Working Paper No. 140

Fix the Game, Not the Dame: Restoring Equity in Leadership Evaluations

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

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 140

Fix the Game, Not the Dame: Restoring Equity in Leadership Evaluations

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

January 2017

This paper was previously circulated under the title "Fix the Game – Not the Dame: A Team Intervention to Restore Gender Equity in Leadership Evaluations" (2017).

Published as: "Fix the Game, Not the Dame: Restoring Equity in Leadership Evaluations." *Journal of Business Ethics, 161*(2020): 497-511. By Jamie L. Gloor, Manuela Morf, Samantha Paustian-Underdahl and Uschi Backes-Gellner.
DOI: https://doi.org/10.1007/s10551-018-3861-y

Fix the Game, Not the Dame: Restoring Equity in Leadership Evaluations

Jamie L. Gloor [1]

Manuela Morf [3]

Samantha Paustian-Underdahl [4]

Uschi Backes-Gellner [2]

Abstract

Female leaders continue to face bias in the workplace compared to male leaders. When employees are evaluated differently because of who they are rather than how they perform, an ethical dilemma arises for leaders and organizations. Thus, bridging role congruity and social identity leadership theories, we propose that gender biases in leadership evaluations can be overcome by manipulating diversity at the team level. Across two multiple-source, multiple-wave, and randomized field experiments, we test whether team gender composition restores gender equity in leadership evaluations. In Study 1, we find that male leaders are rated as more prototypical in male-dominated groups, an advantage that is eliminated in gender-balanced groups. In Study 2, we replicate and extend this finding by showing that leader gender and team gender composition interact to predict trust in the leader via perceptions of leader prototypicality. The results show causal support for the social identity model of organizational leadership and a boundary condition of role congruity theory. Beyond moral arguments of fairness, our findings also show how, in the case of gender, team diversity can create a more level playing field for leaders. Finally, we outline the implications of our results for leaders, organizations, business ethics, and society.

Keywords: Gender, Prototypicality, Trust

1 School of Management, Technical University of Munich, Munich, Germany, Jamie.Gloor@gmail.com

2 Business Administration, University of Zurich, Zurich, Switzerland, Manuela.Morf@unilu.ch

3 Center for Human Resource Management, University of Lucerne, Lucerne, Switzerland, scpaustian@gmail.com

4 College of Business, Florida State University, Tallahassee, USA

Introduction

Despite decades of women's progress in earning college degrees and becoming managers (Perry 2013), women remain a small minority in the highest levels of leadership (i.e., the "glass ceiling"; Catalyst 2018; Morrison et al. 1994). Widespread gender stereotypes and leader prototypes that more closely associate men with leadership can drive differential evaluations of male and female leaders, perpetuating a gender gap in leadership (Eagly and Karau 2002; Eagly et al. 1992; Lord and Hall 2003; Schein 1973, 2001). In response, some women may adopt more masculine or agentic behaviors (e.g., being more voluble; Brescoll 2011) that are congruent with leader prototypes but incongruent with their gender role, leading to negative backlash (Brescoll 2011; Brescoll and Uhlmann 2008; Rudman 1998; Rudman et al. 2012). Organizations may also implement policies and programs ostensibly intended to support and retain female employees, yet these initiatives can counterintuitively create additional barriers for potential female leaders (e.g., gender quotas and maternal leave benefits can highlight women's gender roles and stereotypes; Gloor et al. 2018; Heilman et al. 1997). For these reasons, business ethics scholars consider the persistent and pervasive gender gap in leadership a pressing issue in need of innovative and effective solutions (e.g., Hernandez Bark et al. 2016; O'Neil et al. 2008; Oakley 2000).

We propose a team-level intervention that aims to "fix the game, not the dame" by changing group-level leader prototypes via manipulation of the group gender composition as opposed to women's behaviors. Guided by role congruity theory (Eagly and Karau 2002) and the social identity model of organizational leadership (Hogg 2001; van Knippenberg and Hogg 2003), we propose that the gender composition of the group (i.e., the percentage of women in the team) may weaken or even override team leader prototypes to improve followers' responses to female leaders. Gender stereotypes, defined as widely held persistent and pervasive oversimplified images or ideas of a particular type of person or group (Schein et al. 1996), can be extremely difficult to change (Moss-Racusin et al. 2014). However, group prototypes are more malleable than stereotypes and are largely influenced by contextual norms. Prototypicality is a "set of characteristics possessed by most category members" (Cronshaw and Lord 1987, p. 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, whereas group prototypes are alterable and relevant for leadership effectiveness (e.g., Giessner et al. 2009; also see van Knippenberg 2011), we target team-level prototypes for intervention.

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 that affects the way leader gender relates to team perceptions of leader prototypicality and trustworthiness. According to the social identity theory of leadership, a leader is perceived to be more group prototypical (i.e., embodying the group identity as a team or organization) as he or she builds influence and legitimacy from group members who believe the leader represents what is group-normative (Hogg 2001; van Knippenberg and Hogg 2003). A leader may be viewed as more group-normative to the extent to which he or she shares demographic characteristics with team members. Role congruity theory (Eagly and Karau 2002) suggests that, in general, prejudice toward female leaders results from the perceived incongruity between the characteristics of women and the requirements of leadership roles. This prejudice toward female leaders can vary depending on features of the leadership context and the characteristics of leaders' evaluators (Eagly and Karau 2002). Therefore, prejudice against female leaders may weaken as the gender diversity of the group increases.

We manipulate gender composition at the team level in two randomized, multiple-wave, and multiple-source field experiments[1] to examine whether the male advantage in perceptions of leader prototypicality and trustworthiness can be eliminated in more gender-balanced teams. Our key contributions are threefold. First, we bridge classic work on gender and leadership (Eagly and Karau 2002) with leadership and group prototype research (Hogg 2001; van Knippenberg and Hogg 2003). The idea that group prototypes may override broader societal stereotypes has been theorized (van Knippenberg 2011) and described in the context of organizational compositions or professional stereotypes (Eagly and Karau 2002; Perry et al. 1994). However, to our knowledge, this proposition has not yet been experimentally tested in the field or examined at the team level.

Second, individual and organizational efforts to improve gender equity may focus on females, which can inadvertently create additional barriers to their success (e.g., Brescoll 2011; Brescoll and Uhlmann 2008; Gloor et al. 2018; Heilman et al. 1997; Rudman 1998; Rudman et al. 2012). Thus, typical interventions to improve gender equity may be misdirected or even morally questionable. We aim to improve gender equity in leadership evaluations by adjusting team compositions as opposed to changing female leaders' behaviors in the workplace or creating new company-wide hiring or family policies for female employees.

Finally, ethics and social justice perspectives have been the prevailing arguments for employees' equal employment and career opportunities regardless of their demographic background (e.g., Dwertmann et al. 2016; Eagly and Carli 2007). However, our research goes a step further to causally illustrate why and how this might be true and potentially self-reinforcing in the case of gender. That is, having more gender-diverse teams may be an admirable goal in its own right, but more diverse teams may also create a fairer playing field for female leaders.

In the following, we present relevant literature on gender and leadership with a focus on the context in which leadership is enacted: teams. We define and describe group prototypicality, leadership prototypes, and trust, summarizing the relevant research and the key propositions we aim to test. Finally, we outline several theoretical and practical implications, including the relevance of the team gender composition for leaders, organizations, and business ethics more broadly as well as several specific ideas for future research.

Leadership as a Group Process

Leadership is defined by its context because a leader cannot exist without followers. Given the highly embedded nature of leadership, it is important to remember that leaders not only lead the group, but are also members of the groups they lead (e.g., the President of the USA is also an American; van Knippenberg 2011). This shared identity shapes followers' responses to and evaluations of their leaders, which is a process driven by prototypes.

People are quick to categorize themselves 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, p. 187). Comparable to stereotypes, prototypes serve as mental heuristics that are retrieved in relevant situations to guide perception, self-conception, and eventual action (Cronshaw and Lord 1987; Hogg 2001). However, prototypes also comprise a contextual element that allows them to be responsive to specific social

[1] In accordance with King et al. (2013), we define a field experiment as "a method that uses random assignment to implement a manipulation relevant to working adult participants engaging in genuine tasks or with genuine outcomes in natural settings" (p. 502).

contexts (Hogg et al. 1998, 2006). For example, a liberal, democratic leader may be viewed as more prototypical of her constituents in metropolitan New York, but less prototypical in rural South Carolina.

According to theory, group prototypes are inherently context based (Hogg 2001; van Knippenberg and Hogg 2003); as group composition changes, group prototypes evolve accordingly. In other words, the context shapes what constitutes a typical group member and thus who would be a typical "leader" of this group. These benchmarking processes have direct implications for leadership evaluations because perceptions of group representativeness are strong and positive predictors of the degree to which leaders can influence followers (van Knippenberg 2011). We build on this work to examine the effects of team gender composition on leader prototypes (i.e., mental representations of what constitutes "good" leadership) for male and female leaders. That is, we manipulate leaders' objective group prototypicality in terms of gender to causally examine its effect on team perceptions of how "leader-like" the leader is.

Leader Gender, Team Gender Composition, and Leadership Evaluations

Role congruity theory is based on 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 creates stereotypes that foster gendered responses to leadership (Eagly and Karau 2002). Meta-analytic results bolster this assertion, indicating that men are perceived as more closely fitting stereotypical leadership prototypes and thus are evaluated more favorably than women (Eagly et al. 1992). However, more recently, a meta-analysis has found that gendered responses to leadership vary according to certain contextual moderators, including the percentage of male raters evaluating the leader (Paustian-Underdahl et al. 2014). Indeed, women are viewed as equally effective leaders as men in gender-diverse groups of followers (Paustian-Underdahl et al. 2014).

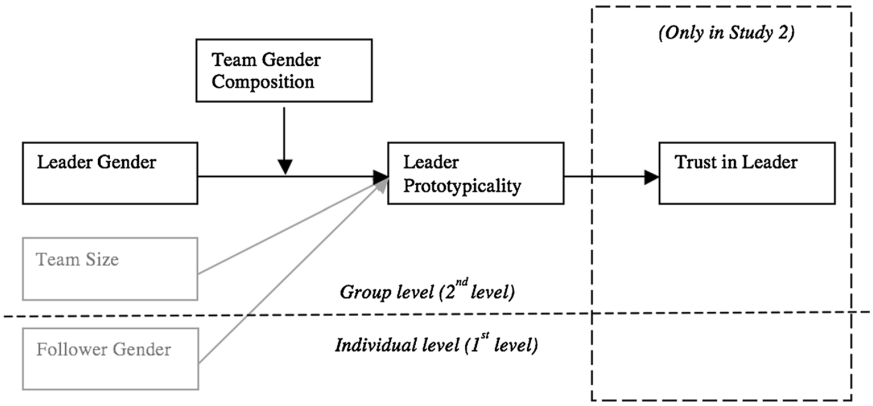
In line with this work, the social identity model of organizational leadership (van Knippenberg and Hogg 2003) argues that leaders' group context (i.e., the team) can influence followers' responses to leadership beyond leader characteristics. In other words, leaders are more effective in mobilizing and influencing followers when the leader's identity more closely reflects that of the team or group (Hogg 2001). Most research has examined subjective characteristics of identity (e.g., attitudes; Giessner et al. 2013, Study 1; Hais et al. 1997; Monzani et al. 2014; van Knippenberg and van Knippenberg 2005); however, initial findings also support this effect for leaders' objective characteristics (e.g., sex). For example, Hogg et al. (2006) compared perceptions of male and female leaders as a function of whether group norms emphasized stereotypically masculine or feminine qualities. They found that gendered group norms relate to leaders' group prototypicality contingent upon leader gender. Social identity and social influence theories (e.g., Ashforth and Mael 1989; Hogg and Abrams 1988; Tajfel and Turner 1986) similarly argue that groups are a critical source of social influence and information.

We build on this work by dovetailing propositions from role congruity (Eagly and Karau 2002) and social identity theories (Hogg 2001; van Knippenberg and Hogg 2003) to propose that when women lead male-dominated groups, they may be viewed as less prototypical leaders compared to male leaders and female leaders of more gender-balanced groups. When female leaders are more representative of their followers, as they would be in gender-balanced teams, equity in evaluations of male and female leaders should be restored. Given the extent to which many organizations and leadership roles continue to be male-dominated (Catalyst 2018; Eagly and Karau 2002; Swiss Federal Statistical Office 2014), leading to negative consequences for female leaders (see Heilman 2012), we propose that team gender composition could be a prime point of intervention. Specifically, team gender composition may reduce or even override the more general, societal leader stereotypes to improve follower responses to female leaders as

the team gender composition shifts from male-dominated to more gender-balanced. In other words, male leaders are likely viewed as more prototypical than female leaders are in male-dominated groups, but this effect should dissipate in more gender-balanced groups.

Hypothesis 1 The male advantage in leader prototypicality is smaller in gender-balanced teams than in male-majority teams.

Fig. 1 Complete theoretical model.



Note: Control variables are in gray; academic major was also a control variable in Study 2

Group Prototypicality and Trust

Leadership scholars view the trust of subordinates as an essential component of effective leadership (e.g., Bennis and Nanus 1985; Braun et al. 2013; Dirks and Ferrin 2002; Zand 1997). For example, Conger and Kanungo (1998, p. 46) state that "…leading implies fostering changes in followers through the building of trust and credibility. In turn, trust enables and builds enduring commitment in the pursuit of a future goal." Indeed, trust is universal to leadership theories (Dansereau et al. 2013, p. 800). Thus, trust in leadership is an important leadership outcome that is likely affected by leaders' group prototypicality.

Social identity leadership scholars argue that prototypicality is an important driver of effective leadership and group performance because more group prototypical leaders are more trusted to pursue the group's interests (e.g., Giessner and van Knippenberg 2008; Giessner et al. 2009; van Knippenberg and Hogg 2003). Indeed, leaders' group prototypicality has consistent positive effects on a range of leadership evaluations, such as performance ratings and effectiveness (see van Knippenberg 2011). We build on this research by arguing that not only leaders' group prototypicality but also their prototypicality as leaders is associated with leader trust and effectiveness.

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

By integrating role congruity (Eagly and Karau 2002) and social identity theories (Hogg 2001; van Knippenberg and Hogg 2003), we propose that team gender composition will interact with leader gender to affect team ratings of leader dissipate, thus increasing gender equity in team leadership evaluations. Formally,

Hypothesis 3 The male leadership advantage in leader trust via leader prototypicality is smaller in gender-balanced teams than in male-majority teams.

In Study 1, we test our first hypothesis regarding the effects of leader gender and team gender composition (i.e., leader group prototypicality) on team ratings of leader prototypicality. Then, in Study 2, we expand on this finding by testing the extent to which leader gender and team gender composition interact to relate to team perceptions of trust in the leader through team leader prototypicality ratings (Fig. 1).

Methods

Data were collected in the same manner and in the same context for Studies 1 and 2, but the data were collected one year apart for each study (2014 and 2015). Additionally, in Study 2, we replicate Study 1 and conceptually extend our model to predict trust in the leader.

Samples and Procedures

We conducted two randomized field experiments among teams of business, economics, and informatics (IT) students at a large university in Western Europe.[2] Followers were incoming first-year students participating in an orientation event starting on the first day of the semester and lasting trust through team ratings of leader prototypicality. Teams consisting of more men are likely to see a male leader as more prototypical than a female leader; thus, the male leader should be viewed as more trustworthy. However, as teams become more gender-balanced, this male advantage should for the duration of the semester. Our specific context was organizational socialization (Fang et al. 2011), wherein our leaders served as "organizational insiders" (Fang et al. 2011) guiding newcomers with institutionalized socialization tactics (i.e., "learning experiences as part of a cohort with clearly defined sequences and timed training and orientation activities"; Jones 1986, p. 131). Leaders were more experienced students enrolled in a leadership and group organization course. Leaders applied for this course and received academic credit upon completion. A total of 35 leaders were selected using systematic criteria (e.g., previous academic performance) and trained for two days. Leaders were not trained in a specific leadership style; instead, they received a general theoretical overview of leadership. This training was intended to develop leaders' skills and effectiveness during the event.

Data were collected via pencil-and-paper surveys distributed 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 six 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 and preparation for an intergroup competition.[3] Thus, leaders' strategic goals included teams' academic and social orientation to campus as well as encouraging followers' creativity and performance for the competition. Leaders also organized subsequent events for their teams; this orientation event was the first, but not last, event during which the leaders, followers, and teams would interact.

Measures

We used a multiple-wave and multiple-source approach to data collection. All perceptual measures

---

[2] The true purpose of our study was unknown to the followers and leaders. 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.

[3] Given the very subjective nature of the competition, which was organized and evaluated by a third party, we were unable to use these data as a team performance outcome.

had 6-point response scales (1 = *does not apply at all* to 6 = *completely applies*). Follower gender (male = 0, female = 1) was self-reported, and leader gender (male = 0, female = 1) was objective information entered by researchers. We manipulated leaders' objective group prototypicality via team gender composition. Leader and follower ratings were both measured by surveys with additional filler items to further 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 Prototype

We assessed followers' perceptions of their leader's embodiment of a prototypical leader with three items (Cronshaw and Lord 1987): *the leader is a typical leader*, *exhibits the behavior of a leader*, and *fits one's image of a leader* ($\alpha = .88$ in Study 1; $\alpha = .89$ in Study 2).

Group Prototypicality

We randomly determined team gender composition (i.e., leaders' objective prototypicality of the group in terms of gender) as male-majority (20% women) or gender-balanced (50% women). Our manipulations were strategically chosen to mirror the current gender composition of women in leadership roles (i.e., 20%; Swiss Federal Statistical Office 2014) and approximately equal group gender compositions (50%). Furthermore, team gender compositions of 20–50% are feasible in modern workplaces given that women have composed at least half of college-degree earners for several decades (Perry 2013). These team gender compositions also echo *skewed* (20% women) and *balanced* (40–60% women) designations from critical mass theory (Kanter 1977).

Due to common issues with field experiments (e.g., new participants and no-shows on the day of the event), team gender compositions were not always exactly 20 and 50%. Thus, we also had leaders report the exact number of men and women in their teams. Since these estimates are more accurate and informative than a dichotomous variable, we used the leader-reported measure of team gender composition in our empirical analyses; however, our results remain largely the same in size and significance when using the dichotomous variable.

Trust in Leader

We assessed followers' trust in their leaders in Study 2. We used the 6-item behavioral trust scale developed by Gillespie and Mann (2004; $\alpha = .89$). Items assessed included the extent to which followers trust their leaders' skills, judgment, and values as well as 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…relying on his/her task-related skills and abilities? …sharing your personal beliefs?*; $\alpha = .82$).

Control Variables and Robustness Tests

According to the relational demography theory (Tsui et al. 1992; Tsui and O'Reilly 1989), women may prefer female leaders and vice versa for men. Thus, we controlled for follower gender (0 = male, 1 = female). Theoretically, group identity may be more cohesive in smaller groups, whereas more practically, followers may interact more with their teammates and leaders in smaller groups. Post (2015) also found a female leadership advantage in larger teams. Thus, we also controlled for group size for theoretical reasons. The results remained unchanged if these two control variables were excluded from

the analysis.

By design, we hold constant potentially meaningful facets of group diversity (e.g., age and tenure), as recommended by team diversity scholars (Jackson et al. 2003; van Knippenberg and Schippers 2007), to better isolate the effect of team gender diversity. Specifically, 97% of the participants were 18–25 years old, and all were new to the university. Three groups in Study 2 were not from business and economics, so we included academic major as a control variable.

To rule out alternative explanations for our findings, we also assessed leaders' self-conceptions as leaders (Epitropaki et al. 2017) as well as leader–follower gender match (Tsui and O'Reilly 1989). Leaders' self-reported prototypicality was measured with the same three items used to assess team ratings (Cronshaw and Lord 1987) but adapted for leaders' self-ratings. Leaders completed these items twice: once before ($\alpha = .84$ in Study 1; $\alpha = .72$ in Study 2) and once after leading their teams ($\alpha = .89$ in Study 1; $\alpha = .82$ in Study 2).

Results

Results of Study 1

The analyses were conducted with Mplus (version 7.4). Continuous predictors were centered prior to the analyses, and unstandardized coefficients are reported. Given our relatively small sample sizes, we calculated one-tailed tests for all directional hypotheses.

Descriptive Statistics

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 ($n = 74$)[4] for a final total of 426 participants (38.5% women). Team gender was randomly assigned as male-majority (20%) or balanced (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), and the 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 that balanced teams comprised a greater share of women ($M = 46.81$, SD = 8.00) than did male-majority teams ($M = 26.87$, SD = 6.28), Cohen's $d = 2.77$. Second, we 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 indicated that followers noticed their group gender composition and scored it in line with our manipulation for balanced ($M = 2.94$, SD = .0.35) and male-majority teams ($M = 2.23$, SD = 0.54), $t(425) = 16.01$, $p < .001$, Cohen's $d = 1.56$. These checks reflect very large effects, indicating that our manipulations were effective.

---

[4] The leaders of these teams did not closely follow the experimenters' randomized assignments of participants to teams; as a result, these teams were also larger than the other 32 teams, $F(1505) = 1640.08$, $p < .001$, $\eta^2 = .77$, providing additional justification to exclude them from Study 1.

Table 1 Study 1 descriptive statistics, correlations, and scale reliability

Variables	*M*	SD	1	2	3	4	5	6	7
1. Leader gender	.52	.50							
2. Team gender composition	31.37	16.15	−.20**						
3. Team size	15.00	4.13	.17**	−.10*					
4. Academic major	.14	.35	.39**	−.12*	.76**				
5. Leader prototypicality	4.76	0.36	−.29**	.22**	−.73**	−.60**			
6. Trust in leader	5.63	0.25	−.35**	.29**	−.76**	.85**	.67**		
7. Follower gender	.31	.46	−.10*	.31**	−.06	−.12*	.10**	.14**	

$N = 434$, nested within 35 teams. Leader and follower gender were coded male = 0, female = 1 and academic major IT = 1, management/economics = 0

*$p < .05$; **$p < .01$

Preliminary Analyses

The data represent 426 participants nested within 32 teams. Our hypotheses propose that leader gender and team composition affect group-level perceptions of the leader. Thus, we computed an intraclass correlation coefficient (ICC1) to show the percentage of total variance in the dependent variable that is between groups (ICC1 = .06). Although not extremely high, this value suggests that there is meaningful variance in group perceptions of leader prototypicality (Bryk and Raudenbush 1992). Additionally, we examined whether aggregation was justified in our data by calculating $r^*_{wg(j)}$ values. Examining $r^*_{wg(j)}$ values suggested that the level of agreement was above the typical cutoff of .70 to support aggregation for leader prototypicality (average $r^*_{wg(j)} = .89$). Given the $r^*_{wg(j)}$ values, our theoretical reasons (Bliese 1998), and an ICC(2) of .46, which also suggested reliable group means (Castro 2002; LeBreton and Senter 2008), we aggregated leader prototypicality. Comparisons of the null model with a model allowing for random intercepts also indicated a significant difference, $\chi^2(1, N = 426) = 5.86$, $p < .01$. Thus, there is a significant intercept variation according to group.

Hypothesis Testing

The linear mixed-effects model indicates that male leaders are not rated as more prototypical leaders than female leaders ($b = -0.12$, $p = .19$; see Table 2). There is also no main effect of team gender composition on leader prototypicality ($b = 0.01$, $p = .39$). However, as expected, these null main effects are qualified by a significant interaction between leader gender and team gender composition ($b = 0.01$, $p = .03$).

Table 2 *Study 1 linear mixed-effects models*

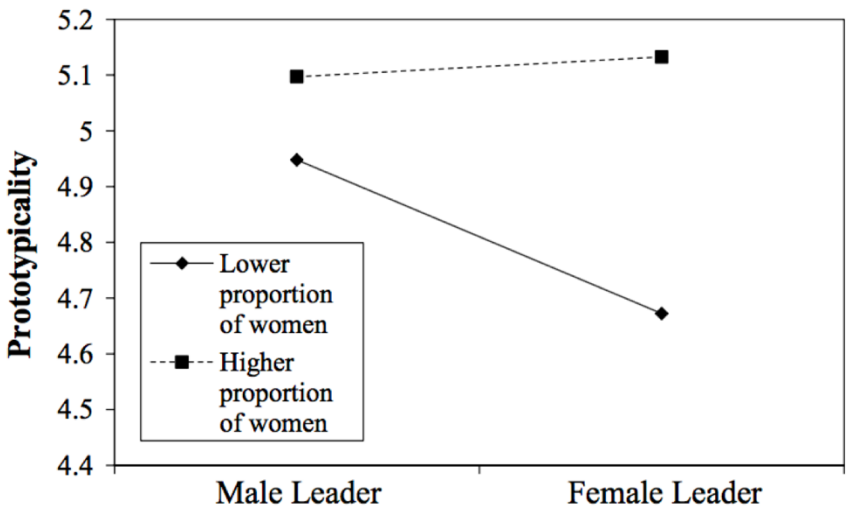
Leader prototypicality				
Variable	Main effects model		Interaction effect model	
	b	SE	*b*	SE
1st level				
Follower gender	−2.96	20.52	−2.89	16.76
2nd level				
Team size	−0.01	0.02	0.01	0.02
Leader gender	−0.14	0.10	−0.12	0.09
Team gender composition	0.01	0.01	0.01	0.01
Team gender × leader gender	–	–	0.01*	0.01
Intercept	4.97**	0.08	4.96**	0.08
Log-likelihood	−433.17		−533.59	

$N = 426$, nested within 32 teams. Leader and follower gender were coded male = 0, female = 1. All coefficients are unstandardized

*$p < .05$; **$p < .01$

As depicted in Fig. 2, when there are more men in the team, female leaders are rated as less prototypical leaders than males. However, this effect is eliminated when there are more women in the team. This pattern of results supports Hypothesis 1.

Fig. 2 *Interaction between leader gender and group gender composition (Study 1).*



Note: The interaction is plotted at ± 1 SD from the mean or 22 and 52% women in the team (respectively)

Results of Study 2

Descriptive Statistics

Descriptive statistics and correlations are displayed in Table 3. From a total of 467 followers, 33 were eliminated due to missing data. The final sample consisted of 434 participants (31.33% women). The team share of women ranged from 0 to 63%, with an average of 31.37% (SD = 16.15). Leaders of balanced teams reported more women on their teams (M = 41.42, SD = 11.18) than did leaders of male-majority teams (M = 17.05, SD = 10.27, Cohen's d = 2.27), as intended by our manipulation.

Table 3 *Study 2 descriptive statistics, correlations, and scale reliability*

Variables	*M*	SD	1	2	3	4	5	6	7
1. Leader gender	.52	.50							
2. Team gender composition	31.37	16.15	−.20**						
3. Team size	15.00	4.13	.17**	−.10*					
4. Academic major	.14	.35	.39**	−.12*	.76**				
5. Leader prototypicality	4.76	0.36	−.29**	.22**	−.73**	−.60**			
6. Trust in leader	5.63	0.25	−.35**	.29**	−.76**	.85**	.67**		
7. Follower gender	.31	.46	−.10*	.31**	−.06	−.12*	.10**	.14**	

N = 434, nested within 35 teams. Leader and follower gender were coded male = 0, female = 1 and academic major IT = 1, management/economics = 0

*p < .05; **p < .01

We conducted a series of t-tests to examine potential differences in sample characteristics between Studies 1 and 2. The only significant difference was found for participant gender: 38% of followers were male in Study 1, whereas only 31% of followers were male in Study 2 (p < .05). The main reason for this was that male followers were overrepresented in the three teams that were excluded from Study 1 due to compromised randomization. It is noteworthy that we also control for follower gender across all analyses.

Preliminary Analyses

The data represent 434 participants nested within 35 teams. An ICC1 was computed to show the percent of total variance in leader prototypicality and trust in a leader 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 and Raudenbush 1992). An examination of $r^*_{wg(j)}$ values suggested that the level of agreement was above the typical cutoff of .70 to support aggregation for leader prototype and trust (average $r^*_{wg(j)}$ = .90 and .94, respectively). The ICC2 values for leader prototype (.62) and trust in the leader (.40) also suggested reliable group means, supporting aggregation (Bliese 1998; Castro 2002; LeBreton and Senter 2008). Model comparisons of the null model with a model allowing for random intercepts indicated a significant difference for leader prototypes [$\chi^2(1, N = 434)$ = 23.11 and p < .001] as well as trust in the leader [$\chi^2(1, N = 434) = 6.64$, $p < .01$].

Hypothesis Testing

As expected, we replicated a significant interaction effect of leader gender and team gender composition on leader prototypicality ($b = 0.02$, $p = .003$; see Table 4). As depicted in Fig. 3, when there are more men in the team, female leaders are rated as less prototypical leaders than male leaders. However, this effect is eliminated with more women in the team. This pattern of effects replicates and further supports Hypothesis 1.

Table 4 *Study 2 linear mixed-effects models*

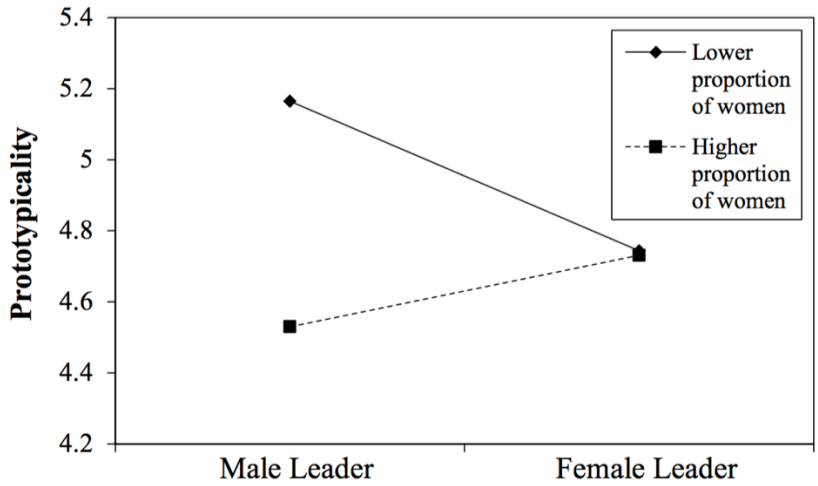
Main effects model interaction effect model								
Variable	Leader prototypicality		Trust in leader		Leader prototypicality		Trust in leader	
	b	SE	*b*	SE	*b*	SE	*b*	SE
1st level								
Follower gender	1.20	1.17	−0.55	0.82	2.32**	0.73	−1.09*	0.58
2nd level								
Team size	−0.06**	0.02	0.02	0.02	−0.06**	0.01	0.01	0.02
Academic major	−0.03	0.24	−0.49**	0.12	0.17	0.20	−0.52**	0.12
Leader gender	−0.11	0.09	0.01	0.05	−0.11	0.08	−0.03	0.06
Team gender composition	−0.00	0.01	0.00	0.00	−0.01	0.00	0.01	0.00
Team gender × leader gender	–	–	–	–	0.02**	0.01	–	–
Leader prototypicality	–	–	0.16	0.20	–	–	0.41*	0.21
Intercept	4.80**	0.04	4.89**	0.96	4.79**	0.05	3.66**	0.98
Log-likelihood	−407.94				−521.45			

$N = 434$, nested within 35 teams. Leader and follower gender were coded male = 0, female = 1 and academic major IT = 1, management/economics = 0. All coefficients are unstandardized

*$p < .05$; **$p < .01$

As an extension of Study 1, we also tested the relation between leader prototypicality and trust in the leader in Study 2. As expected, leader prototypicality was positively associated with leader trust ($b = 0.41$, $p < .05$), supporting Hypothesis 2. To test our proposed mediated moderation effect, we followed the recommendations of Preacher and Hayes (2008) to estimate the indirect effects with 95% Confidence Intervals (CI)—conditional on the moderator. We calculated the conditional indirect effects of leader gender on leader trust via leader prototypicality at three different values of our moderator, team gender composition (the mean ± 1 SD). The results show that female leaders are viewed as less trustworthy (via leader prototypicality) in male-dominated teams ($b = -0.14$, 95% CI = [−.26, −.03], at −1 SD of team gender composition), but this effect becomes nonsignificant with more women in the team ($b = -0.05$, 95% CI = [−.11, .02] and $b = 0.05$, 95% CI = [−.02, .12], at the mean and + 1 SD of team gender composition, respectively). These results support Hypothesis 3 and our overall model (Fig. 1).

Fig. 3 *Interaction between leader gender and group gender composition (Study 2).*



Note: The interaction is plotted at ± 1 SD from the mean or 15 and 47% women in the team (respectively)

Supplementary Analyses

Statistical Power

The number of groups tested in Studies 1 and 2 could be considered small ($N = 32$ and $N = 35$), which could raise concerns about statistical power. However, insufficient power is more of a concern when effects are not found. Thus, insufficient statistical power is less of a threat to our key findings of interest, for which we found repeated empirical support. Insufficient statistical power is also one reason why we did not conduct simple slopes analyses. Simple slopes tests have been shown to be unreliable when sample sizes are small (e.g., Liu et al. 2017). Methods experts (e.g., Dawson 2014) also argue that simple slopes tests are not always necessary. In the following, we describe several robustness checks that were conducted to rule out alternative explanations for our findings. To optimize statistical power, we combined data sets from Studies 1 and 2 for the following analyses.

Leader Ratings

Although there is ample evidence to guide our prediction that team responses to leaders change according to leader gender and team gender composition, it is unclear whether only the followers' perceptions of leaders change or whether the leaders' conceptions of themselves also change. If leaders' self-perceived similarity to the leader prototype also changes based on the teams that they lead, their group-directed actions may be altered. This notion builds on principles from the social identity theory of leadership (Hogg 2001; van Knippenberg and Hogg 2003) such that leaders' self-perceived prototypicality and degree of team identification predict their team-oriented attitudes and behaviors. However, the vast majority of research on leader prototypicality has examined followers' perceptions of leaders' group prototypicality (see van Knippenberg 2011) rather than leader prototypicality. Indeed, a recent review by Epitropaki et al. (2017) has similarly highlighted the sparse work on leadership identity.

Given this paucity of research, we also include an exploratory assessment of leaders' self-reported leader prototypicality. We assessed the relationship between leader gender and leader prototypicality at two points in time: before and after leaders worked with their teams. Only in the latter instance could leaders

have possibly been influenced by their objective group prototypicality because this information was unknown to them before they led their teams.

Across the two studies, half of the leaders were women (51.43%) and half of the groups were gender-balanced (54.29%). Of 70 leaders, 65 (92.9%) returned completed surveys at both time points. These 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).

We ran a series of mixed analyses of variance (ANOVAs) using within-(Time 1, Time 2) and between-subject variables (leader gender: male or female; team gender composition: male majority or gender balanced). The results indicate no significant main effects or interactions apart from an overall increase in self-rated leader prototypicality from Time 1 (M = 4.23, SD = 0.81) to Time 2 (M = 4.48, SD = 0.88; $F(1, 61) = 7.33$ $p = .009$, $\eta^2 = 0.11$); all other ps = .36–.70. Thus, leaders' self-rated prototypicality does not appear to be affected by our intervention of group gender composition.

Leader–Follower Gender Match

Given our knowledge of the relational demography theory and findings (e.g., Tsui and O'Reilly 1989), the gender composition of the leader–follower dyads may affect followers' ratings of leaders in more gender-balanced teams rather than team-based perceptions of prototypicality. That is, with more women in the team, there are more gender-matched pairs, which could provide an alternative explanation for our findings.

To test the potential effect of dyadic similarity, we created a new variable of gender match (1) or mismatch (0) between followers and leaders. We then included this variable in our previous model predicting leader prototypicality. This new variable does not explain additional variance in our model, nor does it change our overall patterns of results. Thus, our results do not seem to be explained by female (male) followers' higher ratings of leaders of the same gender.

Discussion

Our findings provide evidence that the male advantage in leadership ratings (i.e., prototypicality and trustworthiness) can be mitigated in gender-balanced teams. In doing so, we provide causal support for the social identity theory of organizational leadership (Hogg 2001; van Knippenberg and Hogg 2003) and a boundary condition of role congruity theory (Eagly and Karau 2002). We also show evidence against potential alternative explanations of leaders' self-perceptions of prototypicality, which could change based on the teams that they lead, and dyadic gender match (if female followers rate female leaders more positively). In the following, we outline the implications of our findings for theory, practice, business ethics, and society.

Theoretical Implications

We aimed to bridge classic work on gender and leadership (Eagly and Karau 2002) with leadership and group prototypes research (Hogg 2001; van Knippenberg and Hogg 2003) to make several core contributions and outline specific areas for future research. First, the proposition that team contextual features and group prototypes undermine broader societal biases regarding women's incongruity as leaders has been theorized (Eagly and Karau 2002; Hogg 2001; van Knippenberg and Hogg 2003; van Knippenberg 2011). However, to our knowledge, this proposition has not yet been tested. We provide causal evidence for this idea by showing that gender differences in perceptions of leadership

prototypicality dissipate with increasing gender diversity in teams.

However, it is important to note that the teams in our study never exceeded 63% women. It is possible that there may be a critical point at which team gender composition no longer helps and might even be detrimental to female leaders. For example, social psychology research has found a stigma-by-association effect for female leaders who lead majority female teams (Pryor et al. 2012). Field experimental evidence also supports this proposition such that stigma toward individual team members, as well as teams as a whole, increased with the proportion of women on the team (West et al. 2012). Thus, future research should seek to delineate the boundary conditions of the positive effects of team gender composition for teams and female leaders, extend this research from intragroup to intergroup perceptions, and examine other organizational outcomes also influenced by team gender composition (e.g., team performance; Hoogendoorn et al. 2013).

Second, we demonstrated a new type of prototypicality benchmarking. To date, social identity leadership researchers have mostly manipulated leaders' perceived group prototypicality via fabricated feedback about leaders' values or beliefs (e.g., Giessner et al. 2013, Study 1; Hais et al. 1997; Monzani et al. 2014; van Knippenberg and van Knippenberg 2005). This was often done for individual participants who were ostensibly in groups or anticipated group interaction (e.g., Hais, et al. 1997; Hogg et al. 1998, 2006; Monzani et al. 2014), organized in virtual teams, or with virtual leaders (e.g., Giessner et al. 2013, Study 1; van Knippenberg and van Knippenberg 2005). This speaks to the power of group prototypes. However, we found converging effects after manipulating leaders' objective group prototypicality (in terms of gender); our followers were also nested in actual groups, and they interacted with real team members and leaders for several hours. Because leadership is inherently a social process (Chemers 2001), it is perhaps most appropriately or accurately studied through social interactions. However, a key distinction between this past work and the current research is that we did not measure leaders' perceived group prototypicality (i.e., the leader represents what is characteristic about the team), as is common in social identity studies of leadership (see van Knippenberg 2011). Thus, it is unclear whether and how our manipulation would have influenced team perceptions of leaders in this regard.

Third, beyond our implications for female leaders, our findings stimulate avenues of inquiry about male leaders. Interventions such as ours should not affect team ratings of men, because they are similarly prototypical as female leaders in more gender-balanced groups (Hogg 2001; van Knippenberg and Hogg 2003). Our results seem to support this proposition in Study 1 (Fig. 2), but not in Study 2 (Fig. 3). This discrepancy could be data driven given the different values represented in our figures (i.e., lower boundaries of 15 and 22%). That is, a recognizable number of women in the team may be required to influence teams' ratings of their female leaders (i.e., a cross-level effect), and this value could fall between 15 and 22%. This idea echoes themes from tokenism theory such that a "critical mass" requires more than 20% women in a team (Kanter 1977) or a "magic number" of three (Joecks et al. 2013; Konrad et al. 2008; Torchia et al. 2011) to influence teams (i.e., within-level effects) or organizational outcomes. However, this is only one possible post hoc explanation.

Fourth and finally, our findings also have implications for the diversity literature. Scholars have argued that actual diversity is a key facet of diversity climate, because more diversity means a more favorable climate (see Dwertmann et al. 2016). Our findings illustrate why and how this holds true for gender and leadership. Meta-analytic evidence supports a similar idea such that women have fewer leadership disadvantages in settings with more balanced organizational gender demography (Eagly et al. 1995; Paustian-Underdahl et al. 2014). Thus, beyond the ethics and social justice arguments for equal career opportunities (e.g., Dwertmann et al. 2016; Eagly and Carli 2007), more diverse teams and organizations may also have a self-reinforcing effect by creating a more level playing field for female leaders.

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 2018). However, despite their representation at lower levels, women remain a glaring minority in leadership positions (e.g., only 5.2% of CEOs are women; Catalyst 2018). 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.

Furthermore, there are extraordinary costs invested in leadership training programs despite a lack of evidence of transfer (Baldwin and Ford 1988; Burke and Day 1986; Burke and Hutchins 2007). Individual interventions that train women to display more masculine or agentic behaviors also risk backlash (Brescoll 2011; Brescoll and Uhlmann 2008; Rudman 1998; Rudman et al. 2012), whereas organizational interventions may incidentally reinforce gender roles and biases (e.g., see Gloor et al. 2018; Heilman et al. 1997). Thus, we encourage organizations to consider teambased interventions to restore gender equity in leadership evaluations.

In the case that teams are already established or must be constructed based on non-gender-based criteria (e.g., employee education or expertise), practitioners can also use our findings to inform their interpretations of leader evaluations. For example, a woman from a male-majority team may provide similarly negative performance feedback about a woman supervisor as her male teammates. This effect would be unexpected according to relational demography perspectives (Tsui et al. 1992; Tsui and O'Reilly 1989) and could be interpreted as problematic female same-sex interactions in organizations (see Sheppard and Aquino 2017), damaging 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 can easily be assessed by including a single item about team gender.

These practical measures may also help to resolve three particular moral issues relevant to employees, organizations, and society. First, there is a fundamental moral case for equity in leadership appraisals if male and female leaders perform the same behavior but receive different reactions to or ratings of their performance. These ratings are then used to designate penalties and rewards and thus perpetuate the gender gap in leadership. Such a situation would be unfair for competent female leaders and overly reassuring for less competent male leaders. Given general moral preferences for fair performance appraisals (Dusterhoff et al. 2014), the consequences of such inequities could also resonate throughout teams and organizations.

Second, growing evidence indicates that women may be more moral, ethical, and other-oriented socially compassionate leaders than men (see Eagly 2005). Empirical evidence suggests that a larger number of women serving on boards of directors are associated with more ethical firm behavior (Nekhili and Gatfaoui 2013). A recent meta-analysis has also found that female board representation is positively associated with board monitoring and firm profitability (Post and Byron 2015). Thus, female leaders may lead in a more moral manner than men, bringing a more ethically oriented style to their teams and organizations without necessarily sacrificing the bottom line.

Finally, there are ethical implications pertaining to an inefficient use of the labor force. Students invest significant time and effort in their studies, whereas society invests substantial funds in their training and education (OECD 2017). Thus, attracting and retaining trained female talent by recruiting more female leaders and ensuring equitable raings and rewards for existing female leaders also make sense as a societal priority. Such an initiative also has implications for workforce sustainability in the long term, particularly where increasingly common immigration restrictions necessitate a more efficient use of locally trained

talent (including women), especially for highly specialized workers (Dutu 2014).

Strengths, Limitations, and Future Research

Our study is methodologically rigorous, including a replication of large effects and a conceptual extension. King et al. (2013) endorse field experiments such as ours as a gold standard for external validity, especially when examining sensitive topics such as gender bias. We also avoid the threat of common method variance (Podsakoff et al. 2012) by using data collected from different sources (e.g., followers and leaders) and including objective data (e.g., team share of women and leader gender) collected at different times (e.g., before and after the orientation event). Interactions cannot be artifacts of common method variance (Siemsen et al. 2010). Because we intervened and manipulated team gender composition by randomly assigning leaders and followers to teams, we can make a causal claim based on our findings. That is, leader gender predicts team ratings of leader prototypicality and trust depending on the team gender composition. We also provide evidence against several other potential alternate explanations (e.g., leaders' self-conceptions as leaders or increasing shares of female follower–leader dyads drive this effect).

However, as with any study, our research has its limitations. For example, our student sample could limit the generalizability of our results. However, aligned with common conceptions of leadership, our student leaders guided their teams toward shared strategic goals (e.g., academic and social orientation as well as creativity and performance for the intergroup competition). Although our student leaders had no evaluative or disciplinary influence on their followers, they could dismiss individuals from the event. Thus, despite our explicit references to our more senior students as "leaders" in our measures and event organization, such an arrangement may be more representative of modern, flatter hierarchies (e.g., project managers or peer leadership) than more traditional conceptions of leadership, the former of which is becoming increasingly common in today's more interdependent organizations (Rajan and Wulf 2006; Wegman et al. 2016). We recommend that future research replicates and extends our findings by testing these hypotheses within organizational teams to better understand the extent to which team gender composition and leader gender relate to leader evaluations in ongoing work teams. However, leaders and followers would not be randomly assigned to teams, so this type of design could be threatened by endogeneity.

Future research should also examine the effects of interventions such as ours on team ratings of male leaders. Although men should be rated similarly to women in more gender-balanced groups because they are equally prototypical of their groups (Hogg 2001; van Knippenberg and Hogg 2003), our results support this idea only in Study 1. As previously discussed, this could be due to the specific ranges of team gender compositions depicted in our graphs, which differed across studies. However, future research with a more continuous range of team gender compositions is needed to better assess this possibility.

We also encourage studies in other countries to examine whether cultural preferences for gender egalitarianism may influence our findings. The Western European country where our studies occurred has lower ratings for gender egalitarianism in comparison with Eastern or Nordic European countries (House et al. 2004). Similarly, the disciplines studied, management and economics, are more masculine in both stereotypes and actual demographic composition. Thus, future research could also examine team responses to male and female leaders in traditionally female disciplines or occupations such as teaching or nursing. However, these instances do not contribute to the larger patterns of social and economic inequality (Budig 2002) as in the traditionally male contexts that we examined. Finally, we created teams with low or balanced shares of women. Although this allowed 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 actual organizations could help address this concern because many organizations have teams comprising a variety

of gender compositions. However, this option would not solve the potential endogeneity problem.

Conclusions

Our results highlight the benefits of increasingly diverse employees for female leaders and organizations, but only if teams are designed with leaders in mind. Indeed, if the leadership game is rigged in men's favor, women may face a double bind of backlash regardless of their ability or performance, which poses a continuing ethical dilemma for organizations and society. However, according to our reasoning and results, there is hope for restoring gender equity and equality in leadership if we fix the game, not the dame.

Acknowledgements

We thank Steffen Giessner, Levke Henningsen, Alina Hernandez Bark, Steve Karau, Lucas Monzani, Christian Troester, Daan van Knippenberg, and Christian Voegtlin for their comments on previous versions of this manuscript, as well as our editor and three anonymous reviewers. This research was completed as part of the first author's dissertation, which was conducted at the chair of Professor Bruno Staffelbach. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the studies.

References

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.

Bennis, W., & Nanus, B. (1985). *Leaders: The strategies for taking charge*. New York: Harper & Row.

Bliese, P. D. (1998). Group size, ICC values, and group-level correlations: A simulation. *Organizational Research Methods, 1*(4), 355–373.

Braun, S., Peus, C., Weisweiler, S., & Frey, D. (2013). Transformational leadership, job satisfaction, and team performance: A multilevel mediation model of trust. *The Leadership Quarterly*, *24*(1), 270–283. https://doi.org/10.1016/j.leaqua.2012.11.006.

Brescoll, V. L. (2011). Who takes the floor and why: Gender, power, and volubility in organizations. *Administrative Science Quarterly, 56*(4), 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*(3), 268–275.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

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. (2018). *Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity*. Current Population Survey. Retrieved from: www.bls.gov. Accessed 5 April 2018.

Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review, 6*(3), 263–296.

Burke, M. J., & Day, R. R. (1986). A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology, 71*(2), 232–245.

Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlations, rwg(j), hierarchical linear modeling, within- and between-analysis, and random group resampling. *The Leadership Quarterly, 13,* 69–93.

Catalyst. (2018). *Women CEOs of the S&P 500*. Retrieved from www.catalyst.org. Accessed 5 April 2018.

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: Blackwell.

Conger, J. A., & Kanungo, R. N. (1998). *Charismatic leadership in organizations*. Thousand Oaks, CA: Sage.

Cronshaw, S. F., & Lord, R. G. (1987). Effects of categorization, attribution, and encoding processes on leadership perceptions. *Journal of Applied Psychology, 72*(1), 97–106.

Dansereau, F., Seitz, S. R., Chiu, C., Shaughnessy, B., & Yammarino, F. J. (2013). What makes leadership, leadership? Using self-expansion theory to integrate traditional and contemporary approaches. *The Leadership Quarterly, 24,* 798–821.

Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business Psychology*, *29*(1), 1–19. https://doi.org/10.1007/s10869-013-9308-7.

Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology, 87,* 611–628.

Dusterhoff, C., Cunningham, J. B., & MacGregor, J. N. (2014). The effects of performance rating, leader–member exchange, perceived utility, and organizational justice on performance appraisal satisfaction: Applying a moral judgment perspective. *Journal of Business Ethics, 119*(2), 265–273.

Dutu, R. (2014). Women's role in the Swiss economy. *OECD Economics Department Working Papers, No. 1144*, OECD Publishing.

Dwertmann, D. J. G., Nishii, L. H., & van Knippenberg, D. (2016). Disentangling the fairness & discrimination and synergy perspectives on diversity climate: Moving the field forward. *Journal of Management, 42,* 1136–1168.

Eagly, A. H. (1987). *Sex differences in social behavior: A social role interpretation*. Hillsdale, NJ: Earlbaum.

Eagly, A. H. (2005). Achieving relational authenticity in leadership: Does gender matter? *The Leadership Quarterly, 16,* 459–474.

Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders*. Boston, MA: Harvard Business School Press.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Bulletin, 109*(3), 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*(1), 3–22.

Epitropaki, O., Kark, R., Mainemelis, C., & Lord, R. G. (2017). Leadership and followership identity processes: A multilevel review. *The Leadership Quarterly, 28,* 104–129.

Fang, R., Duffy, M. K., & Shaw, J. D. (2011). The organizational socialization process: Review and development of a social capital model. *Journal of Management, 37,* 127–152.

Giessner, S. R., & van Knippenberg, D. (2008). "License to fail": Goal definition, leader group prototypicality, and perceptions of leadership effectiveness after leader failure. *Organizational Behavior and Human Decision Processes, 105*(1), 14–35.

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.

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*(4), 658–667.

Gillespie, N. A., & Mann, L. (2004). Transformational leadership and shared values: The building blocks of trust. *Journal of Managerial Psychology, 19,* 588–607.

Gloor, J. L., Li, X., Lim, S., & Feierabend, A. (2018). An inconvenient truth? Interpersonal and career consequences of "maybe baby" expectations. *Journal of Vocational Behavior, 104,* 44–58.

Hais, 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*(10), 383–

399.

Heilman, M. E., Block, C. J., & Stathatos, P. (1997). The affirmative action stigma of incompetence: Effects of performance information ambiguity. *Academy of Management Journal*, *40*(3), 603– 625. https://doi.org/10.5465/257055.

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, *32*, 113–135. https://doi. org/10.1016/j.riob.2012.11.003.

Hernandez Bark, A. S., Escartin, J., Schuh, S. S., & van Dick, R. (2016). Who leads more and why? A mediation model from gender to leadership role occupancy. *Journal of Business Ethics, 139*(3), 473–483.

Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review, 5*(3), 184–200.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. London: 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*(5), 1248–1263.

Hoogendoorn, S., Oosterbeek, H., & van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science, 59*(7), 1514–1528.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage Publications.

Jackson, S. E., Joshi, A., & Erhardt, N. L. (2003). Recent research on team and organizational diversity: SWOT analysis and implications. *Journal of Management, 29*(6), 801–830.

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.

Jones, G. R. (1986). Socialization tactics, self-efficacy, and newcomers' adjustments to organizations. *Academy of Management Journal, 29,* 262–703.

Kanter, R. M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology, 82*(5), 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*(4), 501–521.

Konrad, A. M., Kramer, V., & Erkut, S. (2008). The impact of three or more women on corporate boards. *Organizational Dynamics, 37*(2), 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.

Liu, Y., West, S. G., Levy, R., & Aiken, L. S. (2017). Tests of simple slopes in multiple regression models with an interaction: Comparison of four approaches. *Multivariate Behavioral Research, 52*(4), 1–20.

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* (pp. 48–64). London: Sage.

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*(4), 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.

Nekhili, M., & Gatfaoui, H. (2013). Are demographic attributes and firm characteristics drivers of gender diversity? Investigating women's positions on French boards of directors. *Journal of Business Ethics, 118*(2), 227–249.

O'Neil, D. A., Hopkins, M. M., & Bilimoria, D. (2008). Women's careers at the start of the 21st century: Patterns and paradoxes. *Journal of Business Ethics, 80,* 727–743.

Oakley, J. G. (2000). Gender-based barriers to senior management positions: Understanding the scarcity of female CEOs. *Journal of Business Ethics, 27,* 321–334.

OECD. (2017). *Education at a glance 2017: OECD Indicators*. Paris: OECD Publishing. Retrieved from http://dx.doi.org/10.1787/ eag-2017-en.

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*(6), 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. Accessed 31 May 2017.

Perry, E. L., Davis-Blake, A., & Kulik, C. T. (1994). Explaining gender-based selection decisions: A synthesis of contextual and cognitive approaches. *The Academy of Management Review, 19*(4), 786–820.

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.

Post, C. (2015). When is female leadership an advantage? Coordination requirements, team cohesion, and team interaction norms. *Journal of Organizational Behavior, 36*(8), 1153–1175.

Post, C., & Byron, K. (2015). Women on the boards and firm financial performance: A meta-analysis. *Academy of Management Journal, 58*(5), 1546–1571.

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*(2), 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*(4), 759–773.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and*

categorization (pp. 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*(3), 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*(4), 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.

Sheppard, L. D., & Aquino, K. (2017). Sisters at arms: A theory of female same-sex conflict and its problematization in organizations. *Journal of Management, 43*(3), 691–715.

Siemsen, E., Roth, A., & Oliveira, P. (2010). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods, 13*(3), 456–476.

Swiss Federal Statistical Office. (2014). *2014–2023 scenarios of higher education institutions-teaching staff.* Retrieved from www.bfs.admin.ch/. Accessed 31 May 2017.

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., pp. 7–24). Chicago: Nelson-Hall.

Torchia, M., Calabro, A., & Huse, M. (2011). Women directors on corporate boards: From tokenism to critical mass. *Journal of Business Ethics, 102,* 299–317.

Tsui, A. S., Egan, T. D., & O'Reilly, C. A. (1992). Being different: Relational demography and organizational attachment. *Administrative Science Quarterly, 37*(4), 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*(2), 402–423.

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, pp. 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: Blackwell.

van Knippenberg, D. (2011). Embodying who we are: Leader group prototypicality and leadership effectiveness. *The Leadership Quarterly, 22,* 1078–1091

van Knippenberg, D., & Hogg, M. A. (2003). A social identity model of leadership effectiveness in organizations. *Research in Organizational Behavior*, *25*, 243–295.

van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology, 58,* 515–541.

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.

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, 44,* 352–386.

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.

Zand, D. E. (1997). *The leadership triad: Knowledge, trust and power*. New York: Oxford University Press.