Reading the face of a leader: women with low facial masculinity are perceived as competitive

Silberzahn, Raphael and Menges, Jochen (2016) Reading the face of a leader: women with low facial masculinity are perceived as competitive. Academy of Management Discoveries, 2 (3). pp. 272-289. ISSN 2168-1007

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/73376/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk
READING THE FACE OF A LEADER: WOMEN WITH LOW FACIAL MASCULINITY ARE PERCEIVED AS COMPETITIVE

RAPHAEL SILBERZAHN
University of Cambridge - Judge Business School

JOCHEN MENGES
WHU - Otto Beisheim School of Management

In competitive settings, people prefer leaders with masculine faces. But is facial masculinity a trait that is similarly desired in men and women leaders? Across three studies, we discovered that people indeed prefer men and women leaders who have faces with masculine traits. But surprisingly, we find that people also prefer women with low facial masculinity as leaders in competitive contexts (Study 1). Our findings indicate that low facial masculinity in women, but not in men is perceived to indicate competitiveness (Study 2). Thus, in contrast to men, women leaders who rate high in facial masculinity as well as those low in facial masculinity are both selected as leaders in competitive contexts. Indeed, among CEOs of S&P 500 companies, we find a greater range of facial masculinity among women CEOs than among men CEOs (Study 3). Our results suggest that traits of facial masculinity in men and women are interpreted differently. Low facial masculinity in women is linked to competitiveness and not only to cooperativeness as suggested by prior research.

Editor’s Comment

The paper by Silberzahn and Menges provides important insights into the effects of masculine and feminine facial features. Do leaders in competitive situations profit from masculine or feminine features, and does this vary across men and women? Based on their intuition and a few available data points, the authors expected mainstream ideas to hold for men but not for women, where those ideas support masculinity for leaders embedded in strongly competitive situations. The authors were right—among men, those with masculine features were favored as leaders and indeed were seen as more competitive, but among women those with strongly masculine features and those with strongly feminine features were favored and were rated as more competitive. Having shared these findings with several executive audiences, this editor can tell you that practicing managers are quite intrigued by the implications. For scholarship, the findings reported in the paper have implications for several research streams and theories.

C. Chet Miller, Action Editor
INTRODUCTION

Yahoo’s Marissa Mayer, Hewlett Packard’s Meg Whitman, and Facebook’s Sheryl Sandberg have in particular three things in common: They are all top-level leaders in highly competitive companies, they are all women, and none of them look particularly masculine. Yet, previous research studies have found that people prefer men and women leaders with masculine faces, particularly in competitive settings (Little, Burriss, Jones, & Roberts, 2007; Little, 2014; Re, DeBruine, Jones, & Perrett, 2013; Spisak, Dekker, Krüger, & van Vugt, 2012; Spisak, Homan, Grabo, & Van Vugt, 2012). Are Mayer, Whitman, and Sandberg then just exceptions to the rule, or is there something more systematic to the success of nonmasculine-looking women in the corporate world?

Women and men whose faces appear masculine are more likely to attain leadership positions, as high facial masculinity signals relevant personality traits such as dominance and agency (Re et al., 2013; Spisak, Dekker, et al., 2012). These traits, according to leadership theory (Hogue & Lord, 2007) and gender research (Eagly & Karau, 2002), are seen as desirable for leaders. Faces of potential leaders are read for cues to masculinity, and these cues appear to be even more influential than the person’s biological sex in influencing leader choice (Spisak, Dekker, et al., 2012). Higher degrees of facial masculinity thus seem to be desired in both men and women leaders.

However, the appearance of men and women is often evaluated differently (Feinberg and Jones, 2005; Rhodes, 2006; Roney, Hanson, Durante, & Maestripieri, 2006; Smith et al., 2006). Women with feminine faces are considered more attractive (Thornhill & Gangestad, 1999), which may help in attaining positions with higher occupational prestige (Jaeger, 2011). Results from simulated political elections suggest preferences for women with less masculine faces (Hehman, Carpinella, Johnson, Leitner, & Freeman, 2014). Hence, although facial masculinity appears to benefit men’s attainment of leadership positions, for women, the case is less clear-cut. This led us to wonder if facial masculinity affects women and men differently in leadership selections?

In a series of three experimental and field studies, we examined whether facial masculinity affects men and women leaders differently, and why this may be the case. We made three discoveries: First, whereas men in competitive settings benefit from high levels of facial masculinity, women fare well when they either look particularly masculine or when they do not look masculine at all. Second, in contrast to prior research, we find that among women, low facial masculinity is perceived to signal competitiveness, whereas among men, competitiveness was inferred only from high facial masculinity and not from low facial masculinity. Third, our findings suggest that the differential effects of facial masculinity on leadership preferences manifest in the corporate world. We discover that in S&P 500 companies, a greater range of facial masculinity is present among women CEOs compared to men CEOs.

The association between low facial masculinity and competitiveness in women, which we detect in our research, adds to and challenges existing theory and research across several literature streams. We extend gender role research by offering a more nuanced understanding of how attributions of competitiveness vary not just between, but also within gender (Eagly & Karau, 2002). Our discoveries are relevant for the further development of theories concerning leadership prototypes (Hogue & Lord, 2007): Because both women with high and low masculinity are perceived favorably as potential leaders, there may be less of a disadvantage for women based on people’s implicit leadership theories than previously assumed. Our research extends and refines biosocial models of leadership (Spisak, Dekker, et al., 2012). It encourages further research linking leadership choice and attributions about facial masculinity with hormonal effects on facial appearance and behavior (Stanton & Schultheiss, 2007, 2009). And, our findings are of interest to personality researchers who have long wondered about the origin and validity of personality attributions made from facial appearance (Penton-Voak et al., 2006; Zebrowitz & Montepare, 2008).

FACIAL Masculinity AND LEADERSHIP PREFERENCES

When choosing a leader, people are inevitably influenced by others’ facial appearance. People process facial information in a split second and form instantaneous impressions (Willis & Todorov, 2006). This process is reinforced by biosocial models of leadership suggesting that a likely leader’s behavior...
is inferred from masculine facial features (Spisak, Dekker, et al., 2012). Facial features are shaped in the womb and during later life by sex hormones (Boothroyd & Perrett, 2006; Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004). The male sex hormone testosterone makes faces more masculine (Lefevre, Lewis, Perrett, & Penke, 2013), and the female sex hormone estrogen makes faces appear more feminine (Smith et al., 2006). But in addition to influencing facial appearance, hormones also affect behavior. Testosterone is associated with status striving and dominance in both men (Pound, Penton-Voak, & Surridge, 2009; Stanton & Schultheiss, 2009) and women (Jiménez, Aguilar, & Alvero-Cruz, 2012). Thus, biosocial models of leadership posit that over the course of evolution, masculine facial features have become linked in people’s brains with positive inferences concerning leadership behavior:

The more masculine a face looks, the more likely the person is believed to engage in behaviors that are typically expected from leaders.

Biosocial models of leadership suggest that people’s perceptions of faces and leadership preferences are linked through implicit leadership theories. Implicit leadership theories operate outside of awareness and facilitate people’s instantaneous assessment of the leadership qualities of another person (Lord & Dinh, 2014). People use these prototypical expectations to automatically assess whether the leader matches their prototype. If there is a match, then the leader is considered more favorably than if there is no match (Lord, Foti, & De Vader, 1984).

Masculine faces are assumed to match with leadership prototypes because facial stimuli activate schemas. Schemas are organized mental representations of life experiences that facilitate the processing of new information (Baldwin, 1992), allowing individuals to make inferences about a person based on minimal information (Neisser, 1976). For example, so-called baby faces activate a schema that attributes incompetency and a lack of maturity to a person, even in the absence of relevant evidence (Zebrowitz & Montepare, 2008). It is the activated schema that fills in the missing information. Similarly, masculine faces activate a schema that evokes dominance and strength (Little, 2014), traits that are characteristic of prototypical leaders (Lord & Dinh, 2014). Masculine faces are likely associated with dominance and power, because in many animal species, including primates (De Waal, 1982) and humans (Murray & Schmitz, 2011), facial masculinity is linked to greater physical size and strength, body features which in many species determine social status and rank (Henrich & Gil-White, 2001; Little et al., 2007; Re et al., 2013). This pattern was also evident in a study of elementary school children who ascribed strength, resourcefulness, independence and, indeed, leadership to those girls whom they categorized as rather masculine (Hemmer & Kleiber, 1981).

The association between masculinity and leadership is captured in the “think manager, think male” phenomenon (Schein, 1973, 2001). Although preferences are changing, the prototypical expectations for leaders still favor men (Newport & Wilke, 2013). In line with these preferences, masculine facial features are seen as fitting the admittedly gender-biased leadership prototype (Hogue & Lord, 2007).

But the link between facial masculinity and leadership preferences also depends on the context of leadership. Implicit leadership theories are not fixed, but rather adjust according to a given context such that different contexts give rise to different prototypical expectations about leaders (Brands, Menges, & Kilduff, 2015). The biosocial model of leadership posits the contingency that preferences for high facial masculinity apply particularly in situations that involve greater conflict, competition, and aggression, whereas preferences for comparatively lower facial masculinity apply in more peaceful, cooperative settings (Little et al., 2007; Spisak, Dekker, et al., 2012). Such settings may cater to a communal, nurturing leadership style aligned with the female gender stereotype, whereas an agentic dominant leadership style appears suited to competitive settings and is aligned with the male gender stereotype (Eagly & Karau, 2002).

Thus, situational demands activate different leadership prototypes, and people consequently prefer leaders whom they expect to have characteristics suited to that situation (Little, 2014). The contingency advanced by the biosocial model of leadership (Spisak, Homan, et al., 2012) therefore suggest a person-in-situation schema (Brands et al., 2015) that favors masculine-looking leaders in competitive contexts. In such contexts, both men and women who lack facial masculinity may fail to meet people’s
assumptions, as suggested by leadership categorization accounts of implicit leadership theories (Hogue & Lord, 2007), gender research and, particularly, by role congruity theory (Eagly & Karau, 2002).

Yet, in three studies, we discovered a more nuanced pattern of linkages between facial masculinity and leadership preferences for women leaders. In Study 1, we asked individuals to choose potential leaders from a set of candidates with a range of modified degrees of facial masculinity. We manipulated leader gender (male/female) and leadership context (competitive/collaborative), and found that preferences for men and women leaders differ. Women scoring high on masculinity but also those with low facial masculinity were preferred in competitive settings. In Study 2, we examined why these effects occur. With a categorization task, we studied what attributes and characteristics people ascribe based on variations in facial masculinity. We found that women leaders who scored high as well as those who scored low in facial masculinity were attributed competitiveness, whereas for men leaders, this was the case only for those high in facial masculinity. If these effects manifest not only in laboratory settings, but also in field settings, we would expect to find greater range of facial masculinity among women than among men leaders. Indeed, in Study 3, this was confirmed using a sample of male and female CEOs of S&P 500 companies. In summary, the findings across the three studies challenge conclusions in existing research on leadership preferences for women leaders. In Study 1, we discovered a more nuanced pattern of linkages between facial masculinity and leadership preferences the phrases “... given that your company competes heavily with other companies” or “... given that your company collaborates well with other companies.”

**STUDY 1**

We examined whether preferences for facial masculinity differ for men and women leaders in competitive and collaborative contexts. We devised a $2 \times 2$ between-subjects experiment in which two factors—leader gender (man versus woman) and context (competitive versus collaborative)—were varied. Prior research suggests that in competitive environments, people would prefer leaders with higher levels of facial masculinity, and that such preferences would be similar for men and women (Spisak, Dekker, et al., 2012). In contrast, our thinking and some related research studies (Hehman et al., 2014) suggest that people might choose women with faces low in masculinity as leaders in competitive environments.

**Participants and Procedures**

The sample consisted of 187 American adults (44 percent women) from different ethnicities (67 percent White, 14 percent Black, 10 percent Asian, 4 percent Latino, and 4 percent other). All participants resided in the United States, and they averaged 33 years of age (standard deviation [SD] = 12.04) and 12 years of work experience (SD = 11.35). Participants were recruited from Microworkers, an online panel.

Participants accessed our website and, after agreeing to take part in this research, were randomly assigned to one of four conditions. Participants would initially read that a company was either competing heavily or collaborating well, and were then asked to make several leadership choices by selecting a suitable leader from a series of images showing faces of women or men. Each image featured the same person’s face but with varied degrees of masculinity.

**Manipulation 1: Leader gender.** Participants in the men leader condition were shown a series of images of men, and participants in the women leader condition were shown a series of images of women.

**Manipulation 2: Leadership context.** Participants in the competition condition read a scenario that described a company that “has many rivals and competes heavily,” and participants in the collaboration condition read about a company that “has many partners and collaborates well.” The scenarios were taken from prior research (Little et al., 2007) and adapted to a corporate context. Each scenario was reinforced by adding to the subsequent questions concerning leader choice the phrases “... given that your company collaborates well with other companies” or “... given that your company competes heavily with other companies.”

**Dependent variable: Leader choice.** We asked candidates four questions concerning whom they would (1) “prefer for your company,” (2) “hire for your company,” (3) find “most effective,” and (4) believe is “most likely to achieve your company’s goal.” Based on sufficient internal consistency ($\alpha = 0.79$), we combined these four items in an overall measure of leader choice. Participants indicated their choice for each item by selecting one image from a series of five images of men (in the men condition) or women (in the women condition). The

---

*We pretested whether the manipulation would lead people to perceive a company as either competitive or collaborative in a sample of 77 American adults. After reading each scenario, participants reported significantly different ratings concerning competition versus collaboration, $t(76) = 2.74$, $p = .02$. Thus, the manipulation proved effective.*
five images\textsuperscript{2} showed the same basic male (female) face of a Caucasian adult. The images featured different degrees of masculinity, forming a scale from low masculinity (1) to high masculinity (5), with each number indicating a 50 percent decrease or increase in masculinity relative to the original composite: 1 = 100 percent masculinity, 2 = 50 percent masculinity, 3 = original composite, 4 = +50 percent masculinity, and 5 = +100 percent masculinity. All five images were displayed in one row after each of the four questions mentioned above, but the order of presentation of the images was randomized to avoid displaying an obviously linear masculinity scale. Based on participants’ responses, we calculated two scores. First, we calculated participants’ overall leader masculinity preference by averaging the responses on the 1–5 masculinity scale across the four leadership choice questions. Second, we recorded participants’ specific leader masculinity preference for each of the five leader images, counting for how many of the four questions participants selected each of the leader images. When a participant chose a specific leader image, the image received a score of 1. Across the four questions, each leader image thus received scores ranging from 0 (the participant chose the image for none of the leader choice questions) to 4 (the participant chose the image for all of the four questions) from each participant. We then calculated a mean value for each leader image, indicating how often participants on average chose this image. (For ease of interpretation, we also report the relative percentages, again indicating how often participants on average chose the leader image.)

Data Quality Check

The data were collected through an online panel, thus outside a supervised laboratory. Although this approach is increasingly common (Bohannon, 2011) and the data thus collected tend to be of similar or better quality than the data collected in laboratory studies (Buhrmester et al., 2011), there is a need to ensure that participants followed the instructions and provided meaningful answers. In line with other recent organizational research using online panels (Brands et al., 2015), we thus checked the quality of the data prior to our analyses, based on four criteria. Specifically, we established that participants’ responses would be excluded if (1) participants’ response time was longer than 2 SD above the mean, as extended response times may imply that distractions occurred during participation; (2) participants had the same IP address, as this would indicate repeated participation; (3) a participant’s response pattern indicated a lack of attention to the questionnaire items (e.g., all images were selected from just one side of the screen); or (4) participants incorrectly answered an attention check\textsuperscript{3} that was included in the assessment of demographics toward the end of the questionnaire. Based on these four criteria, none of the participants had to be excluded.

Results

We were interested in whether preferences for facial masculinity differ for men and women leaders in competitive and collaborative contexts. Thus we began our analyses with an analysis of variance (ANOVA), entering leader gender (women leaders versus men leaders) and leadership context (competition versus collaboration) as fixed factors, and overall leader masculinity preference as the dependent variable. The means and SDs for each condition are listed in Table 1. Participants preferred high facial masculinity more for men leaders than for women leaders, as indicated by a significant main effect of leader gender, $F(1, 183) = 21.34, p < .01$. Participants also preferred high facial masculinity more for competitive contexts than for collaborative contexts, as indicated by a significant main effect of leadership context, $F(1, 183) = 3.83, p = .05$. Participants’ preferences were not dependent on the specific configuration of leader gender

\textsuperscript{2} The images were taken from a previous study with permission of the corresponding author (Re et al, 2013). The images were created with the software Psychomorph following established procedures (Perrett et al., 1998). Differences in masculinity reflect relative distances between facial landmarks of the man and woman’s faces.

\textsuperscript{3} The multiple choice question asked participants for how many hours on average they slept, and offered as answer options: 1 = 0 hours, 2 = 410 hours, 3 = 24 hours. If participants report 1 or 3, they show a lack of attention to the question.

<p>| TABLE 1 |
| Study 1: Overall Leader Masculinity Preferences in Collaborative versus Competitive Contexts |</p>
<table>
<thead>
<tr>
<th>Leader Gender</th>
<th>Leadership Context</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woman leader</td>
<td>Collaboration</td>
<td>3.03</td>
<td>0.84</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Competition</td>
<td>3.14</td>
<td>1.26</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.09</td>
<td>1.08</td>
<td>96</td>
</tr>
<tr>
<td>Man leader</td>
<td>Collaboration</td>
<td>3.52</td>
<td>0.91</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Competition</td>
<td>3.97</td>
<td>0.74</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.76</td>
<td>0.85</td>
<td>91</td>
</tr>
<tr>
<td>Total</td>
<td>Collaboration</td>
<td>3.27</td>
<td>0.91</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Competition</td>
<td>3.54</td>
<td>1.12</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.42</td>
<td>1.03</td>
<td>187</td>
</tr>
</tbody>
</table>
and leadership context: There was no significant interaction effect, $F(1, 183) = 1.48$, $p = .23$. We examined whether the inclusion of participant gender would affect results in an analysis of covariance (ANCOVA). This was not the case.

The results of these analyses are in line with prior theory and research concerning facial masculinity and leadership choice. However, we suspected that these analyses might disguise preferences for women leaders with low facial masculinity in competitive environments. Indeed, a visual inspection of the distribution of specific leader masculinity preferences suggested that participants’ leader choices followed different patterns, depending on leader gender. To test whether the response distributions for leader choices for men and women were significantly different, we applied a Kolmogorov–Smirnov test. This nonparametric test is commonly used to check whether data are normally distributed by testing the difference between a sample distribution and a standard normal distribution. But this test can also be used to examine whether any two distributions are significantly different from one another. We examined whether the distribution of leader choices varied by leader gender in collaborative and competitive contexts. The results indicated that choices for men leaders follow a different distribution than choices for women leaders, both in competitive ($p < .01$) and collaborative settings ($p < .01$).

To examine whether the distributions differ at the low end or the high end of facial masculinity, we plotted the response frequencies for men and women leaders for both the collaborative and the competitive condition in Tables 2 and 3 (see also Figure 1). For men, Table 1 shows the familiar pattern of a preference for higher masculinity in competitive settings. Across the four leader choice questions, participants selected the image of a man with very high facial masculinity ($100\%$) more often in competitive settings ($M = 1.48$; $37$ percent) than in collaborative settings ($M = 0.90$; $23$ percent). The image of a man with very low facial masculinity ($0\%$) was selected more often in collaborative settings ($M = 0.27$; $33$ percent) than in competitive settings ($M = 0.12$; $6$ percent).

### TABLE 2

<table>
<thead>
<tr>
<th>Masculinity</th>
<th>1 ($-100%$)</th>
<th>2 ($-50%$)</th>
<th>3 (original)</th>
<th>4 ($+50%$)</th>
<th>5 ($+100%$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Leadership Context</strong></td>
<td><strong>Collaboration</strong></td>
<td><strong>Percentage</strong></td>
<td><strong>Mean</strong></td>
<td><strong>SD</strong></td>
<td><strong>Competition</strong></td>
</tr>
<tr>
<td><strong>Percentage</strong></td>
<td>7.0</td>
<td>33.0</td>
<td>27.0</td>
<td>16.0</td>
<td>17.0</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.27</td>
<td>1.32</td>
<td>1.09</td>
<td>0.64</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.45</td>
<td>1.23</td>
<td>1.12</td>
<td>1.04</td>
<td>1.18</td>
</tr>
</tbody>
</table>

**Note.** This table shows how often a particular leader was selected in each condition. Percentages, means, and SDs indicate how many times on average each image was selected for the four leader choice questions.

Number of collaboration = 42 participants.

Number of competition = 49 participants.

### TABLE 3

<table>
<thead>
<tr>
<th>Masculinity</th>
<th>1 ($-100%$)</th>
<th>2 ($-50%$)</th>
<th>3 (original)</th>
<th>4 ($+50%$)</th>
<th>5 ($+100%$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Leadership Context</strong></td>
<td><strong>Collaboration</strong></td>
<td><strong>Percentage</strong></td>
<td><strong>Mean</strong></td>
<td><strong>SD</strong></td>
<td><strong>Competition</strong></td>
</tr>
<tr>
<td><strong>Percentage</strong></td>
<td>20.0</td>
<td>18.0</td>
<td>19.0</td>
<td>12.0</td>
<td>30.0</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.81</td>
<td>0.73</td>
<td>0.77</td>
<td>0.48</td>
<td>1.21</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>1.37</td>
<td>1.05</td>
<td>1.08</td>
<td>0.70</td>
<td>1.40</td>
</tr>
</tbody>
</table>

**Note.** This table shows how often a particular leader was selected in each condition. Percentages, means, and SDs indicate how many times on average each image was selected for the four leader choice questions.

Number of collaboration = 44 participants.

Number of competition = 52 participants.
masculinity was rarely chosen, and less often in the competitive settings (M = 0.12; 3 percent) than in collaborative settings (M = 0.31; 8 percent). We conducted $t$ tests to examine whether these results were statistically significant. Men with high facial masculinity (+100 percent) were selected more often as leaders in the competitive than in the collaborative scenario [M = 1.49 versus M = 0.90, $t(89) = 2.122$, $p = .037$]. Men with low facial masculinity (100 percent) were not selected more often; to the contrary, they tended to be selected less often in the competition than in the collaboration scenario, with results trending toward significance [M = 0.31 versus M = 0.12, $t(89) = 1.705$, $p = .092$].

The pattern for women leaders looks remarkably different however. Figure 2 shows that the image of a woman very high in facial masculinity was chosen more frequently across the four leader choice questions in the competitive context (M = 1.21; 30 percent) compared to the collaborative context (M = 0.68; 17 percent); and the image of a woman very low in facial masculinity was also chosen more frequently in the competitive context (M = 0.81; 20 percent) compared to the collaborative context (M = 0.27; 7 percent). $t$ tests that compared how often on average participants selected women with high and low facial masculinity in competitive as opposed to collaborative environments confirmed the pattern of results. Women leaders with high facial masculinity were selected more often in the competitive than in the collaborative condition [M = 1.21 versus M = 0.68, $t(94) = 1.981$, $p = .047$], and women with low facial masculinity were also selected more often in the competition than in the collaboration scenario [M = 0.81 versus M = 0.27, $t(94) = 2.473$, $p = .015$].

Thus the difference between preferences for facial masculinity between men and women leaders manifests at the low end of the facial masculinity scale. When choosing a man leader for a competitive company environment, participants tended to prefer faces that were very high in masculinity, but when choosing a woman leader for such an environment, participants tended to prefer faces that were either very high in masculinity or very low in masculinity.

Discussion

The findings from this study indicate that people systematically select men and women for leadership positions based on facial features. However, 4 As robustness checks, we conducted four regression analyses with participants’ gender as a control variable. Results indicate that women leaders with high ($b = 0.54$, $p = .05$) and with low facial masculinity ($b = 0.51$, $p = .02$) were selected more often in the competitive than in the collaborative scenario. Men leaders with high ($b = 0.57$, $p = .05$) but not those with low masculinity ($b = 0.16$, $p = .04$) were selected more often in the competitive than in the collaborative scenario.
participants’ preferences for facial masculinity follow different patterns for men and women leaders, which stands in contrast to findings from prior research (Spisak, Dekker, et al., 2012). We did indeed find that high masculinity (in men and women) is more preferred in competitive than in collaborative scenarios (Re et al., 2013). Yet, in competitive scenarios, women leaders with low facial masculinity were also preferred more than in collaborative settings. In contrast, for men leaders with low masculinity, this was not the case. This preference for low masculinity among women leaders in a scenario of competition is somewhat surprising, given that low facial masculinity has been linked to perceptions of warmth, cooperativeness, and lower levels of dominance (Perrett et al., 1998).

Why then is it that in this study, women with low facial masculinity are preferred more in competitive than in collaborative settings? Previous research suggests that people choose leaders with facial features that signal traits that are likely beneficial for a particular task (Little, 2014). Given that women with low facial masculinity were more often chosen in competitive settings, we wondered whether women with low facial masculinity are attributed competitiveness. That is what we chose to examine in Study 2.

**STUDY 2**

Do inferences from facial masculinity concerning competitiveness and collaborativeness differ for men and women? In Study 2, we addressed this question through a categorization task. In a prestudy, we first created the stimulus material by asking participants to write essays about men and women leaders. From those essays, we extracted statements, or idea units (Chafe, 1980), describing behaviors, characteristics, and experiences of leaders. Next, we asked independent coders to categorize each extracted statement in terms of whether it was indicative of someone who exhibited competitive behavior.

In the main study, we asked participants (differing from those of the prestudy) to assign each statement to a leader with high facial masculinity, or to one with low facial masculinity, or to neither. We wondered if participants would associate statements indicative of competitive behavior to women with both high and low facial masculinity but only to men high in facial masculinity. This would help us better understand findings from Study 1.

**Prestudy**

The prestudy served to obtain content-rich stimulus material that captured the behaviors, characteristics,
and experiences that ordinary people associate with successful leaders in competitive businesses. We asked 44 American adults (43 percent women; 85 percent White, 7 percent Asian, 4 percent Latino, and 4 percent Black) recruited through Amazon Mechanical Turk to write a success story about a business leader. Participants averaged 33 years of age (SD = 10.19) and 12 years of work experience (SD = 10.15).

After agreeing to take part in our research, participants accessed our website and were instructed to write an essay about the success of the CEO of a fictional company that we named ACME Inc. We described the company as a major movie production company, and continued as follows: “ACME is operating in a highly competitive industry, producing expensive movies to be screened in movie theaters all over the world. When the CEO took over 5 years ago, ACME was in trouble. It lagged behind its competitors, it lacked good movie ideas, and its top managers were estranged. Five years later, these problems are history, and ACME has produced a record profit this year. Your article should describe what led to this success.” We also showed participants an image with the caption “CEO of ACME Inc.” that was taken from Study 1. The image was either a man or a woman, and either high, medium, or low facial masculinity. Thus, we obtained a diverse set of essays about successful leaders, and protagonists were not only men and women leaders with high, but also leaders with medium and low facial masculinity. After viewing the photo and reading the instructions, participants proceeded to the next page, where they were given space to write their essay. Three text boxes gave participants some structure for their text. The first text box was titled “CEO characteristics,” the next “Coworkers,” and the last “Competitors.”

We separated the essays by the gender of the protagonist, according to the image we had shown to participants.6 By keeping texts about men and women protagonists separate, we retained gender-specific meanings and connotations (e.g., being a varsity cheerleader or captain of the soccer team may have different connotations for men than for women). From all essays, we extracted only those statements that contained general descriptions of the protagonist, such as those describing characteristics (e.g., “is very hands on”), experiences (e.g., “valedictorian of her class for high school and college”), or behavior (e.g., advice giving: “many of his friends come to him for business advice”). We ignored other statements that were context specific (e.g., “[His] strategy has been to move away from flagship movies and [he] invested time and resources into R&D”) or that were unrelated to the protagonist (e.g., “ACME had been spreading itself thin in the past, trying to please all audiences from kids to senior citizens”). After obvious duplicates were removed, this process resulted in 211 statements describing men’s behavior, characteristics, and experience, and 250 statements describing women’s behavior, characteristics, and experience in a leadership context.

Coding of Stimulus Material

To code the extent to which the statements captured the central variables of competitive and collaborative qualities, we handed the statements to two independent coders. The focus of this study was on lay perceptions, thus we hired two coders from the general American population through the online platform Odesk (now called Upwork). The coders live in different parts of the United States and offer research support on a freelance basis. One coder was a woman with 16 years of work experience, and the other coder a man with 6 years of work experience. Both were rated highly for the quality of their work on Odesk. The coders were unaware of the purpose of this study. We asked each coder separately whether a given statement described a behavior, characteristic, or experience that was indicative of someone who is good at competing, collaborating, or neither. As categorization was not straightforward for many statements, we decided beforehand that only statements would be retained for which both coders independently agreed on a category. This approach is analogous to scale development, where items that show poor fit are routinely culled (Hinkin, 1995). In our context, coders converged on the same categorization for 126 (60 percent) of the 211 statements about men; and for 167 (67 percent) of the 250 statements about women. Of these statements, 42 were indicative of competition for men, and 62 were indicative of competition for women. The remaining statements were indicative of collaboration (68 for men, 68 for women) or of neither competition nor collaboration (16 for men, 37 for women).

Main Study

We recruited 252 American adults (53 percent women; 74 percent White, 8 percent Asian, 7 percent Latino, 8 percent Black, and other 3 percent) via the online panel Microworkers. Participants averaged 32 years of age (SD = 11.6) and 11 years of work experience (SD = 10.15). After agreeing to take part in this research, participants were randomly allocated to one of four groups. In each group, participants were shown two images of faces and a list of statements. The images were taken from Study 1, and the statements were taken from the prestudy. In group 1, participants saw

---

6 One essay was excluded, because the participant did not correctly recognize the gender of the depicted CEO. Although we had shown the participant the image of a woman, the participant used male pronouns in the essay.
two male faces: One of the faces was the unmodified original composite, while the comparison image had been modified to depict high masculinity (+100 percent). Group 2 was identical to group 1, except that the comparison image depicted a man with low masculinity (100 percent). In group 3, participants saw two female faces: One of the faces was the unmodified original composite, and the comparison image had been modified to depict high masculinity (+100 percent). In group 4, the comparison image depicted a woman with low masculinity (100 percent). The survey website presented statements about men to groups 1 and 2, and statements about women to groups 3 and 4. Participants were asked to assign each statement to a leader with modified masculinity (high in groups 1 and 3, or low in groups 2 and 4), to a leader with unmodified masculinity or to neither leader. Because we were interested in attributions to the modified image featuring high or low masculinity, respectively, we coded such attributions for the purpose of analysis as 1 and other attributions (to the unmodified leader image or neither leader) as 0. Each participant was assigned 80 randomly drawn statements from the overall statement pool to reduce cognitive load. Participant gender was assessed as a control variable.

Data Quality Check

Prior to the analyses, we checked the quality of the data. As in Study 1, we established that participants' responses would be excluded if (1) participants' response times were longer than 2 SD above the mean; (2) participants had the same IP address; or (3) a participant's response pattern was indicative of a lack of attention to the questionnaire items (e.g., all statements were attributed to the same image). To keep the survey short, a separate question to test attention was not included. None of the participants met any of the criteria for exclusion.

Analysis

The research design featured repeated answers from participants who each assessed multiple statements, so we used generalized estimating equations (GEE) to calculate regression parameters (Liang & Zeger, 1986). We were interested in whether participants attributed a statement to a particular image (1) or not (0), thus we specified a binary logistic model. GEE is an extension to generalized linear models that can process binary variables with repeated measurements (Hu, Goldberg, Hedeker, Flay, & Pentz, 1998). The estimated regression coefficients from the GEE method are similar to those from the standard logistic models, but in contrast to standard logistic models, GEE does not produce biased standard errors for data featuring repeated measurements (Hu et al., 1998). An important consideration in GEE is the ratio of Level 1 data (in our case, the number of statements that the participants rated) to Level 2 data (in our case, the number of participants). If this ratio is too low,
GEE models fail to converge (Bell, Ferron, & Kromrey, 2008). The ratio for our dataset across the different conditions was on average 0.57 (SD = 0.04), and all models converged without errors. For each GEE regression model, we report the odds ratios (OR), the lower and upper limit of the confidence intervals, and the p values as indicators of significance (Cumming, 2014). The default value of OR = 1 shows the absence of an effect in OR. Values lower than 1 indicate a negative effect, and values greater than 1 indicate a positive effect.

### Results

To what extent are competitive statements attributed to men and women with low or high facial masculinity? Table 4 shows the results of the regression models, controlling for participant gender. The four regression models predict perceptions of competitiveness for men (Models 1 and 2) and women (Models 3 and 4), with low facial masculinity (Models 1 and 3) and high facial masculinity (Models 2 and 4). First, we examined the results for men. Model 1 indicates that the chances for the man with low masculinity to be ascribed competitive attributes were low, with results tending toward significance (OR = 0.82 [0.66–1.02], p = .07), whereas Model 2 shows that the chances for the man with high masculinity to be attributed competitiveness were high (OR = 1.34 [1.04–1.73], p = .03).

Next, we examined the results for women. Women with both high and low masculinity were ascribed competitiveness. Model 3 shows that participants were likely to attribute competitive statements to the woman with low facial masculinity (OR = 3.39 [2.89–3.97], p < .01), and Model 4 indicates that the woman with high facial masculinity...
also had a high chance to be seen as competitive (OR = 3.41 [2.44–4.76], \( p < .01 \)).

Tables 5 and 6 list the competitive statements that more than 50 percent of participants assigned to the women and men leaders, respectively.

### Study 3

In this study, we showed images of men’s or women’s faces to research participants and asked them to rate those images for masculinity in comparison to a standard face of a man or a woman, respectively. Participants were not informed that the men’s or women’s faces they were rating were, in fact, the faces of CEOs of S&P 500 companies. We wondered: Would participants perceive differences in the facial masculinity of men CEOs and women CEOs?

### Participants and Procedure

We recruited 98 American adults via the online panel Microworkers to participate in a web-based research study. The sample was balanced in terms of gender (52 percent women) and quite diverse in terms of ethnicity (62 percent White, 16 percent Black, 8 percent Asian, 7 percent Latino, and 7 percent other). Participants averaged 35 years of age (SD = 12.03), and 14 years (SD = 11.52) of work experience.

### Table 6

<table>
<thead>
<tr>
<th>Competitive Statements Concerning Men leaders</th>
<th>Low (%)</th>
<th>High (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Coworkers consider him very driven</td>
<td>33</td>
<td>64</td>
</tr>
<tr>
<td>2. Does not tolerate people trying to act like they are smarter or wiser than he is</td>
<td>27</td>
<td>63</td>
</tr>
<tr>
<td>3. Some employees think he is too harsh</td>
<td>31</td>
<td>63</td>
</tr>
<tr>
<td>4. He would handle conflicts with management with a strong will, unwilling to give in</td>
<td>33</td>
<td>63</td>
</tr>
<tr>
<td>5. He treats others with respect to a degree, but mostly believes he is right</td>
<td>31</td>
<td>61</td>
</tr>
<tr>
<td>6. Analyzing his competitors</td>
<td>32</td>
<td>61</td>
</tr>
<tr>
<td>7. Has high expectations</td>
<td>35</td>
<td>59</td>
</tr>
<tr>
<td>8. They would say he is always on top of them</td>
<td>33</td>
<td>59</td>
</tr>
<tr>
<td>9. It is only on rare occasions that he will listen to an alternate viewpoint</td>
<td>29</td>
<td>58</td>
</tr>
<tr>
<td>10. Known to walk out of meetings, red-faced, until he settles down</td>
<td>29</td>
<td>57</td>
</tr>
<tr>
<td>11. He was different than the rest. He wanted perfection</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>12. Academic excelling student</td>
<td>61</td>
<td>24</td>
</tr>
<tr>
<td>13. In college, he began to stand out because of his academic achievements</td>
<td>50</td>
<td>21</td>
</tr>
<tr>
<td>14. Was an excellent student through his high school and college years</td>
<td>57</td>
<td>21</td>
</tr>
<tr>
<td>15. He graduated with a 3.8 GPA</td>
<td>54</td>
<td>20</td>
</tr>
<tr>
<td>16. Graduated with honors</td>
<td>54</td>
<td>17</td>
</tr>
<tr>
<td>17. Two majors in college</td>
<td>61</td>
<td>13</td>
</tr>
<tr>
<td>18. Received high marks in school as well as excelling in his artistic pursuits</td>
<td>61</td>
<td>12</td>
</tr>
</tbody>
</table>

Note. Participants were asked to evaluate a random subset of 80 out of 126 statements and were divided into two groups. Between 40 and 56 participants evaluated each statement for the man low in masculinity (−100 percent). Between 40 and 56 participants evaluated each statement for the man high in masculinity (+100 percent). For reasons of comparability, results are presented as percentages. Percentages indicate how many people agreed that a competitive statement was characteristic of the individual shown in the target image rather than the reference image or neither of the two. This list shows competitive statements which received more than 50 percent agreement.
experience. After agreeing to take part in our research, participants accessed the survey website and were randomly assigned to two groups. One group \((n = 50; 48\% \text{ women})\) was asked to assess the masculinity of men, and the other group \((n = 48; 56\% \text{ women})\) was asked to assess the masculinity of women. We did not mix the two groups because we wanted to ensure that ratings of masculinity of men were independent of, and unaffected by, ratings of masculinity for women, and vice versa. Each group evaluated 25 images in total. The 25 images of men and the 25 pictures of women all featured CEOs of large companies, but importantly, the study participants were not informed that they were rating CEOs.

We asked participants: “Please indicate how masculine each of the following men (women) is compared to the reference image.” Participants indicated their perception of each person’s masculinity compared to a reference image (1 = much less masculine, 2 = less masculine, 3 = just as masculine, 4 = more masculine, 5 = much more masculine). The reference image featured the man’s (woman’s) face unaltered in masculinity that we had used in Studies 1 and 2 (Re et al., 2013). We used the reference image to calibrate participants’ perceptions of masculinity. The reference image was shown repeatedly throughout the questionnaire; thus, as participants scrolled down the page, they were still able to compare images.

Data Quality Check

Prior to the analyses, we applied the same data quality checks as in Study 1. None of the participants were excluded.

Results

We ran GEE regression models to assess whether differences between perceptions of masculinity among women and men CEOs were statistically significant. The models predicted ratings of “very low masculinity” and “very high masculinity” separately at first, and as a robustness check, we specified a multinomial model that included both categories simultaneously.

For the purpose of these analyses, we coded perceptions of masculinity into two binary dependent variables: very low and very high masculinity. Very low masculinity was indicated by a score of 1 on the perception of masculinity scale, but not by scores from 2 to 5 on that scale. Thus we assigned a value of 1 for a score of 1, and a value of 0 for all other values (2–5). Likewise, very high masculinity was indicated by a score of 5 on the perception of masculinity scale, but not by scores from 1 to 4 on that scale. Thus we assigned a value of 1 for a score of 5, and a value of 0 for all other values (1–4). The use of binary dependent variables requires estimation by way of generalized linear models rather than by ordinary linear regression. As multiple raters evaluated the same pictures, the data are non-independent. Thus, as in Study 2, we ran regressions

### TABLE 7

<table>
<thead>
<tr>
<th>Perceived Masculinity of CEO*</th>
<th>Lowerb</th>
<th>95% CI</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant gendera</td>
<td>1.75**</td>
<td>1.43–2.14</td>
<td>1.79**</td>
<td>1.29–2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEO gendera</td>
<td>1.65**</td>
<td>1.34–2.02</td>
<td>2.30**</td>
<td>1.64–3.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness of fit (QICC)</td>
<td>2,414.88</td>
<td></td>
<td>1,156.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlated data</td>
<td>98</td>
<td></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of participants</td>
<td></td>
<td></td>
<td>2,448</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of CEO images</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. OR higher than one indicate a positive effect lower than 1 indicate a negative effect. The corrected quasi likelihood under Independence Model Criterion (QICC) for model fit is in the “small is better format.”

*Woman = 1, man = 0.

**Compared to the reference image.

*p < .01

6 At the time this study was conducted, there were 25 women listed as CEOs of S&P 500 companies (Catalyst, 2014). All these women were included in our study. From the remaining 475 men listed as CEOs of S&P 500 companies, we randomly selected 25 to be included in our study. Specifically, we generated random numbers from 1 to 500 and then selected the CEOs of those companies whose rank in the S&P 500 list was identical to the random number. If, by coincidence, the CEO was a woman (one event) or if we were unable to obtain a verified image of the company CEO (two events), then we selected the CEO of the company listed subsequently on the ranking of the S&P 500 companies. All pictures were obtained through the Internet and formatted such that CEO’s faces were shown in roughly similar size.
using GEE, an extension of generalized linear models for repeated measurements.

Table 7 shows that more women CEOs than men CEOs were perceived to be much lower in masculinity than the reference picture (Model 1, OR = 1.65 [1.34–2.02], p < 0.01) and, also, that more women CEOs than men CEOs were perceived to be much higher in masculinity than the reference picture (Model 2, OR = 2.30 [1.64–3.23], p < 0.01). These findings suggest that among the 50 CEOs sampled from S&P 500 companies, there is greater range of masculinity among women CEOs than among men CEOs. More women CEOs than men CEOs were perceived to be at either end of the masculinity scale, thus having either high or low masculinity.

As a robustness check, we specified a multinomial GEE regression model that included both categories (i.e., “very low masculinity” and “very high masculinity”) together with the reference category. Thus we specified the dependent variable to have three possible values, corresponding to CEO masculinity being perceived as, first, low (1); second, medium (24); or, third, high (5). Results corroborated the findings that women CEOs more often than men CEOs were perceived to have very low and very high rather than medium masculinity (OR = 2.37 [1.91–2.94], p < .01). Results remained statistically significant when participant gender was included as a control variable.

An inspection of the descriptive statistics corroborated the findings. Mean results indicated that on average, perceived masculinity of men and women CEOs was close to the middle of the scale (3), and that the mean ratings of masculinity in the group that compared men CEOs to the reference image (M = 2.92, SD = 1.14) were similar to the mean ratings of masculinity in the group that compared women CEOs to the reference image (M = 2.88, SD = 0.97). The distributions of facial masculinity ratings, however, differed in men and women CEOs, as indicated by a Kolmogorov–Smirnov test (p < .01). Table 8 shows that among women CEOs, participants perceived there to be more extremes—in 21 percent of all cases, women CEOs were seen as being at the extreme ends of the scale (i.e., high [9 percent] or low [12 percent] in masculinity). Only in 11 percent of all cases were men CEOs at the extreme ends of the scale (i.e., high [4 percent] or low [7 percent] in masculinity).

### Discussion

This study investigated whether in a field context, distributions of facial masculinity differ between men and women leaders. Among women CEOs, we find a comparatively greater range of facial masculinity than among men CEOs of S&P 500 companies. Such companies face considerable competition (Friedman, 2005; Gupta, Govindarajan, & Wang, 2008) and Studies 1 and 2 suggest that in such contexts, women with either low or high levels of masculinity would be preferred. In line with this conclusion, the results of this study suggest that not only women with high masculinity but also women with low masculinity become leaders in the corporate world. This finding challenges existing theory and research, which posits a universal advantage of facial masculinity in competitive settings (Re et al., 2013; Spisak, Dekker, et al., 2012).

Somewhat surprisingly, in this study, men CEOs were not perceived to be particularly high in facial masculinity. Based on existing research and Studies 1 and 2, we were expecting people to perceive male CEOs to feature high rather than medium masculinity. But ratings suggested that many CEOs were perceived relatively similarly to the reference image depicting the unmodified original composite man. One possible explanation is that the man shown in the reference image, in the absence of other comparisons, was perceived as quite masculine.

### Table 8

<table>
<thead>
<tr>
<th>Ratinga</th>
<th>Frequencies</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Woman CEOs</td>
<td>Man CEOs</td>
</tr>
<tr>
<td>1 (Much less masculine)</td>
<td>139</td>
<td>92</td>
</tr>
<tr>
<td>2 (Less masculine)</td>
<td>308</td>
<td>340</td>
</tr>
<tr>
<td>3 (Similarly masculine)</td>
<td>370</td>
<td>498</td>
</tr>
<tr>
<td>4 (More masculine)</td>
<td>276</td>
<td>266</td>
</tr>
<tr>
<td>5 (Much more masculine)</td>
<td>107</td>
<td>54</td>
</tr>
<tr>
<td>Number of CEO images</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Number of participants</td>
<td>48</td>
<td>50</td>
</tr>
</tbody>
</table>

*aCompared to the reference image.*

Author’s voice: What was the most difficult or challenging aspect of this research project and paper?
This points to the limitation that the images we used across studies allow inferences concerning relative deviations in masculinity from the baseline original composite, but the absolute masculinity of the baseline original composite is difficult to discern. Therefore, findings from Study 3 cannot be interpreted as evidence that men CEOs are of average masculinity; rather, that men CEOs (in contrast to women CEOs) were perceived more similarly in terms of their facial masculinity.

**GENERAL DISCUSSION**

In a series of three studies, we found that leadership choices based on facial masculinity differ for women and men leaders. Findings from Study 1 indicate that masculine-looking men and women are more preferred in competitive than in collaborative contexts. However, we also discovered that women with low facial masculinity are chosen more often as leaders in competitive than in collaborative scenarios. This is remarkable because previous research has linked low facial masculinity to perceptions of warmth and cooperativeness (Perrett et al., 1998). Study 2 corroborated that, indeed, women with low facial masculinity are attributed competitiveness, whereas men with low facial masculinity are not seen as competitive. Given these findings we expected that in general business settings, we would find men leaders who look rather masculine and women leaders who look either rather masculine or not masculine at all. This notion was supported by Study 3, in which we found a greater range of facial masculinity among women CEOs than among men CEOs of S&P 500 companies.

This research confirms the basic assumption of biosocial models of leadership: People’s perceptions of facial masculinity systematically affect their choices of leaders and the attributes that they assign to leaders. But our findings indicate that leadership choices for men and women differ more than is suggested by prior studies (Re et al., 2013; Spisak, Dekker, et al., 2012). Extant theory and research converge on the notion that both women and men benefit in leadership selection processes from facial masculinity, particularly in competitive settings, but our findings suggest that women can also benefit from low facial masculinity.

Perhaps the most striking discovery of our research was that women with low facial masculinity were ascribed competitiveness. These attributions go beyond the typical gender-biased leadership prototype (Hogue & Lord, 2007). Gender stereotype research suggests that low facial masculinity should facilitate schemas concerning stereotypically female qualities, such as communality, warmth, and solidarity (Zebrowitz & Montepare, 2008). Indeed, it has been argued that followers expect feminine-looking leaders to be cooperative (Spisak, Homan, et al., 2012). But competitiveness among women has long been ignored or overlooked (Clutton-Brock & Huchard, 2013). Our research shows that low facial masculinity in women is linked to competitiveness and not only to cooperativeness as prior research suggests.

Biosocial models of leadership posit that the linkage between facial masculinity and leadership choice is rooted in hormonal processes that shape both facial appearance and behavior (Spisak, Dekker, et al., 2012). While for men the link between testosterone, facial masculinity, and status-striving behavior seems clear (Stanton & Schultheiss, 2009), for women, there is some evidence that dominant behavior is not only due to high levels of testosterone, but is also associated with a potent form of estrogen (Stanton & Schultheiss, 2007, 2009). Facial masculinity results from the interplay of testosterone and estrogen. Relative changes in visual cues indicative of testosterone and estrogen levels are reflected in the facial composites used in this study (Perrett et al., 1998; Re et al., 2013). To perceivers, the image that we used of the woman’s face with low facial masculinity likely signaled relatively higher levels of estrogen (Perrett et al., 1998; Smith et al., 2006). If estrogen makes faces less masculine, but is also linked to dominant behavior (Stanton & Schultheiss, 2007, 2009), then our discoveries that women with low facial masculinity are chosen as leaders and seen as competitive appear quite plausible and in line with recent biological research on hormones. Thus striving for dominance is associated with the hormone testosterone in men, but appears to be linked with the hormone estrogen in women (Schultheiss, 2013).

Our findings in turn suggest that evolutionary-based implicit leadership theories are more sophisticated than previously assumed. Apparently, people do not only take manifestations of testosterone into account when conferring leadership on others, but also manifestations of estrogen. Our research suggests that not only are women with low facial masculinity suitable leaders in communal, cooperative settings.
involving a “niche for feminine leadership” (Re et al., 2013; Spisak, Dekker, et al., 2012: 6), but that they are also seen (Study 1) and employed (Study 3) as leaders in competitive settings. The leadership disadvantages arising from low facial masculinity appear to affect only one gender: men. In line with previous research that suggests that men’s status suffers to the extent that they do not fit with gendered expectations (Moss-Racusin et al., 2010), men with low facial masculinity were not seen as competitive leaders in our research.

This study is an example of how a more differentiated understanding of social and organizational issues—such as the question of who should be the leader—can be advanced through a more fine-grained understanding of the complex biological, and particularly hormonal, processes that underpin overt appearance and behavior (Becker & Menges, 2013). The effects revealed in this research suggest patterns that are much more nuanced than those derived from a binary viewpoint concerning biological sex. Thus, these effects challenge extant gender stereotypes and implicit leadership research that has, for the most part, neglected to take biologically rooted differences within sex categories into account. Future theory and research should move from the examination of differences between men and women to the consideration of differences among men and differences among women.

Our research is limited in its focus on one cultural setting and the use of a relatively restricted set of images. Prior research shows that facial masculinity effects manifest across cultures (Re et al., 2013; Spisak, Dekker, et al., 2012), so we assume the effects would hold outside North America. Nonetheless, we encourage crosscultural research to replicate and extend our findings. The images used in this research were validated in previous research (Re et al., 2013) and were suited to our research context. Future research should aim to create and utilize images that allow for inferences concerning absolute masculinity. Such images would make it possible to resolve the question of why a large proportion of men CEOs in Study 3 were perceived as similar in masculinity to the reference image, even though most research points out that men leaders benefit from high facial masculinity in competitive settings.

Finally, future research should investigate how low facial masculinity among men and women leaders relates to perceived femininity. There may be different types of perceived femininity, signaling either competitiveness (Stanton & Schultheiss, 2007) or cooperativeness (Perrett et al., 1998).

In conclusion, this research suggests that people seem to implicitly understand facial markers of the hormones testosterone and estrogen as signs for competitiveness and leadership suitability in men and women with high facial masculinity, as well as in women with low facial masculinity. What people read in the faces of Yahoo’s Marissa Mayer, Hewlett Packard’s Meg Whitman, and Facebook’s Sheryl Sandberg is that these nonmasculine-looking women are perceived as competitive, and therefore well suited to lead their businesses toward the achievement of organizational goals and future success.

REFERENCES


Willis, J., & Todorov, A. 2006. First impressions: Making up your mind after a 100ms exposure to a face. *Psychological Science*, 17: 592–598.


Raphael Silberzahn is a research associate at the University of Cambridge - Judge Business School, from where he also obtained his PhD.

Jochen Menges is a University Professor of Leadership and Human Resource Management at the WHU - Otto Beisheim School of Management in Düsseldorf, Germany. He obtained his PhD from the University of St. Gallen in Switzerland.