was the first, but certainly not the last event during which the leaders and teams would interact.

Measures

We used a multi-study and multi-source approach. All perceptual measures had 6-point response scales (1 = *does not apply* at all to 6 = *totally & completely applies*). Leader and follower gender (male = 0, female = 1) were collected via self-report from leaders and followers (respectively), although it is arguably an objective measure. We manipulated leaders' objective group prototypicality via team gender composition. Followers' perceptions of leader prototypicality were measured via survey, with additional filler items to disguise our study's purpose, as well as standard orientation day evaluation items (e.g., satisfaction with the amount of information received and the organization of the event) for the dean's office.

Leader prototypicality. We assessed followers' ($N = 426$) perceptions of their leader's embodiment of a prototypical leader with 3 items (Cronshaw & Lord, 1987). These 3 items have also been used to assess perceptions of leadership in a Swiss setting with high reliability ($\alpha = .92$; Antonakis, Fenley, & Liechti, 2011). Items included: *the leader is a typical leader*, *exhibits the behavior of a leader*, and *fits one's image of a leader* ($\alpha = .88$).

Team gender composition. We randomly determined team gender composition as male majority (20% women) or gender-balanced (40-50% women). Our manipulations were strategically chosen to mirror current gender compositions (i.e., 20% of women in leadership; State Secretary for Economic Affairs, 2012) and approximately equal group gender compositions. Furthermore, team gender compositions of 20% to 50% are also feasible in modern workplaces given that women have composed at least half of college degree earners for several decades (Perry, 2013). Finally, these team gender compositions also echo *skewed* (20% women) and *balanced* (40-60% women) designations from critical mass theory (Kanter, 1977). Due to constraints regarding the gender of participants available in our overall sample, while we tried to create precisely male majority (20% women) or gender-balanced (40-50% women) teams, this was not always possible; thus, leaders also reported the actual number of men and women followers in their teams, which we used to measure team gender composition in our empirical analyses.

Control variables. According to relational demography theory (Tsui, Egan, & O'Reilly, 1992; Tsui & O'Reilly, 1989), women may prefer female leaders (and vice versa for men). Thus, we control for follower gender (0 = male, 1 = female). Theoretically, group identity may also be more cohesive in smaller groups, while more practically, followers may interact more with their teammates and leaders in smaller groups. Thus, we also control for group size.

By design, we essentially control for two potentially meaningful facets of group diversity, namely, tenure and area of study (Jackson, Joshi, & Erhardt, 2003; van Knippenberg & Schippers, 2007). Specifically, all followers were new to the group and the organization. Also, random assignment was constrained by academic discipline such that management and economics students were nested within the same groups and information technology students were nested within the same groups. Thus, tenure and area of study are kept constant by design, while other potentially influential facets of group diversity (e.g., age) also do not differ across groups given randomized follower assignment to groups.

RESULTS

Analyses were conducted with Mplus (version 7.4), R (version 0.99.491), and STATA (version 14.1). Predictors were centered prior to analyses. Unstandardized coefficients are reported.

 Insert Table 1 about here

Descriptive Statistics

The descriptive statistics and correlations are displayed in Table 1. From a total of 512 followers, 12 were eliminated due to missing data, and 3 teams were eliminated due to our randomization being compromised ($n = 74$),² for a final 426 participants (38.5% women). Team gender was randomly assigned as male majority (20%) or balanced (40-50%). However, as is common with field experiments, there was slight variation in the actual proportion of women in each group (e.g., no-shows or newcomers who had not signed up for the event); team share of women ranged from 20% to 63.64% ($M = 37.22$, $SD = 12.31$).

To ensure the validity of our manipulations, we conducted a series of manipulation checks. First, leader reports showed balanced teams comprised a greater share of women ($M = 46.81$, $SD = 8.00$) than the male majority teams ($M = 26.87$, $SD = 6.28$), Cohen's $d = 2.77$;

²These 3 teams were also significantly larger than the other 32 teams, $F(1,505) = 1640.08$, $p < .001$, $\eta^2_p = .77$, providing additional justification for excluding them from Study 1.

mirroring aggregated follower reports. Second, we also included a measure of followers' own ratings of their perceived group gender composition, for which followers rated their team composition from 1 (*all men*) to 5 (*all women*). Analysis of this item similarly indicated that followers noticed their group gender composition and scored it in line with our manipulation, for balanced ($M = 2.94$, $SD = .35$) and majority male teams ($M = 2.22$, $SD = .55$), $t(431) = 16.256$, $p < .001$, Cohen's $d = 1.58$. These checks reflect very large effects, indicating that our manipulations remain sound.

Multi-Level Analyses

Preliminary analyses. The data represent 426 participants nested within 32 teams. Given that our hypotheses reflect leader gender and team composition affecting group-level perceptions of the leader, an intraclass correlation coefficient (ICC1) was computed to show the percent of the total variance in the dependent variable that is between groups (ICC1 = .06). While not extremely high, this value does suggest that there is meaningful variance in group perceptions of leader prototypicality (Bryk & Raudenbush, 1992). Additionally, we examined whether aggregation was justified in our data by calculating r^*_{wg} values. Examining r^*_{wg} values suggested that the level of agreement was above the typical cutoff of .70 to support aggregation for leader prototypicality (average $r^*_{wg} = .89$). The ICC(2) for leader prototypicality (.46) also suggested reliable group means, supporting aggregation (Castro, 2002; LeBreton & Senter, 2008). Model comparisons of the null model with a model allowing for random intercepts indicate a significant difference, $\chi^2(1, N = 426) = 5.86$, $p < .01$). Thus, there is significant intercept variation according to group.

Leader prototypicality. The linear mixed effects model indicates that male leaders are not rated as more prototypical than female leaders, $b = -0.018$, $p = .876$ (see Table 2). There is also no main effect of team gender composition on leader prototypicality, $b = 0.009$, $p = .059$. However, as expected, these main and null effects are qualified by an interaction

between leader gender and team gender composition ($b = 0.021, p < .05$; see Figure 2). These results support Hypothesis 1.

Insert Table 2 & Figure 2 about here

Simple slopes analyses indicate that as expected, female leaders are more negatively rated than male leaders in unbalanced teams (-1 $SD = 25%$ female followers; $b = -0.370, SE = 0.143, p < .01$). However there are no differences in prototypicality ratings for gender-balanced teams (+1 $SD = 49%$ female followers; $b = .021, SE = 0.141, p = .881$). On a descriptive level, the gendered difference in ratings of leader prototypicality in male majority teams (Cohen's $d = .86$) is also much larger than in gender-balanced teams (Cohen's $d = .03$), a large versus no effect. Thus, Hypothesis 1 is supported.

STUDY 2

In Study 2, we not only aim to replicate Study 1, but also to extend the model to predict another leadership outcome: trust. Data was collected in the same way and in the same context as in Study 1, only 1 year apart with a new cohort of students.

METHODS

Measures

Leader prototypicality. As in Study 1, we used the same 3-item measure ($\alpha = .89$).

Team gender composition. As in Study 1, we manipulated the proportion of women in the teams.

Trust in leader. We assessed followers' trust in their leaders with the 6-item scale developed by Gillespie and Mann (2004; $\alpha = .89$). Items assessed to what extent followers trust their leaders' skills, judgments, and values, and how willing they are to share their feelings and personal information with the leader (e.g., *To what extent do you trust your leader in regards to sharing your personal beliefs?*; $\alpha = .82$).

Control variables. As in Study 1, we included the same control variables.

RESULTS

Descriptive Statistics

From a total of 467 followers, 33 were eliminated due to missing data. The final sample consisted of 434 participants (31.33% women). Actual team share of women ranged from 0% to 63.00 %, with an average of 31.37% women ($SD = 16.15$).

Leaders of balanced teams reported more women in their teams ($M = 41.42$, $SD = 11.18$) than leaders of male majority teams ($M = 17.05$, $SD = 10.27$, Cohen's $d = 2.27$). Thus, our manipulation is valid.

In total were 54.27% of our leaders were women and approximately half of our groups were gender-balanced (60.00%). Finally, 33 of our 35 leaders (94.29%) returned completed surveys at both time points. Leaders were evenly distributed across leader gender ($n = 35$ or 51.43% women) and team gender conditions ($n = 9$ or 42.25% of female leaders in balanced gender teams). Further descriptives and correlations reported in Table 3.

 Insert Table 3 about here

Multi-Level Analyses

Preliminary analyses. Similarly to Study 1, the data represent 434 participants nested within 35 teams. An ICC1 was computed to show the percent of the total variance in leader prototypicality and trust in the leader that is between groups (ICC1 = .11 and .05, respectively). These values suggest that there is meaningful variance in group perceptions of leader prototypicality and trust in the leader (Bryk & Raudenbush, 1992). Additionally, we examined whether aggregation was justified in our data by calculating r^*_{wg} values. Examining r^*_{wg} values suggested that the level of agreement was above the typical cutoff of .70 to support aggregation for leader prototypicality and trust (average $r^*_{wg} = .90$ and .94, respectively). The ICC(2) for leader prototypicality (.62) and trust in leader (.40) also

suggested reliable group means, supporting aggregation (Castro, 2002; LeBreton & Senter, 2008). Regarding prototypicality, model comparisons of the null model with a model allowing for random intercepts indicate a significant difference, $\chi^2(1, N = 424) = 23.11, p < .001$). Also, the model allowing for random intercepts when predicting trust in leader significantly differs from the null model, $\chi^2(1, N = 424) = 6.64, p < .01$).

Leader prototypicality. As expected, we replicated a significant interaction effect of leader gender and team gender composition on prototypicality ($b = 0.014, p = .006$; Table 4).

 Insert Table 4 about here

Trust in leader. As an extension to Study 1, in Study 2 we tested the relationship between leader prototypicality and trust in leader. As expected, leader prototypicality is positively associated with trust in leader ($b = 0.643, p < .001$), supporting Hypothesis 2.

To further examine our proposed moderation-mediation effect, we followed the recommendations of Preacher and Hayes (2008) and estimated the indirect effects, conditional on our moderator (team gender composition). We examined the indirect effects of leader gender on leader trust via prototypicality at $-1 SD$, the mean, and $+1 SD$ on team gender composition. Results show that female leaders are seen as less trustworthy (via prototypicality) in more male dominated teams ($b = -.13, CI = [-.21, -.05]$), but this effect becomes non-significant as the percent of women in the teams increases.

Because our mediator, leader prototypicality, is not an exogenous variable in our model, we followed the recommendations of Antonakis, Bendahan, Jacquart and Lalive, (2010) and Podsakoff, MacKenzie, and Podsakoff (2012) to test if endogeneity may threaten the stability of our results. Specifically, we estimated an augmented regression (i.e., Durbin-Wu-Hausman tests; Davidson & MacKinnon, 2003), which indicates that endogeneity is not an issue in our data if the residuals for the first stage of the model are not significantly related

to our outcome variable of trust (Antonakis et al., 2010; Podsakoff et al., 2012). As we found no significant relationship ($p = .21$), endogeneity does not bias our results; we are confident that our estimates are stable and common method variance is not a threat to our findings.

Robustness Checks

To optimize statistical power, we combine data sets from Study 1 and Study 2 for the following analyses. These robustness checks are also examined at level 1 due to individual-level nature of the potential alternative explanations.

Leader-follower gender match. First, we found no significant effect of follower gender in predicting prototypicality or trust (see Tables 2 & 4). Second, dyadic similarity might be an alternate explanation for our results, because female followers may rate female leaders more highly (and vice versa for male followers and leaders). To test this proposition, we estimated a model that predicts prototypicality on the individual-level using the same variables as the multi-level models, but including a dichotomous variable representing gender match (1) or mismatch (0) between followers and leaders. This additional variable does not explain additional variance in our model ($p = .274$), nor does it change our overall patterns of results. Thus, our results are not explained by female followers' higher ratings (and by design, there are more female followers in gender-balanced teams), nor are they explained by female followers' higher ratings of leaders of the same gender (i.e., female leaders).

Leader ratings. Approximately half of our leaders were women (51.43%) and approximately half of our groups were gender-balanced (54.29%). Leaders and followers were randomly assigned to teams.³ In total, 65 of our 70 leaders (92.9%) returned completed surveys at both time points. Leaders were evenly distributed across leader gender ($n = 35$ or 53.8% women) and team gender conditions ($n = 33$ or 50.8% balanced gender teams).

³ The true purpose of our study was unknown to followers and leaders. Indeed, we found similar gender compositions of leaders and teams in previous years without intervention, so we have no reason to believe that participants were aware of our study purpose or our manipulations.

Finally, we proposed that it is followers' perceptions of leaders that change depending on leader' objective group prototypicality, not that leaders' behaviors change depending on their group prototypicality. Thus, we also assess leaders' self-ratings of prototypicality before and after the orientation event (i.e., before and after they lead their teams) with a series of mixed ANOVAs using within- (Time 1, Time 2) and between-subjects variables (leader gender: male or female; team gender composition: male majority or gender-balanced). We find no significant main effects or interactions apart from a significant increase in self-rated leader prototypicality from Time 1 ($M = 4.23$, $SD = 0.81$) to Time 2 ($M = 4.48$, $SD = 0.88$; $F(1, 61) = 7.327$, $p = .009$, $\eta^2_p = .107$), all other $ps = .36-.70$.

In summary, we have no evidence that male leaders rated themselves as more prototypical than female leaders or that leaders of male majority groups rated themselves as more prototypical than leaders of gender-balanced groups, either before or after leading their teams. Instead, it seems that our male and female leaders had similar conceptions of themselves as leaders, regardless of their own gender or their team's gender composition. Furthermore, leaders' self-rated leader prototypicality increased with leadership experience in a similar manner for male and female leaders as well as for leaders of male majority and gender-balanced teams.

DISCUSSION

We proposed that the team gender context is an ideal point of intervention to overcome the broader societal leader gender biases that persistently disadvantage female leaders; our research confirms this idea. Drawing from the social identity theory of organizational leadership (van Knippenberg & Hogg, 2003) and role congruity theory (Eagly & Karau, 2002) our findings show that the male advantage in leader perceptions is eliminated in gender-balanced teams. In doing so, we extended the social identity model of leadership theory by showing that leaders' own self-perceptions do not change according to their team's

gender composition, and we ruled out alternative explanations by providing evidence that our effects are not simply driven by female followers or by female followers rating female leaders more positively. Finally, to better isolate the effects of team gender composition, our team-level manipulation of interest, we also controlled for other potentially influential aspects of diversity by design (e.g., tenure, age, and educational background).

Theoretical Implications

We aimed to bridge classic work on gender and leadership (Eagly & Karau, 2002) with leadership and group prototypes research (Hogg, 2001; van Knippenberg & Hogg, 2003) to make several core contributions and outline specific areas for future research. First, the proposition that local group prototypes trump broader societal leadership prototypes has been theorized (Hogg, 2001; van Knippenberg, 2011) and is in-line with similar leadership theories (e.g., see Lord, Foti, & DeVader, 1984; Lord & Hall, 2003), but to our knowledge, has not yet been empirically tested to date. Thus, we provide empirical support for this proposition.

Second, leadership and group prototypes research has largely discussed this proposition pertaining to social identity in general. However, we chose to examine the specific identity of gender given that women in leadership is a topic of modern significance in research and the practice due to the persistent and pervasive glass ceiling (Catalyst, 2015; Morrison et al., 1994). Furthermore, there are extraordinary costs invested in leadership training programs despite a lack of evidence of transfer (Baldwin & Ford, 1988; Burke & Day, 1986; Burke & Hutchins, 2007) and backlash ensuing from women who display more ‘male’, masculine or agentic behaviors (Brescoll, 2012; Brescoll & Uhlmann, 2008; Rudman, 1998; Rudman et al., 2012). Thus, we developed a theoretically-based solution to improve the lack of female representation in leadership and break the glass ceiling (Morrison et al., 1994), but without backlash. Theoretically, there is also no detriment to male leaders incurred by an intervention such as ours, because even in gender-balanced groups, male leaders are similarly

prototypical of the group as female leaders (Hogg, 2001; van Knippenberg & Hogg, 2003); however, future research should examine team gender compositions beyond the mid-point (i.e., with more than 50% women) to more accurately examine this proposition.

Third, our study elucidates a new type of prototypicality benchmarking. To date, researchers have mostly manipulated leaders' perceived group prototypicality via fabricated feedback about leaders' values, styles, or beliefs (e.g., Giessner et al., 2013, Study 1; Haines et al., 1997; Monzani et al., 2014; van Knippenberg & van Knippenberg, 2005). By intervening in actual teams and manipulating leaders' objective group prototypicality via team gender composition, we showed converging effects using a new and perhaps more rigorous method. Previous research has also shown high reliability between actual versus perceived demographic differences in previous research (e.g., Cunningham, 2007), which we also showed in Study 1. However, future research could examine if an objective manipulation of leaders' group prototypicality in terms of other visible traits (e.g., race or ethnicity) similarly influence followers' perceptions of their leaders' group prototypicality.

Finally, previous studies of leader and group prototypicality examined individuals who were ostensibly in groups or anticipated group interaction, but they were tested quickly and completely alone (Haines, et al., 1997; Hogg et al., 1998; Hogg et al., 2006; Monzani et al., 2014), in virtual teams or with virtual leaders (Giessner et al., 2013, Study 1; van Knippenberg & van Knippenberg, 2005). This speaks to the power of prototypes; however, our findings are likely more generalizable given that our followers were nested in actual groups and interacted with real-life followers and leaders for several hours. Indeed, leadership is inherently a social process (Chemers, 2001). King and colleagues (2013) also endorse randomized field experiments such as ours as a gold standard for empirical reasons and external validity—especially when examining sensitive topics such as gender bias.

Practical Implications

Our findings also offer implications for practice, for example, in guiding team formation and leader assignments. Teams are becoming more gender diverse as increasingly more women enter traditionally male-dominated fields (Bureau of Labor Statistics, 2014). But despite their representation at lower levels, women remain a glaring minority in leadership positions (e.g., only 4.4% of CEOs are women; Catalyst, 2015). According to our findings, such demographic changes at the lower level may also benefit female leaders in ways that have been overlooked to date, but only if teams are designed with gender in mind.

In light of our findings and in line with previous research on tokenism and critical mass theory (e.g., Kanter, 1977; Konrad, Kramer, & Erkut, 2008; Sekaquaptewa & Thompson, 2003), it seems the tipping point for equity in responses to leaders occurs between approximately 20-40% women in the team. These results also dovetail with meta-analytic evidence that women have fewer leadership disadvantages in settings with more balanced organizational gender demography (Eagly, Karau, & Makhijani, 1995; Paustian-Underdahl et al., 2014). However, one should not misinterpret these findings as implying regression since Kanter's seminal research on critical mass theory 40 years ago. This stream of research shows a tipping point of team demographic composition with more than 20% women in a team on intragroup interaction processes (Kanter, 1977) or with a 'magic number' of 3 or more women in a board and on organizational performance or innovation (Joecks, Pull, & Vetter, 2013; Konrad et al., 2008; Torchia, Calabro, & Huse, 2011). Yet there is an important distinction between this research and the current study: critical mass research examines the effect of group gender composition—typically women's demographic representation on executive boards or in political parties—on intragroup processes or organizational outcomes (e.g., performance or innovation; Joecks et al., 2013; Torchia et al., 2011). However, in the current study, we extend this research to show that group gender composition not only

influences team-level outcomes, but group gender composition can also extend *across levels* to influence team-level perceptions of and responses to leaders via prototypes.

Yet, there is likely also a critical point where team gender composition fails to provide marginal returns, and might even be detrimental for female leaders. As previously discussed, women are valued less than men, especially in masculine-typed or male-dominated positions. Although this has not yet been examined by social identity leadership theorists, there is relevant evidence from social psychology that suggests a stigma-by-association effect for female leaders who lead majority female teams (Pryor, Reeder, & Monroe, 2012). Field experimental evidence also supports this proposition, such that stigma towards individual team members (as well as teams as a whole) increased with the proportion of women in the team (West, Heilman, Gullett, Moss-Racusin, & Magee, 2012). Thus, future research should seek to delineate the boundary conditions of team gender composition's positive effects for female leaders and extend this research from intragroup to intergroup perceptions. In a similar vein, it would also be fruitful for future research to examine responses to male leaders in traditionally female occupations or jobs (e.g., nursing); However, these particular instances do not contribute to the larger patterns of social and economic inequality like responses to female leaders in traditionally male occupations and jobs.

In the case that teams are already established or must be constructed based on non-gender-based criteria (e.g., employee education or expertise), then practitioners can also use our findings to inform their interpretations of leader evaluations. For example, a woman from a majority male team may provide similarly negative performance feedback about a woman supervisor as her male teammates. This effect would not only be unexpected according to relational demography perspectives (Tsui et al., 1992; Tsui & O'Reilly, 1989), but it may be interpreted in line with the problematization of female same-sex interactions in organizations (see Sheppard & Aquino, 2014) and hurt the case for increasing women in the workforce.

Thus, the potential influence of the team gender context on evaluations such as performance reviews or 360-degree feedback should not be overlooked, and it can be easily assessed by including a single item indicating the gender demography of the leader in question's team.

For female leaders, it also seems plausible that more female representation at the team level may buffer competent female leaders from strategic rejection (see Parks-Stamm, Heilman & Hearn, 2008). This would occur if female leaders are perceived as more prototypical and exemplary of their group members in more gender balanced teams, rather than as competition or threats.

Strengths, Limitations, and Future Research

Our study is methodologically rigorous, including a replication and extension. Specifically, we avoid the threat of common method variance (Podsakoff, MacKenzie, Lee & Podsakoff, 2003) or rather measurement error (Spector, 2006) by using data collected from different sources (e.g., followers and leaders), including objective data (e.g., team member gender, team share of women, and leader gender) collected at different times (e.g., before and after the orientation event). Because we intervened and manipulated team gender composition, randomly assigning leaders and followers to teams, and showed that endogeneity is not a threat to the second-stage of our model, we can make a causal claim based on our findings (see Antonakis, Bendahan, Jacquart, & Lalive, 2010). That is, leader gender predicts leader prototypicality and trust, depending on team gender composition. We also provide evidence against multiple alternate explanations (e.g., leaders' behaviors change or increasing shares of female followers or female follower-leader dyads drive this effect).

However, as with any study, our research has its limitations. For example, leaders had no evaluative or disciplinary influence on followers, with the exception that they could dismiss individuals from the event (as needed). Such an arrangement may be more representative of modern, flatter hierarchies (e.g., project managers or peer leadership).

However, such arrangements are increasingly common in today's more interdependent organizations (Rajan & Wulf, 2006; Wegman, Hoffman, Carter, Twenge, & Guenole, 2016). Despite this, we recommend that future research extend our findings by testing these hypotheses within organizational teams to better understand the extent to which gender team composition and leader gender relate to leader evaluations in ongoing workplace teams.

In addition, we created teams with low or balanced shares of women by design. Although this allows us to maintain generalizability to typical work groups, we were unable to draw conclusions about groups that were all male or all female. Field studies in real organizations would help address this concern as many organizations have teams made up of a large variety of gender compositions but they are not randomly assigned, so they would have the disadvantage of not solving the endogeneity problem.

Finally, our research was conducted in a university in Western Europe. Thus, our conclusions are bound by the cultural context within which we have undertaken our research. Yet, considerable evidence indicates group prototypicality (van Knippenberg, 2011) as well as leader and gender prototypes (Koenig et al., 2011; Schein, 2001) are generally consistent across countries and management contexts (Hernandez Bark, Escartin, & van Dick, 2014).

Conclusions

Our results highlight the potential benefits of recent demographic changes and increasing numbers of women even at the lower level for female leaders and potentially organizations—but only if teams are designed with female leaders in mind. Indeed, if the leadership game is rigged in favor of men, women face a double-bind of backlash regardless of their ability or performance. However, there is hope of restoring gender equality in leadership if we fix the game—not the dame.

REFERENCES

- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2010. On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21:1086-1120.
- Antonakis, J., Fenley, M., & Liechti, S. 2011. Can charisma be taught? Tests of two interventions. *Academy of Management: Learning & Education*, 10: 374-396.
- Ashforth, B. E., & Mael, F. 1989. Social identity theory and the organization. *Academy of Management Review*, 14: 20-39.
- Baldwin, T. T., & Ford, J. K. 1988. Transfer of training: A review and directions for future research. *Personnel Psychology*, 41: 63-105.
- Brescoll, V. L. 2012. Who takes the floor and why: Gender, power, and volubility in organizations. *Administrative Science Quarterly*, 56: 622-641.
- Brescoll, V. L., & Uhlmann, E. L. 2008. Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace. *Psychological Science*, 19: 268-275.
- Budig, M. J. 2002. Male advantage and the gender composition of jobs: Who rides the glass escalator? *Social Problems*, 49: 258-277.
- Bureau of Labor Statistics. 2014. Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity. *Current Population Survey*. Retrieved from: www.bls.gov.
- Burke, M. J., & Day, R. R. 1986. A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology*, 71: 232-245.
- Burke, L. A., & Hutchins, H. M 2007. Training transfer: An integrative literature review. *Human Resource Development Review*, 6: 263-296.
- Bryk, A.S. & Raudenbush, S.W. 1992. *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Catalyst. 2015. Women CEOs of the S&P 500. Retrieved from www.catalyst.org
- Castro, S. L. 2002. Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlations, $r_{wg}(j)$, hierarchical linear modeling, within- and between-analysis, and random group resampling. *The Leadership Quarterly*, 13: 69-93.
- Chemers, M. M. 2001. Leadership effectiveness: An integrative review. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 376-399). Oxford, England: Blackwell.
- Cronshaw, S. F., & Lord, R. G. 1987. Effects of categorization, attribution, and encoding processes on leadership perceptions. *Journal of Applied Psychology*, 72: 97-106.
- Cunningham, G. B. 2007. Perceptions as reality: The influence of actual and perceived demographic dissimilarity. *Journal of Business and Psychology*, 22: 79-89.
- Davidson, R., & MacKinnon, J. G. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.
- Eagly, A. H. 1987. *Sex differences in social behavior: A social role interpretation*. Hillsdale, NJ: Earlbaum.
- Eagly, A. H., & Karau, S. J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Bulletin*, 109: 573-598.
- Eagly, A. H., Karau, S. J., & Makhijani, M. G. 1995. Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, 117, 125-145.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. 1992. Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111: 3-22.
- Giessner, S. R., & van Knippenberg, D. 2007, April. Leading FOR the team: Situational determinants of team-oriented leader behavior. Paper presented at the 2007 Annual Meeting of the *Society of Industrial and Organizational Psychology*, New York.

- Giessner, S. R., & van Knippenberg, D. 2009, June. When does a leader show fair behavior? Influences of group prototypicality and the social context. Paper presented at the *Erasmus Leadership Conference*, Rotterdam, The Netherlands.
- Giessner, S. R., van Knippenberg, D., van Ginkel, W., & Sleebos, E. 2013. Team-oriented leadership: The interactive effects of leader group prototypicality, accountability, and team identification. *Journal of Applied Psychology*, 98: 658-667.
- Giessner, S. R., van Knippenberg, D., & Sleebos, E. 2009. License to fail? How leader group prototypicality moderates the effects of leader performance on perceptions of leadership effectiveness. *The Leadership Quarterly*, 20: 434-451.
- Gillespie, N. A., & Mann, L. 2004. Transformational leadership and shared values: The building blocks of trust. *Journal of Managerial Psychology*, 19: 588-607.
- Hains, S. C., Hogg, M. A., & Duck, J. M. 1997. Self-categorization and leadership: Effects of group prototypicality and leader stereotypicality. *Personality and Social Psychology Bulletin*, 23: 383-399.
- Hernandez Bark, A. S., Escartin, J., & van Dick, R. 2014. Gender and leadership in Spain: A systematic review of some key aspects. *Sex Roles*, 70: 522-537.
- Hogg, M. A. 2001. A social identity theory of leadership. *Personality and Social Psychology Review*, 5: 184-200.
- Hogg, M. A., & Abrams, D. 1988. *Social identifications: A social psychology of intergroup relations and group processes*. London, England: Routledge.
- Hogg, M. A., Fielding, K. S., Johnson, D., Masser, B., Russell, E., & Svensson, A. 2006. Demographic category membership and leadership in small groups: A social identity analysis. *The Leadership Quarterly*, 17: 335-350.
- Hogg, M. A., Hains, S. C., & Mason, I. 1998. Identification and leadership in small groups: Salience, frame of reference, and leader stereotypicality effects on leader evaluations. *Journal of Personality and Social Psychology*, 75: 1248-1263.
- Ito, T. A. & Urland, G. R. 2003. Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality & Social Psychology*, 85: 616-626.
- Joecks, J., Pull, K., & Vetter, K. 2013. Gender diversity in the boardroom and firm performance: What exactly constitutes a “critical mass”? *Journal of Business Ethics*, 118: 61-72.
- Kanter, R. M. 1977. Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, 82: 965-990.
- King, E. B., Hebl, M. R., Botsford Morgan, W., & Ahmad, A. S. 2013. Field experiments on sensitive organizational topics. *Organizational Research Methods*, 16: 501-521.
- 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: 616-642.
- Konrad, A. M., Kramer, V., & Erkut, S. 2008. The impact of three or more women on corporate boards. *Organizational Dynamics*, 37: 145-164.
- LeBreton, J. M., & Senter, J. L. 2008. Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11: 815-852.
- Lord, R. G., Foti, R. J., & DeVader, C. L. 1984. A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, 34: 343-378.
- Lord, R., & Hall, R. 2003. Identity, leadership categorization, and leadership schema. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity Processes in Groups and Organizations*, 48-64. London, England: Sage.

- Macrae, M. C., Milne, A. B., & Bodenhausen, G. V. 1994. Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, 66: 37-47.
- Monzani, L., Hernandez Bark, A. S., van Dick, R., & Peiro, J. M. 2014. The synergistic effect of prototypicality and authenticity in the relation between leaders' biological gender and their organizational identification. *Journal of Business Ethics*, 132: 737-752.
- Morrison, A. M., White, R. P., VanVelsor, E. & the Center for Creative Leadership. 1994. *Breaking the glass ceiling: Can women reach the top of America's largest corporations?* Reading, MA: Addison-Wesley.
- Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. 2014. Scientific diversity interventions. *Science*, 343: 615-616.
- 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.
- Paustian-Underdahl, S. C., Walker, L. S., & Woehr, D. J. 2014. Gender and perceptions of leadership effectiveness: A meta-analysis of contextual mediators. *Journal of Applied Psychology*, 99: 1129-1145.
- Perry, M. J. 2013. Stunning college degree gap: Women have earned almost 10 million more college degrees than men since 1982. *American Enterprise Institute*. Retrieved from www.aei.org.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. & Podsakoff, N. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88: 879-903.
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. 2012. Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63: 539-569.
- Preacher, K. J., & Hayes, A. F. 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40: 879-891.
- Pryor, J. B., Reeder, G. D., & Monroe, A. E. 2012. The infection of bad company: Stigma by association. *Journal of Personality and Social Psychology*, 102: 224-241.
- Rajan, R. G., & Wulf, J. 2006. The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies. *The Review of Economics and Statistics*, 88: 759-773.
- Rosch, E. 1978. Principles of categorization. In E. Rosch, & B. B. Lloyd (Eds.), *Cognition and Categorization*, 27-48. Hillsdale, NJ: Erlbaum.
- 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., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. 2012. Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48: 165-179.
- Schein, V. E. 1973. The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, 57: 95-100.
- Schein, V. E. 2001. A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, 57: 675-688.
- Schein, V. E., Mueller, R., Lituchy, T., & Liu, J. 1996. Think manager-think male: a global phenomenon? *Journal of Organizational Behavior*, 17: 33-41.

- Sekaquaptewa, D. & Thompson, M. 2003. Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, 39: 68-74.
- Sheppard, L. D. & Aquino, K. 2014. Sisters at arms: A theory of female same-sex conflict and its problematization in organizations. *Journal of Management*.
- Spector, P. E. 2006. Method variance in organizational research: Truth or urban legend? *Organizational Research Methods*, 9: 221-232.
- Spybrook, J., Bloom, H, Condon, R., Hill, C., Martinez, A., Raudenbush, S. 2011. Optimal Design Plus Empirical Evidence: Documentation for the "Optimal Design"(Version 3.01) [Software]. Available from <http://hlmsoft.net/od/>.
- State Secretary for Economic Affairs. 2012. European comparison: High flexibility and autonomy of Swiss employees. Retrieved from www.seco.admin.ch
- Tajfel, H., & Turner, J. C. 1986. The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of Intergroup Relations* (2nd ed.: 7-24).
- Torchia, M., Calabro, A., & Huse, M. 2011. Women directors on corporate boards: From tokenism to critical mass. *Journal of Business Ethics*, 102: 299-317.
- Turner, J. C. 1985. Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research*, Vol. 2: 77-122. Greenwich, CT: JAI.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. 1987. *Rediscovering the social group: A self-categorization theory*. Oxford, England: Blackwell.
- Tsui, A. S., Egan, T. D., & O'Reilly, C. A. 1992. Being different: Relational demography and organizational attachment. *Administrative Science Quarterly*, 37: 549-579.
- Tsui, A. S. & O'Reilly, C. A. 1989. Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. *Academy of Management Journal*, 32: 402-423.
- van Knippenberg, D. 2011. Embodying who we are: Leader group prototypicality and leadership effectiveness. *The Leadership Quarterly*, 22: 1078-1091.
- van Knippenberg, B., & van Knippenberg, D. 2005. Leader self-sacrifice and leadership effectiveness: The moderating role of leader prototypicality. *Journal of Applied Psychology*, 90: 25-37.
- van Knippenberg, D., van Knippenberg, B., De Cremer, D. & Hogg, M. A. 2004. Leadership, self, and identity: A review and research agenda. *The Leadership Quarterly*, 15: 825-856.
- van Knippenberg, D. & Hogg, M. A. 2003. A social identity model of leadership effectiveness in organizations. In B. M. Staw & R. M. Kramer (Eds.), *Research in Organizational Behavior*, 25: 243-295.
- van Knippenberg, D., & Schippers, M. C. 2007. Work group diversity. *Annual Review of Psychology*, 58: 515-541.
- Wegman, L. A., Hoffman, B. J., Carter, N. T., Twenge, J. M., & Guenole, N. 2016. Placing job characteristics in context: Cross-temporal meta-analysis of changes in job characteristics since 1975. *Journal of Management*.
- West, T. V., Heilman, M. E., Gullett, L., Moss-Racusin, C. A., & Magee, J. C. 2012. Building blocks of bias: Gender composition predicts male and female group members' evaluations of each other and the group. *Journal of Experimental Social Psychology*, 48: 1209-1212.
- Williams, C. L. 1992. The glass escalator: Hidden advantages for men in the "female" professions. *Social Problems*, 39: 253-267.

TABLE 1
Study 1 Descriptives, Correlations, and Scale Reliability

Variables	<i>M</i>	<i>SD</i>	1	2	3	4
<i>Group level (2nd level)</i>						
1. Leader Gender	.52	.50				
2. Team Gender Composition	37.22	12.31	.10*			
3. Team Size	14.38	2.12	-.15**	-.35***		
4. Leader Prototypicality	4.97	0.28	-.20***	.29***	-.13**	(.88)
<i>Individual level (1st level)</i>						
5. Follower Gender	.38	.49	.03	.25***	-.07	.07

Note. $N = 426$, nested within 32 teams.

* $p < .05$

** $p < .01$

*** $p < .001$

TABLE 2
Study 1 Linear Mixed-Effects Models

Variable	Leader Prototypicality					
	Main Effects			Interaction Effect		
	<i>b</i>	<i>t</i>	<i>SE</i>	<i>b</i>	<i>t</i>	<i>SE</i>
<i>1st level</i>						
Follower Gender	0.000	-0.275	0.000	0.000	-0.285	0.000
<i>2nd level</i>						
Team Size	0.032	1.180	0.026	0.050	1.820	0.027
Leader Gender	-0.018	-2.329	0.111	-0.009	-0.085	0.109
Team Gender Composition	0.009	1.969	0.004	0.009*	2.108	0.004
Team Gender X Leader Gender				0.021*	2.373	0.009
Leader Prototypicality						
Intercept	4.896***	89.896	0.054	4.884***	90.809	0.054
<i>Log-Likelihood</i>	12770.65			12769.65		

Note. $N = 426$, nested within 32 teams.

* $p < .05$

** $p < .01$

*** $p < .001$

TABLE 3
Study 2 Descriptives, Correlations, and Scale Reliability

Variables	<i>M</i>	<i>SD</i>	1	2	3	4	5
<i>Group level (2nd level)</i>							
1. Leader Gender	.52	.50					
2. Team Gender Composition	31.37	16.15	-.20***				
3. Team Size	15.00	4.13	.17**	-.10*			
4. Leader Prototypicality	4.76	0.36	-.29***	.22***	-.73***	(.89)	
5. Trust in Leader	5.63	0.25	-.35***	.29***	-.76***	.67***	(.81)
<i>Individual level (1st level)</i>							
6. Follower Gender	.31	.46	-.10*	.31***	-.06	.10*	.14**

Note. $N = 434$, nested within 35 teams.

* $p < .05$

** $p < .01$

*** $p < .001$

TABLE 4
Study 2 Linear Mixed-Effects Models

Variable	Main Effects			Leader Prototypicality Interaction Effect			Trust in Leader		
	<i>b</i>	<i>z</i>	<i>SE</i>	<i>b</i>	<i>z</i>	<i>SE</i>	<i>b</i>	<i>z</i>	<i>SE</i>
<i>1st level</i>									
Follower Gender	1.217	0.389	1.013	1.809	0.563	1.028			
<i>2nd level</i>									
Team Size	-0.057***	-0.165***	0.010	-0.056***	-0.156***	0.009			
Leader Gender	-0.118	-0.171	0.083	-0.102	-0.142	0.076			
Team Gender Composition	-0.003	-0.150	0.004	-0.004	-0.172	0.004			
Team Gender X Leader Gender				0.014**	0.291**	0.005			
Leader Prototypicality							0.412***	0.642***	0.129
Intercept	4.374***	12.669***	0.322	4.213***	11.804***	0.321	3.682***	16.085***	0.418
<i>R</i> ²	.627**			.829***			.412**		
<i>Log-Likelihood</i>	496.146			613.717			905.353		

Note. *N* = 434, nested within 35 teams.

* *p* < .05

** *p* < .01

*** *p* < .001

FIGURE 1

Complete Theoretical Model

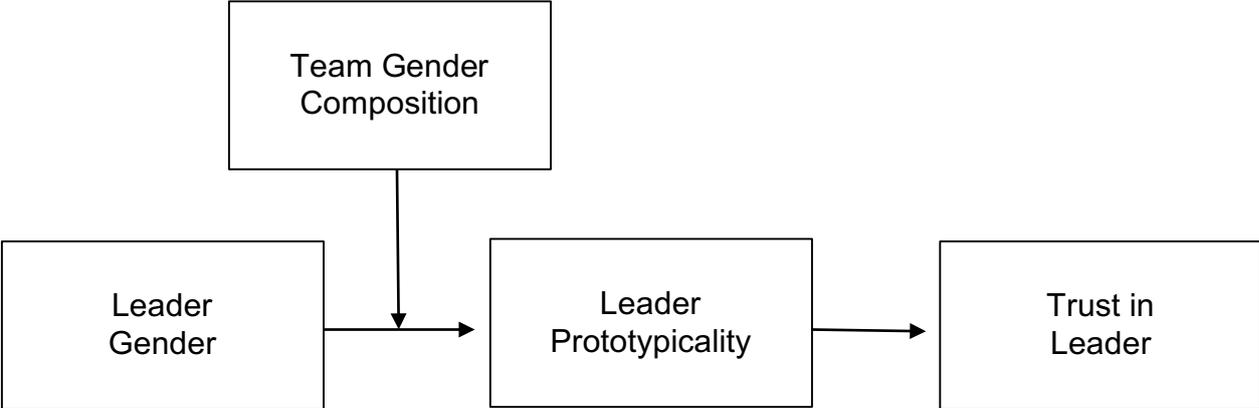


FIGURE 2

First-Stage Interaction Plot (Study 1)

