Routledge
Taylor & Francis Group

Makeup your mind: The impact of styling on perceived competence and warmth of female leaders

Jennifer Klatt[a], Sabrina C. Eimler[b], and Nicole C. Krämer[a]

[a]University of Duisburg-Essen; [b]Hochschule Ruhr West - University of Applied Sciences

ABSTRACT

Women are still underrepresented at the highest management levels. The think-manager-think-male phenomenon suggests that leadership is associated with male rather than female attributes. Although styling has been shown to influence the evaluation of women's leadership abilities, the relevant specific features have been left remarkably unaddressed. In a 2 × 2 × 2 × 2 (skirt/pants, with/without jewelry, loose hair/braid, with/without makeup) between-subjects design, 354 participants evaluated a woman in a photograph. Women with makeup, pants, or with jewelry were rated as more competent than women without makeup, with skirts, or without jewelry. A combination of loose hair and no makeup was perceived as warmest, and women with loose hair were more likely to be hired than those with braids. In sum, even subtle changes in styling have a strong impact on how women's leadership abilities are evaluated.

Only 5% of Fortune 500 companies' CEOs are female (Fortune, 2015). Because women hold more academic degrees in the United States or the European Union than men (European Commission, 2012; National Center for Education Statistics, 2014), this is not an educational problem. Numerous studies identified gender stereotypes as one obstacle for women in leadership, as they influence observers' behavioral expectations and, in turn, person perception (e.g. Deutsch, LeBaron, & Fryer, 1987; Heilman, 2001). Moreover, gender stereotypes also affect the behavior of the target person, who may, for example, strive to conform to stereotypic role norms (e.g. Moss-Racusin & Rudman, 2010; Rudman & Glick, 2001).

Psychological research found that people attribute the same characteristics to men and managers, while these similarities in attributions are less prominent for women and managers. This effect, known as the *think-manager-think-male* phenomenon (Schein, 1973, 1975), has been used to explain the unequal distribution. Therefore, since a male appearance ostensibly has a positive impact, the question arises of whether a masculine styling could be advantageous for women in this regard.

This question has been investigated, among others, by Von Rennenkampff, Kühnen, and Sczesny (2003), who found that men and women who had a masculine styling were regarded as having better leadership abilities than those with a feminine styling. However, so far, research has only used an overall variation, in which numerous aspects such as clothes, hairstyle, or makeup were manipulated simultaneously. What remains unclear is the impact of single, isolated styling elements. Therefore, the goal of this study was to overcome confounding aspects and vary specific elements systematically.

The study at hand used photographs of women in business clothes as stimulus material, in which the styling was varied with regard to their clothes, hair, makeup, and jewelry. In an online questionnaire, participants were asked to rate the women's competence and warmth and then state how likely they would hire them for a leadership position.

CONTACT Jennifer Klatt ✉ jennifer.klatt@uni-due.de 🖵 Department of Social Psychology, University of Duisburg-Essen, Forsthausweg 2, Duisburg 47057, Germany.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/vsoc.

Gender stereotypes

Gender stereotypes are cognitive representations of, for example, attributes that are typically associated with the respective gender, such as personality characteristics, abilities, or typical behavior (Eckes, 2002). Research (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972; Carli & Eagly, 2001) has identified attributes that are seen as typical and desirable for men (e.g. aggressive, clever, dominant, or active) and women (e.g. emotional, friendly, loyal, or warm).

One possible explanation for why these characteristics may be seen as desirable for the respective gender lies in the *social role theory* (Eagly, 1987), which describes that desirable attributes derive from the traditional distribution of men and women in social roles. Women are associated with characteristics like kindness or sensitivity (e.g. in the context of childrearing). These attributes have been termed *communal*. Men, by contrast, tend to occupy *agentic* roles that are said to require energetic and self-confident behavior (e.g. leadership positions or paid jobs, as opposed to housework, in general). Role-congruent behavior is generally evaluated positively, whereas non-congruent behavior leads to negative consequences, such as the degradation of a person (Diekman & Goodfriend, 2006). These mechanisms lead to individuals who learn gendered behavior, which is also described as tertiary sex characteristics (Birdwhistell, 1970), to maintain a dichotomy in the definitions of gender (Lorber, 1994). Although the female role has changed and become more strongly associated with masculine characteristics than in the past (in general: Diekman & Eagly, 2000; in the business context: Bosak & Sczesny, 2011), the question arises whether the desirable attributes have changed accordingly.

Another perspective that needs to be taken into consideration with regard to gender stereotypes is the *stereotype content model* (Fiske, Cuddy, Glick, & Xu, 2002). This model describes two main categories that define the contents of all stereotypes: *competence* and *warmth*. It was found that many stereotypes are ambivalent in the sense that they show a high value on one of the dimensions and a low value on the other. In the context of gender stereotypes, examples are stereotypes regarding housewives and career women. Housewives are seen as having low competence but receive high ratings for their warmth, whereas career women who are perceived as highly competent supposedly lack warmth (Fiske et al., 2002).

Gender stereotypes in the leadership context

The association of male attributes with leaders' attributes is described in the *think-manager-think-male* phenomenon (Schein, 1973, 1975; Sczesny, 2003a). Participants were asked to rate attributes according to how strongly they characterize men in general, managers, and women in general. The results showed that attributions made to men and managers correlated strongly, whereas for women, other attributes were rated as characteristic. In a meta-analysis of studies dealing with leadership and gender stereotypes, the authors found a greater correlation for men and managers than for women and managers, and managers were more strongly linked to agency than to communion (Koenig, Eagly, Mitchell, & Ristikari, 2011).

Heilman asserts that there is a perceived *lack of fit* (Heilman, 1983, 2001) for women aspiring to leadership positions. Males, as well as male leaders and leaders in general, are characterized as task-oriented, whereas women or female leaders are described as person-oriented (Heilman, 1983). The same is true for the perceived leadership style of male and female leaders: Women are perceived as having more person-oriented (e.g. empathy or communication skills) and men as having more task-oriented (e.g. ambition and negotiation skills) leadership abilities (Sczesny, 2003a). With regard to the consequences for the evaluation of a person, it was found that in working situations, non-conforming behavior (e.g. a woman who shows masculine behavior) can lead to decreased acknowledgment of the person's work and a decreased likelihood of being hired, which is called *backlash* (Burgess & Borgida, 1999; Moss-Racusin & Rudman, 2010; Rudman & Glick, 2001).

Styling in the context of gender stereotypes in leadership

It is not only gender itself that can activate gender stereotypes: Deaux and Lewis (1984) describe the significant influence of physical appearance on masculine characteristics, which they even found to explain more variance than a person's gender. According to Von Rennenkampff et al. (2003), there are two types of physical attributes: characteristics of the body that cannot be manipulated and are biologically given, and the changeable elements such as hairstyle and clothes; in the following, we will refer to the latter as styling elements. With respect to the biologically given characteristics, Sczesny, Spreemann, and Stahlberg (2006) found that persons with a masculine appearance were evaluated as having more leadership abilities regardless of their gender (Sczesny & Kühnen, 2004; Sczesny et al., 2006). Furthermore, Rule and Ambady (2009) demonstrated that the competence ratings of CEOs' faces correlated with company profits.

Research with regard to changeable elements of styling also demonstrates an influence on the perception of a person (Von Rennenkampff et al., 2003). In terms of purposefully chosen elements of physical appearance, several studies analyzed the effect of clothes. A positive influence of the professional evaluation, e.g. with regard to competence as well as the likelihood of getting hired of a (female) candidate, could be demonstrated both with regard to the styling's formality (Kwon & Johnson-Hillery, 1998) and grooming (Mack & Rainey, 1990) with a formal and well-groomed styling enhancing the evaluation. A study by Forsythe, Drake, and Cox (1985) asked whether a higher degree of masculinity can also improve the evaluation of a candidate. A positive relationship was found between the masculinity of the outfits and the likelihood of getting hired as a manager. This study result illustrates well that the think-manager-think-male paradigm is also applicable to styling aspects.

The relation between biologically given and self-chosen attributes of physical appearance was examined by Von Rennenkampff et al. (2003). Both factors influenced the perception of leadership abilities. Biologically masculine-looking persons as well as masculine-styled ones were rated more positively than the feminine versions. However, in a second study, the authors found that there is an advantage for feminine-looking persons if the leadership position explicitly requires person-oriented, and therefore stereotypically feminine, abilities (Von Rennenkampff et al., 2003).

In summary, research has demonstrated a huge impact of styling on the perception of women in general and especially of women in a working or leadership context. However, to our knowledge, so far, no experimental research exists on the impact of specific styling elements. When clothing, makeup, jewelry, and hairstyle are changed simultaneously, no conclusions can be drawn on the impact of individual styling elements. Therefore, this study aims to address this research gap.

Research questions

The *lack of fit* (Heilman, 2001) and the *think-manager-think-male* phenomenon (e.g. Schein, 1973) describe empirically well-supported barriers that women may face on their way to leadership positions. Women are not perceived as having the same abilities as men, with the latter being the same abilities that are necessary for and associated with a leadership position. This is surely a disadvantage, which already begins with reading a person's name, e.g. on a job application form. Additionally, research on sex-typed physical appearance also explains disadvantages for women. Deaux and Lewis (1984) even found that the impact of physical appearance is somewhat stronger than that of gender itself. If one combines this information with the findings of Sczesny and colleagues (2006) and Von Rennenkampff and colleagues (2003) regarding the influence of a masculine appearance on perceived leadership abilities, it can be concluded that not only the general impact of the gender, but also the effect of physical appearance, exerts a strong influence on the likelihood of a woman being perceived as a leader.

The study by Von Rennenkampff and colleagues (2003) does not examine which specific styling elements are most influential. In their study, an overall masculine styling is compared to

an overall feminine styling, which varies hairstyle, makeup, and clothes simultaneously. Other studies in this area use similar procedures: For example, Mack and Rainey (1990) vary hairstyle, makeup, clothes, and jewelry. This leads to inconsistent results: For example, in Von Rennenkampff and colleagues' (2003) study, the feature of makeup was used for the feminine styling condition, which was rated as having fewer leadership abilities. By contrast, in the study by Mack and Rainey (1990), makeup was used in the well-groomed condition and led to an increased likelihood of getting hired. Thus, based on these two studies, it is not possible to draw an overall conclusion on whether or not the use of makeup leads to an ascription of leadership qualities. The same problems can be found with regard to jewelry. For clothes, the aspect of fashion also interferes with the transferability of previous research: For example, in Forsythe and colleagues' study (Forsythe et al., 1985) even the most masculine outfit includes a skirt, probably due to business/fashion standards of the 80s. The only element that has no conflicting results so far and is likely to be comparable to today's fashion is the hairstyle: Loose hair was used for the feminine styling, and a braid was considered masculine (Von Rennenkampff et al., 2003). Therefore, we hypothesize:

Hypothesis 1 (H1): Compared to women with a braid, women with loose hair will be...
   (a) ...rated as more feminine.
   (b) ...rated as less masculine.

Additionally, we tested how every single specific styling element influences the perception of a person in an occupational leadership situation. For hairstyle, it can be assumed that the differences with regard to masculinity/femininity lead to different perceived leadership abilities as well. Thus, we derive the following hypothesis from the literature on gender stereotypes in the leadership context, as described in the stereotype content model (Fiske et al., 2002):

H2: Compared to women with a braid, women with loose hair will be...
   (a) ...rated as less competent.
   (b) ...rated as warmer.
   (c) ...less likely to get hired.

However, for the other styling elements and possible interaction effect between the elements, no hypotheses can be stated. Thus, we ask:

Research Question 1 (RQ1): Which styling elements influence...
   (a) ...a woman's perceived competence?
   (b) ...a woman's perceived warmth?
   (c) ...a woman's likelihood of her getting hired?

Method

Design

In a $2 \times 2 \times 2 \times 2$ between-subjects design, one photograph was presented to each participant. Four styling elements were varied (loose hair/braid, no makeup/makeup, skirt/pants and no jewelry/jewelry), which resulted in 16 possible combinations. In order to prevent potential effects from only being valid for one specific woman, four women were styled accordingly and photographed with each of the 16 styling combinations, resulting in 64 photographs in total. Therefore, for each of the styling conditions, one of the four photographs was chosen randomly for presentation. The study was approved by the ethic commission of the University of Duisburg-Essen.

Participants

The sample consisted of 354 participants (212 women) aged between 18 and 55 years ($M = 25.41$, $SD = 8.20$). They were mostly students of psychology and computer science (52.3%) or employees (22%). 28% of the participants were in leadership positions or had already made employee-hiring decisions in the past. The majority (79.7%) of the sample had the general qualification for university entrance or an academic degree. Further, 11.6% had a migration background. Participants were recruited online via message boards both from the university and from websites on hobbies such as movies or clothes, resulting in a sample with both students and a broader range of people who use Internet boards. They had the opportunity to win a voucher as a reward for their participation.

Stimulus material

To prevent individual differences, such as in hair color, from causing an overall distortion of the ratings, and therefore to reach a higher generalizability, we used photographs of more than one woman. To this aim, in a pretest, photographs of 14 women with the same styling were rated regarding their perceived competence and the probability of them being a leader (each rated on a 5-point Likert scale). All women gave informed consent on using their pictures for scientific studies, and some of them agreed as well on having them used in scientific publications. After inspecting the means of the women, 10 out of 14 women had to be dismissed, because one of the variables had an extreme value. The remaining four women were tested for significant differences with a t-test for the two means that showed the greatest difference. They did not differ significantly regarding their perceived competence and probability of being a leader. All women were white and had long hair (so that a braid could be made) with two of them being blonde, one brunette, and one black-haired. Both the hair color and the individual effect of each woman caused no significant differences in the main study, thus, the results of this study should be generalizable to women with similar characteristics.

Sixteen photos of each woman showed all possible combinations of the styling: loose hair/braid, no makeup/makeup, skirt/pants and no jewelry/jewelry. All women had a neutral facial expression and stood in a straight position looking toward the camera in a photo studio. The same light and camera settings were used for the shootings, as well as the same clothes (a white shirt with a black skirt or black pants), same makeup (foundation, eyeliner, mascara, eye shadow, and discreet lipstick), and same jewelry (a black pearl necklace with small earrings). All 16 styling combinations were photographed for each woman (see Figure 1).

Procedure

In the beginning, participants gave informed consent to taking part in a study in which they would have to rate a person in a photograph. For each participant, one photograph was presented without further information to avoid confounding biases, which could be evoked by a specific job title or branch of industry. The woman in the picture was referred to as a "person" or "person in the picture" throughout the questionnaire. In the end, the participants were thanked and given the opportunity to take part in a lottery.

Measures

Competence and warmth

We assessed competence and warmth by means of various scales. Due to the similarity and high correlations of the scales, which are common in this research area, a factor analysis with 59 items was conducted. The items were taken from the following scales: 20 items were taken from the BEM sex role inventory (Bem, 1974; German version: Schneider-Düker & Kohler, 1988), with 10 items each from the masculine (e.g. "determined") and feminine (e.g. "caring") subscales. Twelve items were
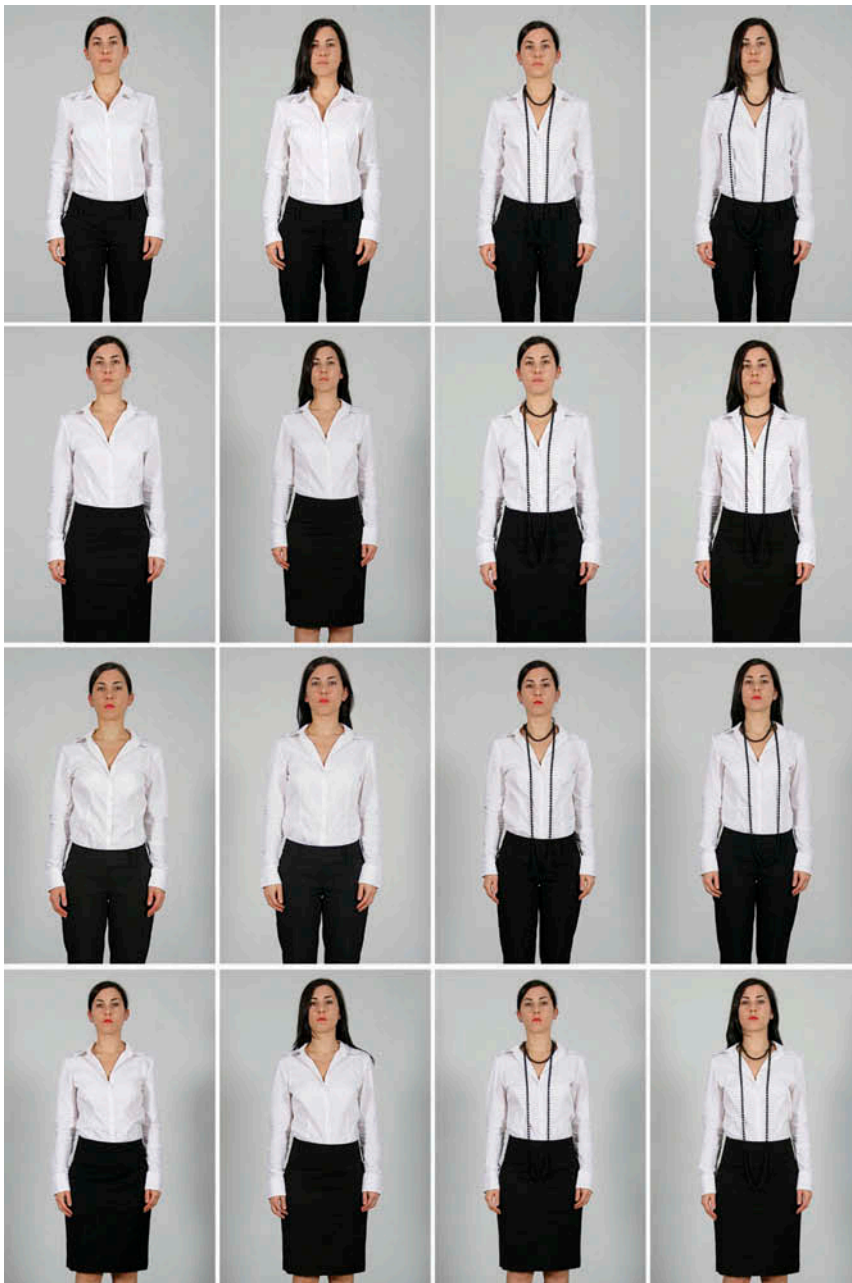
Figure 1. 16 photographs were taken of each woman. In this figure, the 16 possible styling combinations are shown for one of the four women. An example for a masculine styling can be found in the first row, first image from the left: A braid without makeup and pants. An example for a feminine styling is the second image from the left in the first row: Loose hair and no makeup. The combination of loose hair and makeup was rated as arrogant: an example can be found in row four, fourth image from the left.

taken from the stereotype content model (Fiske et al., 2002) with items from three subscales: competence (five items, e.g. "intelligent"), warmth (four items, e.g. "warm") and status (three items, e.g. "How well educated is this person?"). Two items on leadership abilities (e.g. "The person is an effective leader") and seven items each on masculine traits (e.g. "rational") and feminine traits (e.g. "compassionate") were chosen from Bosak and Sczesny (2011), and 16 items were chosen on

task orientation (e.g. "authority") and person orientation (e.g. "empathy"; Sczesny, 2003a; German version: 2003b). All items were measured on a 5-point Likert scale (from 1 = Strongly disagree to 5 = Strongly agree). Duplicate items were removed. A factor analysis with principal axis factoring was selected using the varimax rotation as rotation method. A cutoff value of .7 was used to determine the most influential items in the factors. Two factors resulted from the factor analysis: There were 16 items for the factor *competence* (Cronbach's α = .97) and 13 items for the factor *warmth* (Cronbach's α = .96). They explained 50.13% of the variance. For an overview of the items and the factor loadings, see Table 1. For further analyses, sum scores were calculated.

Likelihood of getting hired

To assess the likelihood of getting hired, an item was constructed ("Would you hire the person you saw in the photograph for a leadership position in a company?"), which was rated on a 5-point Likert scale (from 1 = Under no circumstances to 5 = Most certainly).

Masculinity and femininity of the styling

For the masculinity/femininity of the person in the photograph, two items ("The person appears to be physically rather feminine." and "The person appears to be physically rather masculine.") rated on a 5-point Likert scale were created (from 1 = Strongly disagree to 5 = Strongly agree).

Arrogance

As a special measure for collecting data on a possible backlash effect, the item "arrogant" was rated on a 5-point Likert scale (from 1 = Strongly disagree to 5 = Strongly agree).

Table 1. Original items and factor loading values for the factors competence and warmth (*N* = 354).

Item	Factor Competence	Factor Warmth
Confident	*.870*	−.050
Self-reliant	*.858*	−.098
Strong	*.831*	−.049
Competent	*.827*	.007
Strong personality	*.826*	−.014
Ambitious	*.822*	−.157
Determined	*.813*	−.113
Successful	*.803*	.020
Willing to take a stand	*.793*	−.233
Self-confidence	*.787*	−.047
Forceful	*.784*	.005
Self-assertiveness	*.753*	−.118
Courageous	*.753*	.114
Persuasive	*.723*	.098
Independent	*.714*	−.004
Decisive	*.704*	.053
Performance-oriented	*.704*	−.131
Sensitive to the needs of others	−.089	*.847*
Loving	−.055	*.844*
Caring	−.084	*.833*
Sincere	−.049	*.827*
Compassionate	−.133	*.822*
Kind	−.147	*.818*
Good-Natured	−.153	*.814*
Gentle	−.153	*.801*
Affectionate	−.075	*.798*
Soft-hearted	−.245	*.788*
Warm	−.087	*.750*
Empathy	−.175	*.749*

Note. Factor loadings over .7 appear in italics.

Procedure

In an online questionnaire, participants were randomly assigned to one of the 16 conditions (all possible combinations of makeup vs. no makeup, jewelry vs. no jewelry, skirt vs. pants, and loose hair vs. braid). They were asked to look at the photograph for their respective condition thoroughly and subsequently rate it with regard to the dependent variables. Afterward, data on socio-demographics and on the participants' experience with leadership were collected. Finally, the participants had the opportunity to enter the lottery, read a debriefing, and were thanked for their participation.

Results

A MANOVA was performed to examine the hypothesis and the research question. The four styling elements, clothing, jewelry, hairstyle, and makeup, were used as fixed factors and the described measures as dependent variables. Using Pillai's trace, there was a significant effect for makeup (V = 0.68, F (7, 332) = 3.48, p = .001) and the interaction of makeup and hairstyle (V = 0.46, $F(7, 332)$ = 2.28, p = .028). The other main effects and interactions were not significant. Therefore, the results caused by makeup and by an interaction between makeup and hairstyle should be given extra attention.

In the following, the results for masculinity and femininity of the styling as well as the ratings for arrogance are presented in the section on the general evaluation of the styling. Afterward, the evaluation of leadership abilities contains the results of competence, warmth, and the likelihood of getting hired.

General evaluation of the styling

Masculinity and femininity of each styling

H1 stated that a loose hairstyle should lead to increased femininity (a), and decreased masculinity (b) compared to the braid. Women who had loose hair ($F(1, 354)$ = 6.78, p = .010, η^2 = .02, Cohen's d = 0.29) were rated as more feminine (M = 3.61, SD = 1.04) than those with a braid (M = 3.29, SD = 1.17). Moreover, there was an interaction effect with the hairstyle and the makeup. The most feminine combination was loose hair without makeup, followed by loose hair with makeup and a braid with makeup. The least feminine option was a braid without makeup ($F(1, 354)$ = 5.89, p = .016, η^2 = .02, Cohen's d between 0.02 and 0.53). For mean values, see Figure 2. H1a was therefore supported.

Similar results were found for masculinity. Again, the combination of a braid without makeup was evaluated as most masculine, followed by makeup with loose hair, no makeup and loose hair, and finally a combination of makeup and a braid ($F(1, 354)$ = 5.53, p = .019, η^2 = .02, Cohen's d between 0.03 and 0.46). For means and standard deviations, see Figure 2. Additionally ($F(1, 354)$ = 3.93, p = .048, η^2 = .01, Cohen's d = 0.23), pants were perceived as more masculine (M = 2.45, SD = 1.23) than a skirt (M = 2.18, SD = 1.16). An example for a masculine styling and a feminine styling according to these results can be found in Figure 1. H1b has not been supported.

In summary, hairstyle and makeup seem to have a consistent impact on the perceived masculinity and femininity of a woman's styling because the interaction of hair and makeup was found for both variables. Therefore, the special role of hair and makeup needs to be taken into account when discussing the results.

Relation of styling and arrogance

Regarding main effects, only makeup was found to be influencing the women's arrogance ($F(1, 354)$ = 10.93, p = .001, η^2 = .03, Cohen's d = 0.34), with makeup being evaluated as more arrogant (M = 3.50, SD = 1.14) than no makeup (M = 3.07, SD = 1.35). It should be noted, that this is one of the largest effect sizes in this study. Additionally, there were two interaction effects: Hair and makeup ($F(1, 354)$ = 7.85, p = .005, η^2 = .02, Cohen's d between 0.06 and 0.63) interacted with loose hair and no makeup being especially low,
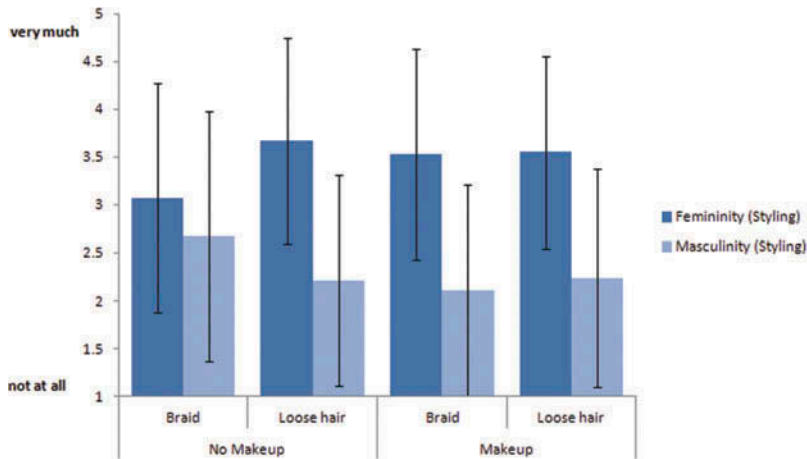
Figure 2. Descriptive statistics of the interaction effect between hairstyle and makeup regarding the femininity and masculinity of the styling.

then followed by a combination of no makeup and a braid, makeup with a braid, and finally loose hair and makeup, which received the highest ratings on arrogance. Further, hairstyle and clothes ($F(1, 354) = 4.51$, $p = .034$, $\eta^2 = .01$, Cohen's d between 0.03 and 0.27) influence the perceived arrogance of a woman in a way that loose hair and a skirt are perceived as the least arrogant, a combination of pants and a braid as the second least arrogant, then a combination of a braid and a skirt while loose hair and pants are rated as the most arrogant combination (for means see Figures 3 and 4).

Evaluation of leadership abilities

Competence
RQ1a and *H2a* address the question of whether the styling causes a difference in the evaluation of the perceived competence of the women. It was found ($F(1, 354) = 4.64$, $p = .032$, $\eta^2 = .01$, Cohen's $d = 0.21$) that women who wore pants were perceived as more competent ($M = 3.38$, $SD = 0.93$) than
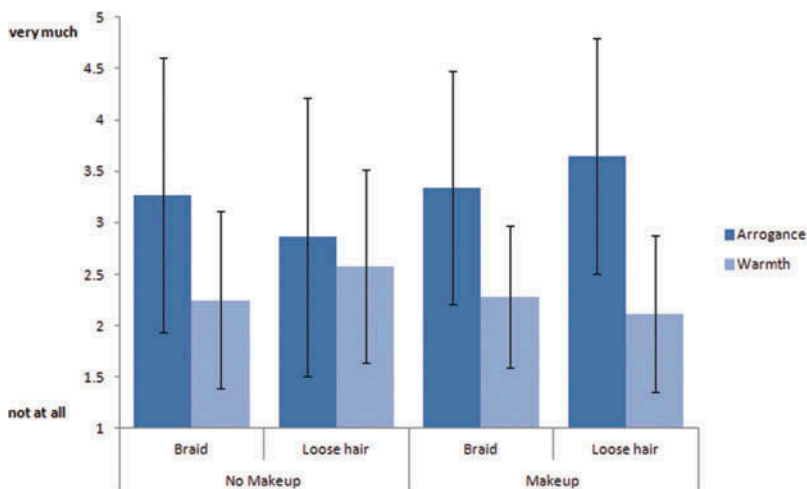

Figure 3. Descriptive statistics of the interaction effect between hairstyle and makeup regarding arrogance and warmth.
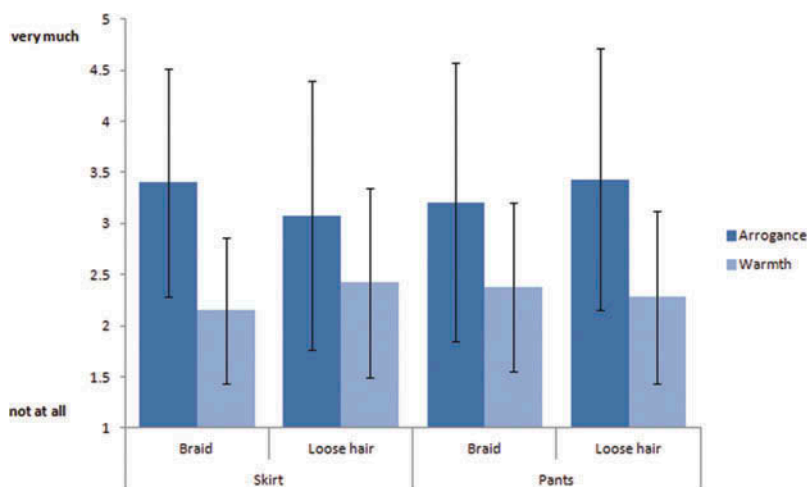
Figure 4. Descriptive statistics of the interaction effect between hairstyle and clothes regarding arrogance and warmth.

those who wore a skirt (M = 3.19, SD = 0.91). Additionally, women with makeup (M = 3.40, SD = 0.89) were rated as more competent ($F(1, 354)$ = 4.95, p = .027, η^2 = .01, Cohen's d = 0.23) than those without makeup (M = 3.19, SD = 0.95). Further ($F(1, 354)$ = 4.08, p = .044, η^2 = .01, Cohen's d = 0.20), women with jewelry (M = 3.38, SD = 0.88) were evaluated as more competent than those without jewelry (M = 3.20, SD = 0.96). No interaction effects occurred. Therefore, regarding *RQ1a*, it can be stated that pants, makeup, and jewelry strengthen the perceived competence of a woman. Hairstyle did not influence competence ratings; *H1a* needs to be rejected.

Warmth

RQ1b and *H2b* asked whether the styling also causes a difference in the perception of the women's warmth. For warmth, a main effect occurred ($F(1, 354)$ = 5.49, p = .020, η^2 = .02, Cohen's d = 0.25): Makeup was perceived as less warm (M = 2.20, SD = 0.73) than no makeup (M = 2.41, SD = 0.92). Two interaction effects were found. Hairstyle and makeup ($F(1, 354)$ = 9.22, p = .003, η^2 = .03, Cohen's d between 0.04 and 0.55): Loose hair and no makeup was perceived as the warmest styling, followed by a braid with makeup, a braid without makeup, and the least warm styling being a combination of loose hair and makeup. This effect has also one of the largest effect sizes in this study and should thus be given special attention. Hairstyle and clothing ($F(1, 354)$ = 4.98, p = .026, η^2 = .02, Cohen's d between 0.05 and 0.33): Loose hair and a skirt were rated as warmest, followed by a braid with pants, loose hair and pants, and a braid with a skirt were rated as least warm (for means, see Figures 3 and 4). Therefore, makeup, as well as combinations of hair and makeup and of hair and clothes, influence a woman's perceived warmth. Hairstyle alone did not, thus, *H2b* is not supported.

Likelihood of getting hired

RQ1c asked whether there is a difference in the ratings for the likelihood of getting hired for a leadership position by means of the styling. *H2c* assumed that women with a braid were more likely to get hired than those with loose hair. Regarding the question of how likely the women were to get hired, a main effect was found for the hairstyle ($F(1, 354)$ = 5.24, p = .023, η^2 = .02, Cohen's d = 0.22): Women with loose hair were more likely to get hired (M = 2.94, SD = 1.03) than those with a braid (M = 2.70, SD = 1.11). No interaction effects emerged from the calculation. Thus, only the hairstyle influences the likelihood of getting hired. However, the effect was exactly opposite to what was expected in *H2c*. Thus, the hypothesis needs to be rejected, too.

Discussion

H1 stated the difference that hairstyling creates with regard to a woman's perceived masculinity/ femininity. However, hairstyle did only influence the ratings for femininity and not masculinity. Von Rennenkampff and colleagues (2003) varied masculinity/femininity, i.e., using the hairstyle, which is partly in line with these findings. Also, there were interaction effects of hairstyle and makeup, in which both the most masculine and the most feminine styling had no makeup, but varied with regard to the hairstyle. This highlights the importance of (1) varying styling elements separately, because otherwise interaction effects, as the above mentioned, cannot occur; and of (2) separate measures for masculinity/femininity, because it was shown here that femininity is not equal to low masculinity or vice versa.

H2 stated effects caused by hairstyle with regard to leadership abilities. The think-manager-think-male phenomenon (Schein, 1973, 1975) was applied to the context of styling for women in leadership before. Studies so far found the concept to be true also for masculine appearances (Forsythe et al., 1985; Von Rennenkampff et al., 2003). However, in this study, none of the effects stated in *H2a–c* were found. In the case of *H2c*, even an opposite effect could be observed: Women with loose hair, who were rated as more feminine-looking, were hired more often than those with a braid. An explanation for this could be some information the participants assumed on the position, because Von Rennenkampff et al. (2003) found that feminine styling is advantageous if the position requires stereotypically feminine abilities. However, we purposefully provided no information on the position whatsoever. Still, we cannot determine now which background participants had in mind when they evaluated the stimulus person. Another explanation addresses again the question of whether the think-manager-think-male phenomenon can really be applied to the styling context. The empirical background on this topic is still pretty scarce. The findings at hand directly contradict the findings of Forsythe et al. (1985), because in their study, the woman with the most masculine-looking styling was hired, and here, the most feminine-looking one was chosen. There did not seem to be a perceived lack of fit (Heilman, 1983, 2001) for the more feminine-looking candidate.

Regarding the *RQ1a–c*, which assumed that, as a result of the styling, differences in the perception of women regarding their competence, warmth, and likelihood of getting hired will be found, several significant results emerged.

Regarding competence, women with makeup, pants, or jewelry were rated higher than those wearing the other options, i.e. no makeup, jewelry, or skirt. These women score high on task-oriented variables such as self-confidence or determination. With regard to jewelry and makeup, this can be also related to Mack and Rainey's (1990) finding that these elements fitted with an overall well-groomed styling, which enhances perceived leadership abilities. With regard to the pants, it can be assumed that they enhance the evaluation of competence because they were rated as more masculine than a skirt, a finding that is in line with Von Rennenkampff et al. (2003). However, not all masculine elements enhance the ratings of competence, and not all elements that enhance the competence ratings are specifically masculine, again not supporting the think-manager-think-male assumption.

Regarding warmth, an effect with one of the largest effect sizes occurred: The combination of loose hair and no makeup received the highest ratings. Therefore, these women are associated most strongly with skills that are person-oriented, such as compassion or sincerity. The interaction of loose hair and no makeup was not only significantly rated as very feminine, but it was also rated as the least masculine (two different interaction effects emerged with the same result). Therefore, the high femininity of the styling might have activated this rather feminine skill set. However, loose hair alone was rated feminine, too, but no evidence for increased ratings of warmth could be found. However, it has to be noted that the interactions have to be interpreted with caution since the sample size was rather small.

Styling variables influenced the perceived arrogance of the women, indicating that certain combinations of these variables might trigger backlash. One of the findings for arrogance, the main effect of makeup being more arrogant than no makeup, had one of the two largest effect

sizes of all results, indicating that this aspect should not be underestimated. In line with this, makeup also received lower ratings regarding warmth, which can be seen as a social penalty for women with makeup as well. This fits the stereotype content model (Fiske et al., 2002) because women with makeup who received high ratings with regard to competence did also receive low ratings for warmth, which is consistent with an ambivalent stereotype (e.g. a "career woman").

Since Von Rennenkampff et al. (2003) found that a rather masculine styling should be beneficial for the evaluation of leadership abilities, it would be reasonable to assume that beneficial styling elements for the perception of competence and the likelihood of getting hired equal a masculine styling. In our study, this was not found to be the case. Masculine styling includes pants and a combination of a braid without makeup, whereas feminine styling includes loose hair and loose hair without makeup. On the one hand, some styling elements, such as makeup or jewelry, are not perceived to be specifically masculine or feminine. On the other hand, loose hair, for example, is perceived as feminine but increases the likelihood of getting hired. Thus, a simple equalization of a masculine styling and perceived leadership skills, as could have been assumed based on previous research (Forsythe et al., 1985; Von Rennenkampff et al., 2003), and the think-manager-think-male phenomenon (Schein, 1973, 1975) cannot be made here. Future research should try to investigate why an overall change of styling, which varies femininity/masculinity ratings, influences perceived leadership abilities, but a single variation of masculine/feminine styling elements does not. The interactions between the elements may be related to other concepts than masculinity/femininity as well (as, for example, the arrogance finding adds another dimension). Future research could investigate this by using a broader methodological spectrum, e.g. qualitative research methods. In general, it is noteworthy that even small changes in the styling lead to significant effects on person perception.

Returning to the question of what women can bear in mind when choosing their styling for a job interview for a leadership position, women can rely on several main effects such as makeup, jewelry, and pants, which each increase perceived competence and loose hair, which in turn increases the likelihood of getting hired. However, specific combinations may also lead to undesirable effects such as the combination of makeup and loose hair, which is associated with high rating of arrogance, indicating the possibility that persons wearing this styling might be more endangered of meeting backlash effects.

Limitations and outlook

This study has several limitations that need to be discussed. Although around 30% of the participants were leaders or had made employee-hiring decisions, a sample with more actual leaders would have made it possible to distinguish whether leaders differ from other people with regard to their evaluation of women. In our sample, no significant systematical differences based on this variable occurred, indicating that students and persons with relevant professional experience have comparable judgments regarding styling in a professional context. Another variable that could be influential is the question of whether the evaluator had ever had contact with a female leader before in a working context. Future research should aim at a broader distributed sample to examine these questions as well. Further, the sample size at hand in combination with the explorative procedure resulting in 70 calculations means that the results should in general be interpreted with caution. This is especially true for null effects, such as the effects for leadership skills caused by hairstyle, which was hypothesized in *H1*. Although no such results were found in this study, it should not be concluded that this connection is definitely non-existing. Future research should try to replicate these findings or further enlarge sample size and reduce calculations. Also, it should be noted that the effect sizes found are rather small, and hence, the question arises whether such small significant effects have any practical consequences. However, Agars (2004) argues that gender effects often get underestimated because of small effect sizes. However, gender is a very prominent and repetitively perceived social stimulus, which cumulates small effects to a significant influence in everyday life.

Furthermore, in this study, specific details on the vacant leadership position were intentionally omitted in order to avoid associations with typical male- or female-dominated fields. This was done because the industrial context determines the desirable skills for the position (Von Rennenkampff et al., 2003). However, we cannot be sure what participants had in mind when answering the questions. It might have been the case that the opposite of the desired measures happened: Maybe a certain styling made participants associate the non-specified position with a rather male or female industry. Future research could add details on the positions and vary them systematically to examine this question.

It cannot be guaranteed that there were no differences between the women in the stimulus material that may have been influential. A huge effort has been made to photograph as many women as possible who were willing to take part in a 1-hour photo shoot. Although the characteristics of interest for this study, e.g. competence, showed no significant differences, there may have been biological differences, which can have a huge impact given, for example, the results of Von Rennenkampff and colleagues (2003). However, before starting the calculations for the research questions, it was checked whether there were any differences that were caused by one of the four randomly used women and no significant effects emerged. In our study, the existing differences between the women (e.g. hair color) enhance generalizability of the results. Nevertheless, in future studies, it would be interesting to systematically vary these additional differences. Additionally, the generalizability does only apply to white women with a comparable body type, age, attractiveness, and without visible stigmata. Although there were several variations of the above-mentioned dimensions in the photo material (such as race, hair length, and body type), the pretest to determine comparable women resulted in four white women of roughly the same age with a slender figure, probably indicating influences of these variables. Thus, a variation of race and other variables would be beneficial to achieve a broader generalizability to more women. Especially in the case of age, this would additionally increase the credibility of the women in the role of female leaders because they were estimated as being younger than 30 years, which is quite young for a person in a leading position.

Future research could also include other styling elements, such as different types of makeup (eye makeup, lipstick, or only foundation), clothes (a jacket could be added), jewelry (large earrings or subtle necklaces), different hair length, and other hairstyles, as well as other accessories, such as glasses, shoes, or a scarf.

Moreover, the participants' gender could be examined because several studies from the think-manager-think-male context suggest that male and female evaluators may vary in the evaluation of a person and his/her abilities (e.g. Brenner, Tomkiewicz, & Schein, 1989). However, not all studies identified these differences (Schein, 1973, 1975).

Conclusion

The aim of the study was to reveal which styling elements have the strongest impact on how women's leadership abilities are perceived. The assumption that a masculine styling is beneficial in general could not be supported, contradicting earlier research in this field. Results showed that makeup, jewelry, and pants seem to have an enhancing effect on the perceived competence. The combination of loose hair and no makeup is rated as warm, and loose hair increased the likelihood of getting hired. Therefore, loose hair, as well as makeup, jewelry, and pants, each seem to be advantageous for a woman in a leadership position.

Acknowledgments

Notes on contributors

Jennifer Klatt has worked as a researcher in the field of social psychology and human computer-interaction at the University of Duisburg-Essen. Her research interests are gender stereotypes in the context of leadership, virtual reality, video games, and virtual humans. *Sabrina C. Eimler* is Professor for Human Factors and Gender Studies at the Hochschule Ruhr West - University of Applied Sciences in Bottrop, Germany. She is especially interested in psychological aspects of human-computer-interaction of all kinds, ranging from social media to autonomous driving. *Nicole C. Krämer* is Professor for Social Psychology: Media and Communication at the University of Duisburg-Essen in Germany. Her research interests include human-technology-interaction, social media, and gender issues.

References

Agars, M. D. (2004). Reconsidering the impact of gender stereotypes on the advancement of women in organizations. *Psychology of Women Quarterly*, *28*, 103–111. doi:10.1111/j.1471-6402.2004.00127.x

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162. doi:10.1037/h0036215

Birdwhistell, R. (1970). *Kinesics and context: Essays on body motion communication*. Philadelphia, PA: University of Pennsylvania Press.

Bosak, J., & Sczesny, S. (2011). Exploring the dynamics of incongruent beliefs about women and leaders. *British Journal of Management*, *22*(2), 254–269. doi:10.1111/j.1467-8551.2010.00731.x

Brenner, O. C., Tomkiewicz, J., & Schein, V. E. (1989). The relationship between sex role stereotypes and requisite management characteristics revisited. *The Academy of Management Journal*, *32*(3), 662–669. doi:10.2307/256439

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, *28*(2), 59–78. doi:10.1111/j.1540-4560.1972.tb00018.x

Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law*, *5*(3), 665–692. doi:10.1037/1076-8971.5.3.665

Carli, L. L., & Eagly, A. H. (2001). Gender, hierarchy, and leadership: An introduction. *Journal of Social Issues*, *57*(4), 629–636. doi:10.1111/0022-4537.00232

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology*, *46*(5), 991–1004. doi:10.1037/0022-3514.46.5.991

Deutsch, F. M., LeBaron, D., & Fryer, M. M. (1987). What is in a smile? *Psychology of Women Quarterly*, *11*(3), 341–352. doi:10.1111/j.1471-6402.1987.tb00908.x

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin*, *26*(10), 1171–1188. doi:10.1177/0146167200262001

Diekman, A. B., & Goodfriend, W. (2006). Rolling with the changes: A role congruity perspective on gender norms. *Psychology of Women Quarterly*, *30*(4), 369–383. doi:10.1111/j.1471-6402.2006.00312.x

Eagly, A. H. (1987). *Sex differences in social behavior: A social role interpretation*. Hillsdale, NY: Lawrence Erlbaum.

Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles*, *47*(3/4), 99–114. doi:10.1023/A:1021020920715

European Commission. (2012). *Women in economic decision-making in the EU: Progress report: A Europe 2020 initiative*. (European Union, Ed.). Retrieved from http://ec.europa.eu/justice/gender-equality/files/women-on-boards_en.pdf

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. doi:10.1037/0022-3514.82.6.878

Forsythe, S., Drake, M. F., & Cox, C. E. (1985). Influence of applicant's dress on interviewer's selection decisions. *Journal of Applied Psychology*, *70*(2), 374–378. doi:10.1037/0021-9010.70.2.374

Fortune. (2015). *Women CEOs in the Fortune 500*. Retrieved from http://fortune.com/2013/05/09/women-ceos-in-the-fortune-500/

Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. V, pp. 269–298). Greenwich, CT: JAI Press.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*(4), 657–674. doi:10.1111/0022-4537.00234

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*(4), 616–642. doi:10.1037/a0023557

Kwon, Y.-H., & Johnson-Hillery, J. (1998). College students' perceptions of occupational attributes based on formality of business attire. *Perceptual and Motor Skills*, *87*, 987–994. doi:10.2466/pms.1998.87.3.987

Lorber, J. (1994). "Night to his day": The social construction of gender. In J. Lorber (Ed.), *Paradoxes of gender* (pp. 13–36). New Haven, CT: Yale University Press.

Mack, D., & Rainey, D. (1990). Female applicants' grooming and personnel selection. *Journal of Social Behavior and Personality*, *5*(5), 399–408.

Moss-Racusin, C. A., & Rudman, L. A. (2010). Disruptions in women's self-promotion: The backlash avoidance model. *Psychology of Women Quarterly*, *34*(2), 186–202. doi:10.1111/j.1471-6402.2010.01561.x

National Center for Education Statistics. (2014). *Percentage of persons 25 to 29 years old with selected levels of educational attainment, by race/ethnicity and sex: Selected years, 1920 through 2013*. Retrieved from https://nces.ed.gov/programs/digest/d13/tables/dt13_104.20.asp

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*(4), 743–762. doi:10.1111/0022-4537.00239

Rule, N. O., & Ambady, N. (2009). She's got the look: Inferences from female chief executive officers' faces predict their success. *Sex Roles*, *61*(9–10), 644–652. doi:10.1007/s11199-009-9658-9

Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, *57*(2), 95–100. doi:10.1037/h0037128

Schein, V. E. (1975). Relationships between sex role stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology*, *60*(3), 340–344. doi:10.1037/h0076637

Schneider-Düker, M., & Kohler, A. (1988). Die Erfassung von Geschlechtsrollen: Ergebnisse zur deutschen Neukonstruktion des Bem Sex-Role-Inventory [Assessment of sex roles: Results of a German version of the Bem sex-role inventory]. *Diagnostica*, *34*(3), 256–270.

Sczesny, S. (2003a). A closer look beneath the surface: Various facets of the think-manager-think-male stereotype. *Sex Roles*, *49*(7/8), 353–363. doi:10.1023/A:1025112204526

Sczesny, S. (2003b). Führungskompetenz: Selbst- und Fremdwahrnehmung weiblicher und männlicher Führungskräfte [Leadership competence: Self-perception and perception of others of female and male leaders]. *Zeitschrift Für Sozialpsychologie*, *34*(3), 133–145. doi:10.1024//0044-3514.34.3.133

Sczesny, S., & Kühnen, U. (2004). Meta-cognition about biological sex and gender-stereotypic physical appearance: Consequences for the assessment of leadership competence. *Personality and Social Psychology Bulletin*, *30*(1), 13–21. doi:10.1177/0146167203258831

Sczesny, S., Spreemann, S., & Stahlberg, D. (2006). Masculine = competent? Physical appearance and sex as sources of gender-stereotypic attributions. *Swiss Journal of Psychology*, *65*(1), 15–23. doi:10.1024/1421-0185.65.1.15

Von Rennenkampff, A., Kühnen, U., & Sczesny, S. (2003). Die Attribution von Führungskompetenz in Abhängigkeit von geschlechtsstereotyper Kleidung [The attribution of leadership competence depending on gender-stereotypical clothing]. In U. Pasero (Ed.), *Gender - from costs to benefits* (pp. 170–182). Wiesbaden, Germany: Westdeutscher Verlag.