

RESEARCH ARTICLE

# Leader evaluation and team cohesiveness in the process of team development: A matter of gender?

Núria Rovira-Asenjo<sup>1☯‡\*</sup>, Agnieszka Pietraszkiewicz<sup>2☯‡\*</sup>, Sabine Sczesny<sup>2</sup>, Tània Gumí<sup>1</sup>, Roger Guimerà<sup>1,3</sup>, Marta Sales-Pardo<sup>1\*</sup>

**1** Department of Chemical Engineering, Universitat Rovira i Virgili, Tarragona, Catalonia, Spain,

**2** Department of Psychology, Division of Social Psychology and Social Neuroscience, University of Bern,

Bern, Switzerland, **3** Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain

☯ These authors contributed equally to this work.

‡ These authors are first authors on this work.

\* [nuria.rovira@urv.cat](mailto:nuria.rovira@urv.cat) (NRA); [apietraszkiewicz@gmail.com](mailto:apietraszkiewicz@gmail.com) (AP); [marta.sales@urv.cat](mailto:marta.sales@urv.cat) (MSP)



**OPEN ACCESS**

**Citation:** Rovira-Asenjo N, Pietraszkiewicz A, Sczesny S, Gumí T, Guimerà R, Sales-Pardo M (2017) Leader evaluation and team cohesiveness in the process of team development: A matter of gender? PLoS ONE 12(10): e0186045. <https://doi.org/10.1371/journal.pone.0186045>

**Editor:** Luís A. Nunes Amaral, Northwestern University, UNITED STATES

**Received:** April 10, 2017

**Accepted:** September 13, 2017

**Published:** October 23, 2017

**Copyright:** © 2017 Rovira-Asenjo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work received support from "Discovery, decomposition and dynamics of complex networks", <https://www.jsmf.org>, <https://www.jsmf.org/grants/2010022/>. The grant was awarded to award R. Guimerà and M. Sales-Pardo. It also received support from "Statistical Mechanics for Big Data: acquisition, analysis and modeling", Funder: MINECO (Spain), Project id: FIS2013-

## Abstract

Leadership positions are still stereotyped as masculine, especially in male-dominated fields (e.g., engineering). So how do gender stereotypes affect the evaluation of leaders and team cohesiveness in the process of team development? In our study participants worked in 45 small teams (4–5 members). Each team was headed by either a female or male leader, so that 45 leaders (33% women) supervised 258 team members (39% women). Over a period of nine months, the teams developed specific engineering projects as part of their professional undergraduate training. We examined leaders' self-evaluation, their evaluation by team members, and team cohesiveness at two points of time (month three and month nine, the final month of the collaboration). While we did not find any gender differences in leaders' self-evaluation at the beginning, female leaders evaluated themselves more favorably than men at the end of the projects. Moreover, female leaders were evaluated more favorably than male leaders at the beginning of the project, but the evaluation by team members did not differ at the end of the projects. Finally, we found a tendency for female leaders to build more cohesive teams than male leaders.

## Introduction

The impact of leader gender on the evaluation of leaders has been studied for quite some time (e.g., [1,2]). While researchers agree that there are negligible gender differences in leaders' performance [3], the debate on whether gender influences leader's self-evaluations and their evaluation by others is still on-going [4–6]. Past research using experimental laboratory studies and organizational studies has documented that female leaders are equally effective as male leaders [3,7], yet female leaders are evaluated slightly less favorably than male leaders (e.g., [1,8,9]).

One of the reasons why female leaders receive less positive evaluations is gender stereotyping: women and men are perceived differently in the context of leadership (e.g., [1–5]).

47532-C3, <http://www.mineco.gob.es/portal/site/mineco/>. The grant was awarded to award R. Guimerà and M. Sales-Pardo. "The roles of legitimacy and authority figure's gender in social perception" from Rectors' Conference of the Swiss Universities (CRUS), <https://www.swissuniversities.ch/de/themen/sciex/> also provided support. The grant was awarded to Agnieszka Pietraszkiewicz. "The Social Perception of Leaders—Does Gender Still Matter?" (#164512), <http://p3.snf.ch/project-164512>, <http://www.snf.ch/en/Pages/default.aspx>, was awarded to Agnieszka Pietraszkiewicz. The doctoral scholarship from the Universitat Rovira i Virgili and the mobility scholarship AEE2013 from the Universitat Rovira i Virgili was awarded to N. Rovira-Asenjo. A Research Fellowship "The roles of legitimacy and authority figure's gender in social perception" (#12.272) from Rectors' Conference of the Swiss Universities (CRUS) was awarded to Agnieszka Pietraszkiewicz. <https://www.swissuniversities.ch/de/themen/sciex/>; and a research grant "The Social Perception of Leaders—Does Gender Still Matter?" (#164512) from the Swiss National Science Foundation was awarded to Agnieszka Pietraszkiewicz. <http://p3.snf.ch/project-164512>, <http://www.snf.ch/en/Pages/default.aspx>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

Women in Western societies are generally seen as more communal (e.g., caring, emotional, helpful) than men, who are perceived as more agentic (e.g., active, decisive, dominant) (e.g., [6,7]). The stereotype of leaders is closer to the male stereotype, that is, more agentic than communal (see [8] for an overview), documented in numerous studies with different methodologies (e.g., [4]) and different sample populations (e.g., [9]). The mismatch between female stereotype and leader role ([1,10]) also has significant consequences for women's self-views. Women see themselves as less agentic and more communal than men (e.g., [11,12]) and as less suited for a leadership position [13]. In contrast to evaluations by others, female leaders tend to consider themselves less effective than male leaders [14].

The present research aims to determine the impact of gender stereotypes on leader evaluation in a real-life process of team development, focusing on the following questions: (1) Do gender stereotypes influence the self-evaluations of leaders? (2) Do gender stereotypes influence the way female and male leaders are evaluated by others? If so, does the impact of gender stereotypes on self-evaluations and evaluations by others remain stable or does it change in the process of team development, as increasing experience with the respective individual leads to a decrease in stereotyping? In addition, we are interested in the structure of the networks that develop around female and male leaders: (3) To what extent does a leader's gender affect social relationships within the team, specifically team cohesiveness? In dealing with these questions, the present research extends past findings in at least three distinct ways:

First, the present study assesses the impact of gender stereotypes on leader evaluation in a *real-life team building process*. The majority of earlier findings are based on laboratory settings (e.g., [15–18]). Only few studies have been conducted in organizational contexts (e.g., [19]), even though research settings are known to have a significant impact on leader evaluation (see [20] for a review). Since increased exposure to a particular person reduces the stereotyping of that person [21], studies based on hypothetical or laboratory leaders may yield stronger gender differences than studies on real-life superiors in actual organizations [19]. The present study therefore investigates leaders and team members that were involved in real projects as part of their academic curriculum. The project outcomes actually affected their academic records, a factor which is absent in laboratory settings. Thus, leader evaluation within this professional setting had ecological validity and can help to understand leader evaluation during team development.

Second, our study examines *longitudinal changes*, both in the self-evaluation of leaders and in the evaluation by members of their teams. The aim is to assess whether the impact of gender stereotypes remains stable or changes during team development. Although gender stereotypes are deeply rooted in society and therefore difficult to overcome [7,22], previous research has documented that increased exposure to an individual results in a decrease of stereotyping (e.g., [21,23]). During impression formation people simplify their evaluation of others by classifying them as members of certain groups (e.g., female individuals are automatically associated with the group of women and male individuals with the group of men, as individuating requires too much mental effort; see [21,24]). As a result, first reactions to individuals are based on stereotypes (e.g., [25,26]). If sufficient time and/or mental resources become available, however, individuation, takes place. Now perceivers base their impressions on attributes of the individuals they have become acquainted with and modify their initial judgments (e.g., [27]). In this way, the influence of gender stereotypes on the evaluation of women and men in leadership positions may decrease over time. To the best of our knowledge, the present study is the first piece of research addressing this question.

Finally, our study extends the present debate on the *role of gender for the functioning of the team*. The existing literature shows that various tasks (e.g., management or research) are increasingly accomplished by groups [28,29], hence it is important to examine the effect of

gender on team processes. Earlier studies on gender differences in leadership have mainly focused on members' evaluations of team leaders (e.g., [16,19]). Our study extends such investigations by examining the quality of relations between leaders and team members (i.e., the structure of teams headed by female and male leaders). Specifically, we determine team cohesiveness, that is, the density of connections between team members, and the clustering of connections around leaders (the formation of triangles of connections within the team that involve the leader). In this way, we assess whether leader gender affects the network structure of teams. So far, there is little research on leadership and social networks [30–32]. But the use of network measures to capture leadership dynamics is a growing field which lends itself to research with an additional gender perspective. The present study is the first that investigates the relationship between leader gender and team cohesiveness, and can therefore contribute to on-going debates on the role of gender in team performance (e.g., [33–35]).

### Leaders' self-evaluation and their evaluation by team members

Earlier research has shown that gender stereotypes may limit women's *self-evaluation*, for instance, their confidence in succeeding at male-typed tasks and occupations (e.g., [36,37]). This may damage their performance or their career aspirations (e.g., [38,39]). For instance, female students were found to underestimate their competence in performing a scientific task in comparison to male students [40]. Also, male leaders tended to overestimate their effectiveness, while female leaders perceived themselves to be at the same level as their subordinates [41,42]. This may be explained by the fact that men tend to have higher self-esteem than women (e.g., [43]) and self-esteem is predictive for how people evaluate their own performance [44,45]. However, leaders' self-evaluation may change during team development: Past research suggests that in organizational settings women gain self-confidence or self-esteem in achievement situations (e.g., [46,47]). Thus, female leaders' self-appraisal may come to be less affected by gender stereotypes in the process of team development and they may evaluate themselves more favorably over time.

**Hypothesis 1.** We assume that, due to the impact of gender stereotypes on women's self-views [40], female leaders evaluate themselves less favorably than male leaders at the beginning of the project (Time I); this gender difference will have disappeared at the end of the project (Time II).

Gender stereotypes also affect the *evaluation by others*. Due to the mismatch between traditional conceptions of leadership and the female stereotype there is a tendency to judge women as less suitable leaders than men [1,8,10,48,49]. Furthermore, female leaders tend to be evaluated less favorably than male leaders (e.g., [50]). The influence of gender stereotypes is especially important in impression formation. Thus, the less time respondents had to read a vignette about hypothetical leaders, the more favorable were their evaluations of male compared to female leaders [51]. Paradoxically, a recent meta-analysis of 95 studies documents that basically female leaders seem to be evaluated as more effective than male leaders [14]. The shifting standards model [52,53] offers an explanation for such favorable evaluations of female leaders. This model posits that people judge members of groups by category-specific standards, that is, they compare them to specific within-category standards. According to gender stereotypes, people judge male leaders as more capable than female leaders (e.g., [54,55]). Consequently, they tend to evaluate the leadership competence of a particular woman in relation to women's supposedly lower standards of competence, and the leadership competence of a particular man in relation to men's supposedly higher standards of competence [56]. As a result, a woman assigned as leader may be evaluated more favorably (due to the lower standards applied) than a man with equal competences in the same situation (due to higher standards).

Based on the recent meta-analytical findings [14], it is reasonable to assume that team members' evaluations of female leaders surpass evaluations of men at the beginning of the project, especially in a male-dominated field such as engineering as in the present research. That is, at the beginning of the projects female leaders may be evaluated based on their group membership (i.e., as tokens). Over time, however, as team members become increasingly familiar with their individual leader, female leaders may no longer benefit from within-category standards. Instead, team members may base their evaluations on between-category standards.

**Hypothesis 2.** We expect team members in a male-dominated field (engineering) to evaluate female leaders more favorably than male leaders at the beginning of the teamwork (Time I), due to within-category judgment [57,58]. This gender difference is assumed to disappear towards the end of the project (Time II) as a result of individuation [23,26,59].

## Cohesiveness in the process of team development

Cohesiveness is “the degree to which members are attracted to a group and motivated to remain part of it” [35]. Cohesiveness was found to be associated with, among other things, a decrease in team conflicts [60,61], innovation in teams [62], and improvement in individual performance [63,64]. Thus, informal connections between individuals can have important implications for the team as a whole, as they limit or facilitate the flow of resources between and within teams [65]. A meta-analysis of 37 studies documents that teams with dense networks of interpersonal ties are more successful in achieving their goals and are more motivated to stay together [66]. Nonetheless, the relationship between the density of connections within a team and team performance is still unclear: while some scholars found that the density of the network of informal social ties was positively associated with team performance [67,68], other studies did not confirm these findings [69].

Numerous authors have claimed that effective leadership combines the ability to “accurately perceive the network relations that connect people, and to actively manage these network relations” [70]. Recent findings have indeed documented an impact of network relationships on team effectiveness (e.g., [66]). Recent research by Post [71] indicates that teams with female leaders are more cohesive and more cooperative than teams with male leaders. This may be due to gender differences in other-orientation, as women were found to be more empathic (e.g., [72–74]), more supportive [75,76], more cooperative, collaborative, more oriented towards enhancing others' self-worth, and at the same time less hierarchically oriented than men [77–79]. Furthermore, female leaders are commonly expected to devote themselves to the common good of a team or a company in [5], an expectation which may contribute to an increased team cohesiveness in female-led teams. Therefore, it is plausible to assume that networks will become more cohesive around female leaders in the process of team development.

In our study, we followed the approach of Quintane and colleagues [80], who argued that there are two aspects that contribute to team cohesiveness: reciprocity in social interactions and closure (in other words, the presence of triangles of reciprocal interactions). Our analysis thus focuses on two factors: the cohesiveness of the network of connections within the team and the relative cohesiveness of the network of connections around the leader. Analysis of the former factor makes our study comparable to that of earlier analyses. Analysis of the latter factor enables us to exploit information on how network connections arrange themselves around the leader and to assess structural differences that result from leadership over time, particularly differences in the structural role of leaders. Analyzing the teams' networks also enables us to go beyond previous studies by assessing whether changes in evaluations are related to changes in the social structure of a team, specifically to changes in the role of the leader. Considering these aspects, we predict the following:

**Hypothesis 3.** Overall team cohesiveness and relative cohesiveness around the leader, as defined by the team's network of social interactions, do not differ for female- and male-led teams at the beginning of the projects (Time I). However, female-led teams become more cohesive around the leader than male led-teams towards the end of the projects (Time II) (see [71]).

## Methods

The data used in the study is exempt from Institutional Review Board review because of the following reasons: i) In the study we do not collect the data ourselves but we use existing data—team members-evaluations (i.e., other-evaluations and network data) and self-evaluations of team leaders (i.e., self-evaluations) that were routinely collected as part of the evaluation process of the Integrated Project (first year mandatory course) and the Team Leadership and Management (senior year course) since 2008 (our data is from 2010 to 2013) at the Universitat Rovira i Virgili, Spain. Students provided consent to participate in the courses mentioned before and to use this data for educational and research purposes; ii). The data was collected by the university that gave us access to an anonymized version of the data. Students cannot be identified, neither directly nor through identifiers linked to them. Their names are anonymized, and we have no demographic information other than gender. We do not have access to any information that could identify individual participants during or after data collection.

Our sample consists of engineering students who were involved in projects at their university. The purpose of these projects was to have the students apply the knowledge and skills acquired in different courses to solve an engineering problem (e.g., the pre-design of a chemical plant) over the course of nine consecutive months. The projects were headed by senior-year students who had been carefully selected for this program (see below). In the framework of the program, team members met with their leaders on a weekly basis to discuss the progress of their work, their achievements and future steps.

As an integral part of project evaluation, team members (including leaders) evaluated other team members and provided self-evaluation in an online survey that was administered twice. Collaboration of the teams always started in September. The first measurement (Time I) took place three months later in December. The second evaluation (Time II) took place in the ninth and final month of the collaboration. The evaluations were part of the course syllabus. Part of this data has already been used for other research purposes, namely for the prediction of conflicts in teams (see [81]).

The list of abbreviations used in the study is presented in [Table 1](#).

## Sample

The sample includes 45 small teams consisting of 3 to 7 student members; most teams had 5–6 members. In total, we analyzed data of 303 individuals, 45 of whom were team leaders (15 females; 30 males) and 258 team members (100 females; 158 males). The average proportion of females per team was 39%. In our sample, the gender composition of teams varied widely (proportion of females ranging from 0% to 75%). Although our sample was too small to investigate the effect of gender composition in a systematic and conclusive manner, we argue that gender composition is unlikely to have affected our findings to a large extent, because (i) we did not find significant differences between the gender composition of male- and female-led teams ( $U_{female} = 195.5$ ;  $p = .30$ ); (ii) the overall team composition is rather balanced within both male and female headed teams ( $Mdn = .5$ ).

Team members were first-year students of chemical engineering at a public Catalan university, enrolled in a mandatory course of their academic curriculum, while leaders were senior-

**Table 1. Table of abbreviations used in the manuscript.**

Abbreviation	Description
ANOVA	Analysis of Variance
<i>C</i>	Clustering coefficient
DP	D'Agostino and Pearson normality test statistic
<i>F</i>	<i>F</i> -ratio statistic (used in ANOVA)
ICC	Intra-class correlation coefficients
<i>M</i>	Mean
<i>Mdn</i>	Median
<i>N</i>	Sample size
<i>p</i>	Probability
<i>r</i>	Pearson's correlation coefficient
<i>RC</i>	Relative clustering coefficient
$r_{WG(j)}$	Inter-rater agreement index
<i>SD</i>	Standard error
<i>t</i>	<i>t</i> -test statistic
<i>U</i>	Mann-Whitney test statistic
<i>Z</i>	Wilcoxon signed-rank test statistic
$\eta^2$	Measure of effect size (used in ANOVA)
$\rho$	Spearman's rank correlation coefficient

<https://doi.org/10.1371/journal.pone.0186045.t001>

year students enrolled in a course on project management. In order to increase chances of success, the senior students underwent a selection process to become team leaders. All teams worked on similar projects. The participants did not know each other at the start of the project.

## Design

In the study, we measured the self-evaluation and other-evaluation of leaders at two stages of team development: at the beginning (Time I) and at the end of team development (Time II). The data was collected via online surveys over three academic years (the project teams worked together during one of the following academic years: 2010–2011, 2011–2012, and 2012–2013). Note that the overall evaluations of leaders were measured only in the last two academic years, hence the data set for this particular variable is smaller ( $N = 29$ ; 11 female leaders; 18 male leaders). All participants logged into a specially designed online platform and completed the survey individually. They were not paid for participation. The raw data is available in the Supporting Information (S1 Dataset).

## Leader evaluation measures

Leadership requires various specific skills such as coordination, time control, conflict management, or motivating team members to achieve objectives [82]. Therefore in the present study, leaders were asked to evaluate themselves regarding four aspects: “Please evaluate the following aspects of yourself: (1) management skills (coordination, task management, time control, etc.), (2) moderation skills (securing participation of all members, conflict management, public recognition of good work, etc.), (3) motivational skills (motivation of team members, stimulating change in team members' behavior to achieve objectives, etc.), and (4) empathy with team members.” Responses to each item were given on 11-point rating scales (0 = non-existent/poor, 10 = completely developed/excellent). We averaged individual responses for each leader and created self-evaluation indices (SE) for Time I and Time II (Cronbach's alpha = .69 and

.78 respectively). A factor analysis (principal axis) using direct oblimin rotation (see [83]) revealed that all items of the respective dimensions loaded sufficiently on the first factor (i.e.,  $> .30$ ), both for Time I and Time II.

We assessed the evaluation of leaders by team members regarding the same four aspects (OE). The instruction was as follows “Please evaluate the following aspects of *name of leader*”. Responses were again marked on 11-point rating scales (0 = non-existent/poor, 10 = completely developed/excellent). Subsequently, we computed the average evaluation by team members for each leader. Then we performed two principal factor analyses on this evaluation, one for Time I and another for Time II. The factor analysis showed that all items loaded sufficiently on one factor (i.e.,  $> .30$ ) for both points of time. We then created two reliable indices of other-evaluations of leaders (OE) for Time I and II (Cronbach’s alpha .95 and .90, respectively). In addition, the participants gave an overall evaluation of their leader (OOE). The question was as follows: “What is your overall evaluation of your leader?” Answers were given on an 11-point scale (0 = poor; 10 = excellent).

Taken together, we had one measure for the self-evaluation of leaders: the Self-Evaluation Index (SE), and two measures for the other-evaluation of leaders: the Other-Evaluation Index (OE) and the Overall Other-Evaluation Index (OOE).

**Internal consistency of teams.** To determine if team-level aggregation of our dependent variables was empirically justifiable, we took two precautions. First, we calculated within-group agreement ( $r_{WG(j)}$ ; [84]). Specifically, we computed the  $r_{WG(j)}$ ’s using the evaluations at Time I and Time II for two cases: (a) assuming an expected uniform distribution of ratings between 0 and 10; and (b) performing a random group resampling for each team [85]. For case (a) we found that for all teams  $r_{WG(j)}$ ’s fall within the desired range ( $Mdn = .985$ ). For case (b) we found that for 40 out of the 45 teams  $r_{WG(j)}$ ’s were within the desired range ( $Mdn = .917$ ). The other five teams had consistently lower-scoring leaders. In these teams, members consistently gave lower scores than in other teams but the ratings displayed a larger variability, therefore we did not exclude these teams from our analysis.

Second, we performed intra-class correlation coefficients tests [86]. For other-evaluation (OE) ICC(1) was .719,  $F(43,43) = 3.55$ ;  $p < .0001$ . Although no absolute standard value for aggregation based on  $r_{WG(j)}$  and ICC has been established, an  $r_{WG(j)}$  equal to or greater than .70 and ICC(1) values exceeding .05 [85] are considered sufficient to warrant aggregation. Based on these indicators, we aggregated the individual-level dependent variable measures to team-level variables [85].

**Network construction.** In the survey, each team member (including the leader) answered one yes/no question for every other team member: “Would you choose this person to work with you in a new team?” We used the answers to this question to construct the network of relationships within each team at Time I and Time II. Specifically, we constructed directed team networks for Times I and II. If team member A answered “yes” when asked about team member B, we drew an edge from A to B; otherwise, we did not. This question served as a measure of team structure.

*Notes on the statistical analysis.* Due to the small size of our sample, we paid particular attention to the statistical treatment of this data, especially to meeting the requirements for analyses of variance (ANOVA) or *t*-tests for related samples for sample comparison. To that end, we applied the D’Agostino and Pearson (DP) normality test [87] to our samples. For overall other evaluation (OOE) we found that for female OOE at Time II and female OOE change did not pass the DP normality test ( $p < .01$ ). For the Other-Evaluation Index (OE), we found that female OEs at Times I and II and male OEs at Time II were non-Gaussian (DP,  $p < .01$ ). All samples for self-reported evaluations passed the normality test. Finally, none of the samples for the network analysis passed the normality test. Therefore, as tests involving non-Gaussian

samples we used the Mann-Whitney U test for independent samples as the appropriate non-parametric test instead of ANOVA, and the Wilcoxon signed-rank test for matched samples instead of the  $t$ -test for related samples. In any case, our results are fairly independent of the tests we used. In order to analyze the correlations between variables, we first set up scatter plots for each pair of variables, most of which showed linear trends. We therefore tested the significance of the linear correlations between different variables by computing Pearson's  $r$  and obtaining its corresponding  $p$ -value. The results we report do not change if we compute the non-parametric Spearman's  $\rho$ .

## Results

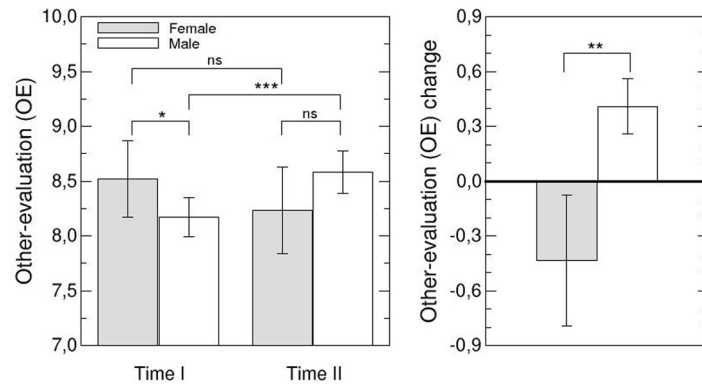
### Evaluations of female and male leaders

First, we analyzed the *self-evaluations of leaders (SE)*. We included the self-evaluations of all leaders who answered the questionnaire at both times of measurement ( $N = 36$ ; 14 female leaders; 22 male leaders). We conducted a 2 (gender of leader: female vs. male)  $\times$  2 (time: Time I vs. Time II) analysis of variance, with SE as dependent variable. Gender of leader was a between-subjects variable, time a within-subjects variable. This analysis revealed a significant main effect of time,  $F(1,34) = 10.112$ ,  $p = .003$ ,  $\eta^2 = .229$ : leaders evaluated themselves more favorably at Time II ( $M = 8.33$ ,  $SD = .80$ ) than at Time I ( $M = 7.94$ ,  $SD = .85$ ). The interaction between gender of leader and time did not reach the conventional level of significance,  $F(1,34) = 3.01$ ,  $p = .092$ ,  $\eta^2 = .081$ . However, we found that female leaders evaluated themselves significantly better at the end of the project ( $M = 8.71$ ,  $SD = .94$ ) than at the beginning ( $M = 8.02$ ,  $SD = .90$ ),  $t(13) = -2.32$ ,  $p = .037$  (see Hypothesis 1), whereas male leaders' self-evaluations did not change over the course of the project,  $t(21) < 1$ . Moreover, female leaders evaluated themselves more favorably at Time II than male leaders did ( $M = 8.71$ ,  $SD = .94$  vs.  $M = 8.1$ ,  $SD = .69$ ),  $t(36) = 2.45$ ,  $p = .019$ . For Time I, there was no significant difference,  $t(36) < 1$  (see Hypothesis 1).

To analyze the *evaluations of leaders by team members (OE and OOE)* we included only those participants who completed the survey at Time I and Time II, to avoid sample bias and to ensure the highest level of reliability in the aggregated data. We analyzed responses of 223 team members from 44 teams (15 headed by a female leader; 29 by a male leader). In order to interpret the other-evaluation of female and male leaders, we compared the other-evaluation index (OE) of male and female leaders for both times of measurement. Team members evaluated female leaders more favorably than male leaders at Time I ( $U = 152.5$ ,  $p = .055$ ) (see Hypothesis 2), but the difference was below the conventional level of significance. At Time II, however, they evaluated female and male leaders similarly ( $U = 189.5$ ,  $p = .248$ ) (see left plot of Fig 1 and Table 2) (see Hypothesis 2). These findings were supported by the analysis of overall other-evaluations (OOE) (using the mean evaluation each leader obtained from all team members,  $N = 29$ , 11 female team members, 18 male team members) at both times of measurement. At Time I, female leaders tended to be evaluated more favorably than male leaders,  $F(1,28) = 3.32$ ,  $p = .0799$ , but there was no gender difference at Time II ( $U = 89.0$ ,  $p = .490$ ) (see left plot of Fig 2 and Table 2) (see Hypothesis 2).

Additionally, we investigated whether self-evaluations were correlated with team member evaluations at Time I and Time II, for female and male leaders separately. For male leaders there was no significant relation between self- and other-evaluations, neither at Time I nor at Time II,  $r(27) = .198$ ;  $p = .36$  for Time I and  $r(27) = .38$ ,  $p = .66$  for Time II. For female leaders there was a significant positive correlation between self- and other-evaluations at Time II,  $r(13) = .537$ ,  $p < .048$ , but not at Time I,  $r(13) = .35$ ,  $p = .21$ . These results are consistent with earlier findings, which indicate that women are generally more accurate in their self-ratings





**Fig 1. Analysis of other-evaluation index (OE).** The left plot shows mean OEs at Time I and Time II for male and female leaders (see legend). Error bars indicate the mean standard error. The right plot shows OE changes for male and female leaders. Significance is defined as follows: ns—not significant; \*  $p < .1$ ; \*\*  $p < .05$ ; \*\*\*  $p < .001$ .

<https://doi.org/10.1371/journal.pone.0186045.g001>

than men, so that there is higher agreement between self- and other-ratings for female managers than for male managers [88].

Furthermore, we investigated changes in the other-evaluation of female and male leaders over time. Changes in the other-evaluation of female and male leaders were significantly different (OE:  $F(1,28) = 4.64, p = .041$  and OOE  $U = 48.0, p = .021$ ). This is due to the fact that male leaders were evaluated better at Time II than at Time I (OE:  $Z = 80.0, p = .005$ , OOE:  $t(17) = -2.18, p = .027$ ), while there were no changes in the evaluation of female leaders from Time I to Time II (OE:  $Z = 33.0, p = .153$ , OOE:  $Z = 13.0, p = .48$ ) (see Hypothesis 2).

### Cohesiveness in the process of team development

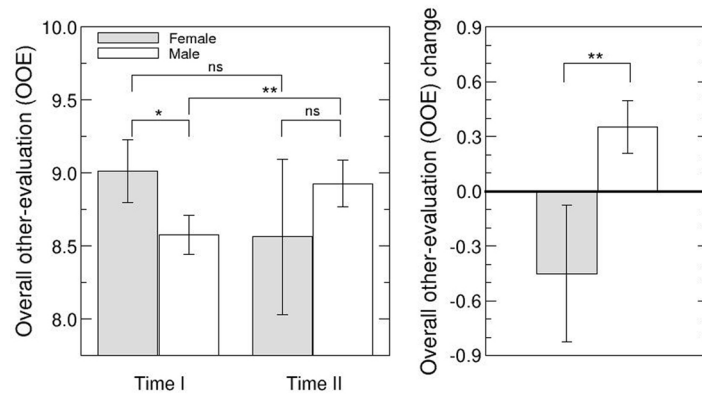
To determine team cohesiveness we analyzed the structure of all teams. Specifically, we examined two aspects: (i) the overall cohesiveness of the team and (ii) the relative cohesiveness of the network around the leader (see Hypothesis 3). For this purpose, we did not use any existing software but wrote our own code in Python and then used the `scipy.stats` module to perform the hypothesis-testing analyses, and the `networkx` module to compute network metrics. As mentioned above, we constructed team networks for Times I and II, using the yes/no answers to the question: “Would you choose this person to work with you in a new team?” To ensure the highest level of reliability in the aggregated data, we included only team members who completed the survey at both times of measurement (223 team members, 44 teams, 15 female leaders, 29 male leaders).

As outlined earlier, team cohesiveness is captured in network terms via reciprocity and closure. Reciprocity is the degree to which pairs of team members, A and B, wish to work with one another. We therefore constructed a network of reciprocal connections for each team (see Fig 3). We connected the team members A and B if there was a connection from A to B and a connection from B to A in the directed network. To quantify reciprocity we computed the

**Table 2. Gender differences in other-evaluation index (OE), overall other-evaluation index (OOE) and relative clustering (RC) at Time I (TI) and Time II (TII).**

Gender	OE TI	SD	OE TII	SD	OOE TI	SD	OOE TII	SD	RC TI	SD	RC TII	SD
Female	8.52	.35	8.23	.40	9.01	.22	8.56	.53	.92	.06	1.03	.13
Male	8.17	.18	8.58	.19	8.58	.13	8.93	.16	1.08	.08	.90	.11

<https://doi.org/10.1371/journal.pone.0186045.t002>

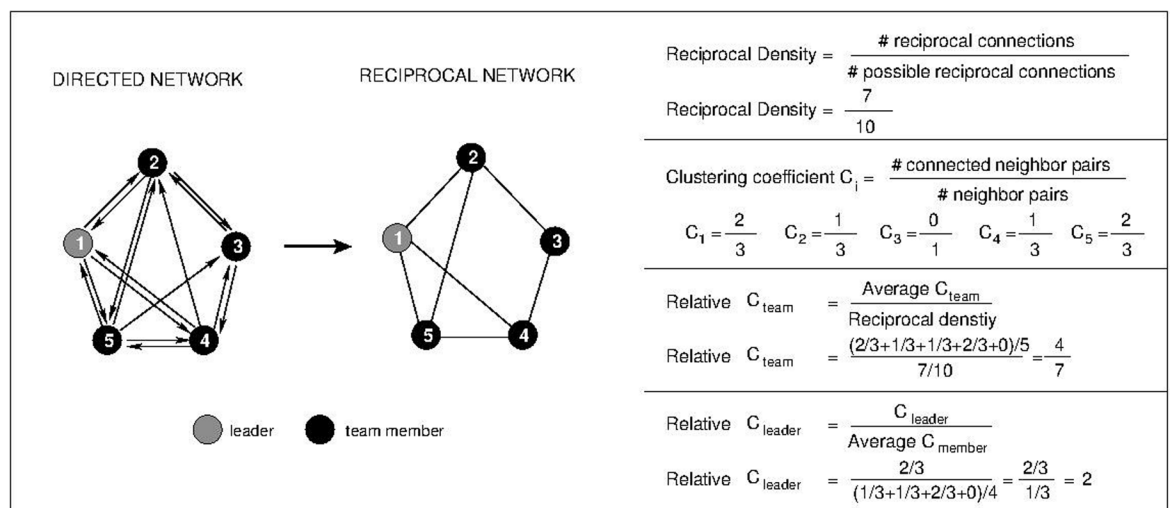


**Fig 2. Analysis of overall other-evaluation (OOE).** The left plot shows means for leaders' OOE at Time I and Time II for male and female leaders (see legend). Error bars indicate the mean standard error. The right plot shows OOE changes for males and females. Significance is defined as follows: ns—not significant; \*  $p < .1$ ; \*\*  $p < .05$ ; \*\*\*  $p < .001$ .

<https://doi.org/10.1371/journal.pone.0186045.g002>

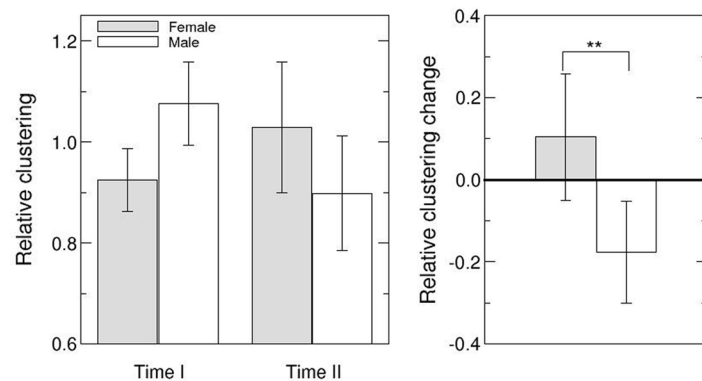
*reciprocal density*, i.e., the number of connected pairs out of all possible pairs of team members [89].

Closure in social networks refers to the existence of triangles in the social network of (reciprocal) interactions. Therefore, we computed the relative clustering coefficient of the team. More specifically, the clustering coefficient of team member A is the number of pairs of neighbors of A that are connected in the proportion of the team [90]. The clustering coefficient of the team is then computed as the average clustering coefficient of all team members. However, the larger the number of connections in a team, the larger is the clustering coefficient. Therefore, we needed to control for the density of connections to compare closure in different



**Fig 3. Illustration of social network analysis.** As an example of a directed network of relationships between pairs of neighbors taken from the survey (left diagram), we show how to obtain the reciprocal network and different network metrics: reciprocal density, the local clustering coefficient  $C_i$  for each team member, the relative clustering coefficient of the team  $C_{\text{team}}$  and the relative clustering of the leader  $C_{\text{leader}}$ . In the left diagram, the directed network shows only positive interactions ('Yes' answers in the survey). Black nodes mark regular team members, whereas the grey node indicates the team leader. If two team members A and B have positive interactions  $A \rightarrow B$  and  $B \rightarrow A$  in the directed network, then there is a reciprocal interaction between A and B (right diagram). From the reciprocal network of interactions, we computed the quantities of interest as indicated in the diagram. The symbol '#' indicates total numbers.

<https://doi.org/10.1371/journal.pone.0186045.g003>



**Fig 4. Analysis of the relative clustering (RC) around leaders.** The left plot shows mean RCs at Time I and Time II for male and female leaders (see legend). Error bars indicate the mean standard error. The right plot shows changes in RC for males and females. Note that the change in leaders' RC is significantly different for males and females ( $p < .05$ ).

<https://doi.org/10.1371/journal.pone.0186045.g004>

teams. To this end, we computed the *relative clustering coefficient of a team* (RCT), which is the average clustering coefficient divided by the reciprocal density (which equals the probability that two team members B and C that are connected with A are also connected with each other).

Interestingly, we found no differences in overall cohesiveness between female-led and male-led teams at Times I and II (Reciprocal density Time I:  $U = 137.5, p = .381$ ; Time II:  $U = 143.5, p = .460$ ; RCT Time I:  $U = 90.0, p = .012$ ; Time II:  $U = 90.0, p = .151$ ). Note that the significant differences at Time I are due to male-led teams having the same relative clustering coefficient of 1, which probably goes back to the low density of some of these teams; however, we believe that this is unlikely to be a general feature of male-led teams, which renders the difference not meaningful. In fact, the change in the relative clustering of teams was very small, both for male- and female-led teams, throughout our analysis (change for female leaders:  $Z = 3.0, p = 1.0$ ; change for male leaders:  $Z = 12.0, p = .735$ ).

Next, we investigated the cohesiveness around each team leader in order to capture differences that might have emerged over time due to a possible link between leader gender and team cohesiveness. Note that investigating team cohesiveness around the leader amounts to determining the centrality of the leader within the team in terms of reciprocity and closure. The former aspect is based on the density of reciprocal leader-member connections and the latter on the clustering of network interactions around the leader, that is, the number of triangles in the network involving the leader. To assess the relative cohesiveness of the network of connections around the leader, we used the *relative clustering coefficient* (RC) of the leader, which is the ratio between the clustering coefficient of the leader  $C_{leader}$  and the *average clustering coefficient* of the remaining team members (see Fig 4). Note that if  $RC > 1$ , interactions cluster more around the leader than around other team members, whereas in case of  $RC < 1$  team interactions cluster more around other team members than the leader. We found a significant difference in the change of relative clustering coefficients for male and female leaders ( $U = 55.5, p = .032$ , cf. Fig 4). The RC of female leaders tended to increase from being lower than that of team members ( $RC < 1$ ) to exceeding that of team members ( $RC > 1$ ), while the RC for male leaders showed the opposite trend. Our results thus indicate that only female leaders build teams in which team member relationships increasingly cluster around the leader in the process of team development.

We did not find significant differences between male and female leaders in any other measure of leader centrality in the network, probably because these measures do not capture traits that are characteristic of leadership styles (Table 2).

## Discussion

The present research provides first insights into the impact of gender stereotypes on self-evaluations of female and male leaders, on their evaluation by team members and on cohesiveness in the processes of team development in a male-dominated field, namely engineering. The strength of the present study lies not only in its longitudinal design, but also in the fact that it was conducted in a well-structured setting that bears resemblance to an authentic professional setting. The teams worked on comparable tasks (creating a chemical plant) over the same period of time (nine months), leaders and team members possessed similar amounts of experience and had access to similar resources offered by the university. Such standardized conditions would be very difficult to find in other organizational settings, where leaders with different experience work on different projects with different time frames (even within the same department).

At the onset of the projects female leaders were expected to evaluate themselves less favorably than male leaders, due to the impact of gender stereotypes on women's self-views [40]; this gender difference was assumed to disappear towards the end of the projects (Hypothesis 1). Contrary to our prediction, however, we did not find a gender difference in leaders' self-evaluations at the start of the project. One possible explanation could be that in our sample all project leaders had to pass a fairly strict selection process to become leaders and were appointed by persons in charge of the program. Being selected and appointed as leader may have strengthened the self-evaluation of female leaders and thus leveled the self-evaluations of women and men. Interestingly, the self-evaluation of female leaders even surpassed that of male leaders at the end of the project. This could be explained with a gain in self-confidence on the part of women. In other words, leading a team successfully may have boosted female leaders' self-esteem and led to the observed improvement in self-evaluation. However, this explanation is of a speculative nature and further research is needed to explain such shifts in female leaders' self-views.

As for the evaluation of leaders by others, team members showed the expected tendency to evaluate female leaders better than male leaders at the start of their joint work (see Hypothesis 2), in accord with the shifting standards model [56,91], whereas evaluations of female and male leaders became similar at the end (Time II; see Hypothesis 2). These results document a certain female advantage at the beginning of the project and confirm earlier findings [92]. Unexpectedly, however, we observed that the evaluation of male leaders by team members improved at the end of the project. There may be at least two explanations for this improvement: On the one hand, male leaders match the stereotypical image of a leader better than female leaders [1,10]. Hence, team members might still evaluate male leaders more favorably, and this male advantage might cumulate and intensify over time. In this case individuation would not have occurred and the influence of the stereotype would have increased (rather than decreased; e.g., [91,93]) in the process of team development. On the other hand, male leaders may not only have been perceived differently but also may have behaved differently, for example, in a more agentic way than female leaders. In male-dominated fields such as engineering, team members may value and expect a more agentic leadership style (e.g., more goal-orientation) (e.g., [94]). Although male leaders were evaluated slightly less favorably at the beginning of the projects in the present research, they may have displayed the agentic behaviors consistent with the leader stereotype and may therefore have received an extra plus in the

final evaluations (e.g., [14]). Anyway, female and male leaders were evaluated similarly at the end of the projects in the present study.

Finally, drawing on earlier research on gender differences in team cohesiveness (e.g., [71]), we expected female leaders to create teams that showed more cohesiveness at the end of the projects (Hypothesis 3). Indeed, we found a tendency for female leaders to get connections to cluster around them more than male leaders—although there were no differences in overall team cohesiveness between teams with male and female leaders. Our findings are in line with previous research [71] and suggest that female leaders may be more focused on building and sustaining relationships [95] and that they promote interaction styles that support and maintain social exchanges [96], which translates into a higher cohesiveness of teams led by female compared to male leaders.

Our findings have broader theoretical implications. So far, there has been little research on long-term changes in the perception of female and male leaders ([91] is an exception). Our study provides empirical evidence for existing theoretical positions by documenting that a female advantage in the evaluation of leaders decreased over time in a natural setting. Furthermore, we found that the teams supervised by female leaders were characterized by more connections between leaders and team members than the teams with male leaders. This may go back to a tendency of female leaders to emphasize collaboration and teamwork and to use a participative and interpersonally oriented leadership style (see the meta-analyses by [79,97,98]). A stronger connection between female leaders and their team members may be especially beneficial when groups have to solve socially complex tasks which, for example, require discussion, negotiation or coordination between functionally diverse or geographically dispersed units ([71]; see also [99,100]). Therefore, our study is a first step towards capturing gender differences in team network structures in a real-life setting. However, more research is needed to examine whether gender differences in team structure exist in other settings as well (e.g., organizational settings).

Moreover, future studies should take leaders' self-construal in their theoretical reasoning into consideration. Nowadays self-evaluations of women and men are becoming more and more similar. For instance, female and male business students were found to ascribe themselves task- and person-oriented traits (agentive and communal traits, respectively) to a similar extent [101]. Therefore, the gender differences in leadership styles exhibited by women and men may have been smaller than expected, and capturing them with the help of network structure may have been difficult. Our assumption that different leadership styles are favored by female and male leaders and that they translate into team structure hence warrants more testing. For instance, previous studies have documented that cohesiveness translates into better team performance (e.g., [102,103]); theoretical explanations should therefore focus more on the conditions under which other- vs. task-orientation yields better communication and cohesiveness [71].

Finally, some practical implications of the study should be mentioned. First, female leaders' self-evaluation was similar to male leaders' self-evaluation at the start of the project and became more favorable in the course of the team development. This suggests that female leaders experienced a gain in self-confidence. Being appointed a leader by third parties (e.g., selection authorities) may not only have prevented women from evaluating themselves less favorably in the beginning, but may also have reduced the influence of gender stereotypes on women's self-views as they developed their leader role. Second, our results show a certain advantage of female leaders at the onset of the projects. In line with the shifting standards model [56], team members seem to have measured the leadership competence of individual female leaders against lower standards of competence and that of individual male leaders

against higher standards of competence. However, once female leaders had taken up their role as leaders, this tendency diminished.

Finally, our findings suggest that female leaders may be more effective in organizational settings requiring high networking skills of leaders. In teams with a history of conflict or communication issues, for example, the female leadership style might be more effective at preserving a close-knit core structure within the team that warrants its functioning.

## Limitations and future studies

The fact that our study was conducted in close collaboration with a public university enabled us to gather data in an ongoing university program over three academic years. However, it also implied some limitations. Firstly, the data relied on the pre-existing measures applied by the university. Future research should examine the role of time in self- and other-evaluations of leaders with the help of established scales (e.g., [104–107]). Secondly, while our study extends knowledge on the role of gender in the process of team development and points to potential underlying mechanisms, subsequent research should systematically investigate possible underlying mechanisms such as self-construal, gender group identification, or leadership styles. Thirdly, due to ethical regulations at the participating university we had no access to any objective measure of effectiveness, neither for the leaders nor for the teams (e.g., grades that leaders received for their work). Thus, we do not know whether leaders' self-evaluation and their evaluation by team members or cohesiveness in the team were associated with a better outcome of the projects. Fourthly, due to the similarity in the gender composition of all teams (i.e., slightly more males) and the relatively small sample, we were not able to analyze the impact of the teams' gender composition on the self- and other-evaluation of leaders and on team cohesiveness. Such an impact can be expected, because the little research there indicates differences in how female and male subordinates evaluate female and male leaders (e.g., [50,108]), with men being more likely than women to attribute successful manager characteristics to women [109]. Previous findings also suggest that gender diversity can have negative (e.g., [110,111]) as well as positive effects (e.g., [112–114]) on team processes and outcomes.

## Conclusions

The present research indicates that team leaders' gender plays a role for certain evaluations: In our study, female leaders tended to evaluate themselves more favorably at the end of a project than at the beginning, and their self-evaluations were more favorable than those of men at the end. Team members tended to evaluate female leaders better than male leaders at the onset of the collaboration. In the process of team development, however, this tendency faded away. At a later stage, leaders' evaluations by team members seemed to be determined by individuation rather than gender stereotyping [23,26,59]. In addition, networks headed by women showed a slightly higher cohesiveness than those headed by men. These findings expand our knowledge on stereotypical biases in the perception of female and male leaders: We were able to demonstrate that the influence of stereotypes on self- and other-evaluations of leaders changes in the process of team development. Analyzing network structure, we also showed that female leaders, even in a highly male-dominated context, may be more successful in building teams that are more cohesive.

## Supporting information

**S1 Dataset. Raw data obtained from the responses of leader evaluations by other team members (spreadsheet 1) and from the self-evaluation survey of the leaders (spreadsheet**

2). Spreadsheet 3 contains the dictionary of the data. (XLS)

## Author Contributions

**Conceptualization:** Núria Rovira-Asenjo, Agnieszka Pietraszkiewicz, Sabine Sczesny, Tània Gumí, Roger Guimerà, Marta Sales-Pardo.

**Data curation:** Núria Rovira-Asenjo, Marta Sales-Pardo.

**Formal analysis:** Núria Rovira-Asenjo, Agnieszka Pietraszkiewicz, Marta Sales-Pardo.

**Investigation:** Núria Rovira-Asenjo, Tània Gumí, Roger Guimerà, Marta Sales-Pardo.

**Methodology:** Núria Rovira-Asenjo, Sabine Sczesny, Tània Gumí, Roger Guimerà, Marta Sales-Pardo.

**Project administration:** Agnieszka Pietraszkiewicz.

**Resources:** Núria Rovira-Asenjo, Marta Sales-Pardo.

**Supervision:** Sabine Sczesny, Marta Sales-Pardo.

**Writing – original draft:** Núria Rovira-Asenjo, Agnieszka Pietraszkiewicz, Sabine Sczesny, Marta Sales-Pardo.

**Writing – review & editing:** Núria Rovira-Asenjo, Agnieszka Pietraszkiewicz, Sabine Sczesny, Marta Sales-Pardo.

## References

1. Heilman ME. Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *J Soc Issues*. 2001; 57: 657–674. <https://doi.org/10.1111/0022-4537.00234>
2. Eagly AH, Karau SJ. Gender and the emergence of leaders: A meta-analysis. *J Pers Soc Psychol*. 1991; 60: 685–710. <https://doi.org/10.1037/0022-3514.60.5.685>
3. Eagly AH, Sczesny S. Stereotypes about women, men, and leaders: Have times changed? The glass ceiling in the 21st century: Understanding barriers to gender equality. Washington: American Psychological Association; 2009. pp. 21–47. <https://doi.org/10.1037/11863-002>
4. Koenig AM, Eagly AH, Mitchell AA, Ristikari T. Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychol Bull*. 2011; 137: 616–642. <https://doi.org/10.1037/a0023557> PMID: 21639606
5. Ryan MK, Haslam SA, Hersby MD, Bongiorno R. Think crisis-think female: the glass cliff and contextual variation in the think manager-think male stereotype. *J Appl Psychol*. American Psychological Association; 2011; 96: 470–484. <https://doi.org/10.1037/a0022133> PMID: 21171729
6. Abele AE. The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *J Pers Soc Psychol*. American Psychological Association; 2003; 85: 768–776. <https://doi.org/10.1037/0022-3514.85.4.768> PMID: 14561129
7. Fiske ST, Stevens LE. What's so special about sex? Gender stereotyping and discrimination. In: Oskamp S, Costanzo M., editors. *Gender issues in contemporary society: Applied social psychology annual*. Newbury Park, CA: SAGE Publications; 1993. pp. 173–196.
8. Eagly AH, Karau SJ. Role congruity theory of prejudice toward female leaders. *Psychol Rev*. 2002; 109: 573–598. <https://doi.org/10.1037/0033-295X.109.3.573> PMID: 12088246
9. Dodge K, Gilroy F, Fenzel L. Requisite management characteristics revisited: Two decades later. *J Soc Behav*. 1995; Available: <http://search.proquest.com/openview/3ae8b0a91d924160bbac527082cb8d46/1?pq-origsite=gscholar&cbl=1819046>
10. Heilman ME. Sex bias in work settings: The Lack of Fit model. *Res Organ Behav*. 1983; 5: 269–298. Available: <http://psycnet.apa.org/psycinfo/1984-10927-001>
11. Bem SL. The measurement of psychological androgyny. *J Consult Clin Psychol*. 1974; 42: 155–62. Available: <http://www.ncbi.nlm.nih.gov/pubmed/4823550> PMID: 4823550

12. Spence JT, Buckner CE. Instrumental and Expressive Traits, Trait Stereotypes, and Sexist Attitudes: What Do They Signify? *Psychol Women Q.* 2000; 24: 44–53. <https://doi.org/10.1111/j.1471-6402.2000.tb01021.x>
13. Bosak J, Sczesny S. Gender Bias in Leader Selection? Evidence from a Hiring Simulation Study. *Sex Roles.* 2011; 65: 234–242. <https://doi.org/10.1007/s11199-011-0012-7>
14. Paustian-Underdahl SC, Walker LS, Woehr DJ. Gender and Perceptions of Leadership Effectiveness: A Meta-Analysis of Contextual Moderators. *J Appl Psychol.* American Psychological Association; 2014; 99: 1129–1145. <https://doi.org/10.1037/a0036751> PMID: 24773399
15. Garcia-Retamero R, López-Zafra E. Prejudice against women in male-congenial environments: Perceptions of gender role congruity in leadership. *Sex Roles.* Kluwer Academic Publishers-Plenum Publishers; 2006; 55: 51–61. <https://doi.org/10.1007/s11199-006-9068-1>
16. Heilman M. E., Wallen A. S., Fuchs D. & Tamkins MM. Penalties for Success: Reactions to Women Who Succeed at Male Gender-Typed Tasks. *J Appl Psychol.* 2004; 89: 416–427. <https://doi.org/10.1037/0021-9010.89.3.416> PMID: 15161402
17. Johnson SK, Murphy SE, Zewdie S, Reichard RJ. The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organ Behav Hum Decis Process.* 2008; 106: 39–60. <https://doi.org/10.1016/j.obhdp.2007.12.002>
18. Parks-Stamm EJ, Heilman ME, Hearn KA. Motivated to Penalize: Women's Strategic Rejection of Successful Women. *Personal Soc Psychol Bull.* 2008; 34: 237–247. <https://doi.org/10.1177/0146167207310027> PMID: 18212332
19. Elsesser KM, Lever J. Does gender bias against female leaders persist? Quantitative and qualitative data from a large-scale survey. *Hum Relations.* SAGE PublicationsSage UK: London, England; 2011; 64: 1555–1578. <https://doi.org/10.1177/0018726711424323>
20. Ragins BR. Gender effects in subordinate evaluations of leaders: Real or artifact? *J Organ Behav.* 1991; 12: 259–268. <https://doi.org/10.1002/job.4030120308>
21. Fiske ST. Stereotyping, prejudice, and discrimination. In: DT Gilbert & ST Fiske, editor. *The handbook of social psychology.* 1997. pp. 357–411. 10.4324/9780203567708
22. Spence J. Gender-related traits and gender ideology: evidence for a multifactorial theory. *J Pers Soc Psychol.* 1993; Available: <http://psycnet.apa.org/journals/psp/64/4/624/>
23. Brewer MB. A dual process model of impression formation. *Adv Soc Cogn.* 1988; 1: 1–36.
24. Miller AG. *In the Eye of the Beholder: Contemporary Issues in Stereotyping.* Praeger Publishers, editor. New York, NY; 1982. <https://doi.org/10.2307/2294751>
25. Tajfel H. Cognitive Aspects of Prejudice. *J Soc Issues.* 1969; 25: 79–97. <https://doi.org/10.1111/j.1540-4560.1969.tb00620.x>
26. Fiske ST, Neuberg SL. A continuum model of impression formation: Influences of information and motivation on attention and interpretation. *Advances in experimental social psychology.* 1990. pp. 1–74. Available: <http://www.sciencedirect.com/science/article/pii/S0065260108603172>
27. Heilman M. Information as a deterrent against sex discrimination: The effects of applicant sex and information type on preliminary employment decisions. *Organ Behav Hum Perform.* 1984; Available: <http://www.sciencedirect.com/science/article/pii/0030507384900199>
28. Wuchty S, Jones BF, Uzzi B. The increasing dominance of teams in production of knowledge. *Science (80-).* 2007; 316: 1036–9. <https://doi.org/10.1126/science.1136099> PMID: 17431139
29. Gowers T, Nielsen M. Massively collaborative mathematics. *Nature.* 2009; Available: <http://www.nature.com/nature/journal/v461/n7266/full/461879a.html>
30. Hoppe B, Reinelt C. Social network analysis and the evaluation of leadership networks. *Leadersh Q.* 2010; 21: 600–619. <https://doi.org/10.1016/j.leaqua.2010.06.004>
31. Friedrich T, Griffith J, Mumford M. Collective leadership behaviors: Evaluating the leader, team network, and problem situation characteristics that influence their use. *Leadersh Q.* 2016; Available: <http://www.sciencedirect.com/science/article/pii/S104898431600014X>
32. Brass DJ, Galaskiewicz J, Greve HR, Tsai W. Taking stock of networks and organizations: A multilevel perspective. *Academy of Management Journal.* 2004. pp. 795–817. <https://doi.org/10.2307/20159624>
33. Bear JB, Woolley AW. The role of gender in team collaboration and performance. *Interdiscip Sci Rev.* Taylor & Francis; 2011; 36: 146–153. <https://doi.org/10.1179/030801811X13013181961473>
34. Appelbaum SH, Audet L, Miller JC. Gender and leadership? Leadership and gender? A journey through the landscape of theories. *Leadersh Organ Dev J.* MCB UP Ltd; 2003; 24: 43–51. <https://doi.org/10.1108/01437730310457320>
35. Wendt H, Euwema MC, van Emmerik IJH. Leadership and team cohesiveness across cultures. *Leadersh Q.* 2009; 20: 358–370. <https://doi.org/10.1016/j.leaqua.2009.03.005>



36. Heatherington L, Daubman KA, Bates C, Ahn A, Brown H, Preston C. Two investigations of “female modesty” in achievement situations. *Sex Roles*. 1993; 29: 739–754. <https://doi.org/10.1007/BF00289215>
37. Sieverding M. Frauen unterschätzen sich: Selbstbeurteilungs-Biases in einer simulierten Bewerbungssituation. *Zeitschrift für Sozialpsychologie ZFSP*. 2003; Available: <http://archiv.ub.uni-heidelberg.de/volltextserver/19703/>
38. Davies PG, Spencer SJ, Steele CM. Clearing the Air: Identity Safety Moderates the Effects of Stereotype Threat on Women’s Leadership Aspirations. *J Pers Soc Psychol*. 2005; 88: 276–287. <https://doi.org/10.1037/0022-3514.88.2.276> PMID: 15841859
39. Davies PG, Spencer SJ, Quinn DM, Gerhardstein R. Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally. *Personal Soc Psychol Bull*. 2002; 28: 1615–1628. <https://doi.org/10.1177/014616702237644>
40. Ehrlinger J, Dunning D. How chronic self-views influence (and potentially mislead) estimates of performance. *J Pers Soc Psychol*. 2003; 84: 5–17. <http://dx.doi.org/10.1037/0022-3514.84.1.5> PMID: 12518967
41. Vecchio RP, Anderson RJ. Agreement in self-other ratings of leader effectiveness: The role of demographics and personality. *Int J Sel Assess*. 2009; 17: 165–179. <https://doi.org/10.1111/j.1468-2389.2009.00460.x>
42. Brutus S, Fleenor JW, McCauley CD. Demographic and personality predictors of congruence in multi-source ratings. *J Manag Dev*. 1999; 18: 417–435. <https://doi.org/10.1108/02621719910273569>
43. Kling KC, Hyde JS, Showers CJ, Buswell BN. Gender differences in self-esteem: a meta-analysis. *Psychol Bull*. 1999; 125: 470–500. <https://doi.org/10.1037/0033-2909.114.1.29> PMID: 10414226
44. Lindeman M, Sundvik L, Rouhiainen P. Under- or overestimation of self? Person variables and self-assessment accuracy in work settings. *J Soc Behav Personal*. 1995; 10: 123–134. Available: <http://search.proquest.com/openview/334ea902e9bca5141e0997eab2d165a0/1?pq-origsite=gscholar&cbl=1819046>
45. Jussim L, Coleman L, Nassau S. The influence of self-esteem on perceptions of performance and feedback. *Soc Psychol Q*. 1987; 50: 95–99. <https://doi.org/10.2307/2786894>
46. McCarty PA. Effects of Feedback on the Self-Confidence of Men and Women. *Acad Manag J*. 1986; 29: 840–847. <https://doi.org/10.2307/255950>
47. Johnson M, Helgeson VS. Sex Differences In Response To Evaluative Feedback: A Field Study. *Psychol Women Q*. 2002; 26: 242–251. <https://doi.org/10.1111/1471-6402.00063>
48. Eagly AH, Johannesen-Schmidt MC. The Leadership Styles of Women and Men. *J Soc Issues*. 2001; 57: 781–797. <https://doi.org/10.1111/0022-4537.00241>
49. Bosak J, Sczesny S. Am I the Right Candidate? Self-Ascribed Fit of Women and Men to a Leadership Position. *Sex Roles*. 2008; 58: 682–688. <https://doi.org/10.1007/s11199-007-9380-4>
50. Eagly AH, Makhijani MG, Klonsky BG. Gender and the Effectiveness of Leaders: A Meta-Analysis. *Psychol Bull*. 1992; 111: 3–22. <https://doi.org/10.1037/0033-2909.117.1.125>
51. Martell RF. Sex Bias at Work: The Effects of Attentional and Memory Demands on Performance Ratings of Men and Women. *J Appl Soc Psychol*. Blackwell Publishing Ltd; 1991; 21: 1939–1960. <https://doi.org/10.1111/j.1559-1816.1991.tb00515.x>
52. Biernat M. Toward a Broader View of Social Stereotyping. *Am Psychol*. 2003; 58: 1019–1027. <https://doi.org/10.1037/0003-066X.58.12.1019> PMID: 14664690
53. Biernat M. Standards and expectancies: Contrast and assimilation in judgments of self and others. *Essays in social psychology*. 2005. <https://doi.org/10.4324/9780203338933>
54. Eagly AH, Karau SJ, Makhijani MG. Gender and the effectiveness of leaders: A meta-analysis. *Psychol Bull*. American Psychological Association; 1995; 117: 125–145. <https://doi.org/10.1037/0033-2909.117.1.125> PMID: 7870858
55. Eagly A, Carli L. The female leadership advantage: An evaluation of the evidence [Internet]. *The leadership quarterly*. 2003. Available: <http://www.sciencedirect.com/science/article/pii/S1048984303000584>
56. Biernat M, Fuegen K. Shifting Standards and the Evaluation of Competence: Complexity in Gender-Based Judgment and Decision Making. *J Soc Issues*. 2001; 57: 707–724. <https://doi.org/10.1111/0022-4537.00237>
57. Biernat M, Manis M, Nelson TE. Stereotypes and standards of judgment. *J Pers Soc Psychol*. 1991; 60: 485–499. <https://doi.org/10.1037/0022-3514.60.4.485>
58. Biernat M, Manis M. Shifting standards and stereotype-based judgments. *J Pers Soc Psychol*. 1994; 66: 5–20. <https://doi.org/10.1037/0022-3514.66.1.5> PMID: 8126651

59. Neuberg SL, Fiske ST. Motivational influences on impression formation: outcome dependency, accuracy-driven attention, and individuating processes. *J Pers Soc Psychol.* 1987; 53: 431–44. <https://doi.org/10.1037/0022-3514.53.3.431> PMID: 3656080
60. Jehn KA, Mannix EA. The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Acad Manag J. Academy of Management;* 2001; 44: 238–251. <https://doi.org/10.2307/3069453>
61. Nibler R, Harris KL. The Effects of Culture and Cohesiveness on Intragroup Conflict and Effectiveness. *J Soc Psychol.* 2003; 143: 613–631. <https://doi.org/10.1080/00224540309598467> PMID: 14609056
62. Mumford MD, Hunter ST. Innovation in Organizations: A Multi-Level Perspective on Creativity. *Research in Multi-Level Issues.* 2005. pp. 9–73. [https://doi.org/10.1016/S1475-9144\(05\)04001-4](https://doi.org/10.1016/S1475-9144(05)04001-4)
63. Langfred C. Is group cohesiveness a double-edged sword? An investigation of the effects of cohesiveness on performance. *Small Gr Res.* 1998; Available: <http://sgr.sagepub.com/content/29/1/124.short>
64. Chang A, Bordia P. A Multidimensional Approach to the Group Cohesion-Group Performance Relationship. *Small Gr Res.* 2001; 32: 379–405. <https://doi.org/10.1177/104649640103200401>
65. Brass DJ. Being in the Right Place: A Structural Analysis of Individual Influence in an Organization. *Adm Sci Q.* 1984; 29: 518–539. <https://doi.org/10.2307/2392937>
66. Balkundi P, Harrison DA. Ties, leaders, and time in teams: strong inference about network structure's effects on team viability and performance. *Acad Manag J.* 2006; 49: 49–68. <https://doi.org/10.5465/AMJ.2006.20785500>
67. Cummings JN, Cross R. Structural Properties of Work Groups and Their Consequences for Performance. *Soc Networks.* 2003; 25: 197–210. Available: <http://www.sciencedirect.com/science/article/pii/S0378873302000497>
68. Reagans R, Zuckerman EW. Networks, Diversity, and Productivity: The Social Capital of Corporate R & D Teams. *Organ Sci.* 2001; 12: 502–517. <https://doi.org/10.1287/orsc.12.4.502.10637>
69. Sparrowe RT, Liden RC, Wayne SJ, Kraimer ML. Social networks and the performance of individuals and groups. *Acad Manag J.* 2001; 44: 316–325. <https://doi.org/10.2307/3069458>
70. Balkundi P, Kilduff M. The ties that lead: A social network approach to leadership. *The Leadership Quarterly.* 2006. pp. 419–439. <https://doi.org/10.1016/j.leaqua.2006.01.001>
71. Post C. When is female leadership an advantage? Coordination requirements, team cohesion, and team interaction norms. *J Organ Behav.* 2015; 36: 1153–1175. <https://doi.org/10.1002/job.2031>
72. Fukushima H, Hiraki K. Perceiving an opponent's loss: gender-related differences in the medial-frontal negativity. *Soc Cogn Affect Neurosci.* 2006; 1: 149–157. <https://doi.org/10.1093/scan/nsi020> PMID: 18985125
73. Feingold A. Gender differences in personality: A meta-analysis. *Psychol Bull.* 1994; 116: 429–456. <https://doi.org/10.1037/0033-2909.116.3.429> PMID: 7809307
74. Cundiff NL, Komaraju M. Gender differences in ethnocultural empathy and attitudes toward men and women in authority. *J Leadersh Organ Stud.* 2008; 15: 5–15. <https://doi.org/10.1177/1548051808318000>
75. Tannen D. Gender differences in topical coherence: Creating involvement in best friends' talk. *Discourse Process.* Taylor & Francis Group; 1990; 13: 73–90. <https://doi.org/10.1080/01638539009544747>
76. Aries E. Men and women in interaction: Reconsidering the differences. Oxford University Press, editor. Oxford University Press; 1996.
77. Gabriel S, Gardner WL. Are there "his" and "hers" types of interdependence? The implications of gender differences in collective versus relational interdependence for affect, behavior, and cognition. *J Pers Soc Psychol.* 1999; 77: 642–655. <https://doi.org/10.1037/0022-3514.77.3.642> PMID: 10510513
78. Book EW. *Why the Best Man for the Job is a Woman—The Uniques Female Qualities of Leadership.* New York, NY: Harper Collins; 2009.
79. Eagly AH, Johnson BT. Gender and leadership style: A meta-analysis. *Psychol Bull.* 1990; 108: 233–256. <https://doi.org/10.1037/0033-2909.108.2.233>
80. Quintane E, Pattison PE, Robins GL, Mol JM. Short- and long-term stability in organizational networks: Temporal structures of project teams. *Soc Networks.* 2013; 35: 528–540. <https://doi.org/10.1016/j.socnet.2013.07.001>
81. Rovira-Asenjo N, Gumí T, Sales-Pardo M, Guimerà R. Predicting future conflict between team-members with parameter-free models of social networks. *Sci Rep.* 2013; 3: 1999. <https://doi.org/10.1038/srep01999> PMID: 23770685

82. Nohria N, Khurana R. Advancing leadership theory and practice. *The handbook for leadership theory and practice*. Harvard Business Press; 2010. pp. 3–26. Available: <http://www.hbs.edu/faculty/Pages/item.aspx?num=37554>
83. Fabrigar LR, Maccallum RC, Wegener DT, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods*. 1999; 4: 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
84. James LR, Demaree RG, Wolf G. Estimating Within-Group Interrater Reliability With and Without Response Bias. *J Appl Psychol*. 1984; 69: 85–98. <https://doi.org/10.1037/0021-9010.69.1.85>
85. Bliese PD. Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In: Klein KJ, Kozlowski SWJ, editors. *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. 2000. pp. 349–381. Available: <http://psycnet.apa.org/psycinfo/2000-16936-008>
86. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. SAGE Publications; 2002.
87. D’Agostino R, Pearson ES. Tests for departure from normality. Empirical results for the distributions of  $b^2$  and  $\sqrt{b^2}$ . *Biometrika*. Oxford University Press; 1973; 60: 613–622. <https://doi.org/10.1093/biomet/60.3.613>
88. Yammarino FJ, Atwater LE. Do managers see themselves as other see them? Implications of self-other rating agreement for human resources management. *Organ Dyn*. 1997; 25: 35–44. [https://doi.org/10.1016/S0090-2616\(97\)90035-8](https://doi.org/10.1016/S0090-2616(97)90035-8)
89. Wasserman S, Faust K. *Social network analysis: Methods and applications*. Cambridge Univ Press. 1994; 1: 116. <https://doi.org/10.1525/ae.1997.24.1.219>
90. Watts DJ, Strogatz SHH. Collective dynamics of “small-world” networks. *Nature*. Nature Publishing Group; 1999; 393: 440–442. <https://doi.org/10.1038/30918>
91. Biernat M, Crandall CS, Young L V, Kobryniewicz D, Halpin SM. All that you can be: stereotyping of self and others in a military context. *J Pers Soc Psychol*. 1998; 75: 301–317. <https://doi.org/10.1037/0022-3514.75.2.301> PMID: 9731310
92. Eagly AH, Carli LL. The female leadership advantage: An evaluation of the evidence. *Leadersh Q*. 2003; 14: 807–834. <https://doi.org/10.1016/j.leaqua.2003.09.004>
93. Oakes PJ, Haslam SA, Morrison B, Grace D. Becoming an In-Group: Reexamining the Impact of Familiarity on Perceptions of Group Homogeneity. *Soc Psychol Q*. 1995; 58: 52. <https://doi.org/10.2307/2787143>
94. Davison HK, Burke MJ. Sex discrimination in simulated employment contexts: A meta-analytic investigation. *J Vocat Behav*. 2000; 56: 225–248. <https://doi.org/10.1006/jvbe.1999.1711>
95. Trinidad C, Normore AH. Leadership and gender: a dangerous liaison? *Leadersh Organ Dev J*. 2005; 26: 574–590. <https://doi.org/10.1108/01437730510624601>
96. Maccoby EE. Gender and relationships: A developmental account. *Am Psychol*. American Psychological Association; 1990; 45: 513–520. <https://doi.org/10.1037/0003-066X.45.4.513> PMID: 2186679
97. Van Engen ML, Willemsen TM. Sex and leadership styles: a meta-analysis of research published in the 1990s. *Psychol Rep*. 2004; 94: 3–18. <https://doi.org/10.2466/pr0.94.1.3-18> PMID: 15077742
98. Eagly AH, Johannesen-Schmidt MC, Van Engen ML. Transformational, Transactional, and Laissez-Faire Leadership Styles: A Meta-Analysis Comparing Women and Men. *Psychol Bull*. 2003; 129: 569–591. <https://doi.org/10.1037/0033-2909.129.4.569> PMID: 12848221
99. Wood W, Wendy. Meta-analytic review of sex differences in group performance. *Psychol Bull*. American Psychological Association; 1987; 102: 53–71. <https://doi.org/10.1037/0033-2909.102.1.53> PMID: 3615706
100. Wood W, Polek D, Aiken C. Sex Differences in Group Task Performance. *J Pers Soc Psychol*. 1985; 48: 63–71. <https://doi.org/10.1037/0022-3514.48.1.63>
101. Sczesny S, Bosak J, Neff D, Schyns B. Gender stereotypes and the attribution of leadership traits: A cross-cultural comparison. *Sex Roles*. 2004; 51: 631–645. <https://doi.org/10.1007/s11199-004-0715-0>
102. Chioocchio F, Essiembre H. Cohesion and Performance A Meta-Analytic Review of Disparities Between Project Teams, Production Teams, and Service Teams. *Small Group Research*. 2009. <https://doi.org/10.1177/1046496409335103>
103. Beal DJ, Cohen RR, Burke MJ, McLendon CL. Cohesion and Performance in Groups: A Meta-Analytic Clarification of Construct Relations. *J Appl Psychol*. 2003; 88: 989–1004. <https://doi.org/10.1037/0021-9010.88.6.989> PMID: 14640811

104. Rosen B, Jerdee TH. Effects of applicant's sex and difficulty of job on evaluations of candidates for managerial positions. *J Appl Psychol*. 1974; 59: 511–512. <https://doi.org/10.1037/h0037323>
105. Rice RW, Instone D, Adams J. Leader Sex, Leader Success, and Leadership Process: Two Field Studies. *J Appl Psychol*. 1984; 69: 12–31. <https://doi.org/10.1037/0021-9010.69.1.12>
106. Bartol KM, Butterfield DA. Sex effects in evaluating leaders. *J Appl Psychol*. 1976; 61: 446–454. <https://doi.org/10.5465/AMBPP.1974.17530791> PMID: 947881
107. Jago AG, Vroom VH. Sex Differences in the Incidence and Evaluation of Participative Leader Behavior. *J Appl Psychol*. 1982; 67: 776–783. <https://doi.org/10.1037/0021-9010.67.6.776>
108. Powell GN, Butterfield DA, Parent JD. Gender and Managerial Stereotypes: Have the Times Changed? *J Manage*. 2002; 28: 177–193. <https://doi.org/10.1177/014920630202800203>
109. Duehr EE, Bono JE. Men, women, and managers: Are stereotypes finally changing? *Pers Psychol*. 2006; 59: 815–846. <https://doi.org/10.1111/j.1744-6570.2006.00055.x>
110. Van Dijk H, Van Engen ML, Van Knippenberg D. Defying conventional wisdom: A meta-analytic examination of the difference between demographic and job-related diversity relationships with performance. *Organ Behav Hum Decis Process*. 2012; 119: 38–53.
111. Allmendinger J, Hackman JR. The More, the better? A four-nation study of the inclusion of women in symphony orchestras. *Soc Forces*. 1995; 74: 423–460. <https://doi.org/10.1093/sf/74.2.423>
112. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* (80-). 2010; 330: 686–688. <https://doi.org/10.1126/science.1193147> PMID: 20929725
113. Post C, Byron K. Women on boards and firm financial performance: A meta-analysis. *Acad Manag J*. 2015; 58: 1546–1571. <https://doi.org/10.5465/amj.2013.0319>
114. Kochan T, Bezrukova K, Ely R, Jackson S, Joshi A, Jehn K, et al. The effects of diversity on business performance: Report of the diversity research network. *Hum Resour Manage*. 2003; 42: 3–21. <https://doi.org/10.1002/hrm.10061>