Evolution and Human Behavior 48 (2027) 107011



Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens



Social judgments of faces across world regions are predicted better by distinctiveness than masculinity

Victor Shiramizu [a,*], Junzhi Dong [a], Kathlyne Leger [a], Anthony J. Lee [b], Alex L. Jones [c], Yasaman Rafiee [a], Zuzana Elliott [a], Lisa M. DeBruine [d], Benedict C. Jones [a]

[a] Department of Psychological Sciences & Health, University of Strathclyde, UK
[b] Psychology, Faculty of Natural Sciences, University of Stirling, UK
[c] School of Psychology, Swansea University, UK
[d] School of Psychology & Neuroscience, University of Glasgow, UK

A B S T R A C T

Social judgments of faces influence important social outcomes. Although many researchers have argued that facial masculinity plays a key role in perceptions of both valenced social judgments and dominance, whether these effects are consistent among people from different world regions is highly contentious. Consequently, we investigated possible relationships between masculinity and ratings of 120 faces made by 11,484 participants from eleven world regions (Africa, Asia, Australia and New Zealand, Central America and Mexico, Eastern Europe, Middle East, Scandinavia, South America, United Kingdom, United States and Canada, Western Europe). Surprisingly, masculinity did not significantly predict either perceived valence or dominance. By contrast, facial distinctiveness (i.e., atypicality) was significantly and negatively correlated with valenced perceptions across world regions and predicted valenced perceptions better than did masculinity. Collectively, our results suggest that distinctiveness (or correlates thereof), but not masculinity, shows a consistent association with social judgments of faces across different world regions, potentially due to an "anomalous-is-bad" stereotype whereby people ascribe negative traits to individuals with atypical face shapes.

1. Introduction

Social judgments of faces (i.e., judgments of traits, such as trustworthiness, attractiveness, or dominance based on facial appearance) influence important social outcomes, including hiring decisions, voting decisions, and partnership choices (Olivola, Funk, & Todorov, 2014; Rhodes, 2006). Although many researchers have proposed that effects of sexually dimorphic shape characteristics play a critical role in social judgments of faces (e.g., Little, Jones, & DeBruine, 2011; Perrett et al., 1998), both the veracity of this claim and the extent to which these effects exhibit regional variability are currently unclear. For example, Perrett et al. (1998) reported that both Japanese and UK participants ascribed positive traits (e.g., trustworthiness, attractiveness) to individuals with more feminine face shapes and perceived individuals with more masculine face shapes to be more dominant. Such results have been interpreted as evidence that social judgments of faces are, at least in part, a consequence of psychological adaptations that facilitate effective social interactions and transcend cultural influences (Perrett et al., 1998). By contrast, Scott et al. (2014) reported that the direction of the effects of masculine face shapes on perceptions of attractiveness, dominance, and aggressiveness varied markedly across world regions. Although findings such as these have had a profound influence on theories of appearance-based stereotyping, there are three potentially serious methodological problems with previous studies on this topic.

First, like Perrett et al. (1998), most studies testing for consistency in the effects of masculine shape characteristics on social judgments of faces compared judgments made by participants from only two world regions (e.g., Penton-Voak et al., 1999; Stephen et al., 2012). However, this sampling strategy (comparing participants from only two world regions) limits the conclusions that can be made. Other work has tested participants from countries that are underrepresented in the literature (e.g., Arab countries, Alharbi, Holzleitner, Lee, Saribay, & Jones, 2020, African countries, Phalane, Tribe, Steel, Cholo, & Coetzee, 2017, or Asian countries, Komori, Kawamura, & Ishihara, 2011). However, because these studies have all used different sets of face stimuli, comparing results across these studies is problematic. Scott et al. (2014) analysed social judgments of faces made by participants from twelve countries, but data from more than a hundred participants were analysed for only one of these countries (the UK) and data from fewer than fifty participants were analysed for eight of the twelve countries.

Second, many studies testing for consistency in the effects of masculine shape characteristics on social judgments of faces have used

* Corresponding author.
E-mail address: Victor.Shiramizu@strath.ac.uk (V. Shiramizu).

https://doi.org/10.1016/j.evolhumbehav.2026.107011
Received 17 January 2026; Received in revised form 16 September 2026; Accepted 18 November 2026
Available online 22 November 2026
1090-5138/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

paradigms in which stimuli with shape characteristics that had been experimentally manipulated using computer graphics methods were presented in forced-choice tasks (e.g., Alharbi et al., 2020; DeBruine, Jones, Crawford, Welling, & Little, 2010; Marcinkowska et al., 2019; Penton-Voak et al., 1999; Perrett et al., 1998; Scott et al., 2014). Many researchers have recently criticised this method, noting that it appears to have poor ecological validity (Dong et al., 2023; Jones & Jaeger, 2019; Jones, Jones, & Shiramizu, 2023; Lee, De La Mare, Moore, & Umeh, 2021; Lewis, 2017, 2020; Satchell, Jaeger, Jones, López, & Schild, 2023; Scott, Clark, Boothroyd, & Penton-Voak, 2013). Moreover, recent studies demonstrate that findings for the effects of shape characteristics on social judgments of faces obtained using this method are qualitatively different from those obtained when participants rated natural (i.e., unmanipulated) face images (Dong et al., 2023; Jones et al., 2023; Jones & Jaeger, 2019; Lee et al., 2021; Leger, Dong, DeBruine, Jones, & Shiramizu, 2023; Scott, Pound, Stephen, Clark, & Penton-Voak, 2010). Thus, the methods used in most previous studies have investigated whether participants in different world regions *can* use sexually dimorphic shape characteristics to inform social judgments of faces under highly constrained viewing conditions but do not test whether participants from different world regions actually *do* use those characteristics to inform social judgments of faces under more naturalistic viewing conditions (for a detailed discussion of the importance of this distinction, see Satchell et al., 2023).

Third, studies testing for consistency in the effects of masculine shape characteristics on social judgments of faces have typically investigated only a relatively small number of, often somewhat arbitrarily selected, traits (e.g., Alharbi et al., 2020; DeBruine et al., 2010; Marcinkowska et al., 2019; Penton-Voak et al., 1999; Scott et al., 2014). However, Oosterhof and Todorov (2008) reported that, when US participants rated White faces on a broad range of traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, unhappiness, intelligence, meanness, responsibility, sociability, trustworthiness, weirdness) and these ratings were subjected to Principal Component Analysis (PCA), two components were produced that explained most of the variance in ratings (see also, e.g., Wang et al., 2019). The first component, labelled Valence, was highly correlated with ratings of faces for valenced traits, such as trustworthiness and attractiveness, and is thought to reflect impressions of an individual's *intention* to inflict harm on others. The second component, labelled Dominance, was highly correlated with ratings of faces for the traits dominance and aggressiveness, and is thought to reflect impressions of an individual's *capacity* to inflict harm on others. More recently, Jones et al. (2021) reported results of a large-scale test of this Valence-Dominance model of social judgments of faces, finding that PCA of ratings of an ethnically diverse set of 120 natural (i.e., unmanipulated) face images by 11,484 participants from forty-five countries grouped into eleven world regions (Africa, Asia, Australia and New Zealand, Central America and Mexico, Eastern Europe, Middle East, Scandinavia, South America, United Kingdom, United States and Canada, Western Europe) revealed, in each world region, Valence and Dominance components that were extremely similar to those reported by Oosterhof and Todorov (2008). Consequently, Leger et al. (2023) suggested that researchers investigating potential shape correlates of social judgments of faces should use this data-reduction approach to identify the core dimensions that underpin social judgments of faces, rather than studying shape correlates of individual traits.

In the current study, we used the Valence and Dominance principal components produced by Jones et al.'s (2021) Principal Component Analyses to investigate whether masculine shape characteristics show similar relationships with social judgments of faces made by participants from different world regions. Masculinity of face shape was objectively measured from the open-access (and ethnically diverse) stimuli Jones et al. (2021) employed using a facial-metric technique that has been used in many previous studies of social judgments of faces. Previous secondary analyses used this dataset to test for relationships between social judgments of faces and other aspects of face shape (i.e., facial width-to-height ratio, Durkee & Ayers, 2021). While studies such as Perrett et al. (1998) suggest that masculine face shapes will be negatively correlated with Valence and positively correlated with Dominance and that these relationships will be somewhat consistent across world regions, studies such as Scott et al. (2014) suggest that relationships between face-shape masculinity and Valence and Dominance will vary markedly across world regions. Importantly, our study will be the first to address each of the methodological limitations of earlier studies that we described above. No aspects of our analyses were preregistered.

Although most of the research investigating the role of shape characteristics in social judgments of faces has focussed on possible effects of masculinity, other work has reported that faces rated as highly trustworthy tend to be less distinctive (Leger et al., 2023; Ryali, Goffin, Winkielman, & Yu, 2020; Sofer, Dotsch, Wigboldus, & Todorov, 2015). Indeed, Leger et al. (2023) found that Western participants' trustworthiness ratings of white faces were negatively correlated with their shape distinctiveness, but not shape masculinity. Leger et al. (2023) interpreted these results as evidence that perceptions of trustworthiness are influenced by the 'anomalous-is-bad' stereotype (the tendency for perceivers to erroneously ascribe antisocial personality traits to individuals with atypical physical appearances, Workman et al., 2021). Although tests of the 'anomalous-is-bad' stereotype have generally investigated possible effects of prominent facial anomalies (Jamrozik, Ali, Sarwer, & Chatterjee, 2019; Workman et al., 2021; Workman, Smith, Apicella, & Chatterjee, 2022), recent results linking trustworthiness to facial distinctiveness suggest that this stereotype might also extend to the effects of more subtle aspects of face distinctiveness (i.e., subtle deviations from a prototypical face shape, Leger et al., 2023). Consequently, in addition to testing for possible effects of masculine shape characteristics on Valence and Dominance, we also tested the alternative (but not necessarily mutually exclusive) hypothesis that face-shape distinctiveness would be negatively correlated with scores on the Valence component for each world region.

2. Methods

2.1. Deriving Valence and Dominance principal component (PC) scores

Valence and Dominance PC scores were taken from the results of Jones et al.'s (2021) Principal Components Analyses (PCA) of trait ratings of 60 male and 60 female faces (25% Black, 25% White, 25% East Asian, and 25% Latinx in each sex) from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) that had been rated for a wide range of traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, unhappiness, intelligence, meanness, responsibility, sociability, trustworthiness, weirdness) by 11,484 raters from eleven world regions (see Table 1 for a list of world regions, countries in each world region, and number of raters from each world region). In each world region, individual participants were randomly allocated to rate the faces for one of the traits tested. Instructions were translated from the original English-language version into the appropriate languages for each test site using a parallel back-translation method to ensure accuracy of translations. A definition for each trait was also given to ensure the translated traits accurately captured the intended concepts. The original (English-language) trait definitions were taken from Bainbridge, Isola, and Oliva (2013). Trait definitions were also translated into the appropriate languages for each test site using the same parallel back-translation method used to translate the other instructions.

Inter-rater agreement was high for all combinations of trait and world region (Cronbach's alphas ranging from 0.77 to 0.99). Jones et al. (2021) then subjected these ratings to PCA, conducting separate PCAs for each world region. The proportion of variance in trait ratings that was explained by the Valence and Dominance components for each world region, together with the countries included in each world region

Table 1
World regions (and individual countries assigned to each world region), number of raters who provided face ratings in each world region (after exclusions following preregistered data quality checks), and proportion of variance in ratings explained by Valence and Dominance components.

World region	Countries	Raters (N)	Proportion of variance in ratings explained by the Valence component	Proportion of variance in ratings explained by the Dominance component
Africa	Kenya, Nigeria, South Africa	521	0.57	0.19
Asia	China, India, Malaysia, Taiwan, Thailand	780	0.64	0.18
Australia and New Zealand	Australia, New Zealand	1044	0.63	0.17
Central America and Mexico	El Salvador, Mexico	443	0.56	0.18
Eastern Europe	Hungary, Lithuania, Poland, Russia, Serbia, Slovakia	809	0.65	0.16
Middle East	Iran, Israel, Turkey	414	0.54	0.25
Scandinavia	Denmark, Finland, Norway, Sweden	653	0.64	0.18
South America	Argentina, Brazil, Chile, Colombia, Ecuador	1283	0.59	0.21
United Kingdom	England, Scotland, Wales	361	0.60	0.18
United States and Canada	Canada, United States	3313	0.66	0.17
Western Europe	Austria, Belgium, France, Germany, Greece, Italy, Netherlands, Portugal, Spain, Switzerland	1863	0.65	0.19

and the total number of participants who provided trait ratings for each world region, are given in Table 1. Allocation of individual countries to world regions was set as part of the Stage 1 submission of the Jones et al. (2021) Registered Report (i.e., prior to data collection). For full details of how trait-ratings were collected and analysed to produce these PC (principal component) scores, see Jones et al. (2021). For all world regions, Valence PC scores were highly positively correlated (i.e., all r values >0.60) with ratings of faces for traits such as trustworthiness, attractiveness, and emotional stability. For all world regions, Dominance PC scores were highly positively correlated (i.e., all r values >0.60) with the traits dominance and aggressiveness. Correlations between individual traits and both Valence and Dominance PC scores are reported in full in Jones et al. (2021).

2.2. Calculating distinctiveness face-shape scores

Distinctiveness scores were calculated from each face photograph rated in Jones et al. (2021) using the facefuns package (Hozleitner & DeBruine, 2021) in R (R Core Team, 2021) and open-access image templates for the Chicago Face Database (Singh, Gambrell, & Correll, 2023). This method has been used to measure face-shape distinctiveness in many previous studies (e.g., Cai et al., 2019; Holzleitner et al., 2019; Lee et al., 2016; Leger et al., 2023). This method uses the shape principal components described in the previous section of our methods to measure the distance each face lies from the mathematical average shape for the sample of faces of the same sex. That is, the average shape values for the same-sex sample were calculated and, for each image, the Euclidean distance from the average was derived (e.g., Cai et al., 2019; Holzleitner et al., 2019; Lee et al., 2016; Leger et al., 2023). Higher scores indicate that the face lies a further distance away from the average (i.e., had a more distinctive shape). We measured distinctiveness scores for male and female faces of each ethnicity separately in light of behavioural evidence that faces are primarily processed relative to sex-specific (e.g., Little, DeBruine, & Jones, 2005; Rhodes et al., 2011) and ethnicity-specific (e.g., Jaquet, Rhodes, & Hayward, 2007, Jaquet & Rhodes, 2008; Little, DeBruine, Jones, & Waitt, 2008) prototypes. Higher scores indicate more distinctive (i.e., less prototypical) face shapes. One face had a value more than three standard deviations from the mean value for the sample. The score for this face was winsorized (i.e., adjusted to be three standard deviations from the mean value for the sample) prior to analyses. All values were converted to z-scores (separately for male and female faces) prior to analyses.

2.3. Calculating masculinity face-shape scores

Masculinity of face shape was objectively measured for each of the 60 male and 60 female face images rated in Jones et al. (2021) using the facefuns package (Hozleitner & DeBruine, 2021) in R (R Core Team, 2021) and open-access image templates for the Chicago Face Database (Singh et al., 2023). This geometric morphometric method that has been used in many previous studies to assess masculinity of face shape (e.g., Bartlome & Lee, 2023; Cai et al., 2019; Dong et al., 2023; Holzleitner et al., 2019; Komori et al., 2011; Leger et al., 2023). Shape components were first derived from Principal Component Analysis (PCA) of 132 Procrustes-aligned landmark points (see Holzleitner et al., 2019 for a diagram showing these facial landmarks) on each of the 60 male and 60 female face images. Scores representing masculinity of face shape were then calculated from each photograph using a vector analysis method (e.g., Bartlome & Lee, 2023; Cai et al., 2019; Dong et al., 2023; Holzleitner et al., 2019; Komori et al., 2011; Leger et al., 2023). This method uses the shape principal components to locate each face on a female-male continuum. The female-male continuum was defined by calculating the average shape information of the female faces and the average shape information of the male faces. Masculinity scores were then derived by projecting each image onto this female-male vector. We measured masculinity scores for faces of each ethnicity separately in light of behavioural evidence that faces are primarily processed relative to ethnicity-specific prototypes (Jaquet et al., 2007, Jaquet & Rhodes, 2008; Little et al., 2008). Higher scores indicate more masculine (i.e., less feminine) face shapes. One face had a value more than three standard deviations from the mean value for the sample. The score for this face was winsorized (i.e., adjusted to be three standard deviations from the mean value for the sample) prior to analyses. Consistent with previous studies using different sets of face images (e.g., Leger et al., 2023), distinctiveness and masculinity scores were not significantly correlated ($r = 0.008$, $N = 120$, $p = .935$). All scores were converted to z-scores (separately for male and female faces) prior to analyses.

3. Results

We analysed Valence and Dominance PC scores using linear mixed effects models in which scores from all world regions were included in the same model. Valence and Dominance PC scores were analysed in separate models. In both models, PC scores served as the outcome measure, with distinctiveness and masculinity entered as predictors, along with face sex (effect coded so that −0.5 corresponded to female faces and + 0.5 corresponded to male faces), the interaction between masculinity and face sex, and the interaction between distinctiveness and face sex. The models also included by-region and by-stimulus random intercepts and by-region random slopes for masculinity,

distinctiveness, face sex, and the interactions between masculinity and face sex and between distinctiveness and face sex. These analyses were conducted using the packages tidyverse 1.3.1 (Wickham, 2021), lmerTest 3.1–3 (Kuznetsova, Brockhoff, & Christensen, 2017), and jtools 2.2.3 (Long, 2022). Full results and analysis code for both analyses are available at https://osf.io/ms93u/.

For our analysis of Valence PC scores, there was a significant negative main effect of distinctiveness, indicating that more distinctive faces scored higher on the Valence component (see Table 2). The only other significant effect in this model was an effect of face sex, whereby female faces tended to score higher on the Valence component than did male faces (see Table 2). Estimated between-region variation in the effects of both masculinity and distinctiveness was effectively zero (SDs < 0.001), and the partially pooled regional estimates were therefore identical to the overall estimates across regions (see analysis output on the OSF page). The corresponding analysis of Dominance PC scores showed no significant effects (see Table 3). Between-region variation in the effects of masculinity and distinctiveness was small (SD = 0.021 and SD = 0.016, respectively). Partially pooled regional estimates ranged from 0.106 to 0.139 for masculinity and from −0.124 to −0.097 for distinctiveness (see analysis output on the OSF page).

We fitted a regularized Bayesian multilevel model because several of the between-region variance components were estimated at or very close to zero in the frequentist model. Weakly informative priors were used to regularize these variance components and quantify uncertainty in the between-region variation. Full details of the prior specification and model fitting are available on the OSF page (https://osf.io/ms93u/). For Valence PC scores, distinctiveness showed a negative average association (estimate = −0.21, 95% CrI [−0.37, −0.05]), whereas the average association with masculinity was smaller and uncertain (estimate = −0.08, 95% CrI [−0.24, 0.09]). The only other association with a 95% credible interval excluding zero was the effect of face sex, whereby female faces tended to score higher on the Valence component than male faces (estimate = −0.39, 95% CrI [−0.71, −0.06]). The interactions between masculinity and face sex (estimate = 0.10, 95% CrI [−0.21, 0.42]), and between distinctiveness and face sex (estimate = 0.11, 95% CrI [−0.21, 0.43]), were both uncertain. Estimated between-region variation in the focal associations was very small for both masculinity, SD = 0.01, 95% CrI [0.00, 0.04], and distinctiveness, SD = 0.01, 95% CrI [0.00, 0.03]. The partially pooled regional estimates were also very similar across regions, ranging from −0.088 to −0.071 for masculinity and from −0.214 to −0.207 for distinctiveness (see Fig. 1). Additionally, we computed from this model the average marginal effect (AME) of distinctiveness and masculinity, which represents the average change in valence loadings with unit increases in distinctiveness and masculinity, respectively. As the model is Bayesian, we obtain a distribution of credible values for these estimates. For distinctiveness, the AME was −0.22 [−0.39, −0.06], with a 99.3% probability of being negative. For masculinity, this was −0.08 [−0.25, 0.08], with an 83.5% probability of being negative. We then took the difference between these two estimates, subtracting the masculinity estimate from distinctiveness. The difference in coefficients was −0.14 [−0.38, 0.09], with an 86.9% probability that distinctiveness was larger than masculinity; or equivalently, the odds distinctiveness has a larger association than masculinity is 6.63 (see https://osf.io/ms93u/ for full details of this analysis).

Table 2
Results of our linear mixed effects model analysis of Valence PC scores for all world regions. Note that distinctiveness, but not masculinity, significantly predicted Valence PC scores.

	Estimate	Standard Error	t	df	p
Intercept	−0.001	0.083	−0.018	120.008	0.986
Masculinity	−0.083	0.083	−1.002	120.008	0.318
Face sex	−0.435	0.168	−2.588	126.204	0.011
Distinctiveness	−0.220	0.085	−2.582	120.008	0.011
Masculinity x face sex	0.115	0.166	0.692	120.008	0.490
Distinctiveness x face sex	0.127	0.172	0.739	123.178	0.461

Table 3
Results of our linear mixed effects model analysis of Dominance PC scores for all world regions. Note that this analysis showed no significant effects.

	Estimate	Standard Error	t	df	p
Intercept	0.001	0.085	−0.005	119.998	0.996
Masculinity	0.125	0.086	1.455	120.950	0.148
Face sex	0.159	0.174	0.918	125.756	0.360
Distinctiveness	−0.109	0.088	−1.237	120.462	0.218
Masculinity x face sex	−0.031	0.171	−0.179	119.998	0.858
Distinctiveness x face sex	0.063	0.177	0.356	121.671	0.722

The corresponding Bayesian analysis of Dominance PC scores showed no clear average associations, with 95% credible intervals for all fixed effects including zero. Estimated between-region variation in the associations of masculinity and distinctiveness was small (SD = 0.02, 95% CrI [0.00, 0.06], for both). Partially pooled regional estimates ranged from 0.095 to 0.130 for masculinity and from −0.127 to −0.092 for distinctiveness, with all corresponding 95% credible intervals including zero (see Fig. 1).

As a robustness check, we repeated all analyses using alternative sets of faces to construct the sex- and ethnicity-specific prototypes from which distinctiveness and masculinity were calculated. The results showed the same overall pattern, suggesting that the findings were not dependent on the particular faces used to construct the prototypes. These analyses are reported in full on the OSF page (https://osf.io/ms93u/).

4. Discussion

We investigated possible relationships between aspects of face shape (masculinity, distinctiveness) and Valence and Dominance components derived from ratings of faces made by participants in eleven world regions. Shape distinctiveness was negatively and significantly correlated with scores on the Valence component, and the estimated between-region variation in this association was very small, indicating that this association was similar across world regions. This finding is consistent with those from recent studies that reported negative correlations between distinctiveness and trustworthiness ratings in western samples (Leger et al., 2023; Ryali et al., 2020; Sofer et al., 2015). Moreover, our results for distinctiveness and Valence are consistent with the existence of an anomalous-is-bad stereotype, whereby people tend to ascribe negative traits to individuals with atypical physical appearances (Workman et al., 2021). Thus, our results suggest that shape distinctiveness is consistently associated with Valence judgments across participants from different world regions, in a pattern consistent with the anomalous-is-bad stereotype.

While previous studies reported that experimentally manipulated masculinity in face images had a positive effect on perceptions of dominance and/or negative effects on valenced perceptions (e.g., Alharbi et al., 2020; Penton-Voak et al., 1999; Perrett et al., 1998), relationships between masculinity and Valence or Dominance were not significant for any of the world regions we tested. These null results are consistent with results of other recent work suggesting that the large effects of masculinity on social judgments of faces obtained using stimuli in which sexually dimorphic shape characteristics were experimentally manipulated do not typically occur when natural (i.e., unmanipulated faces) faces are rated and sexually dimorphic shape characteristics are measured objectively from the face images (Dong et al., 2023; Jones et al., 2023; Jones & Jaeger, 2019; Leger et al., 2023; Scott et al., 2010). Indeed, other recent studies using similar methods to those

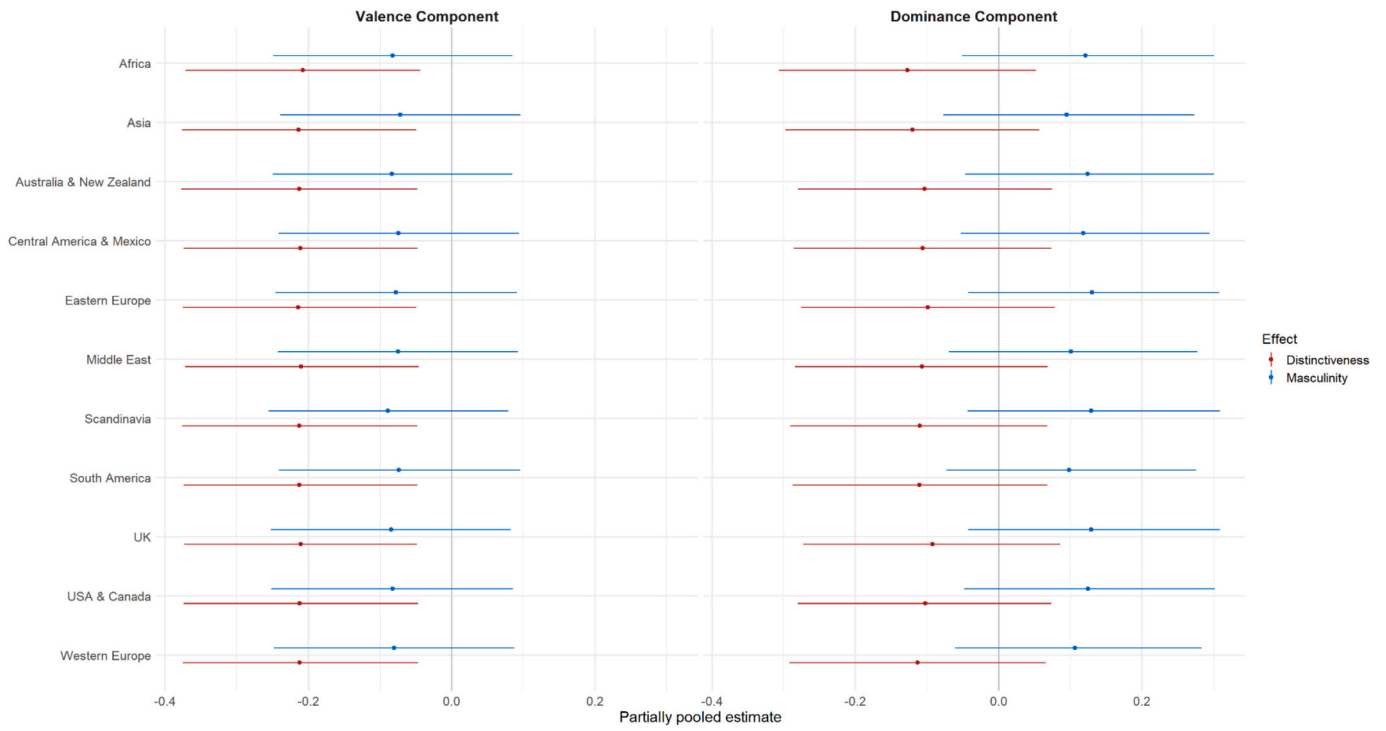


Fig. 1. Partially pooled regional estimates of the associations of facial masculinity and distinctiveness with Valence and Dominance component scores from the regularized Bayesian multilevel models. Points represent posterior mean estimates and horizontal lines represent 95% credible intervals.

that we employed in the current study, but different face-image sets, also found that masculinity did not reliably predict perceptions of dominance (Dong et al., 2023) or trustworthiness (Leger et al., 2023). Thus, shape masculinity appears relatively unimportant for social judgments of faces. While we cannot rule out the possibility that relationships between masculinity and social judgments of faces exist that are too small for our study to detect (note that the critical unit of analysis for these tests was the 120 face images), collectively our results suggest that distinctiveness is a better predictor of valenced perceptions of faces among participants from different world regions than is masculinity. Indeed, the results of our Bayesian analysis of the Valence component supports this interpretation.

To summarise, we tested for evidence that facial masculinity and distinctiveness contribute to valenced perceptions of faces and perceptions of dominance made by participants from eleven world regions. By contrast with the influential claim that masculinity is critical for social judgments of faces, we found little evidence that masculinity was associated with either Valence or Dominance. By contrast, distinctiveness (i.e., atypicality) was negatively associated with Valence, and the direct comparison indicated a stronger association for distinctiveness than for masculinity. While further work is needed to establish whether these results generalise to small-scale societies (see Apicella, Little, & Marlowe, 2007 for preliminary evidence that distinctiveness may predict social judgments of faces in hunter-gatherers), our results suggest that distinctiveness, but not masculinity, is consistently associated with valenced social judgments of faces among people from different world regions.

CRediT authorship contribution statement

Victor Shiramizu: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Junzhi Dong:** Writing – review & editing, Writing – original draft. **Kathlyne Leger:** Writing – review & editing, Writing – original draft. **Anthony J. Lee:** Writing – review & editing, Writing – original draft, Formal analysis. **Alex L. Jones:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Yasaman Rafiee:** Writing – review & editing, Writing – original draft. **Zuzana Elliott:** Writing – review & editing, Writing – original draft. **Lisa M. DeBruine:** Writing – review & editing, Writing – original draft, Validation, Investigation, Formal analysis. **Benedict C. Jones:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by ESRC grant ES/X000249/1 awarded to BCJ and University of Strathclyde Global Research Awards to KL and JD. For the purpose of Open Access, the authors have applied a Creative Commons Attribution (CC BY) to any Author Accepted Manuscript (AAM) version arising from this submission.

References

Alharbi, S. A., Holzleitner, I. J., Lee, A. J., Saribay, S. A., & Jones, B. C. (2020). Women's preferences for sexual dimorphism in faces: Data from a sample of Arab women. *Evolutionary Psychological Science, 6*, 328–334.

Apicella, C. L., Little, A. C., & Marlowe, F. W. (2007). Facial averageness and attractiveness in an isolated population of hunter-gatherers. *Perception, 36*(12), 1813–1820.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology. General, 142*(4), 1323.

Bartlome, R. I., & Lee, A. J. (2023). Facial attractiveness, but not facial masculinity, is used as a cue to paternal involvement in fathers. *Adaptive Human Behavior and Physiology*, 1–16.

Cai, Z., Hahn, A. C., Zhang, W., Holzleitner, I. J., Lee, A. J., DeBruine, L. M., & Jones, B. C. (2019). No evidence that facial attractiveness, femininity, averageness,

or coloration are cues to susceptibility to infectious illnesses in a university sample of young adult women. *Evolution and Human Behavior, 40*(2), 156–159.
DeBruine, L. M., Jones, B. C., Crawford, J. R., Welling, L. L., & Little, A. C. (2010). The health of a nation predicts their mate preferences: Cross-cultural variation in women's preferences for masculinized male faces. *Proceedings of the Royal Society Series B Biological Sciences, 277*(1692), 2405–2410.
Dong, J., Leger, K., Shiramizu, V. K., Marcinkowska, U. M., Lee, A. J., & Jones, B. C. (2023). The importance of face-shape masculinity for perceptions of male dominance depends on study design. *Scientific Reports, 13*(1), 12620.
Durkee, P. K., & Ayers, J. D. (2021). Is facial width-to-height ratio reliably associated with social inferences? *Evolution and Human Behavior, 42*(6), 583–592.
Holzleitner, I. J., Lee, A. J., Hahn, A. C., Kandrik, M., Bovet, J., Renoult, J. P., & Jones, B. C. (2019). Comparing theory-driven and data-driven attractiveness models using images of real women's faces. *Journal of Experimental Psychology. Human Perception and Performance, 45*(12), 1589.
Hozleitner, I. J., & DeBruine, L. M. (2021). facefuns (Version 0.0.0.900) [Computer software]. Retrieved from https://iholzleitner.github.io/facefuns/index.html.
Jamrozik, A., Ali, M. O., Sarwer, D. B., & Chatterjee, A. (2019). More than skin deep: Judgments of individuals with facial disfigurement. *Psychology of Aesthetics, Creativity, and the Arts, 13*(1), 117.
Jaquet, E., & Rhodes, G. (2008). Face aftereffects indicate dissociable, but not distinct, coding of male and female faces. *Journal of Experimental Psychology. Human Perception and Performance, 34*(1), Article 101.
Jaquet, E., Rhodes, G., & Hayward, W. G. (2007). Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences. *The Quarterly Journal of Experimental Psychology, 60*(11), 1457–1467.
Jones, A. L., & Jaeger, B. (2019). Biological bases of beauty revisited: The effect of symmetry, averageness, and sexual dimorphism on female facial attractiveness. *Symmetry, 11*(2), Article 279.
Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., & Sirota, M. (2021). To which world regions does the valence–dominance model of social perception apply? *Nature Human Behaviour, 5*(1), 159–169.
Jones, B. C., Jones, A. L., & Shiramizu, V. (2023). Mapping physical characteristics in face images to social judgements. *British Journal of Psychology, 114*(2), 498–500.
Komori, M., Kawamura, S., & Ishihara, S. (2011). Multiple mechanisms in the perception of face gender: Effect of sex-irrelevant features. *Journal of Experimental Psychology. Human Perception and Performance, 37*(3), 626.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13
Lee, A. J., De La Mare, J. K., Moore, H. R., & Umeh, P. C. (2021). Preference for facial symmetry depends on study design. *Symmetry, 13*(9), Article 1637.
Lee, A. J., Mitchem, D. G., Wright, M. J., Martin, N. G., Keller, M. C., & Zietsch, B. P. (2016). Facial averageness and genetic quality: Testing heritability, genetic correlation with attractiveness, and the paternal age effect. *Evolution and Human Behavior, 37*(1), 61–66.
Leger, K., Dong, J., DeBruine, L. M., Jones, B. C., & Shiramizu, V. K. (2023). Assessing the roles of shape prototypicality and sexual dimorphism in ratings of the trustworthiness of faces. *Scientific Reports, 13*(1), 15662.
Lewis, M. B. (2017). Fertility affects asymmetry detection not symmetry preference in assessments of 3D facial attractiveness. *Cognition, 166*, 130–138.
Lewis, M. B. (2020). Challenges to both reliability and validity of masculinity-preference measures in menstrual-cycle-effects research. *Cognition, 197*, Article 104201.
Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society B, 272*(1578), 2283–2287.
Little, A. C., DeBruine, L. M., Jones, B. C., & Waitt, C. (2008). Category contingent aftereffects for faces of different races, ages and species. *Cognition, 106*(3), 1537–1547.
Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 366*(1571), 1638–1659.
Long, J. A. (2022). jtools: Analysis and presentation of social scientific data. (Version 2.2.0). [Computer software]. Retrieved from https://cran.r-project.org/web/packages/jtools.
Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122–1135.
Marcinkowska, U. M., Rantala, M. J., Lee, A. J., Kozlov, M. V., Aavik, T., Cai, H., & Dixson, B. J. (2019). Women's preferences for men's facial masculinity are strongest under favorable ecological conditions. *Scientific Reports, 9*(1), 1–10.
Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences, 18*(11), 566–570.
Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087–11092.
Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature, 399*(6738), 741–742.
Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*(6696), 884–887.
Phalane, K. G., Tribe, C., Steel, H. C., Cholo, M. C., & Coetzee, V. (2017). Facial appearance reveals immunity in African men. *Scientific Reports, 7*(1), 7443.
R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.
Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.
Rhodes, G., Jaquet, E., Jeffery, L., Evangelista, E., Keane, J., & Calder, A. J. (2011). Sex-specific norms code face identity. *Journal of Vision, 11*(1), 1.
Ryali, C. K., Goffin, S., Winkielman, P., & Yu, A. J. (2020). From likely to likable: The role of statistical typicality in human social assessment of faces. *Proceedings of the National Academy of Sciences, 117*(47), 29371–29380.
Satchell, L., Jaeger, B., Jones, A., López, B., & Schild, C. (2023). Beyond reliability in first impressions research: Considering validity and the need to "mix it up with folks". *Social Psychological and Personality Science, 18*, 1–21.
Scott, I. M., Clark, A. P., Boothroyd, L. G., & Penton-Voak, I. S. (2013). Do men's faces really signal heritable immunocompetence? *Behavioral Ecology, 24*(3), 579–589.
Scott, I. M., Clark, A. P., Josephson, S. C., Boyette, A. H., Cuthill, I. C., Fried, R. L., & Penton-Voak, I. S. (2014). Human preferences for sexually dimorphic faces may be evolutionarily novel. *Proceedings of the National Academy of Sciences, 111*(40), 14388–14393.
Scott, I. M., Pound, N., Stephen, I. D., Clark, A. P., & Penton-Voak, I. S. (2010). Does masculinity matter? The contribution of masculine face shape to male attractiveness in humans. *PLoS One, 5*(10), Article e13585.
Singh, B., Gambrell, A., & Correll, J. (2023). Face templates for the Chicago Face Database. *Behavior Research Methods, 55*(2), 639–645.
Sofer, C., Dotsch, R., Wigboldus, D. H., & Todorov, A. (2015). What is typical is good: The influence of face typicality on perceived trustworthiness. *Psychological Science, 26*(1), 39–47.
Stephen, I. D., Scott, I. M., Coetzee, V., Pound, N., Perrett, D. I., & Penton-Voak, I. S. (2012). Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evolution and Human Behavior, 33*(4), 260–267.
Wang, H., Han, C., Hahn, A. C., Fasolt, V., Morrison, D. K., Holzleitner, I. J., et al. (2019). A data-driven study of Chinese participants' social judgments of Chinese faces. *PLoS ONE, 14*(1), Article e0210315. https://doi.org/10.1371/journal.pone.0210315
Wickham, H. (2021). Tidyverse: Easily install and load the 'Tidyverse' (Version 1.3.1). [Computer software]. Retrieved from https://cran.rproject.org/web/packages/tidyverse/index.html.
Workman, C. I., Humphries, S., Hartung, F., Aguirre, G. K., Kable, J. W., & Chatterjee, A. (2021). Morality is in the eye of the beholder: The neurocognitive basis of the "anomalous-is-bad" stereotype. *Annals of the New York Academy of Sciences, 1494*(1), 3–17.
Workman, C. I., Smith, K. M., Apicella, C. L., & Chatterjee, A. (2022). Evidence against the "anomalous-is-bad" stereotype in Hadza hunter gatherers. *Scientific Reports, 12*(1), 8693.