



Women Who Emerge as Leaders in Temporarily Assigned Work Groups: Attractive and Socially Competent but Not Babyfaced or Naïve?

Freya M. Gruber, Carina Veidt and Tuulia M. Ortner*

Department of Psychology, University of Salzburg, Salzburg, Austria

OPEN ACCESS

Edited by:

Alice H. Eagly,
Northwestern University,
United States

Reviewed by:

Leslie Zebrowitz,
Brandeis University, United States
Minna Paunova,
Copenhagen Business School,
Denmark

*Correspondence:

Tuulia M. Ortner
tuulia.ortner@sbg.ac.at

Specialty section:

This article was submitted to
Personality and Social Psychology,
a section of the journal
Frontiers in Psychology

Received: 18 May 2018

Accepted: 28 November 2018

Published: 19 December 2018

Citation:

Gruber FM, Veidt C and
Ortner TM (2018) Women Who
Emerge as Leaders in Temporarily
Assigned Work Groups: Attractive
and Socially Competent but Not
Babyfaced or Naïve?
Front. Psychol. 9:2553.
doi: 10.3389/fpsyg.2018.02553

The underrepresentation of women in top positions has been in the spotlight of research for decades. Prejudice toward female leaders, which decreases women's chances of emerging as leaders, has been discussed as a potential reason. Aiming to investigate the underlying mechanisms of this prejudice, we focused on the question of how facial characteristics might influence women's leadership emergence. Because other research has related ascribed social competence and ascribed naïveté to attractiveness and babyfacedness, respectively, we hypothesized that ascribed social competence would mediate the impact of ascribed attractiveness on leadership emergence and that ascribed naïveté would mediate the impact of ascribed babyfacedness on leadership emergence. In a pilot study, we analyzed data from 101 participants of a women's leadership contest held in 2015 in Germany. We then confirmed these results in a methodologically improved main study on other women who participated in the contest in one of two other years: 2016 and 2017 ($N = 195$). Women applied to participate in the contest by recording their answers to several questions in a video interview. In the contest, they were assigned to teams of about ten women each and worked on several assessment-center-like tasks. After each task, each member of each team nominated the three women they believed showed the best leadership potential in their group. We operationalized women's leadership emergence as the number of nominations received. We measured participants' facial attractiveness, babyfacedness, social competence, and naïveté by having raters follow a specifically developed rating manual to rate the answers the women gave in the video interviews. In both studies, the results indicated that women with higher ascribed facial attractiveness had higher ascribed social competence, which significantly predicted leadership emergence in the contest. Likewise, women with higher ascribed babyfacedness had higher ascribed naïveté, which significantly, albeit only slightly, negatively predicted leadership emergence. We discuss the implications of the results for personnel selection.

Keywords: leadership emergence, babyfacedness, attractiveness, social competence, naïveté

INTRODUCTION

In modern Western societies, sex-related barriers to occupational success still exist and are reflected in economic data. Women hold only 28% of the CEO positions in the United States (U.S. Bureau of Labor Statistics, 2017) and 16% of the 1,500 board seats of the S&P (Ernst and Young LLP, 2015). Women occupy fewer top executive positions than men in the European Union (Bourgeois, 2017) and fewer full professorships worldwide (Johnson, 2017). Because a lack of qualified women can no longer explain these circumstances (European Commission, 2012; Johnson, 2017), it is important to ask why women rarely hold these top leadership positions. One possible explanation for this inequity can be found in differences in leadership emergence. Taggar et al. (1999) defined leadership emergence as the process by which one team member becomes the leader of an initially leaderless group without holding the position by title or by being previously assigned to it. According to Paunova (2015), one or more leaders may emerge within a group when they are perceived and acknowledged as a leader, for instance, due to the positive effect of their leader behavior on the group during a task (Bass and Bass, 2008). Systematic gender differences in leadership emergence may create obstacles for women because being perceived as not able to emerge as a leader may prevent them from being hired or promoted (e.g., Eagly and Karau, 1991, 2002; Finkelstein et al., 2018). Therefore, it is important to further investigate and understand the underlying mechanisms that can help women emerge as leaders or prevent them from doing so.

Scientists have formulated several theories to explain differences in perceptions of the leadership potential of women and men and differences in their leadership emergence: As a basic principle, social role theory (Eagly, 1987) suggests that differences in the behavior of women and men and also differences in how women and men are perceived in a society stem primarily from the different distributions of men's and women's social roles. In this theory, different roles occupied by men and women have an impact on sex-related stereotypes and ascribed aspects of personality because people tend to make inferences from these roles to associated characteristics. In fact, various studies of different fields have demonstrated relations between social roles and the ascribed characteristics of people (e.g., Diekmann and Schneider, 2010; Ortner et al., 2011; Koenig and Eagly, 2014). Following this reasoning, Eagly and Karau (2002) formulated role congruity theory to further explain women's greater difficulties in reaching leadership positions. According to this theory, individuals receive positive evaluations by others when their characteristics are estimated as confirming the group's typical social roles (Eagly and Diekmann, 2005). In fact, "many people perceive [incongruity] between the characteristics of women and the requirements of leader roles" (Eagly and Karau, 2002, p. 574). This perceived incompatibility between stereotypes of a social group and the requirements for fulfilling a specific role is supposed to result in prejudice, backlash effects (Rudman, 1998), and low evaluations of women as actual or potential leaders (see Eagly and Karau, 2002).

Another related, albeit more general theory that has contributed to explaining inequities in society on the basis of

social group processes is the expectation states theory (Wagner and Berger, 1997; Weyer, 2007). This theory addresses small, task-oriented work groups and highlights the relevance of status characteristics (e.g., gender, age, ethnicity) but also physical attributes for the ascription of task-relevant characteristics. Carli and Eagly (2001) employed this approach to explain differences in the leadership emergence of men and women, proposing that there are biases in evaluations of female leaders. Ridgeway (2001) described this approach as applying expectation states theory in a new context. Accordingly, gender roles are enrooted in the social hierarchy and in leadership because the core of gender stereotypes are cross-cultural schemas about women's status positions, also called status beliefs. These status beliefs have been proposed to be a basic reason for lower leadership emergence in women than in men because men, rather than women, are associated with general competence and status worthiness. However, a recent analysis integrating 16 different opinion polls revealed an increase for women's ascribed competence over the past years and concluded gender equality in ascribed competence for women and men in the United States (see Eagly et al., 2018).

Nevertheless, several studies have provided empirical support for the validity of these theoretical frameworks in explaining differences in the performances of women and men with reference to leadership: For example, Garcia-Retamero and López-Zafra (2006) conducted an experiment to test how prejudice against female leaders in different work environments stems from people's expectations. To do so, they asked non-leader participants to rate vignettes of hypothetical candidates for a leadership position. The data indicated prejudice against women except when the woman worked "in an industry congruent with her gender role" (p. 58). Further results indicated effects of social attribution by illustrating that incongruence between masculine task demands and the female gender role reduced leadership emergence even in dominant women (see Ritter and Yoder, 2004).

A more complex and new model combining aspects of personality traits with behavioral mechanisms was recently proposed along with a meta-analysis on leadership emergence: Badura et al. (2018) addressed the question of why men more frequently emerge as leaders than women. For this purpose, the authors proposed gender, individual characteristics, behavioral aspects, as well as situational factors as determinants of the relation between gender and leadership emergence. In order to test their gender-agency/communion-participation (GAP) model, they coded 1,632 effect sizes. Results indicated a beneficial effect of people's gender on leadership emergence through agency (e.g., assertiveness and dominance), but a detrimental effect through communion (e.g., kindness and nurturance). Furthermore, the relations between these traits and leadership emergence were mediated through participative behavior in group discussions. Moreover, they hypothesized that gender differences in leadership emergence depended on situational factors such as the study's setting (business settings, lab settings, classroom settings), gender egalitarianism in the nation where the data were collected, the length of interaction, and the degree of social complexity of the task. Their results indicated that the gender gap in leadership emergence was significantly

lower in business settings compared with lab settings and significantly larger when the interaction time in the task was shorter.

Besides the influence of agency, communion, and behavioral aspects, physical appearance may also contribute to leadership emergence. Much research has revealed that a person's physical appearance is a relevant characteristic in various areas of life, for example, partner selection (Langlois et al., 2000; for a review, see Rhodes, 2006), political elections (Olivola and Todorov, 2010), judicial decisions (Efran, 1974; Zebrowitz and McDonald, 1991), decisions regarding personnel selection (Zebrowitz et al., 1991; Shannon and Stark, 2003), and also ascribed leadership competences (Sczesny and Kühnen, 2004) or leadership selection processes (Stoker et al., 2016). Attractiveness represents one of the most salient aspects of physical appearance and, therefore, affects the ways in which individuals perceive and evaluate others (Little et al., 2011). Several theoretical explanations have also been proposed to account for the relevance of physical attractiveness in the field of leadership: Implicit personality theory by Bruner and Tagiuri (1954) postulated that people unconsciously build hypothetical constructs of trait elements and inferential relations between these attributes when assessing other people (see Schneider, 1973). On the basis of their meta-analysis of 76 studies, Eagly et al. (1991) identified the association between "beauty" and "goodness" as a physical attractiveness stereotype. They reported a significantly higher attribution of favorable characteristics to attractive individuals in comparison with unattractive individuals. Highly relevant for the domain of leadership emergence, individuals rated as attractive received higher evaluations with reference to social competence, adjustment, potency, and intellectual competence. This finding is in line with a study that reported findings that individuals perceived targets whose faces they had rated as attractive with respect to facial symmetry and proportions as more successful, satisfied, and likable than people whose faces they had rated as unattractive (Braun et al., 2001). In fact, studies have revealed positive effects of ascribed attractiveness on life and business outcomes such as career success and management level (Dietl, 2013), income (Judge et al., 2009), and also ascribed leadership status (Cherulnik et al., 1990). According to Cherulnik (1995), the positive impact of facial attractiveness on leader emergence even accumulates over a lifetime because others' attributions create a favorable distinctive social environment that provides special experiences and opportunities. Individuals judged as attractive were also found to be more confident in their social skills than individuals judged as unattractive (Cherulnik, 1995). Moreover, Sobieraj and Krämer (2014) demonstrated a strong relation between ascribed social competence and ascribed physical attractiveness in a study with experimentally varied virtual avatars. On the basis of these findings, we hypothesized that *ascribed facial attractiveness would positively predict leadership emergence (Hypothesis 1)*. Moreover, we hypothesized that *social competence would positively mediate the impact of facial attractiveness on leadership emergence (Hypothesis 2)* because social competence represents a determinant of leadership skills (Riggio et al., 2003; Groves, 2005; Riggio and Reichard, 2008).

However, not only facial attractiveness affects the ways in which individuals perceive and describe others. The anatomy literature indicates sex differences in facial features (e.g., Enlow, 1982; Gray and Standring, 2008) with female faces having more in common with infantile, babyface-like facets compared with male faces (Guthrie, 1976). Lorenz (1943) defined this set of immature facial characteristics (e.g., big round eyes, a big head, a round face, or a small nose) as the so-called baby schema (German: "Kindchenschema") or babyfacedness. In an evolutionary approach, Perrett et al. (1998) attributed the preference of humans for feminine facial features to the correlations of estrogen-dependent traits with aspects of health and reproductive fitness. Nevertheless, the costs of this form of attractiveness were addressed by investigating whether the more babyish faces of women compared with men further contribute to the attribution of sex-related characteristics and stereotypes. Presenting equally mature-faced male and female faces weakened stereotyped ascriptions and led to the conclusion that women's average facial features promote ascriptions of stereotyped characteristics (Friedman and Zebrowitz, 1992).

In line with these findings, Eagly and Karau's (2002) role congruity theory provides a framework that can be applied to formulate hypotheses on the effects of babyfacedness on leadership emergence: Whereas the leader role is predominantly characterized by agentic traits, such as being competitive, self-confident, aggressive, objective, and ambitious (see Koenig et al., 2011), people possessing a babyfaced physiognomy were described as less physically strong, less socially dominant, less astute (Zebrowitz McArthur and Apatow, 1983/1984), more naïve, and weaker compared with people possessing more mature faces (Keating et al., 2003; for a review, see Montepare and Zebrowitz, 1998). According to these findings, possessing a babyface is inconsistent with the characteristics required by a leadership role. A growing number of studies focusing on babyfacedness have supported this assumption: For instance, studies have indicated a negative relation between babyfacedness and inferred competence among politicians (Poutvaara et al., 2009), and a negative relation between babyfacedness and leadership status in a high school senior class (Cherulnik et al., 1990). Furthermore, Cherulnik (1995) found in an experiment that facial maturity was a significant predictor of leadership emergence. Investigations of hiring preferences have revealed similar findings (Zebrowitz et al., 1991): Undergraduate participants preferred male and mature-faced applicant targets for leadership positions over female and babyfaced applicants, especially for higher-status positions.

It seems plausible that a person with a high level of ascribed naïveté (defined as a collective term that summarizes how naïve, inexperienced, credulous, and generally intelligent an individual is perceived to be) due to his or her facial appearance would find it difficult to meet leader role expectations. Women may be particularly inclined to suffer from this misperception because they are more likely to encounter prejudice due to role incongruence in the first place (Eagly and Karau, 2002). In addition, women are more likely to be judged as babyfaced (Friedman and Zebrowitz, 1992; Chiao et al., 2008; Olivola and Todorov, 2010). In the second part of this study, we therefore

hypothesized that *ascribed babyfacedness would negatively predict leadership emergence (Hypothesis 3)*. Moreover and analogous to Hypothesis 2, we investigated whether these babyface-specific attributions could explain this negative association: Therefore, we hypothesized that *ascribed naïveté would negatively mediate the impact of ascribed babyfacedness on leadership emergence (Hypothesis 4)*.

To summarize the aims of the present study, we intended to investigate the impact of facial characteristics on women's chances of emerging as a leader, and further, whether the ascription of other attributes that are based on facial characteristics would mediate this effect in the setting of a women's leadership contest. On the one hand, we hypothesized that ascribed attractiveness would be a positive predictor of leadership emergence and that ascribed social competence would mediate this relation. On the other hand, we hypothesized that ascribed babyfacedness would be a negative predictor of leadership emergence and that the ascription of naïveté would mediate this relation. Data collected during contests that were held exclusively for women in three consecutive years provided a basis of the analyses: 2015–2017. The predictor variables (ascribed attractiveness and ascribed babyfacedness) and the mediator variables (ascribed social competence and ascribed naïveté) comprised raters' evaluations of women's appearance in video interviews from self-recorded videos from the contest application phase. Peer nominations provided by group members after participating in assessment-center-like tasks at the contests served as the dependent variable leadership emergence. Therefore, we present a pilot study, which uses data from the 2015 contest, and then the main study containing a larger sample (data from the 2016 and 2017 contests) and a revised method.

GENERAL METHODS

The analyses based on data collected across three consecutive years of a women's leadership contest conducted in Germany.¹ Data collection took place in two phases: In the *application phase*, women applied to participate in the contests by submitting their curriculum vitae and letters of recommendation as well as by using a professional software platform to record a video interview consisting of their answers to five or six questions. For each question, applicants could prepare their answer within 2 min. Moreover, participants were instructed to record each of their responses in a maximal duration of 3 min. The women who passed this application phase on the basis of the quality and persuasive power of all of their application documents were invited to participate in the actual leadership contest. At the *leadership contest* – between attendance at keynote presentations, workshops, and networking slots – contest participants took part in 2 to 3 assessment-center-like tasks in which they engaged in group tasks or group discussions or prepared speeches in groups of 7 to 12 women each. After each task, the women nominated the three group members who had been most convincing as

leaders. Contest participants who received the most nominations by their peers were rewarded with vouchers, for example, for travel tickets or career coaching. Participants' data were included in our study only when they agreed to the anonymous analysis of their data at the time when they applied to participate in the contest.

Measures and Variables – General Information

Leadership Emergence

Contest participants competed for peer nominations in each group task in short-term work groups. After each task, women completed a questionnaire to (a) nominate and (b) describe the three group members who had convinced them the most with respect to the group member's potential to act as a leader. Leadership emergence was operationalized as the total number of nominations each woman received from her team members across the entire contest. The more nominations a woman received in the contest, the higher her leadership emergence index. However, we controlled for group size, that is, peer nominations possible in the contest, in the main study.

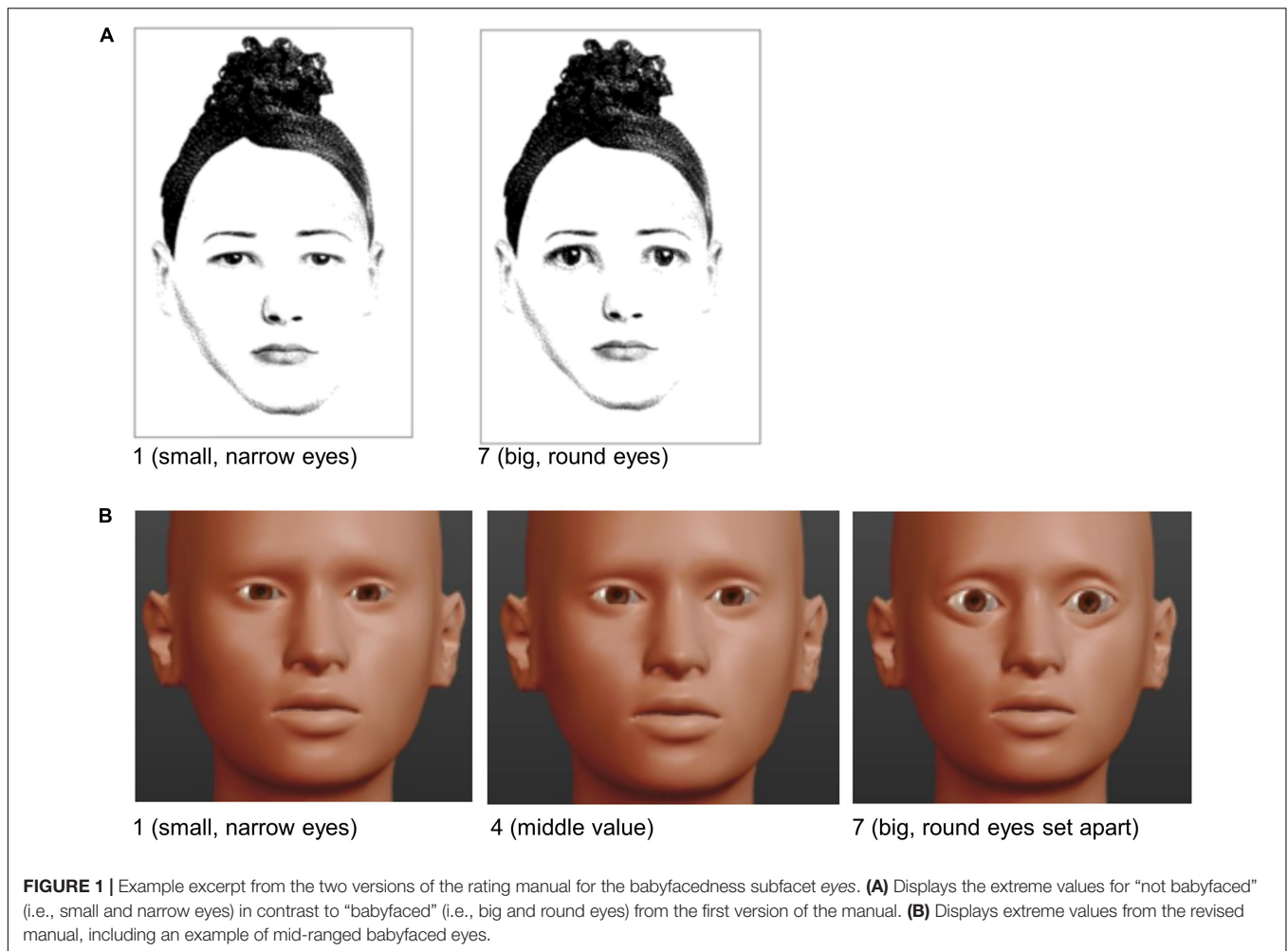
Ascribed Babyfacedness and Ascribed Facial Attractiveness

Raters (three in the pilot study, nine in the main study) evaluated women's faces by watching their application videos. Raters watched a 25-s cut of one interview question without sound in order to focus solely on the participants' outward appearance and to avoid any undue influence of other aspects (e.g., spoken content or voice). The length of the video sequence was determined on the basis of pretests as well as literature suggesting that a rating of appearance was possible after this amount of time (see Ambady and Rosenthal, 1993).

After the video sequence had finished, raters assessed the women's facial attractiveness and babyfacedness. For each babyfacedness item, a rating manual depicted example pictures for the extreme values in advance (see **Figure 1** for an example of the eyes facet). The means of the attractiveness items (three in the pilot study: attractive, good looking, and pleasing; two in the main study: attractive, good looking; all ranging from 1 = low value to 10 = high value) averaged across all raters served as the predictor variable *ascribed facial attractiveness*. The predictor variable *ascribed babyfacedness* resulted of the mean ratings of the four babyfacedness items (nose, eyes, face form, overall impression, ranging from 1 = not babyfaced to 7 = very babyfaced) averaged across all raters.

Before raters gave their evaluations, they were trained on the basis of a rating manual. In the pilot study, this manual contained black-and-white pictures for the extreme values of the babyfacedness items (for an example, see **Figure 1A**). Before conducting the main study, the manual was revised in accordance with information obtained from the pilot study. Changing the instructions from “How would you evaluate the women shown in the videos. . .” to “How would most people evaluate the women shown in the videos compared with women of the same age?” targeted to reduce subjectivity in the raters' impressions. The revised manual further contained higher quality pictures for the

¹ Visit www.we-are-panda.com for more information about the contests.



visual attributes of the babyfacedness items by using the freeware program MakeHuman (version 1.1.0, MakeHuman team, 2016) replacing lower quality pictures from the former manual version. This procedure allowed the image of detailed face avatars to represent the extreme and middle values of each item (for an example, see **Figure 1B**). We further revised or removed single items (see below; see the revised manual here²).

Ascribed Social Competence and Ascribed Naïveté

After having evaluated the women’s facial characteristics, the raters watched the videos (see section “Ascribed Babyfacedness and Ascribed Facial Attractiveness”) a second time, uncut and with the sound on, and estimated the women’s social competence and naïveté. The mean of all items per scale averaged across all raters served as variables further used to test for mediation effects. Ascribed social competence consisted of six items in the pilot study (sociable, confident, warm-hearted, popular, empathic, and socially competent) and of five items in the main study (verbally skilled, sociable, confident, anxious [negatively coded], socially competent; all ranging from 1 = low value to 7 = high value). Ascribed naïveté consisted of three items in the pilot study (naïve,

mature [negatively coded], critical thinker [negatively coded]), and of four items in the main study (naïve, inexperienced, gullible, generally intelligent [negatively coded]; all ranging from 1 = do not agree at all to 7 = strongly agree).

Demographic Data

Age, professional experience, and leadership experience in years were obtained from the CVs the women submitted with their applications.

PILOT STUDY

Methods

Participants

In sum, 109 women who participated in the contest in 2015 (out of a total of 187 women) gave their consent to use information from their CVs and interviews for research purpose and were included. Videos of eight contest participants were excluded due to insufficient picture or sound quality. The ages of the final sample of 101 women ranged from 22 to 53 years ($M = 31.96$, $SD = 6.68$). On average, these women had 9.73 years ($SD = 6.04$) of professional experience and 6.22 years ($SD = 4.54$) of leadership

²<https://osf.io/sf72x/>

experience. Furthermore, 79.2% were German citizens, 5.0% other, and 15.8% did not provide information about their nationality.

Materials

In the pilot study, three raters evaluated facial features and characteristics using the interview sequences from the contest application phase in 2015.³

Procedure

Leadership emergence

In the leadership contest, the women worked on three assessment-center-like tasks. The participants were randomly and differently assigned to groups of 10 to 12 women per each task. The challenges in the tasks were (a) to engage in a 50-min group discussion about the skills needed for successful leadership, (b) to take 80 min to build a wooden construction, including several predefined intermediate goals, and (c) to take 50 min to prepare a political speech for a female executive politician. Each participant gave written nominations of the three most convincing group peers after each task (see section “Leadership Emergence”). Leadership emergence was calculated on the basis of the General Information (see section “Leadership Emergence”).

Facial characteristics and ascribed traits

The ascribed facial attractiveness scale ($\alpha = 0.98$), the ascribed babyfacedness scale ($\alpha = 0.78$), the ascribed social competence scale ($\alpha = 0.85$), and the ascribed naïveté scale ($\alpha = 0.88$) all demonstrated sufficient internal consistency. The ICC values were acceptable, ranging from 0.63 to 0.82 (see section “Measures and Variables – General Information” for all items and procedure).

Statistical Analyses

To test Hypotheses 1–4, we calculated mediation analysis using the “lavaan” package (version 0.6-1, Rosseel, 2018) in R (version 3.5.0). In order to control for non-normality of our data, we conducted robust ULS estimators. Data and R scripts are published in the Open Science Framework (OSF⁴).

Results

Descriptive statistics for the dependent and independent variables are presented in **Table 1**, and the correlations between the variables are presented in **Table 2**. Results of the mediation analysis computed to test Hypotheses 1 and 2, displayed in **Figure 2A**, revealed that ascribed facial attractiveness was a significant predictor of ascribed social competence ($\beta = 0.52$, $SE = 0.08$, $p < 0.001$) and that ascribed social competence was a significant predictor of leadership emergence ($\beta = 0.28$, $SE = 0.12$, $p = 0.018$). A bootstrapping estimation approach with 1,000 bootstrapped samples indicated a significant indirect effect ($\beta_{\text{indirect}} = 0.15$, 95% CI [0.02, 0.27], $p = 0.025$). Yet, ascribed facial attractiveness did not significantly predict leadership emergence before ($\beta_{\text{total}} = 0.04$, $SE = 0.10$, $p = 0.706$) or after ($\beta_{\text{direct}} = -0.11$, $SE = 0.11$, $p = 0.351$) adding the mediator, ascribed social

competence. Because there was a significant indirect effect (i.e., the higher a contest participant’s ascribed facial attractiveness, the higher her ascribed social competence; and the higher her ascribed social competence, the higher her leadership emergence), we concluded that ascribed social competence acted as a mediator between ascribed facial attractiveness and leadership emergence, even though the analysis did not reveal a significant total effect.⁵

The results of the mediation analysis for Hypotheses 3 and 4 indicated that ascribed babyfacedness was a significant predictor of ascribed naïveté ($\beta = 0.26$, $SE = 0.10$, $p = 0.008$) and that ascribed naïveté was a significant predictor of leadership emergence ($\beta = -0.33$, $SE = 0.09$, $p < 0.001$). The bootstrapping estimation approach with 1,000 bootstrapped samples indicated a significant indirect effect ($\beta_{\text{indirect}} = -0.09$, 95% CI [-0.17, -0.02], $p = 0.035$). Ascribed babyfacedness did not significantly predict leadership emergence before ($\beta_{\text{total}} = -0.10$, $SE = 0.10$, $p = 0.336$) or after adding the mediator, ascribed naïveté ($\beta_{\text{direct}} = -0.01$, $SE = 0.10$, $p = 0.920$). The results indicated a mediating effect of ascribed naïveté, which means that the higher a participant’s ascribed babyfacedness, the higher her ascribed naïveté; and the higher her ascribed naïveté, the lower her leadership emergence, as **Figure 2B** illustrates.

MAIN STUDY

Methods

Participants

We included data from a total of 195 contest participants from 2016 (111 of the 178 participants gave consent) and 2017 (84 of the 114 participants gave consent) in our main study. The ages of the contest participants ranged from 21 to 52 ($M = 32.81$, $SD = 6.69$). On average, these women had 9.61 years ($SD = 6.12$) of professional experience and 4.69 years ($SD = 4.33$) of leadership experience (for information separated by year, see **Table 3**). In 2016, 75.6% were German citizens, 2.4% other, and 22% did not provide information about their nationality. In 2017, 18.2% were German citizens, 9.1% other, and 71.7% did not provide information about their nationality.

Materials

For the 2016 and 2017 samples used in the main study, nine raters analyzed the contest participants’ self-recorded video interviews.⁶

Procedure

Leadership emergence

In the leadership contests in 2016 and 2017, in groups that were each comprised of 7 to 10 women, the women worked on three assessment-center-like group tasks in 2016 and on only two tasks in 2017. We excluded one of the tasks used in the 2016 contest from our analysis to enhance comparability between the

⁵Counter to previous standards (see Baron and Kenny, 1986), researchers have suggested that it is acceptable to interpret a mediation effect on the basis of a significant indirect effect even in the absence of a significant direct or total effect (e.g., MacKinnon et al., 2007; Hayes, 2018).

⁶Detailed interview questions can be obtained from the first author.

³Detailed interview questions can be obtained from the first author.

⁴<https://osf.io/kxm4w/>

TABLE 1 | Descriptive statistics for the variables for both the pilot study and the main study, presented separately by contest year.

	Pilot study		Main study (N = 195)			
	2015 (N = 101)		2016 (N = 111)		2017 (N = 84)	
	M	SD	M	SD	M	SD
Leadership emergence	9.34	5.69	6.12	3.79	4.88	3.27
Ascribed facial attractiveness [1–10]	4.98	1.72	5.92	1.27	6.00	1.38
Ascribed social competence [1–7]	4.47	0.65	4.78	0.58	4.82	0.65
Ascribed babyfacedness [1–10]	3.79	1.05	3.95	0.84	3.99	0.87
Ascribed naïveté [1–10]	3.63	0.99	3.38	0.60	3.44	0.67

The more nominations (leadership emergence) a participant received in the contest the higher her leadership emergence. Ascription ratings were on rating scales [anchor in brackets] on which higher scores indicate a higher degree of each ascribed characteristic. N, sample size; M, mean; SD, standard deviation.

TABLE 2 | Correlations between leadership emergence (dependent variable), ascribed babyfacedness and ascribed attractiveness (predictors), and ascribed social competence and ascribed naïveté (mediators) for both the pilot study and the main study.

r	1	2	3	4	5
(1) Leadership emergence		0.04	0.23*	−0.10	−0.33***
(2) Ascribed attractiveness	0.12		0.52***	0.37***	0.23*
(3) Ascribed social competence	0.27***	0.20**		0.20*	−0.02
(4) Ascribed babyfacedness	−0.02	0.52***	0.07		0.26**
(5) Ascribed naïveté	−0.25***	0.13	−0.72***	0.17*	

Higher variable scores indicate a higher degree of each ascribed characteristic. *r* = Pearson correlation coefficients. The values from the pilot study are presented above the diagonal, and the values from the main study (using centered variables) are presented below the diagonal. ****p* < 0.001, ***p* < 0.01, **p* < 0.05.

two samples. In 2016, the challenges in the tasks were (a) to engage in a 45-min group discussion and (b) to spend 40 min preparing an industrial lobby presentation. In 2017, the women were asked (c) to spend 60 min preparing a huge event and (d) to spend 60 min preparing for a public debate on the gender issue. Each participant provided written nominations of the three most convincing group peers after each task (see section “Leadership Emergence”). We summed the number of nominations for each participant. For the analyses, we divided this sum of nominations by the number of peer nominations possible (i.e., group size per task minus one) in order to control for contaminations due to effects of group size.

Facial characteristics and ascribed traits

In the main study, the ascribed facial attractiveness scale ($\alpha = 0.98$), the ascribed babyfacedness scale ($\alpha = 0.82$), the ascribed social competence scale ($\alpha = 0.93$), and the ascribed naïveté scale ($\alpha = 0.90$) all demonstrated sufficient internal consistency. ICC values for all ratings in the main study ranged from 0.88 to 0.89 for the ascribed facial attractiveness items, from 0.81 to 0.89 for the ascribed babyfacedness items, from 0.71 to 0.84 for the ascribed social competence items, and from 0.68 to 0.83 for the ascribed naïveté items (see section “Measures and Variables – General Information” for all items and the procedure).

Statistical Analyses

Before we calculated the mediation analyses in a manner that was similar to the analyses used in the pilot study, we centered the predictor, mediator, and dependent variables on the mean of

the respective contest year in order to enhance comparability for the total sample in the main study. For example, a participant’s ascribed facial attractiveness may have been high in comparison with the entire female population but not in comparison with her competitors in the contest, if all contest participants were also rated as highly attractive. By centering the variables on the respective sample means, we corrected for such potential group characteristics.

As we did in the pilot study, we ran mediation analyses using the “lavaan” package in R. Seven participants had to be excluded from these analyses due to missing information on group size. We further controlled for possible issues of non-normality by, again, using robust ULS estimators. To control for the high correlations between the predictor variables ascribed facial attractiveness and ascribed babyfacedness as well as between the mediator variables, ascribed social competence and ascribed naïveté, revealed by the a priori analyses, we ran further analyses in which we included all variables simultaneously in one path model. As postulated by Preacher and Hayes (2008), this procedure allowed to identify the unique capacity of each mediator to account for the impact of the respective predictor on leadership emergence beyond the shared contribution of the two mediators. Data and R scripts are published in the OSF⁷.

Results

Contest participants in the main study received an average of 5.58 nominations ($SD = 3.62$) across the two tasks. On average (and before the variables were centered), the mean of ascribed facial

⁷<https://osf.io/kxm4w/>

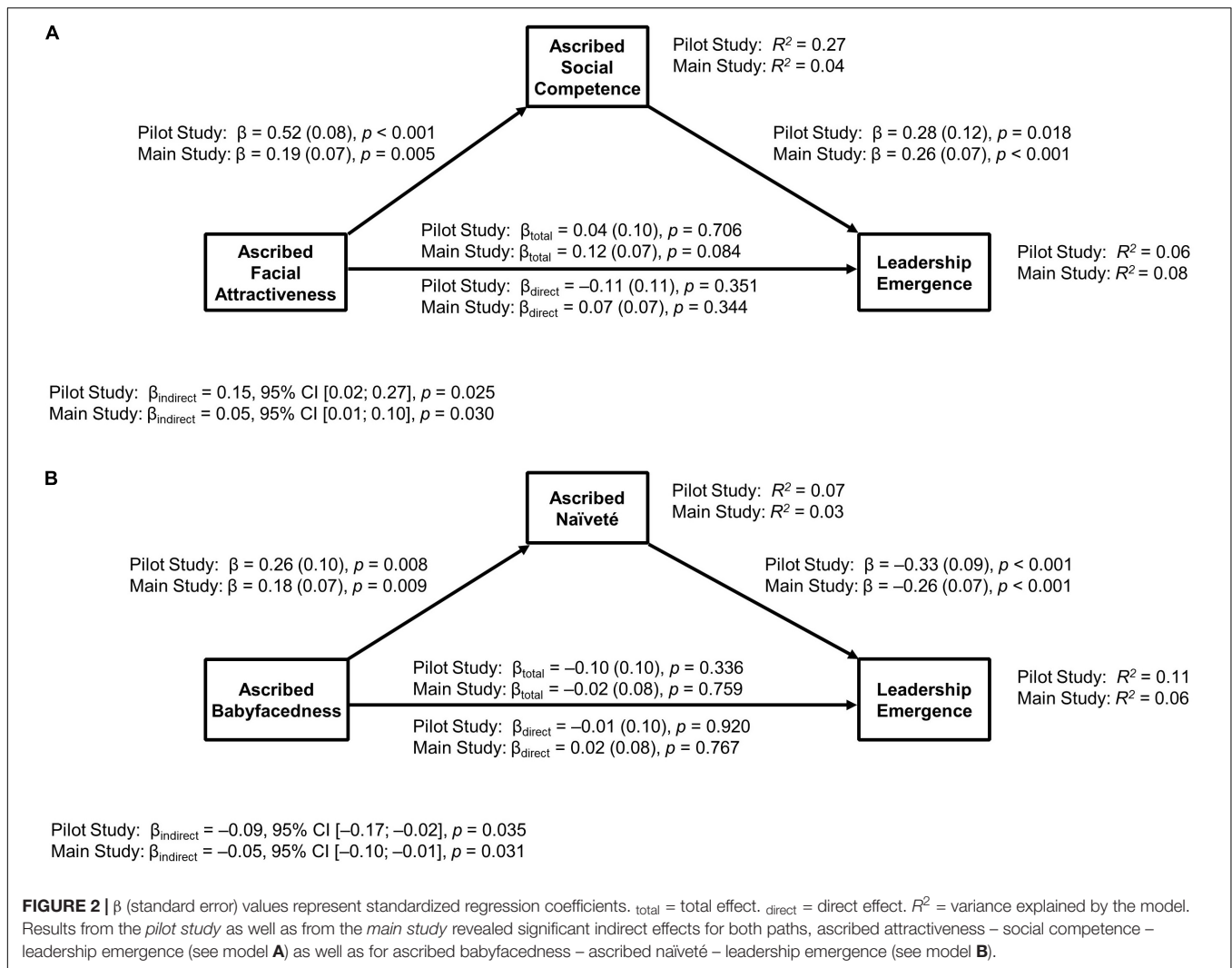


TABLE 3 | Age and years of professional and leadership experience of women in the main study sample, presented separately by contest years and overall.

Main study sample		2016	2017	overall
Sample size	N	111	84	195
Age	<i>M</i> (<i>SD</i>)	33.39 (6.98)	30.20 (4.48)	32.81 (6.69)
	Min – Max	21–52	24–42	21–52
Professional experience in years	<i>M</i> (<i>SD</i>)	10.07 (6.31)	8.43 (5.48)	9.61 (6.12)
Leadership experience in years	<i>M</i> (<i>SD</i>)	4.61 (4.48)	4.99 (3.83)	4.69 (4.33)

N, sample size; *M*, mean; *SD*, standard deviation.

attractiveness was 5.96 ($SD = 1.31$), and the mean of ascribed social competence was 4.80 ($SD = 0.61$). The mean of ascribed babyfacedness was 3.97 ($SD = 0.85$), and the mean of ascribed naïveté was 3.41 ($SD = 0.63$). See **Table 1** for values separated by contest year, and **Table 2** for intercorrelations between the variables.

As hypothesized, the data analyses indicated that ascribed social competence mediated the influence of ascribed facial attractiveness on leadership emergence. Confirming the results

of the pilot study, the higher a contest participant’s ascribed facial attractiveness, the higher her ascribed social competence ($\beta = 0.19$, $SE = 0.07$, $p = 0.005$), and the higher her ascribed social competence, the higher her leadership emergence ($\beta = 0.26$, $SE = 0.07$, $p < 0.001$). The bootstrapping estimation approach with 1,000 bootstrapped samples indicated a significant indirect effect ($\beta_{indirect} = 0.05$, 95% CI [0.01, 0.10], $p = 0.030$). However, ascribed facial attractiveness did not significantly predict leadership emergence, neither before ($\beta_{total} = 0.12$,

$SE = 0.07, p = 0.084$) nor after adding the mediator, ascribed social competence ($\beta_{\text{direct}} = 0.07, SE = 0.07, p = 0.344$; see **Figure 2A**).

The second mediation analysis revealed that ascribed naïveté functioned as a mediator of the influence of ascribed babyfacedness on leadership emergence. According to our results, ascribed babyfacedness significantly predicted ascribed naïveté ($\beta = 0.18, SE = 0.07, p = 0.009$). Further, ascribed naïveté significantly predicted leadership emergence ($\beta = -0.26, SE = 0.07, p < 0.001$), supporting the results of the pilot study as well. In other words, the higher a participant's ascribed babyfacedness, the higher her ascribed naïveté. However, the higher her ascribed naïveté, the lower her leadership emergence. The bootstrapping estimation approach with 1,000 bootstrapped samples indicated a significant indirect effect ($\beta_{\text{indirect}} = -0.05, 95\% \text{ CI } [-0.10, -0.01], p = 0.031$). Further, in line with the results of the pilot study, ascribed babyfacedness did not predict leadership emergence, neither before ($\beta_{\text{total}} = -0.02, SE = 0.08, p = 0.759$) nor after adding the mediator, ascribed naïveté ($\beta_{\text{direct}} = 0.02, SE = 0.08, p = 0.767$; see **Figure 2B**).

We further tested the unique ability of each mediator to mediate the corresponding relation between the predictor and leadership emergence in one joint path model. In this model, we controlled for the correlations between the predictor variables and the mediator variables in order to identify the unique impact of each mediator. Overall, the model showed a very good fit, $\chi^2(2, N = 188) = 0.43, CFI = 1.00, RMSEA < 0.001, SRMR = 0.012$, and is displayed in **Figure 3** with all individual regression coefficients. The unique indirect effects were all non-significant. More specifically, neither the effect of ascribed facial attractiveness on leadership emergence through ascribed social competence ($\beta_{\text{indirect}} = 0.03, 95\% \text{ CI } [-0.01, 0.08], p = 0.271$), nor the effect of ascribed facial babyfacedness on leadership emergence through ascribed naïveté ($\beta_{\text{indirect}} = -0.03, 95\% \text{ CI } [-0.09, 0.03], p = 0.334$) reached the conventional thresholds of significance. Furthermore, there were no unique direct effects of ascribed facial attractiveness ($\beta_{\text{direct}} = 0.14, SE = 0.08, p = 0.053$), ascribed babyfacedness ($\beta_{\text{direct}} = 0.09, SE = 0.10, p = 0.371$), ascribed social competence ($\beta = 0.16, SE = 0.12, p = 0.209$), or ascribed naïveté ($\beta = -0.14, SE = 0.12, p = 0.264$) on leadership emergence. However, when we combined the unique direct and indirect effects of ascribed facial attractiveness, we found a significant unique total effect on leadership emergence ($\beta_{\text{total}} = 0.17, SE = 0.07, p = 0.018$), suggesting that ascribed facial attractiveness had an effect on leadership emergence above and beyond ascribed babyfacedness and ascribed naïveté.

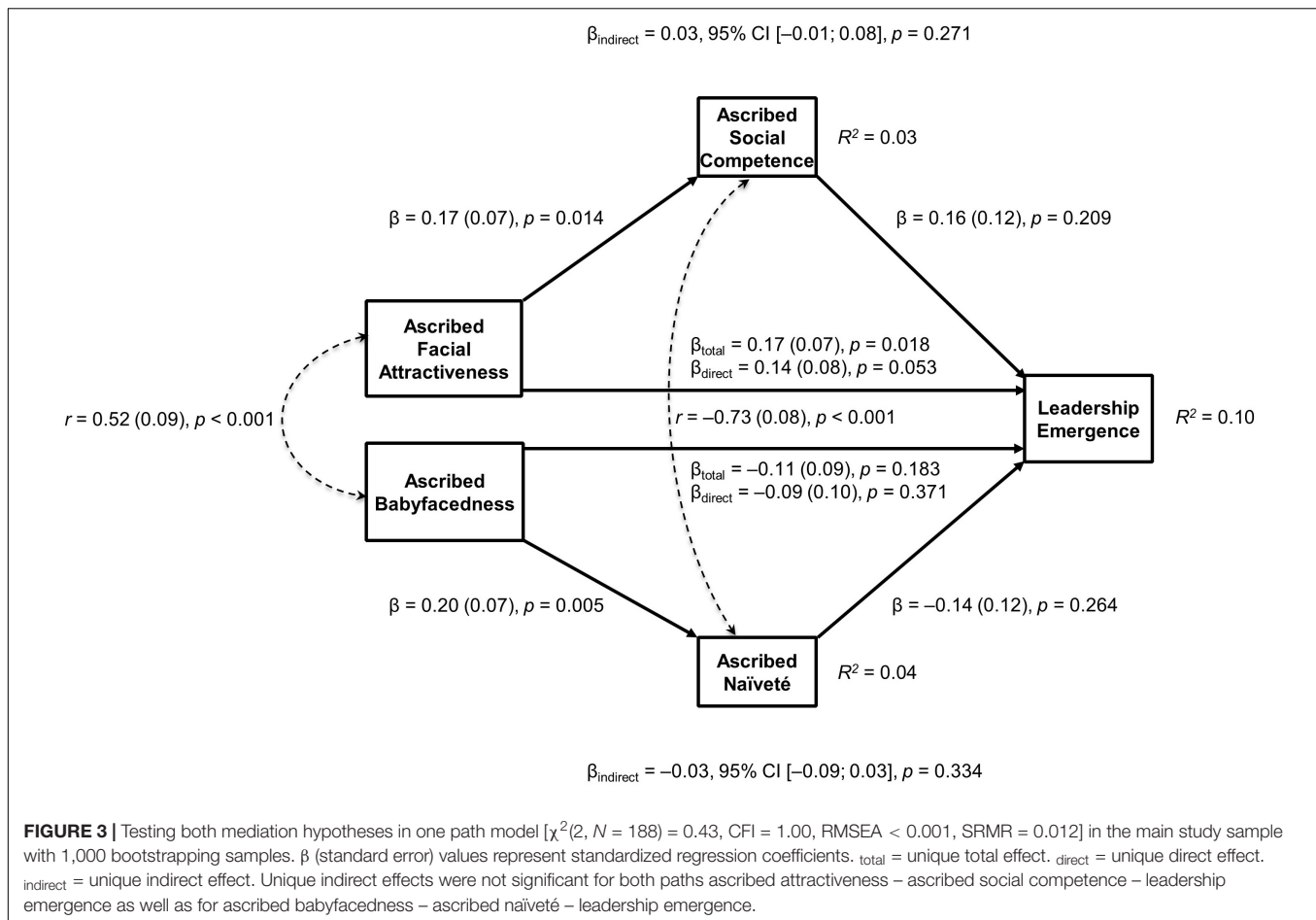
GENERAL DISCUSSION

Aim of the present study was to increase the understanding of the effects of facial appearance and ascribed characteristics on leadership emergence by employing data collected at a women's leadership contest. We captured characteristics of facial appearance by rating participants' application videos as they responded to standardized interview questions. Peer nominations of the women from two exercises that were part of the contest served as a measure of leadership emergence. Analyses

of data from both a pilot study and a main study revealed a significant indirect effect of ascribed facial attractiveness and babyfacedness ratings made by independent raters on leadership emergence. Ascribed social competence mediated the relation between ascribed facial attractiveness and leadership emergence, and ascribed naïveté mediated the relation between babyfacedness and leadership emergence. However, the results did not support Hypotheses 1 and 3 because neither ascribed facial attractiveness nor ascribed babyfacedness significantly predicted leadership emergence. When we tested the hypotheses separately, these findings held both, independent from the inclusion or exclusion of the mediators. Therefore, the analyses in this study revealed that both the women's perceived attractiveness and their babyfacedness were not directly related to their leadership emergence but had an impact through ascribed aspects that were explicitly or implicitly related to these characteristics.

Following the recommendations of a reviewer, we also tested the unique effects of our two mediators and predictors on leadership emergence. Analyses revealed a significant unique effect on leadership emergence only for ascribed facial attractiveness after controlling for ascribed babyfacedness and ascribed naïveté. This finding suggests that the previously reported indirect effects of ascribed social competence and ascribed naïveté showed a large overlap of variance and that the shared and not the unique portion of ascribed personality traits variance mediated the effect of ascribed facial appearance on leadership emergence. Therefore, this finding could signify that an underlying latent variable, namely, ascribed leadership personality, is mediating the effect between ascribed facial appearance and leadership emergence. However, our study was not equipped to handle such a latent variable model due to an insufficient sample size and number of indicators. Future research should investigate the unique and common portions of indirect effects of ascribed facial appearance on leadership emergence through ascribed personality traits by including more indicators for ascribed facial appearance and ascribed personality traits.

These findings also suggest a higher relevance of facial attractiveness for leadership emergence because babyfaced women were perceived as more attractive. However, literature indicated babyfacedness to be related to attractiveness, especially in women (Zebrowitz McArthur and Apatow, 1983/1984). The model employed cannot be used to determine such a halo effect (i.e., possessing babyface causing higher attractiveness ratings in women) because we had defined attractiveness and babyfacedness as equal predictors. In the rating procedure, raters judged facial attractiveness directly before assessing babyfacedness; thus, the latter could not have influenced the attractiveness ratings directly. However, the results of the joint model could further indicate a compensation for the negative influence of babyfacedness on leadership emergence at the same time: In the joint model, partialling out the variance of babyfacedness may explain the significant effect of facial attractiveness on leadership emergence. Babyfacedness has been described as being detrimental to leadership emergence (Cherulnik et al., 1990), but also as positively correlated with facial attractiveness in women (Zebrowitz McArthur and Apatow, 1983/1984). After



controlling for babyfacedness in facial attractiveness, the negative effect of babyfacedness via facial attractiveness was no longer relevant. Thus, the 'pure' facial attractiveness revealed an effect on leadership emergence (i.e., without the confounding effect of babyfacedness). Facial attractiveness was thus no longer influenced by babyfacedness but by other characteristics as, for instance, the symmetry, healthy appearance and other aspects, which we did not explicitly assess.

Further bearing in mind the relation between babyfaced features and femininity (Friedman and Zebrowitz, 1992), the question arises as to whether higher babyfacedness (and thus higher attractiveness) accents the feminine gender role in women and, hence, increases discrimination against them (Zebrowitz et al., 1991). Concluding from our results revealing that babyfacedness weakened the positive effect of attractiveness, it seems that the part of women's attractiveness associated with femininity is not the one promoting but rather preventing their leadership emergence. Further, would babyfaced women therefore have to increase their agentic behaviors to compensate for their even more accentuated gender role in order to emerge as leaders, as indicated by role congruity theory (see Eagly and Karau, 2002; Eagly and Carli, 2007)? Because features of appearance and accordingly ascribed traits depict only some relevant aspects that predict leadership emergence, future

research should also include measures of behaviors and other personality traits, for instance, perceived agency and communion (Johnson et al., 2008; Schock et al., 2018) or task-specific aspects (e.g., Eagly and Karau, 1991; Badura et al., 2018).

Ascribed social competence, related to facial attractiveness, indicated an advantage for obtaining a large number of nominations in the women's leadership contest. On the other hand, ascribed naïveté, related to babyfacedness, indicated a disadvantage. The socially enhancing effects of aspects of attractiveness were well in line with the literature (e.g., Eagly et al., 1991; Langlois et al., 2000). Moreover, a childlike appearance and ascribed naïveté represent characteristics that are counter to the characteristics typically ascribed to leaders (e.g., Zebrowitz McArthur and Apatow, 1983/1984; Zebrowitz et al., 1991; Friedman and Zebrowitz, 1992; Masip et al., 2004; Riggio and Riggio, 2010).

These results are particularly remarkable because the predictor and criterion variables were assessed with different methods. Although both kinds of variables were assessed with ratings, the ratings were made by different samples: The raters of the videos were students, whereas the raters in the contest were competing peers. Furthermore, the predictor variables consisted of ratings of recorded video interviews, whereas the criterion variables based upon real-life interactions in a competitive situation. As the two

kinds of raters may have perceived the same targets of rating in different contexts, somewhat different processes – or at least differing on the level of awareness – may have taken place. For instance, the participants of the contest may have focused more on the competition and their own performance, whereas the raters in the lab explicitly concentrated on the features of facial appearance and personality they had to assess. The finding that ratings of video recordings were related to ratings from a real-life situation may hint at two possible mechanisms: First, the impressions that the raters got from the video recordings were similar to the impressions that the peers had during the contest, the latter of which may have shaped the interactions and ascriptions to a certain extent (see Antonakis and Eubanks, 2017). Therefore, the results may suggest that the ascribed social competence as well as the ascribed naïveté of the women may reflect the important and consensual information that was relevant for raters in both situations (i.e., the application and the contest).

Nevertheless, our results are well in line with previous research on the relevance of appearance for election outcome predictions in political leadership (e.g., Todorov et al., 2005; Antonakis and Dalgas, 2009). In other words, ratings of aspects of physical appearance made by individuals who are not voting in the actual leader election (or as in our study, in the leadership contest) could in part predict the outcome of a leader election (or as in our study, a leader competition).

Second and alternatively, these results may also indicate that consistency in first impressions based on facial appearance may lead people to treat other individuals in certain ways that subsequently shape these individuals' outcomes and behaviors (see Rule and Ambady, 2010; Lukaszewski and Roney, 2011). Further, such impressions and associations based on facial appearance may be extended to the interactional mechanism that underlies a self-fulfilling prophecy (Snyder et al., 1977; further, see Zebrowitz and Montepare, 2008; Todorov et al., 2015). Acting on the assumption that the nominations made by the peers in the contest as well as the video ratings made by external raters reveal real aspects of the participants, according to the self-fulfilling prophecy mechanism, the participants may have developed their characteristics (e.g., social competence or naïveté) on the basis of how they were treated due to their facial appearance, attractiveness, and babyfacedness, respectively. Analogously, at the contest, women may have treated an attractive peer who they therefore perceived as socially competent in a way that increased the peer's confidence in acting like a leader. This, in turn, may have led to more actual leadership behavior, thus convincing the other contestants to nominate her as a convincing leader. For example, the fact that attractiveness may have served as a leadership cue may have led participants to pay more attention to particular women in the contest, which in turn may have increased the abilities of these particular women to excel in the group task (see Gerpott et al., 2017).

Another advantage of this study is that the data collection was part of a real contest, and the sample consisted of real women who participated in the study to win the contest and enhance their reputations. This competitive setting could be interpreted

as comparable to an actual personnel selection situation (e.g., an assessment center) where performance is required within a short period of time. However, the rather unconventional study setting somewhat restricts the interpretation with regard to generalizability: Peers from group tasks usually do not make hiring decisions in personnel selection, but rather, external observers or assessment experts do. Nevertheless, we aimed to contribute to knowledge about which variables are relevant for leadership emergence. Conclusions cannot be drawn with respect to leadership ascription or leadership performance in the medium or long term, especially in comparison with day-to-day teamwork without a competitive character, even though the contest participants had to balance competitiveness with cooperation in working toward a common goal during the tasks. However, our research revealed that appearance is especially important when no information or only a little other information is available (Olivola et al., 2012). In line with previous research, we would expect the impact of facial appearance on leader selection to decrease when information about competence is available because competence-related information should have a greater impact on selection decisions (Kaufmann et al., 2017). Although facial appearance had an impact on women's chances of getting more nominations in the leadership contest in this study, we would expect this effect to diminish in the medium and long terms after women have entered the work place.

Although the setting allowed obtaining an unusual sample of women, future research should include male participants as well as other ethnicities (e.g., Blacks) to investigate interactions between sex and other personal attributes. For instance, previous research indicated that attractiveness judgments were better predictors of election success for women than for men (Berggren et al., 2010). Furthermore, whereas the current study indicated a disadvantage for individuals possessing facial physiognomy reflecting babyfacedness, Livingston and Pearce (2009), for example, found that a larger number of black male CEOs had babyfaced characteristics in their study compared with white male CEOs. A newer study addressing political voting decisions in an Asian culture revealed that babyfacedness was the strongest predictor of percentages of votes beyond competence, attractiveness, warmth, and background characteristics (Chang et al., 2017).

Our findings provide support for earlier studies (e.g., Sczesny and Kühnen, 2004; Olivola and Todorov, 2010; Re and Perrett, 2014) that found that aspects of physical appearance such as facial attractiveness and babyfacedness affect the attribution of job-relevant characteristics in women. This effect may also play an important role in personnel selection. Future studies should investigate the particular relevance of physical appearance with regard to job profiles and job domains. Furthermore, future studies should investigate how other aspects of physical appearance (e.g., body size) impact this effect and should further investigate strategies for overcoming possible disadvantages related to facial appearance, for example, via clothing or hairstyle.

Visual appearance often influences the first impression a person makes, especially in personnel selection (e.g., based on photographs in CVs) and therefore affects the attributions

made to an individual. Under a more practical perspective, we encourage organizations and hiring experts to enhance their awareness of such ascription- and stereotype-inducing mechanisms, helping women to overcome these gender-related obstacles to achieve leadership positions equally to their male counterparts. The use of standardized assessment criteria that are determined in advance (from the side of human resources) and the decision not to include a photo in a job application (from the side of an applicant) may be first steps to diminish such effects, at least with regard to aspects of (facial) appearance.

DATA AVAILABILITY

Datasets are in a publicly accessible repository: the datasets analyzed for this study can be found in the Open Science Framework (<https://osf.io/kxm4w/>).

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the statutes of the University of Salzburg (part XI), Ethic Committee of the Paris Lodron University of Salzburg. The protocol was approved by the Ethic Committee of the Paris Lodron University of Salzburg. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

FG collected the data for the pilot study, co-revised the rating manual, and supervised and organized the data collection

REFERENCES

- Ambady, N., and Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of behavior and physical attractiveness. *J. Pers. Soc. Psychol.* 64, 431–441. doi: 10.1037/0022-3514.64.3.431
- Antonakis, J., and Dalgas, O. (2009). Predicting elections: child's play! *Science* 323, 1183–1183. doi: 10.1126/science.1167748
- Antonakis, J., and Eubanks, D. L. (2017). Looking leadership in the face. *Curr. Dir. Psychol. Sci.* 26, 270–275. doi: 10.1177/0963721417705888
- Badura, K. L., Grijalva, E., Newman, D. A., Yan, T. T., and Jeon, G. (2018). Gender and leadership emergence: a meta-analysis and explanatory model. *Pers. Psychol.* 71, 335–367. doi: 10.1111/peps.12266
- Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51, 1173–1182. doi: 10.1037/0022-3514.51.6.1173
- Bass, B. M., and Bass, R. (2008). *The Bass Handbook of Leadership: Theory, Research, and Managerial Applications*, 4th Edn. New York, NY: Free Press.
- Berggren, N., Jordahl, H., and Poutvaara, P. (2010). The looks of a winner: beauty, gender and electoral success. *J. Public Econ.* 94, 8–15. doi: 10.1016/j.jpubeco.2009.11.002
- Bourgeois, V. (2017). *Pressemitteilung: Nur 1 von 3 Führungskräften in der EU ist eine Frau [Press release: Only 1 out of 3 leaders is female within the EU]*. Available at: <http://ec.europa.eu/eurostat/documents/2995521/7896995/3-06032017-AP-DE.pdf/b49bc03b-00be-448c-b12a-318012c61cda>
- Braun, C., Gründl, M., Marberger, C., and Scherber, C. (2001). *Beautycheck - Ursachen und Folgen von Attraktivität. [Beautycheck - Causes and consequences of human facial attractiveness]*. Available at: <http://psydok.psycharchives.de/jspui/handle/20.500.11780/45>
- Bruner, J. S., and Tagiuri, R. (1954). "The perception of people," in *Handbook of Social Psychology*, Vol. 2, ed. G. Lindzey (Boston, MA: Addison-Wesley), 634–654.
- Carli, L. L., and Eagly, A. H. (2001). Gender, hierarchy, and leadership: An introduction. *J. Soc. Issues* 57, 629–636. doi: 10.1111/0022-4537.00232
- Chang, C.-T., Lee, Y.-K., and Cheng, Z.-H. (2017). Baby face wins? Examining election success based on candidate election bulletin via multilevel modeling. *Asian J. Soc. Psychol.* 20, 97–112. doi: 10.1111/ajsp.12172
- Cherulnik, P. D. (1995). Physical appearance, social skill, and performance as a leadership candidate. *Basic Appl. Soc. Psychol.* 16, 287–295. doi: 10.1207/s15324834basps1603_2
- Cherulnik, P. D., Turns, L. C., and Wilderman, S. K. (1990). Physical appearance and leadership: exploring the role of appearance-based attribution in leader emergence. *J. Appl. Soc. Psychol.* 20, 1530–1539. doi: 10.1111/j.1559-1816.1990.tb01491.x
- Chiao, J. Y., Bowman, N. E., and Gill, H. (2008). The political gender gap: gender bias in facial inferences that predict voting behavior. *PLoS One* 3:e3666. doi: 10.1371/journal.pone.0003666
- Diekmann, A. B., and Schneider, M. C. (2010). A social role theory perspective on gender gaps in political attitudes. *Psychol. Women Q.* 34, 486–497. doi: 10.1111/j.1471-6402.2010.01598.x
- Dietl, E. (2013). *Durch Kinder und Erwachsene beurteilte erste Eindrücke aus Gesichtern, Führung, Core Self-Evaluations und Karriereerfolg [First facial impressions, leadership, core self-evaluations and career success judged by children and adults]*. Available at: http://hss.ulb.uni-bonn.de/diss_online

and analyses for the main study. She mainly contributed to writing the Method, Results, and Discussion sections of the manuscript and revised earlier drafts of the entire manuscript. CV conceptualized the research, data collection and analyses of both the pilot study and the main study. She mainly contributed to writing the first versions of the Introduction and Discussion sections as well as the Results section of the pilot study. TO supervised the conceptualization of this research regarding the operationalization and methods. She revised earlier drafts of the manuscript, in particular providing theoretical input for the Introduction and Discussion sections.

ACKNOWLEDGMENTS

We thank Anne-Kathrin Schock and Thomas Scherndl for their supervisory support regarding conceptualization and analyses. We further thank the student assistants Moritz Baumgärtner, Lennart Dialer, Julian Fuchs, Verena Eberle, Lena Kastner, Anna Krämer, Dominik Mengin, Corinna Pfannenstern, Martina Rogalla, and David Schmid for their help with coding and ratings. We thank Manfred Schmitt and Anna Baumert for great theoretical advice, Tobias Koch for methodological input, and Jane Zagorski for helpful suggestions. We further thank the editors Alice H. Eagly and Sabine Szcesny, as well as Leslie Zebrowitz and Minna Paunova for thoughtful and supportive reviews. We thank the University of Salzburg for financial funding of the Open Access Publication fee. Finally, we thank Isabelle Hoyer and Stuart Cameron for allowing us to conduct research in the context of their leadership contest for female professionals, and the participants of the 2015, 2016, and 2017 contests who agreed to share their data for research purposes.

- Eagly, A. H. (1987). *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Eagly, A. H., Ashmore, R. D., Makhijani, M. G., and Longo, L. C. (1991). What is beautiful is good, but...: a meta-analytic review of research on the physical attractiveness stereotype. *Psychol. Bull.* 110, 109–128. doi: 10.1037/0033-2909.110.1.109
- Eagly, A. H., and Carli, L. L. (2007). *Through the Labyrinth: The Truth about How Women become Leaders*. Boston, MA: Harvard Business School Press.
- Eagly, A. H., and Diekmann, A. B. (2005). "What is the problem? Prejudice as an attitude-in-context," in *On the Nature of Prejudice: Fifty Years after Allport*, eds J. Dovidio, P. Glick, and L. Rudman (Malden, MA: Blackwell Publishing), 19–35.
- Eagly, A. H., and Karau, S. J. (1991). Gender and the emergence of leaders: a meta-analysis. *J. Pers. Soc. Psychol.* 60, 685–710. doi: 10.1037/0022-3514.60.5.685
- Eagly, A. H., and Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* 109, 573–598. doi: 10.1037//0033-295X.109.3.573
- Eagly, A. H., Nater, C., Miller, D., Kaumann, M., and Sczesny, S. (2018). *Gender Stereotypes have Changed: A Cross-Temporal Meta-Analysis of U.S. Public Opinion Polls from 1946 to 2018*. Evanston, IL: Northwestern University.
- Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *J. Res. Pers.* 8, 45–54. doi: 10.1016/0092-6566(74)90044-0
- Enlow, H. D. (1982). *Handbook of Facial Growth*. Philadelphia, PA: W B Saunders.
- Ernst and Young, L. L. P. (2015). *Women on US-boards: What are we seeing?*. Available at: http://www.ey.com/Publication/vwLUAssets/EY_-_Women_on_US_boards:_what_are_we_seeing/%24FILE/EY-women-on-us-boards-what-are-we-seeing.pdf
- European Commission (2012). *She figures: Gender in Research and Innovation*. Available at: http://ec.europa.eu/research/science-society/document_library/pdf_06/she-figures-2012_en.pdf
- Finkelstein, L. M., Costanza, D. P., and Goodwin, G. F. (2018). Do your high potentials have potential? The impact of individual differences and designation on leader success. *Pers. Psychol.* 71, 3–22. doi: 10.1111/peps.12225
- Friedman, H., and Zebrowitz, L. A. (1992). The contribution of typical sex differences in facial maturity to sex role stereotypes. *Pers. Soc. Psychol. Bull.* 18, 430–438. doi: 10.1177/0146167292184006
- Garcia-Retamero, R., and López-Zafra, E. (2006). Prejudice against women in male-congenial environments: perceptions of gender role congruity in leadership. *Sex Roles* 55, 51–61. doi: 10.1007/s11199-006-9068-1
- Gerpott, F. H., Lehmann-Willenbrock, N., Silvis, J. D., and Van Vugt, M. (2017). In the eye of the beholder? An eye-tracking experiment on emergent leadership in team interactions. *Leadersh. Q.* 29, 523–532. doi: 10.1016/j.leaqua.2017.11.003
- Gray, H., and Standring, S. (2008). *Gray's Anatomy: The Anatomical basis of Clinical Practice*. Edinburgh: Churchill Livingstone.
- Groves, K. S. (2005). Gender differences in social and emotional skills and charismatic leadership. *J. Leadersh. Organ. Stud.* 11, 30–46. doi: 10.1177/107179190501100303
- Guthrie, R. D. (1976). *Body Hot Spots: The Anatomy of Human Social Organs and behavior*. New York, NY: Van Nostrand Reinhold.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analyses: A regression-based approach*, 2nd Edn. New York, NY: The Guilford Press.
- Johnson, H. L. (2017). *Pipelines, Pathways, and Institutional Leadership: An Update on the Status of Women in Higher Education*, Washington, DC. Available at: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/84062/UpdateWomenHigherEducation.pdf?sequence=1&isAllowed=y>
- Johnson, S. K., Murphy, S. E., Zewdie, S., and Reichard, R. J. (2008). The strong, sensitive type: effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organ. Behav. Hum. Decis. Process.* 106, 39–60. doi: 10.1016/j.obhdp.2007.12.002
- Judge, T. A., Hurst, C., and Simon, L. S. (2009). Does it pay to be smart, attractive, or confident (or all three)? Relationships among general mental ability, physical attractiveness, core self-evaluations, and income. *J. Appl. Psychol.* 94, 742–755. doi: 10.1037/a0015497
- Kaufmann, M. C., Krings, F., Zebrowitz, L. A., and Sczesny, S. (2017). Age bias in selection decisions: the role of facial appearance and fitness impressions. *Front. Psychol.* 8:2065. doi: 10.3389/fpsyg.2017.02065
- Keating, C. F., Randall, D. W., Kendrick, T., and Gutshall, K. A. (2003). Do babyfaced adults receive more help? The (cross-cultural) case of the lost resume. *J. Nonverbal Behav.* 27, 89–109. doi: 10.1023/A:1023962425692
- Koenig, A. M., and Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: observations of groups' roles shape stereotypes. *J. Pers. Soc. Psychol.* 107, 371–392. doi: 10.1037/a0037215
- Koenig, A. M., Eagly, A. H., Mitchell, A. A., and Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychol. Bull.* 137, 616–642. doi: 10.1037/a0023557
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., and Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* 126, 390–423. doi: 10.1037//0033-2909.126.3.390
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Livingston, R. W., and Pearce, N. A. (2009). The teddy-bear effect does having a baby face benefit black chief executive officers? *Psychol. Sci.* 20, 1229–1236. doi: 10.1111/j.1467-9280.2009.02431.x
- Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung [The hereditary forms of feasible experience]. *Z. Tierpsychol.* 5, 235–409. doi: 10.1111/j.1439-0310.1943.tb00655.x
- Lukaszewski, A. W., and Roney, J. R. (2011). The origins of extraversion: joint effects of facultative calibration and genetic polymorphism. *Pers. Soc. Psychol. Bull.* 37, 409–421. doi: 10.1177/0146167210397209
- MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.* 58, 593–614. doi: 10.1146/annurev.psych.58.110405.085542
- MakeHuman team (2016). *MakeHuman Version 1.1.0*. Available at: <http://www.makehuman.org/>
- Masip, J., Garrido, E., and Herrero, C. (2004). Facial appearance and impressions of 'credibility': the effects of facial babyishness and age on person perception. *Int. J. Psychol.* 39, 276–289. doi: 10.1080/0020759044000014
- Montepare, J. M., and Zebrowitz, L. A. (1998). Person perception comes of age: the salience and significance of age in social judgments. *Adv. Exp. Soc. Psychol.* 30, 93–161. doi: 10.1016/S0065-2601(08)60383-4
- Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., and Todorov, A. (2012). Republicans prefer republican-looking leaders: political facial stereotypes predict candidate electoral success among right-leaning voters. *Soc. Psychol. Pers. Sci.* 3, 605–613. doi: 10.1177/1948550611432770
- Olivola, C. Y., and Todorov, A. (2010). Elected in 100 milliseconds: appearance-based trait inferences and voting. *J. Nonverbal Behav.* 34, 83–110. doi: 10.1007/s10919-009-0082-1
- Ortner, T. M., Müller, S. M., and Garcia-Retamero, R. (2011). Estimations of parental and self intelligence as a function of parents' status: a cross-cultural study in Germany and Spain. *Soc. Sci. Res.* 40, 1067–1077. doi: 10.1016/j.ssresearch.2011.03.006
- Paunova, M. (2015). The emergence of individual and collective leadership in task groups: a matter of achievement and ascription. *Leadersh. Q.* 26, 935–957. doi: 10.1016/j.leaqua.2015.10.002
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Poutvaara, P., Jordahl, H., and Berggren, N. (2009). Faces of politicians: babyfacedness predicts inferred competence but not electoral success. *J. Exp. Soc. Psychol.* 45, 1132–1135. doi: 10.1016/j.jesp.2009.06.007
- Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40, 879–891. doi: 10.3758/BRM.40.3.879
- Re, D. E., and Perrett, D. I. (2014). The effects of facial adiposity on attractiveness and perceived leadership ability. *Q. J. Exp. Psychol.* 67, 676–686. doi: 10.1080/17470218.2013.825635
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208
- Ridgeway, C. L. (2001). Gender, status, and leadership. *J. Soc. Issues* 57, 637–655. doi: 10.1111/0022-4537.00233
- Riggio, H. R., and Riggio, R. E. (2010). Appearance-based trait inferences and voting: evolutionary roots and implications for leadership. *J. Nonverbal Behav.* 34, 119–125. doi: 10.1007/s10919-009-0083-0
- Riggio, R. E., and Reichard, R. J. (2008). The emotional and social intelligences of effective leadership: an emotional and social skill approach. *J. Manage. Psychol.* 23, 169–185. doi: 10.1108/02683940810850808

- Riggio, R. E., Riggio, H. R., Salinas, C., and Cole, E. J. (2003). The role of social and emotional communication skills in leader emergence and effectiveness. *Group Dyn.* 7, 83–103. doi: 10.1037/1089-2699.7.2.83
- Ritter, B. A., and Yoder, J. D. (2004). Gender differences in leader emergence persist even for dominant women: an updated confirmation of role congruity theory. *Psychol. Women Q.* 28, 187–193. doi: 10.1111/j.1471-6402.2004.00135.x
- Rosseel, Y. (2018). Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.3389/fpsyg.2014.01521
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management. *J. Pers. Soc. Psychol.* 74, 629–645. doi: 10.1037/0022-3514.74.3.629
- Rule, N., and Ambady, N. (2010). First impressions of the face: predicting success. *Soc. Pers. Psychol. Compass* 4, 506–516. doi: 10.1111/j.1751-9004.2010.00282.x
- Schneider, D. J. (1973). Implicit personality theory: a review. *Psychol. Bull.* 79, 294–309. doi: 10.1037/h0034496
- Schock, A.-K., Gruber, F. M., Scherndl, T., and Ortner, T. M. (2018). Tempering agency with communion increases women's leadership emergence in all-women groups: evidence for role congruity theory in a field setting. *Leadersh. Q.* (in press). doi: 10.1016/j.leaqua.2018.08.003
- Szczesny, S., and Kühnen, U. (2004). Meta-cognition about biological sex and gender-stereotypic physical appearance: consequences for the assessment of leadership competence. *Pers. Soc. Psychol. Bull.* 30, 13–21. doi: 10.1177/0146167203258831
- Shannon, M. L., and Stark, C. P. (2003). The influence of physical appearance on personnel selection. *Soc. Behav. Pers.* 31, 613–623. doi: 10.2224/sbp.2003.31.6.613
- Snyder, M., Tanke, E. D., and Berscheid, E. (1977). Social perception and interpersonal behavior: on the self-fulfilling nature of social stereotypes. *J. Pers. Soc. Psychol.* 35, 656–666. doi: 10.1037/0022-3514.35.9.656
- Sobieraj, S., and Krämer, N. C. (2014). "What is beautiful in cyberspace? Communication with attractive avatars," in *Social Computing and Social Media*, Vol. 8531, ed. G. Meiselwitz (Cham: Springer).
- Stoker, J. I., Garretsen, H., and Spreuwers, L. J. (2016). The facial appearance of CEOs: faces signal selection but not performance. *PLoS One* 11:e0159950. doi: 10.1371/journal.pone.0159950
- Taggar, S., Hackew, R., and Saha, S. (1999). Leadership emergence in autonomous work teams: antecedents and outcomes. *Pers. Psychol.* 52, 899–926. doi: 10.1111/j.1744-6570.1999.tb00184.x
- Todorov, A., Mandisodza, A. N., Goren, A., and Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science* 308, 1623–1626. doi: 10.1126/science.1110589
- Todorov, A., Olivola, C. Y., Dotsch, R., and Mende-Siedlecki, P. (2015). Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annu. Rev. Psychol.* 66, 519–545. doi: 10.1146/annurev-psych-113011-143831
- U.S. Bureau of Labor Statistics (2017). *Employed Persons by Detailed Occupation, Sex, Race, and Hispanic or Latine Ethnicity*. Available at: <https://www.bls.gov/cps/cpsaat11.htm>
- Wagner, D. G., and Berger, J. (1997). Gender and interpersonal task behaviors: status expectation accounts. *Soc. Perspect.* 40, 1–32. doi: 10.2307/1389491
- Weyer, B. (2007). Twenty years later: explaining the persistence of the glass ceiling for women leaders. *Women Manag. Rev.* 22, 482–496. doi: 10.1108/09649420710778718
- Zebrowitz, L. A., and McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law Hum. Behav.* 15, 603–623. doi: 10.1007/BF01065855
- Zebrowitz, L. A., and Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Soc. Pers. Psychol. Compass* 2, 1497–1517. doi: 10.1111/j.1751-9004.2008.00109.x
- Zebrowitz, L. A., Tenenbaum, D. R., and Goldstein, L. H. (1991). The impact of job applicants' facial maturity, gender, and academic achievement on hiring recommendations. *J. Appl. Soc. Psychol.* 21, 525–548. doi: 10.1111/j.1559-1816.1991.tb00535.x
- Zebrowitz McArthur, L., and Apatow, K. (1983/1984). Impressions of baby-faced adults. *Soc. Cogn.* 2, 315–342. doi: 10.1521/soco.1984.2.4.315

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Gruber, Veidt and Ortner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.