



## Does Bem's Psychological Androgyny map on gender or sex differences in faces?

Tristan MARHENKE<sup>1</sup> y Roland IMHOFF<sup>2</sup>

<sup>1</sup>Psychologischer Psychotherapeut in Erfstadt, Judenstraße

<sup>2</sup>Johannes Gutenberg-Universität Mainz, Mainz (JGU)

(Received on July 8, 2018; Accepted on October 10, 2018)

**ABSTRACT:** Sarah Bem introduced the concept of androgyny, which disconnects sex and gender and includes a continuous representation of gender. What has not been investigated so far is whether the particular qualities postulated by Bem are qualities of gender rather than sex-associated traits. In the present study, the reversed correlation task as a data driven approach was used to determine the implicit gender stereotypes across the faces of men and women and to create an ideal prototype of feminine and masculine faces. Then it was measured which impressions these faces evoke. Two studies and a pilot study (N=514) were conducted. The present study showed that gender and not sex is crucial for the attribution of social characteristics. Pictures of stereotypical faces have been found to be highly suitable for measuring masculinity and femininity. The continuous properties of masculinity and femininity, as outlined by Bem (1974), are still appropriate to differentiate between the stereotypical ideas of men and women.

**Keywords:** Reversed correlation task, face perception, gender, androgyny

*¿Dispone Bem de un mapa psicológico andrógino sobre las diferencias de género o de sexo en los rostros?*

**RESUMEN:** Sarah Bem introdujo el concepto de androginia, que desconecta el sexo y el género e incluye una representación continua de género. Lo que no se ha investigado hasta ahora es si las cualidades particulares postuladas por Bem son cualidades de género en lugar de rasgos asociados al sexo. En el presente estudio, la tarea de inversión de la correlación como enfoque basado en datos se utilizó para determinar los estereotipos de género implícitos en los rostros de hombres y mujeres con el fin de crear un prototipo ideal de rostros femeninos y masculinos. Luego se midió qué impresiones evocan estos rostros. Se realizaron dos estudios y un estudio piloto (N = 514). El presente estudio muestra que el género y no el sexo es crucial para la atribución de las características sociales. Se ha descubierto que las imágenes de rostros estereotipados son muy adecuadas para medir la masculinidad y la feminidad. Las propiedades continuas de la masculinidad y la feminidad, tal como las esbozó Bem (1974), siguen siendo apropiadas para diferenciar entre las ideas estereotipadas de hombres y mujeres.

**Palabras clave:** Trabajo de correlación inversa, percepción facial, género, androginia

*Correspondencia:* Roland Imhoff. 2Johannes Gutenberg-Universität Mainz, Mainz (JGU).  
[roland.imhoff@uni-mainz.de](mailto:roland.imhoff@uni-mainz.de)

## Introduction

Gender is an almost ubiquitous category. On a daily basis, categorizing humans around us in “men” and “women” happens more or less spontaneously (Taylor, Fiske, Etcoff, & Ruderman, 1978) from relatively early age onwards (Bennett, Sani, Hopkins, Agostini, & Malucchi, 2000; Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002). One could say that of the many possibilities to differentiate among humans, gender is one of the most pervasive and frequent: people judge others’ intellectual abilities differently based on their presumed gender (Bian, Leslie, Cimpian, 2017), people use different standards to evaluate performance of men and women (Biernat & Manis, 1994), and are willing to pay one gender less for the same work (Auspurg, Hinz, & Sauer, 2017). Likewise, people “do” gender on a continuous basis: “Men” wear suits and ties, “women” wear skirts and make-up. People use gendered bathrooms and people often behave in gender consistent ways (masculine for men, feminine for women). In fact, people infer presumed gender-typical facial features from gender-typical behavior: Compared to teachers of maths, teachers of arts are imagined more likely as women than men and more likely as feminine than as masculine women (Degner, Mangels, & Zander, 2018). In the present research, we sought explore a similar relation in the opposite direction: Do perceivers infer masculine and feminine behavior traits from masculine and feminine facial features?

It is only a relatively recent insight that there is no deterministic relationship between a person’s sex and their gender (Fausto-Sterling, 2000). Most explicitly voiced by post-structural feminists, the notion that gendered attributes and behavior are non-deterministically related to sex already resonated in Sandra Bem’s (1974) influential work on androgyny. In her work, Bem started from the assumption that both men and women have feminine and masculine traits and that integrating both masculine and feminine traits (i.e., being androgynous) may be particularly beneficial. Until Bem’s (1974) groundbreaking work, sex and gender were treated synonymously in the sense that biological men were expected to be masculine and women to be feminine. Accordingly, a person had to be either masculine or feminine, but not both. In societies with a narrow gender self-concept people might inhibit behaviors that are not compatible with the stereotypical sexual self-concept. Therefore, Bem introduced the concept of androgyny, which disconnects sex and gender and includes a continuous representation of gender, which might allow for “an individual to freely engage in both masculine and feminine” behaviors (1974, p. 155). This allows a treatment of the two as independent: men as well as women can differ in their degrees of masculinity and femininity. In the present research we built on this differentiation between sex (men vs. women) and gender (masculine vs. feminine) to test which of the two factors has more weight in impressions.

We focus on faces here because they have tremendous weight in daily interactions. Typically, before we speak with someone we see their face. We even see the faces of people who we never speak to and we still make quick inferential judgments about the person based on their faces: We decode emotions, but also make judgments of presumed trustworthiness within split seconds (Wills & Todorov, 2006). Thus, the face has tremendous weight in daily human interaction, a fact still not fully recognized by social psychological research that too often relies on verbal material – a modality that is much less ubiquitous in everyday interactions.

Nevertheless, there is already some research on gender and faces. Much like the “masculine” and “feminine” traits in Bem’s Sex Role Inventory are those traits that are seen as stereotypically desirable for men and women respectively, many studies have explored what constituted “ideal” (i.e., stereotypically desirable) faces for men and women respectively. Regarding the desirability of men's faces, some authors (DeBruine, Jones, Smith, & Little, 2010; Perrett et al., 1998; Rhodes, Hickford, & Jeffery, 2000; Penton & Voak, 1999) found that women preferred male subjects with feminine facial features; in other studies a preference for above-average masculine faces was identified (DeBruine et al., 2006; Johnston, Hagel, Franklin, Fink, & Grammer, 2001). Some authors (Reed & Blunk, 1990) have found that some women perceived that men’s facial hair adds to attractiveness while others (DeBruine et al., 2010; Perrett et al., 1998) feel that it decreases attractiveness. Overall, the findings concerning men’s faces seem to be ambiguous. Puts, Jones and DeBruine (2012) suggest that women prefer either masculine faces or slightly feminine male faces on average.

Regarding the desirability of women's faces, a number of studies have found slightly less ambivalent results. Men prefer feminine faces. The femininity and attractiveness of women's faces seems to be important especially when men are seeking long-term partnerships (Confer, Perilloux, & Buss, 2010; Rhodes, 2006; Rhodes, Hickford, & Jeffery, 2000). However, when women have been asked to rate the faces of other women, they have judged women with attractive faces as being more attractive to men as well as more promiscuous and flirtatious (Puts et al., 2011; Brewer & Archer, 2007). When women were asked to assess themselves, women with feminine faces had lower values in terms of social dominance and influence (Quist, Watkins, Smith, DeBruine, & Jones, 2011).

Another approach to determine masculinity and femininity in the face might be to approach the question not from desirability but from stereo-typicality: what are typical features of men and women? When people see each other for the first time, they instantly and automatically draw conclusions about others’ personalities (Willis & Todorov, 2006) and group memberships (Martin & Mcrae, 2007). For example, stereotypes about gender roles are seen in the categorization of feminine-looking women as “warm” (Oosterhof & Todorov, 2008), as the classification of a person's sex as female leads to consensual assumptions about gender-specific characteristics. Similarly, masculine-looking people are automatically judged as “colder” and “more competent” (Walker & Wänke, 2017). These automatic conclusions can be viewed as a form of overgeneralization (Zebrewitz, 2010).

In the present research we took an alternative approach. Based on the venture point of the independence of sex (men vs. women) and gender (masculine vs. feminine), we created idealized images of what masculine, respectively feminine men and women were expected to look like (pilot study). To do so, we employed a data-driven Reverse Correlation Image Classification technique (Mangini & Biedermann, 2004). These images were then shown to other, hypothesis-blind raters and judged on several rating dimensions (Study 1), among them the Bem Sex Role Inventory items (Study 2). Doing so allowed us to isolate the effects of sex and gender on these impressions.

Various approaches have been developed in recent years to systematically represent mental concepts in faces (Cheng, O’Toole, & Abdi, 2001; Blanz & Vetter, 1999; Mangini & Biederman, 2004). Of these we employed in the current study the reversed correlation task

(later cited as RCT) approach in which observers must assess faces that are in high levels of visual noise, a random dot pixel pattern (Mangini & Biederman, 2004). Avoiding presumptions was central to planning the investigation; to achieve this goal, the method is fully data-driven and bottom-up in the sense that no "objective" reality of femininity and masculinity was presented. Instead, the sum of implicit gender stereotypes was examined.

In the present paper, we create an ideal prototype of feminine and masculine faces for a male and a female model. We will then see how this maps on impressions these faces evoke. Two studies and a pilot study were conducted; in the pilot study the stimuli were developed. Study 1 focusses on global attributes, whereas study 2 specifically targets attributes of masculinity and femininity as proposed by Bem (1974) and whether the characteristics are associated with the gender or sex of the faces. All materials can be found on OSF at [https://osf.io/6ea3z/?view\\_only=c9de483bb8994eea80497e359422c328](https://osf.io/6ea3z/?view_only=c9de483bb8994eea80497e359422c328)

### ***Pilot study: Stimulus development***

The pilot study was a pretest focused on creating the stimuli for the two subsequent studies. Images were created of faces that were intended to represent mental representations of masculinity and femininity; these representations should be tentatively distinguished from each other. The RCT allows an estimate, however subtle, of what is in the subject's (rather than the experimenter's) head (Mangini, & Biedermann, 2004).

## **Method**

### ***Sample***

A group of six undergraduate students at the University of Cologne participated in each of the two tasks in exchange for extra credit in psychology courses. All subjects were unaware of the purpose of the experiment.

## **Measures**

### ***Basic faces***

Two basic faces were created. In a first step, sixteen male faces and sixteen female Caucasian faces were selected from the Radboud Face Database (Langner et al., 2010). In these pictures, all persons had a neutral facial expression, their hair was combed back and they were photographed frontally. Their portraits were then merged gradually using the morphing program Fantamorph (AbrosoftFantamorph version 5). This process resulted in two faces (Table 4). In the next step, the images were converted into grayscale images and superimposed with a filter that left the faces recognizable but removed the contours to create base face for the Reverse Correlation Image Classification Task (note that the male base face was previously used in Imhoff, Woelki, Hanke, & Dotsch, 2013).

### ***Reversed Correlation Image Classification Task (RCT)***

A RCT is a task in which observers classify faces while experiencing high levels of visual noise (Mangini, & Biederman, 2004; for its use in social psychological research see Dotsch, Wigboldus, Langner, & van Knippenberg, 2008; Imhoff, Dotsch, Bianchi, Banse, &

Wigboldus, 2011; Imhoff & Dotsch, 2013). The visual noise is a random dot pixel pattern, that looks like analog video and television when no transmission signal is obtained by the antenna receiver. The task was used in the current research to assess the implicit concept of a feminine respectively masculine person. In the RCT, subjects saw two faces over which a filter with a black-and-white noise was placed. Subjects had to decide to which of the two faces a certain concept (e.g. masculinity) applied. Participants were unaware that the same face was presented throughout a task and that only the noise rendered the face to fit more in one category or the other. The differences between the average noise patterns for each classification decision provided an estimate of the information mediating these classifications. When the noise was combined with the underlying face, the resultant images were prototypes of their respective classes.

### ***Procedures***

Subjects participated in four different discrimination tasks. They are always shown basic faces plus the noise. In the four tasks, they identified a feminine woman's face (FF), a masculine woman's face (MF), a feminine man's face (FM), and a masculine man's face (MM). At the beginning of the experiment, the participants were told that they would see 2 pictures each of women (FF, MF) and men (FM, MM). Then they were told to pick the image that they felt was feminine (FF, FM) or masculine (MF, MM). The subjects were then instructed to press the left red button for the left and the right red button for the right image. No feedback was provided, as no responses were correct or incorrect. After participants were given the brief instruction, noisy images were presented two at a time. After the subjects decided on one of the two pictures, the next picture couple was shown directly. After the fourth block of one category, the next category started directly afterwards. There were 4 blocks per category, each with 100 face pairs, i.e. each subject categorized 400 noisy faces for FF, MF, FM and MM. Overall, each subject categorized 1600 faces. The entire experiment lasted approximately 15 minutes and took place in a lab.

## **Results**

### ***Images***

The resulting classification images calculated for all six subjects' data are presented in table 4. Adding or subtracting the classification image to or from the base image resulted in faces that appeared to be effective prototypes of their classes.

The masculine man's face differed from the feminine man's face, and the masculine woman's face differed from the feminine woman's face from a subjective perspective. This discovery illustrated that four very different faces, a feminine man face, a masculine man face, a feminine women face, and a masculine women face, could be formed.

**Table 4.** *Base Images and Classification Image of the RCT*

	Gender	
base face	Masculine	feminine



### Study 1

Study 1 focused on three objectives. First, the faces created in the stimulus development phase were tested to determine whether they differed in terms of femininity and masculinity by a larger sample. Secondly, it is determined which other characteristics are associated with prototypically feminine or masculine faces. Thirdly, it tests whether the characteristics are associated with the gender or the sex of the faces.

### Method

#### *Sample*

A total of 278 participants (186 women, 85 men, mean age = 24.3 years,  $SD = 4.5$ ) were recruited from different German universities and branches of study through a mailing system. Participation was voluntary and took place at home. All but three students had university entrance qualifications or a university degree.

### Measures

#### *Masculine-Feminine Faces*

The subjects evaluated the four faces formed in the RCT (see Table 4) on the basis of nine categories on a 7-point Likert scale. The questions were how feminine or masculine a person was, how attractive this person was, how good a partner this person would be, how socially dominant this person would be, whether this person would be a good spouse, how much the subject would like to get to know this person, how much the person would help in the household, how fond of children this person was, and how old this person was.

#### *Procedures*

At the beginning of the examination, the subjects were told that they were going to see different faces and asked to answer questions about the faces spontaneously and quickly.

The pictures were then presented to participants in a randomized order together with one question each.

### Results

The significant differences between all four faces in terms of variable 'masculinity' confirmed that, as expected, the men's faces were considered to be more masculine than the women's faces; more importantly, the feminized faces were also considered to be more feminine than the masculinized faces. This result was a prerequisite for further analysis. Nine one-way within-subjects analyses of variance were conducted to compare the scores of the four faces. FF was rated most attractive, while (somewhat surprisingly) FF was also perceived most dominant. FM was rated most as fond of children, but MM was rated oldest. Statistical characteristics are presented in table 5.

Nine two-way within-groups analyses of variance were conducted to explore the impact of sex and gender on the different qualities. Over nine comparisons, the main effects of gender were significant, with only one comparison failing to achieve a conventional level of significance (Table 5). Femininity was associated with higher attractiveness, younger age, greater desirability as a partner, social dominance, kindness and greater housekeeping qualities.

The main effects of sex were less prevalent (Table 5) and restricted to the fact that women were seen as more attractive, more desirable partners and older. For most of the variables, however, there were significant interactions indicating that the positive aspect of femininity (attractiveness, desirability as partner) were particularly pronounced for women. Only one of the interactions was a clear-cut ordinal interaction whereby masculine men and feminine women were seen as more dominant than feminine men and masculine women, potentially pointing to a stigma of (facial) gender nonconformity.

**Table 5.** Descriptives and Results of a Two-way within-Subject ANOVA and a one-way within-subjects ANOVA

	Feminine Male (FM)		Masculine Male (MM)		Feminine Female (FF)		Masculine Female (MF)		Main effect Gender		Main effect Sex		Interaction Gender x Sex		Error df
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	F	P	F	P	F	p	
Masculinity/Femininity	4.05 <sup>c</sup>	1.18	5.95 <sup>a</sup>	1.03	1.95 <sup>d</sup>	1.07	5.37 <sup>b</sup>	1.14	1375.23	<.001	358.91	<.001	170.92	<.001	221
Attractiveness	3.30 <sup>b</sup>	1.22	2.10 <sup>c</sup>	1.06	4.98 <sup>a</sup>	1.12	1.76 <sup>d</sup>	0.86	1265.75	<.001	141.27	<.001	378.23	<.001	252
Being Partnered	3.56 <sup>b</sup>	1.60	2.47 <sup>c</sup>	1.48	4.97 <sup>a</sup>	1.51	2.01 <sup>d</sup>	1.15	447.63	<.001	26.90	<.001	118.65	<.001	203
Dominance	3.14 <sup>b</sup>	1.27	4.85 <sup>a</sup>	1.61	4.81 <sup>a</sup>	1.27	3.30 <sup>b</sup>	1.55	0.63	.430	0.57	.450	320.50	<.001	188
Spouse	4.75 <sup>a</sup>	1.48	2.71 <sup>b</sup>	1.26	4.67 <sup>a</sup>	1.32	3.02 <sup>c</sup>	1.30	309.00	<.001	2.67	.104	4.57	.034	178
To Get to Know Someone	4.25 <sup>b</sup>	1.60	2.31 <sup>d</sup>	1.39	4.75 <sup>a</sup>	1.54	2.09 <sup>c</sup>	1.21	474.32	<.001	2.46	.119	19.67	<.001	174
Housekeeper	4.99 <sup>a</sup>	1.23	2.78 <sup>d</sup>	1.35	4.57 <sup>b</sup>	1.40	3.17 <sup>c</sup>	1.40	241.75	<.001	0.02	.878	21.64	<.001	173
Fond of Children	5.32 <sup>a</sup>	1.35	2.80 <sup>d</sup>	1.35	4.68 <sup>b</sup>	1.37	3.09 <sup>c</sup>	1.34	286.70	<.001	4.66	.032	35.25	<.001	172
Age	23.03 <sup>d</sup>	3.97	29.17 <sup>a</sup>	5.88	26.99 <sup>c</sup>	3.86	27.72 <sup>b</sup>	6.29	98.68	<.001	23.55	<.001	111.27	<.001	171

The letters (A, B, C, D) indicate whether significant mean differences were found between the variables. If the same letters are used for variables then there is no significant difference between the variables. For better comprehension, the values of masculinity and femininity have been reversed.



Bonferroni-corrected ( $p = .00139$ ) control analyses were conducted to check for sex differences in the 36 judgments. Male participants thought that the feminine woman was a better partner ( $M = 5.55$ ,  $SD = 1.26$ ) than female participants did ( $M = 4.66$ ,  $SD = 1.57$ ,  $t(205) = 4.03$ ,  $p < .001$ ). Men also wanted to get to know her ( $M = 5.35$ ,  $SD = 1.51$ ) more than women did ( $M = 4.46$ ,  $SD = 1.50$ ,  $t(173) = 3.63$ ,  $p < .001$ ). All other simple tests failed to reach the adjusted alpha level, and these gender differences were not given further attention.

## Discussion

The current study showed that the faces formed in the during stimulus generation differed significantly in their femininity and masculinity. Therefore, it can be assumed that the faces represent mental representations of masculinity and femininity, thus achieving the first goal. The second aim of the study was to determine other characteristics associated with typically feminine or masculine faces: several associations became clear. It has repeatedly been shown that feminized faces by trend are associated with positive characteristics and that this effect cannot be attributed to the sex of the rater. Sex and gender interacted insofar that feminine facial features strengthen the differences in evaluation of sex differences. Third, the comparison between gender and sex of faces clearly showed that gender is more important than sex. The socially constructed idea of masculinity and femininity is clearly associated with several qualities. With the biological sex, however, there were only a few connections. The results indicate that the beliefs of masculinity and femininity are variable and thus culturally shaped.

### *Study 2*

The objective of Study 2 was to test whether the Bem properties (1974) reflect gender differences or sex differences. Additionally, Study 2 sought to check with which properties implicit mental representations of femininity and masculinity are associated.

## Method

### *Sample*

A total of 230 participants (152 men, 78 women, mean age = 22.4 years,  $SD = 4.5$ ) were recruited from different German universities and branches of study through a mailing system. Participation was voluntary and took place at respondents' homes.

## Measures

### *Bem Social Role Inventory (BSRI)*

The Bem Social Role Inventory is a measure of masculinity and femininity as well as gender roles; it assesses how people identify themselves psychologically (Bem, 1974). The tool includes 60 items in three construct categories: masculine items, feminine items, and neutral items (Table 6). The stereotypical properties are all positively connoted. In the present study, 7 items per category were used. The properties of the experimental images were evaluated on a 5-point scale ranging from suitable (1) to unsuitable (5).

### *Masculine and Feminine Faces*

The same four faces were used as in study 1. These faces were created using the RCT, and they represent the mental conceptions of masculinity and femininity (Mangini & Biederman, 2004). For a detailed description of the task and the faces created, see the stimulus development section.

#### *Procedures*

At the beginning of the study, subjects were shown four different faces. They were asked to examine the faces closely and evaluate the psychological properties (BSRI) of the faces.

#### **Results**

Twenty-four two-way within-groups analyses of variance were conducted to explore the impact of sex and gender on levels of the total values of masculinity, femininity and neutral properties, as well as each of the seven individual properties (Table 6). Gender had an overwhelmingly strong main effect. In each of the 21 individual characteristics, feminized faces differed from masculinized faces. In addition, a major gender effect was found in the overall values of masculinity, femininity, and neutral characteristics. Feminized faces had more "feminine" characteristics than masculinized faces as well as significantly higher levels of "neutral" properties. Feminized faces had also significantly more "masculine" characteristics than masculinized faces. However, this difference was less pronounced. In contrast, only a few properties had a main sex effect. No significant major effect for sex was found on the overall scores for femininity, masculinity, and neutrality.

In masculine and feminine traits, a significant interaction was found between gender and sex. In feminine traits feminization of men's faces leads to a large increase in female characteristics. The increase in female characteristics in women's faces, however, is only moderate. With masculine characteristics you can observe a contrary trend. The feminization of women's faces leads to a strong increase of 'masculine' characteristics, whereas in males' faces feminization leads to a moderate decrease of 'masculine' characteristics.

Twenty-four one-way within-subjects analyses of variance were conducted to compare the scores of the four faces. FF were rated most masculine, while (somewhat surprisingly) MF was perceived least masculine. FM was rated most feminine, while (unsurprisingly) MM was rated least feminine.

**Table 6.** Descriptive Statistics of Classified Images Regarding Gender Role Properties and Results of a Two-way within Subject ANOVA and a one-way within-subjects ANOVA

	Feminine Male (FM)		Masculine Male (MM)		Feminine Female (FF)		Masculine Female (MF)		Main Effect Gender		Main Effect Sex		Interaction Effect		Error df
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	
<b>Masculinity</b>															
Has Leadership Abilities	2.51 <sup>b</sup>	0.75	2.64 <sup>b</sup>	1.10	3.68 <sup>a</sup>	0.84	1.94 <sup>c</sup>	0.90	104.15	<.001	12.26	.001	211.41	<.001	159
Ambitious	3.14 <sup>b</sup>	0.92	3.31 <sup>b</sup>	0.99	4.09 <sup>a</sup>	0.77	2.47 <sup>c</sup>	0.92	94.57	<.001	0.59	.442	168.37	<.001	159
Factual	3.43 <sup>b</sup>	0.90	2.80 <sup>c</sup>	1.09	3.76 <sup>a</sup>	0.76	2.91 <sup>c</sup>	0.96	79.05	<.001	10.58	.001	2.37	.126	155
Forceful	2.57 <sup>c</sup>	1.01	4.17 <sup>a</sup>	0.81	3.20 <sup>b</sup>	0.92	3.38 <sup>b</sup>	1.08	117.92	<.001	1.36	.245	80.80	<.001	156
Consistent	3.23 <sup>b</sup>	0.78	3.88 <sup>a</sup>	0.97	4.01 <sup>a</sup>	0.75	3.00 <sup>b</sup>	1.00	5.01	.027	0.66	.420	167.31	<.001	155
Analytical	3.08 <sup>b</sup>	0.84	2.98 <sup>b</sup>	1.00	3.69 <sup>a</sup>	0.85	2.40 <sup>c</sup>	0.89	78.42	<.001	0.14	.714	85.30	<.001	159
Willing to Take Risks	2.82 <sup>c</sup>	1.13	4.23 <sup>a</sup>	0.74	3.29 <sup>b</sup>	0.92	2.96 <sup>c</sup>	1.06	49.83	<.001	34.42	<.001	103.85	<.001	158
<b>Masculinity Total</b>	2.98 <sup>c</sup>	0.52	3.41 <sup>b</sup>	0.59	3.67 <sup>a</sup>	0.52	2.73 <sup>d</sup>	0.59	23.01	<.001	0.02	.896	248.04	<.001	152
<b>Femininity</b>															
Dependent	3.08 <sup>a</sup>	0.91	2.56 <sup>b</sup>	1.04	2.46 <sup>b</sup>	0.83	3.32 <sup>a</sup>	0.92	4.65	.033	1.08	.301	93.16	<.001	155
Sensitive	3.49 <sup>a</sup>	1.01	1.66 <sup>d</sup>	0.71	3.16 <sup>b</sup>	0.83	2.06 <sup>c</sup>	0.80	398.13	<.001	0.30	.582	38.44	<.001	159
Affectionate	3.34 <sup>a</sup>	0.96	1.78 <sup>d</sup>	0.76	2.98 <sup>b</sup>	0.89	2.28 <sup>c</sup>	0.91	1.34	<.001	1.34	.250	53.98	<.001	155
Passionate	2.66 <sup>b</sup>	0.88	2.55 <sup>b</sup>	1.08	3.21 <sup>a</sup>	0.84	2.09 <sup>c</sup>	0.85	70.48	<.001	0.46	.496	63.93	<.001	155
Modest	3.59 <sup>a</sup>	0.93	2.31 <sup>d</sup>	0.80	2.58 <sup>c</sup>	0.76	3.10 <sup>b</sup>	0.90	27.11	<.001	3.43	.066	169.58	<.001	152
Sacrifices Oneself	3.10 <sup>a</sup>	0.96	2.42 <sup>c</sup>	0.99	2.67 <sup>b</sup>	0.83	2.73 <sup>b</sup>	1.00	17.25	<.001	0.76	.385	31.03	<.001	157
Yielding	3.37 <sup>a</sup>	0.90	1.82 <sup>d</sup>	0.79	2.43 <sup>c</sup>	0.82	2.79 <sup>b</sup>	0.99	65.56	<.001	0.04	.842	183.39	<.001	158
<b>Femininity Total</b>	3.24 <sup>a</sup>	0.54	2.17 <sup>d</sup>	0.51	2.78 <sup>b</sup>	0.51	2.61 <sup>c</sup>	0.55	158.72	<.001	0.02	.883	138.81	<.001	144
<b>Neutral</b>															
Trustworthy	3.54 <sup>a</sup>	0.75	1.87 <sup>c</sup>	0.79	3.44 <sup>a</sup>	0.86	2.27 <sup>b</sup>	0.82	390.54	<.001	7.65	.006	19.78	<.001	157
Sociable	3.47 <sup>a</sup>	0.93	2.34 <sup>b</sup>	0.86	3.34 <sup>a</sup>	0.78	2.35 <sup>b</sup>	0.91	240.21	<.001	0.87	.351	0.94	.334	156
Diligent	3.56 <sup>b</sup>	0.81	2.89 <sup>c</sup>	0.91	3.92 <sup>a</sup>	0.75	2.87 <sup>c</sup>	1.00	116.88	<.001	7.82	.006	11.17	.001	155
Conscientious	3.75 <sup>a</sup>	0.81	2.56 <sup>b</sup>	0.88	3.80 <sup>a</sup>	0.68	2.76 <sup>b</sup>	0.88	242.78	<.001	5.23	.024	1.59	.210	154
Reliable	3.67 <sup>a</sup>	0.79	2.81 <sup>b</sup>	0.91	3.73 <sup>a</sup>	0.67	2.90 <sup>b</sup>	0.90	120.43	<.001	2.07	.152	0.09	.765	153
Healthy	3.91 <sup>a</sup>	0.85	3.21 <sup>b</sup>	0.87	3.89 <sup>a</sup>	0.72	2.91 <sup>c</sup>	0.87	184.55	<.001	8.68	.004	7.37	.007	158
Mindful	3.75 <sup>a</sup>	0.79	3.07 <sup>b</sup>	0.79	3.82 <sup>a</sup>	0.63	2.76 <sup>c</sup>	0.85	149.55	<.001	4.53	<.001	10.42	<.001	157
<b>Neutral Total</b>	3.68 <sup>a</sup>	0.46	2.69 <sup>b</sup>	0.53	3.71 <sup>a</sup>	0.43	2.68 <sup>b</sup>	0.57	426.44	<.001	0.19	.661	0.42	.517	146

The letters (A, B, C, D) indicate whether significant mean differences were found between the variables. If the same letters are used for variables then there is no significant difference between the variables. For better comprehension, the values of masculinity and femininity have been reversed

Participants completed a total of 21 ratings for each of the four faces adding up to 84 judgments. A properly Bonferroni-adjusted alpha level used to test for gender differences would thus be at  $p = .0006$ . Adopting even a less conservative p-value of .01 yielded no gender differences on the individual items level or the aggregate level of scales. Based on these results, in the analyses, no further attention was paid to participants' gender.

### Discussion

Study 2 again showed consistent effects of target face gender. Feminized faces had a higher expression of masculine, feminine, and neutral characteristics. A reason for this may be that all words of the BSRI reflect positive characteristics. Masculine and feminine properties are formulated in a positive way, and neutral properties have characteristics that characterize them as socially desirable properties. Since we know from study 1 that feminization leads to an increase in attractiveness, and physical attractiveness is associated with other independent properties (Dion, Berscheid, & Walster, 1972), it can be assumed that feminization also leads to an increase in other positive properties. In contrast to the strong main gender effect, only marginal differences were found when comparing the characteristics of different sexes. This finding suggests that biological sex has little impact on social judgments. These findings replicate and validate the conclusions made in Study 1.

### General Discussion

The present paper presented two studies that examined whether the continuous properties of masculinity and femininity, as outlined by Bem (1974), are appropriate to differentiate between the visual stereotypes of masculine and feminine men and women. Results from two studies converged in remarkably strong inferences raters seemed to draw from these faces. Overall, the dimension masculinity–femininity seemed to have more weight in these impressions than whether the base face was composed of male or female faces.

The study of faces in sex-gender differentiation was done in this study because faces are very important in day-to-day interaction. It is all the more surprising that such visual material is relatively seldom used in social psychological research. In the present study, pictures of faces in general and pictures of stereotypical faces in particular have been found to be highly suitable for research on masculinity and femininity. With the Reversed Correlation Task as a relatively simple technique that is intuitively comprehensible for the subjects, it was possible to strongly induce a concept. This induction strength was shown by the fact that the gendered faces differed in all 21 properties according to Bem (1974). Another advantage of the RCT in the present study is that it measures implicit attitudes. Perceptions of "masculinity" and "femininity" are strongly influenced by social desirability and political correctness, so that a direct questioning can only get a glossed image of the convictions of the people. This legitimate and principled objection could be met here elegantly.

As a potential limitation of our current research, all of our studies were conducted in WEIRD (Henrich, Heine, & Norenzayan, 2010) countries and our results can thus not be easily generalized to other cultural contexts. Future research might address this limitation in a more inclusive cross-cultural approach to examine to what extent cultural patterns of attribution of properties to the gender can be shown.

Out of the plethora of possibilities to categorize people, gender respectively sex is one of the most used categories and for most people a familiar way of organizing their social environment. Therefore, it is all the more significant that the present study has shown that this

trait is probably based on social agreement and that other attributes that are supposed to be "male" and "female" are the result of a tacit social agreement. The Reversed Correlation Task proved to be an adequate method for determining the implicit attitudes and stereotypes of "masculinity" and "femininity." The development of this new technique allowed biological sex to be separated from the social gender and thus to make unique comparisons.

### References

- Auspurg, K., Hinz, T., & Sauer, C. (2017). Why should women get less? Evidence on the gender pay gap from multifactorial survey experiments. *American Sociological Review*, *82*, 179-210.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*, 155-162.
- Bennett, M., Sani, F., Hopkins, N., Agostini, L., & Malucchi, L. (2000). Children's gender categorization: An investigation of automatic processing. *British Journal of Developmental Psychology*, *18*, 97-102.
- Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*, 389-391.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*, 5-20.
- Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Proceedings of the 26th Annual Conference of Computer Graph Interact Techiques*. *19*, 187-194.
- Cheng, Y. D., O'Toole, A. J., & Abdi, H. (2001). Classifying adults' and children's faces by sex: Computational investigations of subcategorical feature encoding. *Cognitive Science*. *25*, 819-838.
- Confer, J. C., Perilloux, C., & Buss, D. M. (2010). More than just a pretty face: men's priority shifts toward bodily attractiveness in short-term versus long-term mating contexts. *Evolution and Human Behavior*, *31*, 348-353.
- DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., et al. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, *273*, 1355-1360.
- DeBruine, L. M., Jones, B. C., Smith, F. G., & Little, A. C. (2010). Are attractive men's faces masculine or feminine? The importance of controlling confounds in face stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 751-758.
- Degner, J., Mangels, J., & Zander, L. (2018). *Think Physics, think Male? Visualizing Gender-stereotypic Representations of Male and Female Teachers*. Unpublished Manuscript.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*, 285-290.
- Fausto-Sterling, A. (2000). *Sexing the Body: Gender Politics and the Construction of Sexuality*. New York: Basic Books.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 1-75.
- Imhoff, R., & Dotsch, R. (2013). Do we look like me or like us? Visual projection as self- or ingroup-projection. *Social Cognition*, *31*, 806-816.
- Imhoff, R., Dotsch, R., Bianchi, M., Banse, R., & Wigboldus, D. H. J. (2011). Facing Europe: Visualizing spontaneous in-group projection. *Psychological Science*, *22*, 1583-1590. doi: 10.1177/0956797611419675

- Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in Psychology, 4*, 386. doi: 10.3389/fpsyg.2013.00386
- Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness: Evidence of hormone-mediated adaptive design. *Evolution and Human Behavior, 22*, 251–267.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*, 1377–1388.
- Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: estimating the information employed for face classifications. *Cognitive Science, 28*, 209–226.
- Martin, D., & Macrae, C. N. (2007). A face with a cue: Exploring the inevitability of person categorization. *European Journal of Social Psychology, 37*, 806–816.
- Oosterhof, N. N., Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Science, 105*, 11087–11092.
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887.
- Puts, D. A., Jones, B. C., & DeBruine, L. M. (2012). Sexual selection on human faces and voices. *Journal of Sex Research, 49*, 227–243.
- Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31*, 1109–1121.
- Quist, M. C., Watkins, C. D., Smith, F. G., DeBruine, L. M., & Jones, B. C. (2011). Facial masculinity is a cue to women's dominance. *Personality and Individual Differences, 50*, 1089–1093.
- Reed, J. A., & Blunk, E. M. (1990). The influence of facial hair on impression formation. *Social Behavior and Personality: An International Journal, 18*, 169–175.
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.
- Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex-typicality and attractiveness: Are supermale and superfemale faces superattractive? *British Journal of Psychology, 91*, 125–140.
- Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778–793.
- Walker, M., & Wänke, M. (2017). Caring or daring? Exploring the impact of facial masculinity/femininity and gender category information on first impression. *PLoS ONE, 12*(10), 1–19.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592–598