PROCEEDINGS B

royalsocietypublishing.org/rspb

CC BY

Research

Cite this article: Wassiliwizky E, Zietsch BP, Kleisner K, Ullén F. 2026 The gender attractiveness gap. *Proc. R. Soc. B* **293**: 20260362. https://doi.org/10.1098/rspb.2026.0362

Received: 30 June 2025
Accepted: 9 April 2026

Subject Category:
Behaviour

Subject Areas:
behaviour, evolution, cognition

Keywords:
facial attractiveness, gender gap, evolutionary psychology, mate preference, sexual dimorphism

Author for correspondence:
Eugen Wassiliwizky
e-mail: eugen.wassiliwizky@ae.mpg.de

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8476530.

THE ROYAL SOCIETY PUBLISHING

The gender attractiveness gap

Eugen Wassiliwizky[1], Brendan P. Zietsch[2], Karel Kleisner[3] and Fredrik Ullén[1,4]

[1]Cognitive Neuropsychology, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany
[2]Deprtment of Psychology, University of Queensland, Brisbane, Queensland, Australia
[3]Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czechia
[4]Department of Neuroscience, Karolinska Institute, Stockholm, Stockholm County, Sweden

EW, 0000-0002-1087-1657; BPZ, 0000-0003-0274-6140

Writers from Darwin to Dawkins have noted that, in humans, women are considered the 'beautiful sex', whereas in most species, it is males who display more elaborate, visually striking traits. This reversal of typical sex roles has long been debated in the literature but never empirically verified. Here, we present a large-scale cross-cultural meta-analysis of same- and opposite-sex ratings of facial attractiveness. Our findings provide robust evidence for a gender attractiveness gap (GAP), with female faces rated more attractive than male faces across rater sexes, cultures, races and age groups. Surprisingly, the GAP is more pronounced among female raters, who rate other women substantially higher than male faces, while male faces receive similarly low ratings from both sexes; the GAP is absent in self-ratings. Morphometric analyses further show that sexual shape dimorphism accounts for a substantial portion of the GAP. Additionally, we observe a tendency for greater stringency among male raters. This study draws on the largest worldwide collection of facial-attractiveness data assembled to date, which we make openly available to facilitate future research. Together, these results demonstrate that facial-attractiveness judgements extend well beyond heterosexual mate choice and reflect the joint influence of facial structure, rater characteristics and sociocultural context.

1. Introduction

The assessment of facial attractiveness is a fundamental aspect of human interaction, often operating subconsciously and continuously in everyday life [1–4]. In particular, the evaluation of facial attractiveness plays a key role in shaping human mate choice [5–7]. A central idea in evolutionary psychology proposes that preferences commonly observed across individuals and populations may have evolved to enhance inclusive fitness—the overall genetic success of an individual through the survival and reproduction of their offspring and kin [3,8]. Selecting mates with higher genetic quality increases the likelihood of producing healthier and more viable offspring. One prominent example of such a preference is facial averageness, which reflects how closely an individual's structural facial features align with the population-wide average. Facial averageness is believed to signal heterozygosity (i.e. greater genetic diversity) which is associated with improved immunocompetence [6–8], developmental stability [9] and overall biological quality [10]. Thus, by favouring faces with higher averageness, individuals indirectly are assumed to select mates with higher genetic quality, thereby enhancing their reproductive success. As a result, the association of these structural facial features with aesthetic preferences and motivational tendencies observed in contemporary datasets is likely to reflect evolutionary processes shaped over millennia of human history.

Another important line of thought in evolutionary theory highlights that the sexes have different strategies in mate selection, arising from differences in parental investment [11,12]. In most sexually reproducing species, females

© 2026 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

have evolved to be the choosy sex, placing selective pressure on males to compete for their attention. Female choice effectively determines which males reproduce and which do not, driving competition among males. This dynamic explains why, in species with sexual dimorphism—where the sexes differ in appearance—males are typically more ornamented and visually striking [13,14]. Darwin's theory of sexual selection was, in fact, developed to explain such phenomena as the extravagant plumage of male peacocks. Beyond visual traits, these selective pressures also shape elaborate male behaviours, such as singing, courtship dances and other complex displays intended to captivate potential mates.

However, Darwin himself observed that, in contrast to examples from the animal kingdom, this dynamic appears to be reversed in humans [15, p. 618]. In other words, it is women, not men, who are often regarded as the 'beautiful sex'. This peculiarity of the human species has been explored by numerous scholars, including Dawkins [16] and Symons [17], and has long been considered an evolutionary riddle [18]. Common expressions idealizing female beauty in various languages—such as 'the fair sex' in English, 'das schöne Geschlecht' in German, 'le beau sexe' in French, 'прекрасный пол' in Russian—seem to support this notion.

Several biological explanations have been proposed for this apparent sex-role reversal in humans. One line of reasoning builds on the link between facial averageness and attractiveness. In species with mating systems characterized by high variance in reproductive success (e.g. polygynous species), selection tends to favour increased variability in the sex with higher reproductive variance [15]. This leads to intense male competition and evolution of greater male trait variability. Human facial morphology appears to follow a similar pattern: male faces are more variable than female faces [19]. In contrast, women display higher facial averageness [20], making them statistically more likely to be perceived as attractive.

Another biological account posits that, unlike species with pronounced male ornamentation, humans are characterized by mutual mate choice rather than exclusively choosy females and intense male–male competition. Mutual mate choice is associated with sex-differentiated signalling across trait domains [21]. In this context, the emphasis on female physical attractiveness may reflect domain-specific ornamentation, with cues to fertility and health prioritized in women, whereas male signalling may be more strongly expressed through non-facial traits such as status or resource acquisition.

A further line of work focuses on developmental mechanisms, specifically the influence of sex hormones on facial morphology, which has been well documented, particularly in relation to sexual dimorphism. While femininity is broadly and consistently preferred in female faces, the preference for masculinity in male faces is far less robust. This may be because higher testosterone levels in men, while signalling dominance and immunocompetence, are also associated with increased risk of aggression, dishonesty, or reduced paternal investment [22]. Female faces, by contrast, tend to retain neotenous traits into adulthood, such as larger eyes, smaller noses, convex foreheads and rounder facial contours (babyfacedness [23]), which are known to elicit caregiving responses, attract visual attention and enhance perceptions of youth and beauty [24,25]. Further sexually dimorphic traits may also contribute to the perception of women as the more attractive sex. Specifically, across multiple populations in Africa (e.g. in Sudan, Senegal, Mauritania, Cameroon, Ethiopia and Somalia [26,27]), as well as in Asia (e.g. Vietnam) [28] and Europe [29], women have been reported to have lighter skin tones than men, which is frequently preferred in women. Similar preferences have also been claimed for culturally independent, pre-colonial societies [30].

Surprisingly, despite its longstanding presence in scientific and cultural discourse—and an extensive body of research on facial attractiveness—the existence of what we term the gender attractiveness gap (GAP) has never been systematically verified. While many theorists (predominantly men) have assumed a shared consensus that women are inherently more attractive than men, this assumption raises a critical question: do both men and women actually perceive women as more attractive, or is this notion simply a reflection of gendered bias among these theorists? Without direct empirical evidence, both the true nature and even the very existence of the GAP remain uncertain.

To place the issue on a solid empirical foundation, we conducted a comprehensive meta-analytic and pooled analysis of facial-attractiveness datasets published over the past decade, leveraging the increased availability of shared data through open-science initiatives. Using a PRISMA-compliant systematic literature search [31], we identified studies with ratings from male and female raters evaluating same- and opposite-sex facial images. Applying predefined inclusion criteria, we compiled and harmonized a large global dataset comprising 52 studies from 76 countries, roughly 28 500 raters, 17 000 facial stimuli and 1.5 million individual ratings. To our knowledge, this constitutes the largest collection of facial-attractiveness data assembled to date. Using this dataset, we rigorously assessed the existence of the GAP and precisely estimated its magnitude with robust cross-cultural validity.

For the majority of studies, we additionally obtained the original facial stimuli from the facial databases and conducted large-scale geometric morphometric analyses, enabling a direct test of whether structural facial features account for the GAP. Finally, we developed a fully automated data-construction pipeline that retrieves each original dataset from its source, standardizes and integrates the data into a harmonized database and links these data to our morphometric measures. This framework ensures full reproducibility of all reported results and provides a resource for future large-scale, cross-cultural research on facial attractiveness and morphometry.

2. Methods

(a) Study identification and data integration

We conducted a systematic literature search of peer-reviewed studies to identify published datasets suitable for large-scale synthesis. Full details on the procedure—including search terms, databases, inclusion and exclusion criteria, the PRISMA flow diagram (electronic supplementary material, figure S1), and a comprehensive table of all included studies (electronic

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

supplementary material, table S1) with detailed information on sample characteristics, stimulus sources, age ranges and data structure—are provided in the electronic supplementary material. These also include a cross-tabulated frequency table (electronic supplementary material, table S2) reporting the number of stimuli by rater cultural background and stimulus race.

From these identified studies, we included attractiveness ratings for front-view facial photographs of real individuals, depicting natural or neutral expressions, presented alone and without face coverings. In some cases, this required selecting data subsets, such as excluding ratings for faces with emotional expressions, masks or digital manipulations. The included studies were originally conducted to address a wide range of research questions, none of which explicitly targeted sex differences in attractiveness. Because the studies differed in response scale ranges and levels of data aggregation, all attractiveness ratings were *z*-standardized within each study across both stimulus sexes prior to harmonization. Figure 1 displays the resulting distribution of attractiveness scores, aggregated at the stimulus level, for female (n = 7383) and male (n = 7992) faces across the entire dataset. Even at this descriptive level, the distribution for male faces is shifted leftward relative to zero on the x-axis, indicating the presence of the GAP. Notably, despite this overall leftward shift, the upper tail of the male distribution is relatively pronounced, suggesting a disproportionate representation of highly attractive male faces compared with what would be expected under a normal distribution.

(b) Statistical analyses

Our analytic strategy proceeded in two main steps, reflecting heterogeneity in the level of data aggregation across studies (electronic supplementary material, table S1). The datasets comprised raw individual-level data, stimulus-level data aggregated separately for female and male raters and stimulus-level data aggregated across rater sex. Across analyses, we consistently used the most inclusive data representation compatible with the respective statistical model, in order to maximize statistical power and coverage.

(i) Meta-analysis

Using all datasets irrespective of aggregation level, we derived study-specific estimates of the GAP, defined as the marginal mean difference in attractiveness between female and male faces obtained from a regression model appropriate to the data structure. When individual-level ratings were available, we fitted linear mixed-effects models (LMMs) with stimulus sex (FaceSex) as a fixed effect and random intercepts for raters and stimuli. When RaterSex was available, it was included (with its interaction with FaceSex) to obtain the marginal F–M contrast averaged across RaterSex levels. For aggregated datasets, analogous models were fitted at the stimulus level using LMMs or linear models as appropriate. In all cases, the squared standard error of the F–M contrast served as the sampling variance. Study-specific estimates were combined in a random-effects meta-analysis using restricted maximum likelihood estimation (REML), allowing for between-study heterogeneity.

(ii) Pooled interaction analyses

To characterize how the GAP varied between female and male raters, we analysed datasets providing stimulus-level attractiveness estimates separately by RaterSex. To maximize data coverage, we computed RaterSex-specific stimulus means for datasets with raw ratings and harmonized these with datasets providing aggregated data. We then fitted LMMs predicting attractiveness from FaceSex, RaterSex and their interaction, with random intercepts for stimuli and study to account for clustering. To verify that aggregation did not affect the results, we re-analysed datasets with raw ratings using an individual-level model that additionally included random intercepts for raters (electronic supplementary material, figure S2, and table S3, right side).

(iii) Pooled stability analyses

We next examined the stability of the GAP across a range of rater- and stimulus-level moderators using individual-level raw data, which were required to retain rater-specific information. On the rater side, we tested moderation by raters' cultural background (based on the World Cultural Map [32]; see electronic supplementary material, table S4, for the mapping of countries to the eight cultural regions), sexual orientation, age and sample type (e.g. student sample, online sample, community sample). On the stimulus side, we examined moderation by stimulus race, stimulus source (e.g. facial databases, laboratory photographs, web profiles) and stimulus age. Moderator information was extracted directly from the published data or the corresponding study descriptions. Not all studies provided information for all moderators; in such cases, the respective study was excluded from the corresponding analysis. Nevertheless, given the size of our harmonized dataset, each moderator analysis was still based on more than half a million individual observations (933 088 on average). For all models, raters, stimuli and studies were included as random-intercept factors. Pairwise *post hoc* comparisons were Bonferroni-corrected to control for multiple testing.

Since six studies (amounting to approx. 86 000 observations from 31 countries and spanning all eight regions of the World Cultural Map) provided data on self-rated attractiveness, we additionally tested whether a GAP emerged when raters evaluated their own attractiveness rather than that of presented stimuli.

(iv) Morphometric mediation analysis

In an additional series of analyses, we examined the extent to which the GAP was mediated by structural features that define male and female faces. To this end, we used morphometric measures provided in a subset of studies (electronic supplementary

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

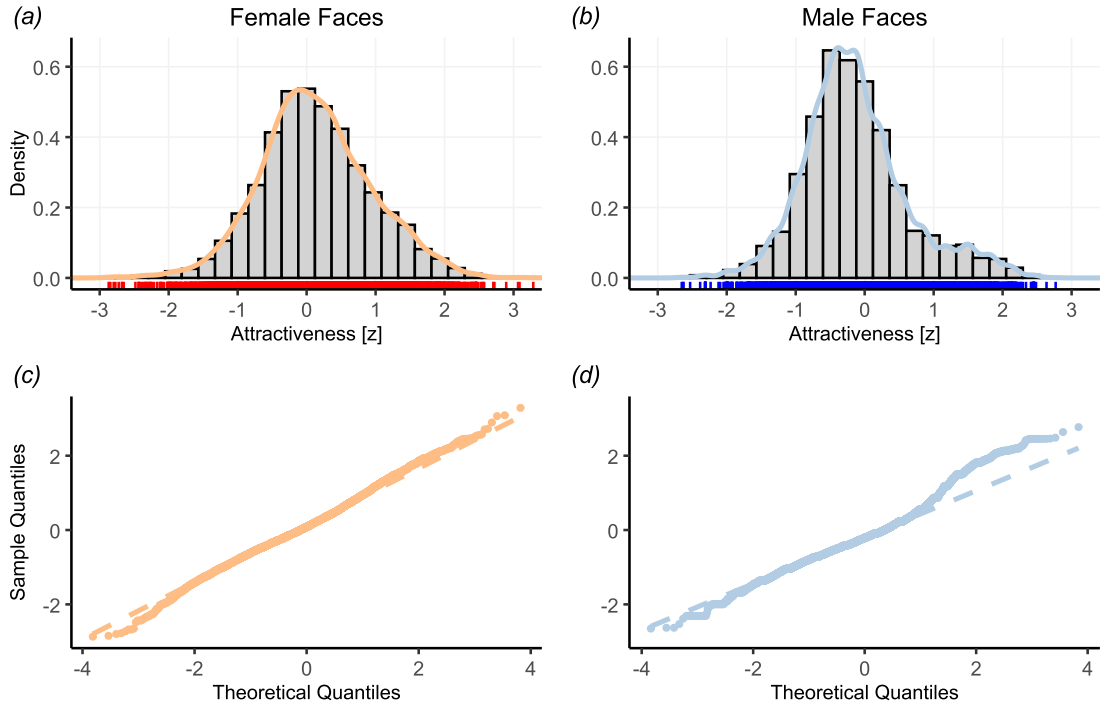

Figure 1. Distribution and normality assessment of attractiveness scores in the harmonized dataset. The upper row displays histograms with overlaid density plots for female (*a*) and male (*b*) stimuli with each stimulus represented by a single aggregated mean. Rug plots (red and blue ticks along the *x*-axis) represent data points for individual stimuli. The lower row presents *Q–Q* plots assessing normality for female (*c*) and male (*d*) stimuli, with sample quantiles plotted against theoretical normal quantiles. Dashed lines indicate expected normal distributions. While both distributions align closely with normality in the central range, male stimuli exhibit a right-tail deviation.

material, table S1). The key metric for our case was sexual shape dimorphism (SShD), a holistic geometric measure of femininity/masculinity. It is calculated using statistical shape analysis of facial landmarks, assessing the Procrustes fit of individual faces with population-based male and female reference configurations [33,34]. The resulting score places a face along a female–male continuum, with more negative values indicating a more feminine face, more positive values a more masculine face and values around zero reflecting androgyny (see electronic supplementary material, figure S3, for the distribution of SShD across female and male stimuli in the present dataset). In general, sex-typicality is preferred by the opposite sex [35–37], although this relationship varies across cultural and socioeconomic contexts, particularly with regard to female preference for masculinity in male faces [38–42].

Given that SShD is associated with both stimulus sex and perceived attractiveness, we conducted a mediation analysis to assess the extent to which SShD accounted for the GAP. To isolate the mediating role of SShD, we first removed the influence of two other well-established influences on attractiveness: youthfulness [11,42,43] and facial averageness [44,45], with the latter quantified as an objective morphometric measure, by projecting individual facial landmarks onto a population-based average configuration and computing the resulting Procrustes residuals [44]. To partial out these two factors from attractiveness, which were not of interest for the present research question, we fitted a model predicting attractiveness through averageness and age and used the resulting residuals as the dependent variable in the mediation analysis (hereafter referred to as *residualized attractiveness*). Consequently, neither averageness nor age could explain variance in residualized attractiveness. Notably, SShD and averageness are uncorrelated (also in our data: $r = -0.03$), such that residualizing one does not undermine the interpretability of the other.

To maximize data coverage, we additionally computed the same morphometric measures for studies that used facial databases as stimuli. Specifically, we collected the corresponding facial images, reconstructed the original stimulus identifiers as they appear in the respective databases and subjected the images to a morphometric analysis based on a deep neural network [46], which derives objective indices of facial femininity–masculinity and averageness equivalent to landmark-based measures. Stimulus age was obtained from the lookup tables typically provided with the respective facial databases. By combining datasets with provided morphometric measures and our self-computed equivalent measures, we analysed a dataset comprising more than one million observations.

Morphometric analyses were conducted using individual-level attractiveness ratings, with random intercepts for stimuli, study and raters to account for repeated measurements and clustering. Model assumptions were evaluated using visual inspection of residual and *Q–Q* plots, which indicated approximately normally distributed residuals with minor deviations in the tails and no substantial heteroscedasticity. Before conducting the mediation analysis, we confirmed that all three measures (SShD, averageness, age) were robust predictors of attractiveness and each hierarchically improved model fit (electronic supplementary material, table S6a and S6b).

All analyses were conducted in R, v. 4.4.1 [47], using the *lme4* package, v. 1.1.35.5 [48]. Parameter estimates were obtained via REML, and *p*-values were calculated using the *lmerTest* package [49]. Profile likelihood 95% confidence intervals (CIs), shown in square brackets following each parameter, were computed using the *confint* function in lme4. Effect sizes are reported as

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

Westfall's *d* measures, calculated as the fixed effect estimate divided by the square root of the pooled random-effects variance [50]. Estimated marginal means (EMMs) were computed with *emmeans*, v. 1.10.3 [51], and used for visualization, with 95% CIs displayed as error bars. Random-effects meta-analyses were conducted using *metafor* [52].

3. Results

(a) Meta-analysis

A random-effects meta-analysis across all available datasets revealed a robust GAP favouring female faces. Across 59 independent GAP estimates from 52 studies, the pooled effect size was 0.36, 95% CI [0.29, 0.42] (figure 2). The overall effect was highly significant ($z = 10.49$, $p < 0.0001$), and the majority of estimates were positive. Between-study heterogeneity was substantial ($\tau^2 = 0.05$, $I^2 = 89.86\%$, $Q_{58} = 439.096$, $p < 0.0001$), justifying the use of a random-effects model.

(b) Pooled interaction analysis (FaceSex × RaterSex)

To examine how the GAP varied as a function of RaterSex, we analysed all datasets that provided this information using a pooled interaction model. This revealed three significant findings (figure 3*a*; electronic supplementary material, table S3, left). First, there was a main effect of FaceSex, indicating that male faces generally received lower attractiveness ratings than female faces ($\beta = -0.377$ [−0.405, −0.349], $t = -26.56$, $p < 0.001$, $d = -0.53$), in line with the meta-analytic results. Second, we observed a main effect for RaterSex, with male raters giving lower attractiveness ratings than female raters ($\beta = -0.133$ [−0.149, −0.117], $t = -15.84$, $p < 0.001$, $d = -0.18$). Depending on the reference group, this bias can be described as a *male stringency bias* and mirrors the recently reported gender rating gap (GRG) in a large-scale analysis of online reviews [99], where men consistently provided more critical evaluations than women. Third, there was a significant FaceSex × RaterSex interaction, indicating that the magnitude of the GAP depended on the rater's own sex ($\beta = 0.162$ [0.137, 0.187], $t = 12.75$, $p < 0.001$, $d = 0.22$). Specifically, women showed a larger difference in their ratings of male versus female faces than men, leading to a more pronounced GAP (figure 3*b*). Modelling the raw data yielded identical result patterns (electronic supplementary material, figure S2, table S3, right side), confirming that the aggregation procedure used to maximize data coverage did not affect the findings.

(c) Pooled stability analyses

Given the heterogeneity observed in the meta-analysis, we next examined the stability of the GAP across a battery of rater- and stimulus-level moderators using pooled individual-level analyses. Across all models, FaceSex remained a highly robust main effect, with large test statistics in every analysis (all *F*-values >100, all *p*-values < 0.001), indicating a reliable GAP. Interactions between FaceSex and the respective moderator were significant for most stimulus- and rater-related characteristics, indicating systematic variation in the magnitude of the GAP across contexts. Notable exceptions were rater age and sampling context, for which no moderation was observed, and stimulus age, for which a statistically significant but modest interaction was found. Full model specifications, estimates and inferential statistics for all stability analyses are reported in electronic supplementary material, table S5a–c.

An overview of the results for the three stimulus-level moderators, the three rater-level moderators, and sampling context is shown in figure 4 (panels a–g). Across all moderator levels, the GAP was positive. Notable exceptions were observed for Black faces and elderly faces, for which the estimated GAP was close to zero. Although the magnitude of the GAP varied across moderator levels, no condition showed a significant reversal of the effect. As a complementary analysis, we examined sex differences in self-rated attractiveness (figure 4*h*). In contrast to evaluations of facial stimuli, no corresponding GAP emerged when participants evaluated their own attractiveness.

(d) Morphometric mediation analysis

We examined whether the relationship between FaceSex and residualized attractiveness was mediated by SShD, while accounting for RaterSex. Figure 5 summarizes the mediation results using three complementary panels: the total effect of FaceSex (a), the direct effect after adjusting for SShD (b) and the resulting attenuation of the GAP (c).

In the total-effect model (path c in mediation nomenclature; figure 5*a*), male faces were rated as less attractive than female faces for both female raters ($\beta = -0.336$ [−0.377, −0.295], $z = -16.16$, $p < 0.001$) and male raters ($\beta = -0.26$ [−0.301, −0.219], $z = -12.44$, $p < 0.001$), consistent with the interaction analysis performed on the pooled dataset with non-residualized attractiveness (figure 3*a*). FaceSex strongly predicted SShD (path a; $\beta = 1.363$, $t = 45.53$, $p < 0.001$ [1.304, 1.422], $d = 1.39$), confirming that the metric effectively captured femininity/masculinity (electronic supplementary material, figure S4a, table S7). SShD was, in turn, a robust predictor of residualized attractiveness when controlling for RaterSex (path b; $\beta = -0.157$ [−0.172, −0.115], $t = -15.161$, $p < 0.001$, $d = -0.2$; electronic supplementary material, figure S4b and table S7). Together, these results satisfied the conditions for a mediation analysis.

After including SShD as a mediator, the direct effect of FaceSex on residualized attractiveness (path c′) was substantially reduced but remained significant (female raters: $\beta = -0.224$ [−0.283, −0.165], $z = -7.438$, $p < 0.001$; male raters: $\beta = -0.148$ [−0.207, −0.089], $z = -4.91$, $p < 0.001$; figure 5*b*), indicating partial mediation. Quantifying this attenuation (figure 5*c*), SShD accounted for 33.25% of the total GAP among female raters and 42.98% among male raters, indicating a larger relative contribution for male

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

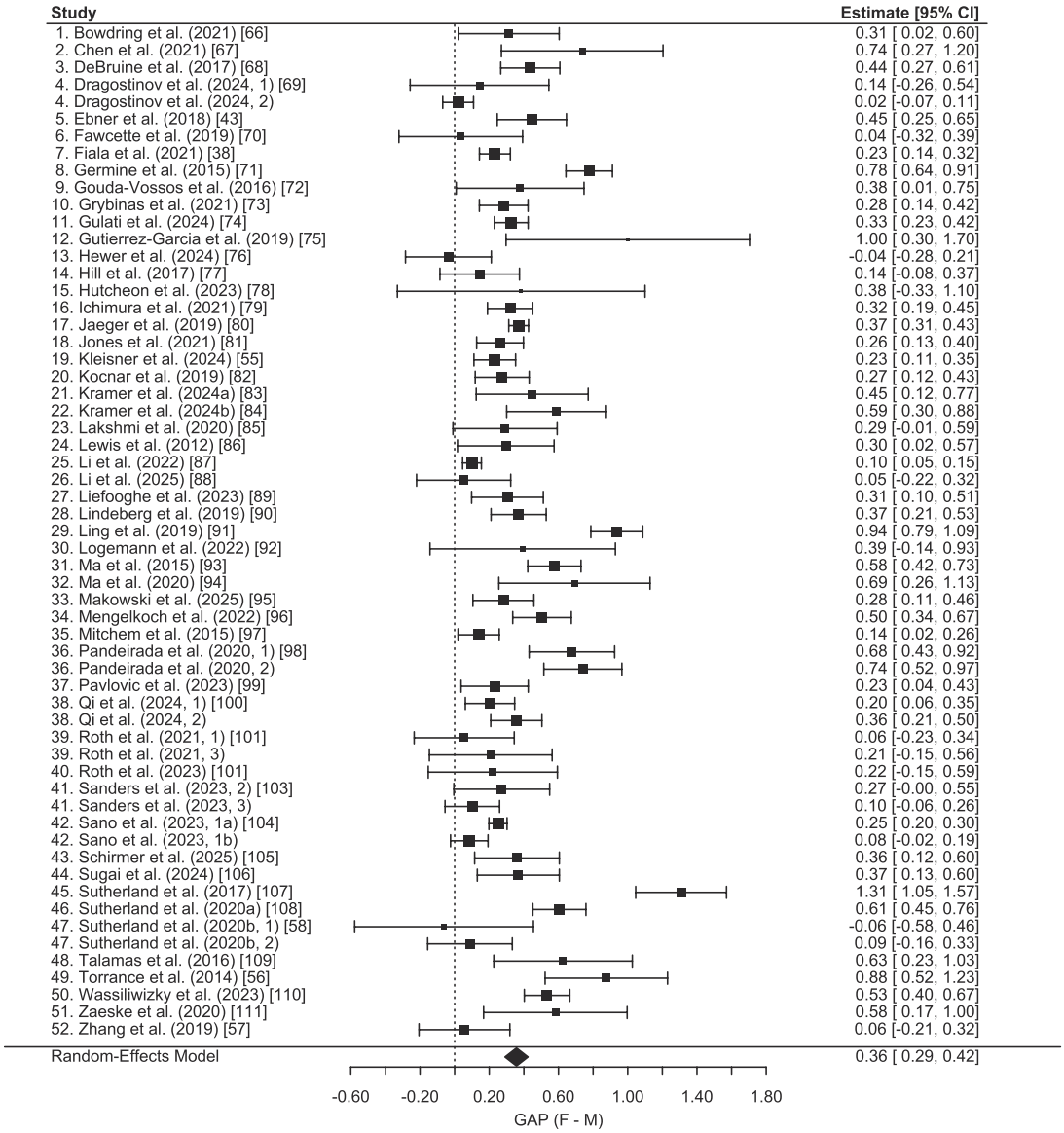

Figure 2. Meta-analysis of the GAP. Forest plot of study-specific GAP estimates (F – M) with 95% CIs. Numbers in parentheses after the year denote distinct studies within multi-study papers. Square sizes of individual GAP estimates reflect inverse-variance weight. The diamond indicates the pooled random-effects estimate with its 95% CI. [53–98].

raters despite a smaller overall GAP. Bootstrap-based confidence intervals based on 1000 resamples confirmed the robustness of this attenuation for both female raters (−0.111, 95% CI [−0.147, −0.101]) and male raters (−0.111, 95% CI [−0.147, −0.101]).

4. Discussion

Our analyses offer robust empirical confirmation of a long-speculated phenomenon: female faces are rated as more attractive than male faces, with an average advantage of 0.36 s.d. This GAP is consistently observed across a wide range of rater and stimulus characteristics and sampling environments, reflecting a systematic pattern in aesthetic judgement, which appears to be confined to evaluations of others, as women and men do not differ in their self-reported attractiveness judgements (figure 4*h*). Notably, we cannot safely conclude that sexual selection is responsible for the GAP, as this would risk overstretching the available evidence. Our analyses capture attractiveness judgements, which constitute only the first step in the causal chain linking perception to evolutionary outcomes (i.e. judgement, behaviour, reproduction and ultimately evolution), and therefore do not permit inferences about selection processes acting on reproductive success. Moreover, although the stability of the GAP across cultures constrains purely sociocultural explanations, such factors may nonetheless contribute to, amplify or stabilize the GAP, as discussed later.

Before turning to theoretical interpretations, however, we first evaluate whether methodological artefacts could explain the GAP. One potential concern is selection or self-selection bias in the construction of the stimulus sets. Could it be that more attractive women are systematically overrepresented among facial stimuli and therefore do not accurately reflect the underlying population? For example, in many studies on facial attractiveness, individuals volunteer as models for stimulus sets. If attractive women were disproportionately likely to participate, whereas men did not show comparable self-selection, this could create a spurious GAP.

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

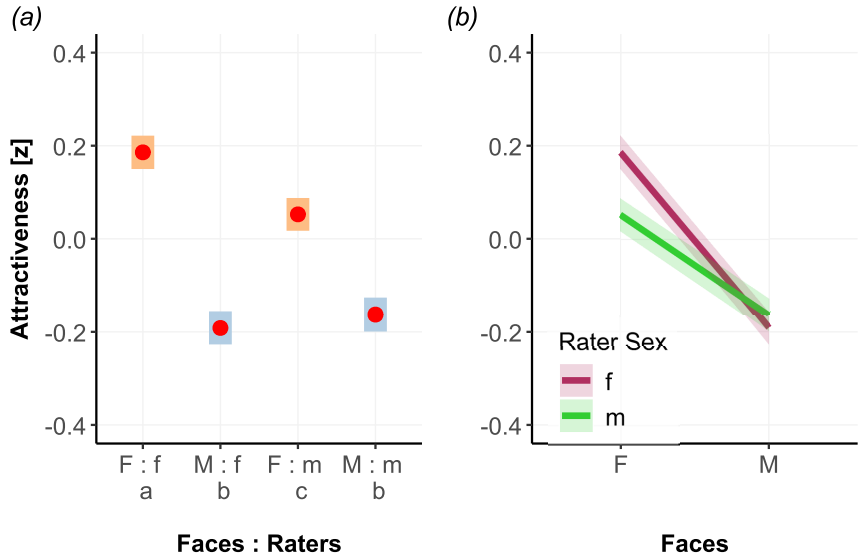

Figure 3. Pooled interaction analysis (FaceSex × RaterSex). (*a*) EMMs of attractiveness by FaceSex (F, M) and RaterSex (f, m). For example, 'F : m' denotes female faces rated by male raters. Note that: (1) male faces (blue) generally receive lower attractiveness ratings (GAP), (2) male raters assign lower ratings overall (GRG), and (3) the discrepancy between female and male faces is more pronounced among female raters. Within each panel, conditions marked with different subscript letters differ significantly ($p < 0.001$). (*b*) Interaction plot showing predicted attractiveness scores by FaceSex and RaterSex. Both lines slope downward from female to male faces, reflecting the overall GAP; the steeper slope for female raters indicates stronger differentiation between male and female faces. Shaded ribbons indicate 95% CIs.

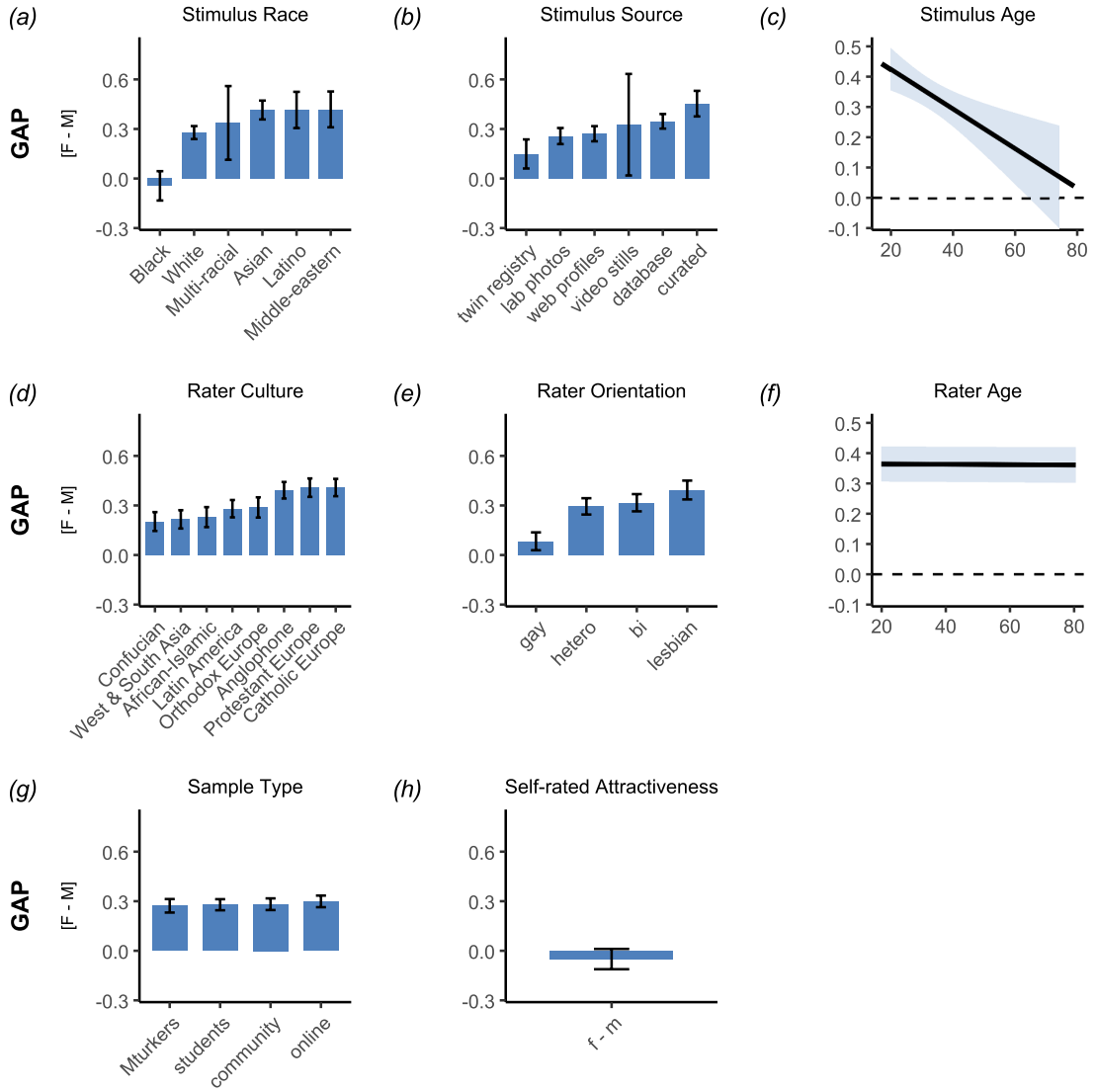

Figure 4. Stability of the GAP. Panels (*a–c*) reflect stimulus-level moderators. Panels (*d–g*) reflect rater-level moderators, including sampling context. Panel (*h*) shows the sex difference in self-rated attractiveness (f – m). Bars depict EMMs of the GAP, expressed in *z*-scored attractiveness units. Error bars and shaded ribbons indicate 95% CIs. Wider confidence intervals reflect smaller representation of specific moderator levels in the dataset.

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

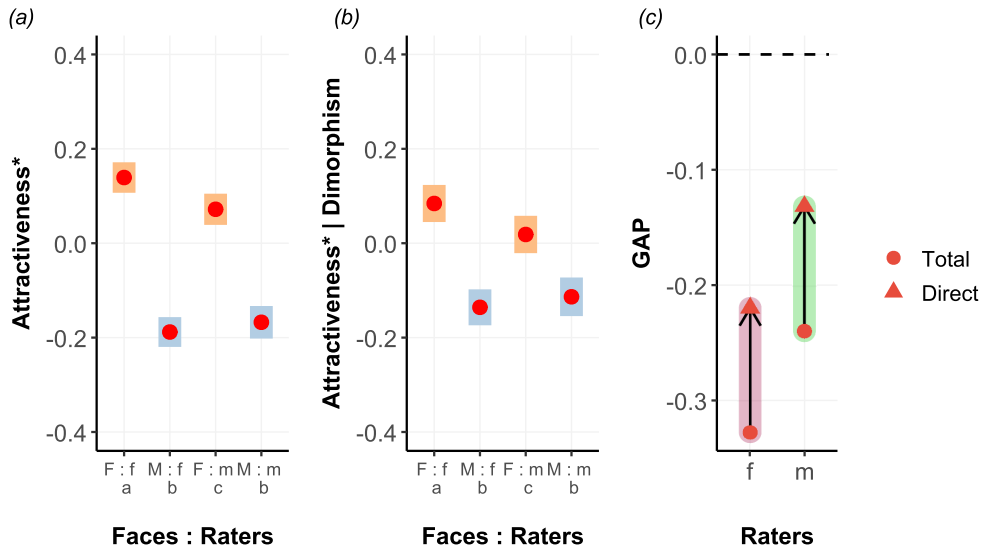

Figure 5. Morphometric mediation of the GAP by SShD. Multilevel mediation models testing whether the effect of FaceSex on attractiveness is mediated by SShD, while controlling for RaterSex. Attractiveness was residualized for averageness and age, as indicated by the asterisk. (*a*) Total effect of FaceSex on attractiveness* (in *z*-units), illustrating the GAP prior to accounting for SShD. (*b*) Direct effect of FaceSex after adjusting for SShD. Bars depict EMMs; error bars indicate 95% CIs. Conditions marked with different subscript letters differ significantly ($p < 0.001$). (*c*) Attenuation of the GAP attributable to SShD, quantified as the difference between the total and direct effects; arrows visualize the reduction in the GAP after adjusting for SShD for female and male raters.

To address this possibility, we examined more closely the distribution of attractiveness scores in our harmonized dataset (figure 1). If a self-selection bias were present, one would expect the lower end of the female distribution to be truncated or the male distribution to be inflated. Neither pattern is observed. Instead, both distributions show continuous, bell-shaped curves with substantial representation across the full attractiveness range, as also indicated by the rug plots. *Q*–*Q* plots further suggest that the distributions closely approximate normality around the central range. An inflation in the upper tail for male faces suggests that, if anything, highly attractive men are somewhat overrepresented (assuming normality of attractiveness in the population). Interestingly, ratings for male faces are also more sharply peaked, indicating a tighter clustering and convergence, while ratings for female faces are more variable. Levene's test confirmed a significant variance difference ($F_{1,15373} = 5.18$, $p < 0.001$), with greater variability among female faces (s.d.= 0.807) than male ones (s.d. = 0.762). These findings suggest that the GAP reflects systematic properties of aesthetic judgement rather than artefacts of sampling or stimulus construction. Consistent with this interpretation, formal moderation analyses revealed a positive GAP across all stimulus sources (figure 4*b*).

A related concern might be grooming. Women generally invest more in grooming than men [100], which might bias attractiveness ratings. However, most studies use standardized photo protocols requiring the removal of makeup and accessories (e.g. [43,101–103]). Although some long-term practices like eyebrow shaping or skincare may not be fully controlled, their impact is likely minimal.

All of this raises the central question: what explains the GAP? While evolutionary frameworks have traditionally shaped the theorizing about facial attractiveness, typically emphasizing opposite-sex attraction, mate selection and reproductive success, several present findings are difficult to reconcile with explanations based solely on sexual attraction. Notably, the GAP persists in female ratings, holds across sexual orientations and is stable across rater age. Cultural influences may also play a role, since female beauty is idealized in many societies and reinforced through media and advertising. Such internalized standards may foster unconscious biases and inflate explicit ratings by blending aesthetic responses to facial features with culturally shaped expectations. Together, these patterns suggest that biological predispositions alone are unlikely to provide a complete explanation and that broader sociocultural influences shaping aesthetic judgement warrant consideration.

A particularly intriguing aspect of the GAP is its modulation by rater sex: the GAP is larger for female raters. This pattern highlights an asymmetry in rating behaviour, especially between same-sex versus opposite-sex evaluations. While men, as expected, rate female faces more favourably than male faces, women show an even stronger preference for female faces. As a result, both sexes rate male faces similarly, but diverge markedly in their judgements of female faces, with women being notably more generous toward their same-sex peers, even after adjusting for facial structure (figure 5*b*). These findings raise two key questions: Why does female rating behaviour not align with a simple sexual-attraction account? And why are women so generous in their evaluations of other women?

One explanation for the first question is that, compared to men, women rely less exclusively on visual cues when assessing male attractiveness. Their evaluations tend to incorporate additional trait inferences, such as perceived dominance or trustworthiness [104], and are further shaped by broader contextual factors, including resource scarcity and exposure to violent settings [105,106]. Social norms and self-presentation concerns may further play a role. In many societies, women are discouraged from openly expressing desire, especially in settings where responses are recorded and scrutinized. This may prompt more reserved or socially constrained female evaluations of opposite-sex faces. These factors are not mutually exclusive; they may interact and jointly contribute to the relatively lower ratings women assign to male faces. More broadly, our findings pertain specifically to facial attractiveness and do not capture other mate preferences, such as body shape, social status or intelligence.

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

The female-to-female generosity may be better explained by sociocultural norms. Across many societies, girls and women are socialized from an early age to place high value on physical appearance and are more frequently exposed than men to media that reinforce idealized standards of female beauty [107,108]. As a result, women may be culturally permitted, or even encouraged, to evaluate one another's appearance more favourably. This may also reflect a sense of peer solidarity and mutual recognition of the shared pressures women face regarding physical appearance.

In any case, the modulation of the GAP by rater sex underscores that same-sex and opposite-sex ratings differ fundamentally in nature. While male ratings appear to align with romantic, heterosexual attraction, this link is less clear for female raters. Technically, this challenges the principle of measurement invariance: attractiveness ratings may tap different underlying constructs depending on the rater's sex. This complicates interpretation and highlights the need for future research. One potential improvement could be to rephrase rating prompts from 'How attractive is this face?' to 'How much do you feel physically attracted to this person?'.

Our results show that men rate facial attractiveness more stringently than women, although this GRG is modest (d = −0.18) compared to the GAP (d = −0.52). This finding aligns with previous reports of male stringency across various domains, including life satisfaction [109], personality judgements [110] and online reviews [99]. The latter is particularly relevant given its similarity to aesthetic evaluation. A recent large-scale study found that men consistently gave lower ratings than women on platforms like Amazon, IMDb and TripAdvisor. Notably, the GRG in our study (d = −0.18) exceeds those reported for products (d = −0.05) and travel destinations (d = −0.01) and matches rather those for movies (d = −0.14) and music (d = −0.16). To assess the consistency of the GRG, we conducted stability analyses using the same pooled moderation framework as for the GAP. In contrast to the GAP, the GRG showed pronounced context dependence, with substantial variation across both stimulus- and rater-related characteristics (electronic supplementary material, figure S5), consistent with previous evidence for the context sensitivity of the GRG [99]. While the magnitude, and in some cases even the direction, of the effect varied across most moderator levels, one notable and consistent pattern emerged: unlike the GAP, the GRG was stable across face ages but increased linearly with rater age, indicating progressively greater stringency among male raters relative to female raters with increasing age.

Our morphometric mediation analysis revealed that structural facial dimorphism accounts for a substantial, but incomplete, portion of the GAP. Importantly, the contribution of SShD differs by rater sex. While the absolute attenuation of the GAP after adjusting for SShD was identical for female and male raters (−0.14 z-units), its *relative* explanatory power was larger for male raters (43%) than for female raters (33%), because the GAP among male raters was smaller to begin with. Importantly, adjusting for SShD led female faces to regress downward towards zero and male faces to regress upward toward zero, indicating that the GAP shrinks from both sides. At the same time, a considerable portion of the GAP remains after controlling for facial shape, particularly among female raters (although non-shape components of sexual dimorphism, such as skin texture, colour and facial hair, remain uncontrolled). The remaining GAP after controlling for SShD is driven primarily by women still assigning higher ratings to female than male faces ('F : f' versus 'M : f' in figure 5*b*), indicating persistent peer generosity. Likewise, male raters remain stricter in judging female faces ('F : f' versus 'F : m'), with no comparable stringency bias observed for male faces ('M : f' versus 'M : m'), implying that their evaluative criteria differ by the sex of the face being judged.

Notably, the only clear exception to the otherwise robust GAP was observed for faces of African descent. Additional rater-sex-specific analyses showed that male raters exhibited no reliable GAP for Black faces (GAP = 0.02 s.d., p = 0.74), whereas female raters rated Black male faces approximately 0.25 s.d. lower than Black female faces (p < 0.001), revealing a pronounced rater-sex asymmetry. Including SShD as a mediator largely eliminated this GAP among female raters (residual GAP = 0.05 s.d., p = 0.58), indicating strong attenuation. Previous findings further suggest that certain African populations, such as Cameroonians, exhibit lower facial SShD compared to European and Latin American populations [111]. Reduced shape dimorphism weakens structural cues that differentiate male and female faces in terms of attractiveness. In such contexts, non-shape cues, especially skin colour, play a more prominent role in attractiveness judgements [28]. Further research is needed to disentangle these effects and to more fully incorporate culturally specific beauty standards across African populations.

What have we learned from this study? We provide extensive cross-cultural empirical evidence for a robust GAP, observed across diverse cultural, racial, personal and methodological contexts, underscoring that facial attractiveness judgements extend well beyond heterosexual mate choice. The GAP's modulation by rater sex—driven by female raters' generosity in same-sex evaluations and unexpectedly low ratings of male faces—adds complexity to how gender shapes aesthetic assessments. These patterns are consistent with the influence of internalized cultural norms linking femininity with beauty, alongside a sense of peer solidarity among women. Morphometric analyses further indicate that sexual shape dimorphism provides a common structural basis for facial attractiveness judgements, accounting for a substantial portion of the GAP, with the explanatory contribution of shape dimorphism depending on rater sex.

Together, these findings advance our understanding of how sex, facial structure and broader sociocultural context jointly shape perceptions of attractiveness. They hold important implications for research in evolutionary psychology, empirical aesthetics, anthropology, sociology and gender studies, particularly with respect to study design, rater selection and interpretation of attractiveness ratings. Finally, our results call for further investigation into the cultural factors that shape evaluations of human attractiveness.

Ethics. All methods and procedures were approved by the Ethics Committee of the Max Planck Society (approval number 2017_12).

Data accessibility. All scripts and data required to reproduce the results, figures and tables are available on the Open Science Framework (https://osf.io/98vrd/), including the automated data-construction script used to generate the harmonized dataset from the original sources, as well as our self-computed morphometric measures for all facial databases included in this article.

Supplementary material is available online [112].

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

Declaration of AI use. ChatGPT was used exclusively to proofread our written text, using the prompt '[proofread:]'.

Authors' contributions. E.W.: conceptualization, data curation, formal analysis, investigation, visualization, writing—original draft, writing—review and editing; B.P.Z.: writing—review and editing; K.K.: writing—review and editing; F.U.: writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. K.K. was supported by the Czech Science Foundation (GA24-11735S).

Acknowledgements. We thank Brenda Edosomwan for assistance with the literature search, Rakat Murshed for support with the analyses, and the corresponding authors who provided their raw data.

References

1. Chatterjee A, Thomas A, Smith SE, Aguirre GK. 2009 The neural response to facial attractiveness. *Neuropsychology* **23**, 135–143. (doi:10.1037/a0014430)
2. Kim H, Adolphs R, O'Doherty JP, Shimojo S. 2007 Temporal isolation of neural processes underlying face preference decisions. *Proc. Natl Acad. Sci. USA* **104**, 18253–18258. (doi:10.1073/pnas.0703101104)
3. Langlois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. 2000 Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* **126**, 390–423. (doi:10.1037/0033-2909.126.3.390)
4. Wassiliwizky E, Menninghaus W. 2021 Why and how should cognitive science care about aesthetics? *Trends Cogn. Sci.* **25**, 437–449. (doi:10.1016/j.tics.2021.03.008)
5. Gangestad SW, Scheyd GJ. 2005 The evolution of human physical attractiveness. *Annu. Rev. Anthropol.* **34**, 523–548. (doi:10.1146/annurev.anthro.33.070203.143733)
6. Little AC, Jones BC, DeBruine LM. 2011 Facial attractiveness: evolutionary based research. *Phil. Trans. R. Soc. B.* **366**, 1638–1659. (doi:10.1098/rstb.2010.0404)
7. Rhodes G. 2006 The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* **57**, 199–226. (doi:10.1146/annurev.psych.57.102904.190208)
8. Kościński K. 2007 Facial attractiveness: general patterns of facial preferences. *Anthropol. Rev.* **70**, 45–79. (doi:10.2478/v10044-008-0001-9)
9. Polak M. 2003 *Developmental instability: causes and consequences*. Oxford, UK: Oxford University Press.
10. Thornhill R, Gangestad SW. 1993 Human facial beauty: averageness, symmetry, and parasite resistance. *Hum. Nat.* **4**, 237–269. (doi:10.1007/BF02692201)
11. Buss DM. 1989 Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures. *Behav. Brain Sci.* **12**, 1–14. (doi:10.1017/s0140525x00023992)
12. Kokko H, Jennions MD. 2008 Parental investment, sexual selection and sex ratios. *J. Evol. Biol.* **21**, 919–948. (doi:10.1111/j.1420-9101.2008.01540.x)
13. Prum RO. 2017 *The evolution of beauty: how Darwin's forgotten theory of mate choice shapes the animal world*. New York, NY: Doubleday.
14. Ryan M. 2018 *A taste for the beautiful*. Princeton, NJ: Princeton University Press. (doi:10.1515/9781400889150)
15. Darwin C. 1874 *The descent of man: and selection in relation to sex*, 2nd edition. New York, NY: Merrill and Baker. (doi:10.5962/bhl.title.54341)
16. Dawkins R. 1976 *The selfish gene*. Oxford, UK: Oxford University Press.
17. Symons D. 1979 *The evolution of human sexuality*. Oxford, UK: Oxford University Press.
18. Gottschall J. 2007 Greater emphasis on female attractiveness in *Homo sapiens:* a revised solution to an old evolutionary riddle. *Evol. Psychol.* **5**, 347–357. (doi:10.1177/147470490700500209)
19. Archer J, Mehdikhani M. 2003 Variability among males in sexually selected attributes. *Rev. Gen. Psychol.* **7**, 219–236. (doi:10.1037/1089-2680.7.3.219)
20. Kleisner K *et al*. 2024 Distinctiveness and femininity, rather than symmetry and masculinity, affect facial attractiveness across the world. *Evol. Hum. Behav.* **45**, 82–90. (doi:10.1016/j.evolhumbehav.2023.08.003)
21. Kraaijeveld K, Kraaijeveld-Smit FJL, Komdeur J. 2007 The evolution of mutual ornamentation. *Anim. Behav.* **74**, 657–677. (doi:10.1016/j.anbehav.2006.12.027)
22. Swaddle JP, Reierson GW. 2002 Testosterone increases perceived dominance but not attractiveness in human males. *Proc. R. Soc. B.* **269**, 2285–2289. (doi:10.1098/rspb.2002.2165)
23. Lorenz K. 1943 Die angeborenen Formen möglicher Erfahrung. *Z. Für Tierpsychol.* **5**, 1943. (doi:10.1111/j.1439-0310.1943.tb00655.x)
24. Keating CF, Randall DW, Kendrick T, Gutshall KA. 2003 Do babyfaced adults receive more help? The (cross-cultural) case of the lost resume. *J. Nonverbal Behav.* **27**, 89–109. (doi:10.1023/A:1023962425692)
25. Zheng W, Luo T, Hu CP, Peng K. 2018 Glued to which face? Attentional priority effect of female babyface and male mature face. *Front. Psychol.* **9**, 286. (doi:10.3389/fpsyg.2018.00286)
26. Kleisner K *et al*. 2025 Sexual dimorphism in facial morphology and skin color in the horn of Africa. *Sex. Dev.* 1–21. (doi:10.1159/000549901)
27. Kleisner K, Kočnar T, Tureček P, Stella D, Akoko RM, Třebický V, Havlíček J. 2017 African and European perception of African female attractiveness. *Evol. Hum. Behav.* **38**, 744–755. (doi:10.1016/j.evolhumbehav.2017.07.002)
28. Pokorný Š, Pavlovič O, Kleisner K. 2024 Sexual dimorphism: the interrelation of shape and color. *Arch. Sex. Behav.* **53**, 3255–3265. (doi:10.1007/s10508-024-02738-9)
29. Frost P. 2022 European hair, eye, and skin color: solving the puzzle. *Evol. Hum. Behav.* **27**, 85–103. (doi:10.1016/j.evolhumbehav.2005.07.002)
30. Wagatsuma H. 1967 The social perception of skin color in Japan. *Daedalus* **96**, 407–443.
31. Page MJ *et al*. 2021 The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71. (doi:10.1136/bmj.n71)
32. Inglehart R, Welzel C. 2005 *Modernization, cultural change, and democracy: the human development sequence*. Cambridge, UK: Cambridge University Press.
33. Holzleitner IJ, Hunter DW, Tiddeman BP, Seck A, Re DE, Perrett DI. 2014 Men's facial masculinity: when (body) size matters. *Perception* **43**, 1191–1202. (doi:10.1068/p7569)
34. Lee AJ, Mitchem DG, Wright MJ, Martin NG, Keller MC, Zietsch BP. 2014 Genetic factors that increase male facial masculinity decrease facial attractiveness of female relatives. *Psychol. Sci* **25**, 476–484. (doi:10.1177/0956797613500713)
35. Penton-Voak IS, Jones BC, Little AC, Baker S, Tiddeman B, Burt DM, Perrett DI. 2001 Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proc. R. Soc. B* **268**, 1617–1623. (doi:10.1098/rspb.2001.1703)
36. Perrett DI, Lee KJ, Penton-Voak I, Rowland D, Yoshikawa S, Burt DM, Henzi SP, Castles DL, Akamatsu S. 1998 Effects of sexual dimorphism on facial attractiveness. *Nature* **394**, 884–887. (doi:10.1038/29772)
37. Rhodes G, Hickford C, Jeffery L. 2000 Sex-typicality and attractiveness: are supermale and superfemale faces super-attractive? *Br. J. Psychol.* **91**, 125–140. (doi:10.1348/000712600161718)
38. Fiala V. 2021 Facial attractiveness and preference of sexual dimorphism: a comparison across five populations. *Evol. Hum. Sci.* **3**, 38. (doi:10.1017/ehs.2021.38)
39. DeBruine LM, Jones BC, Crawford JR, Welling LLM, Little AC. 2010 The health of a nation predicts their mate preferences: cross-cultural variation in women's preferences for masculinized male faces. *Proc. R. Soc. B* **277**, 2405–2410. (doi:10.1098/rspb.2009.2184)

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

40. Lee AJ, McGuire NKJ. 2023 Women's preferences for masculinity in male faces are predicted by material scarcity, but not time or psychological scarcity. *Evol. Psychol.* **21**, 14747049231175073. (doi:10.1177/14747049231175073)
41. Marcinkowska UM *et al*. 2019 Women's preferences for men's facial masculinity are strongest under favorable ecological conditions. *Sci. Rep.* **9**, 10. (doi:10.1038/s41598-019-51464-3)
42. Deffenbacher KA, Vetter T, Johanson J, O'Toole AJ. 1998 Facial aging, attractiveness, and distinctiveness. *Perception* **27**, 1233–1243. (doi:10.1068/p271233)
43. Ebner NC, Luedicke J, Voelkle MC, Riediger M, Lin T, Lindenberger U. 2018 An adult developmental approach to perceived facial attractiveness and distinctiveness. *Front. Psychol.* **9**, 561. (doi:10.3389/fpsyg.2018.00561)
44. Lee AJ. 2016 Facial averageness and genetic quality: testing heritability, genetic correlation with attractiveness, and the paternal age effect. *Evol. Hum. Behav.* **37**, 61–66. (doi:10.1016/j.evolhumbehav.2015.08.002)
45. Rhodes G, Yoshikawa S, Clark A, Lee K, McKay R, Akamatsu S. 2001 Attractiveness of facial averageness and symmetry in non-western cultures: in search of biologically based standards of beauty. *Perception* **30**, 611–625. (doi:10.1068/p3123)
46. Zhao AAZ, Zietsch BP. 2024 Deep neural networks generate facial metrics that overcome limitations of previous methods and predict in-person attraction. *Evol. Hum. Behav.* **45**, 106632. (doi:10.1016/j.evolhumbehav.2024.106632)
47. R Core Team. 2024 *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. See https://www.R-project.org/.
48. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 01. (doi:10.18637/jss.v067.i01)
49. Kuznetsova A, Brockhoff PB, Christensen RHB. 2017 lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82** 1–26. . (doi:10.18637/jss.v082.i13)
50. Westfall J, Kenny DA, Judd CM. 2014 Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *J. Exp. Psychol. Gen.* **143**, 2020–2045. (doi:10.1037/xge0000014)
51. Lenth RV, Bolker B, Buerkner P. 2023 *R package emmeans: estimated marginal means*. See https://cran.r-project.org/web/packages/emmeans/index.html.
52. Viechtbauer W. 2010 Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48. (doi:10.18637/jss.v036.i03)
53. Bowdring MA, Sayette MA, Girard JM, Woods WC. 2021 In the eye of the beholder: a comprehensive analysis of stimulus type, perceiver, and target in physical attractiveness perceptions. *J. Nonverbal Behav.* **45**, 241–259. (doi:10.1007/s10919-021-00371-6)
54. Chen JM, Norman JB, Nam Y. 2021 Broadening the stimulus set: Introducing the American Multiracial Faces Database. *Behav. Res. Methods* **53**, 371–389. (doi:10.3758/s13428-020-01455-4)
55. DeBruine LM, Jones BC. 2017 Face Research Lab London Set. figshare. (doi:10.6084/m9.figshare.5047666)
56. Dragostinov Y, Booth T. 2025 The (un)attractiveness of Dark Triad personalities: assessing fictitious characters for short- and long-term relationships. *J. Pers.* **93**, 1097–1111. (doi:10.1111/jopy.12994)
57. Fawcett TW, Ewans J, Lawrence A, Radford AN. 2019 Attractiveness is positively related to World Cup performance in male, but not female, biathletes. *Behav. Ecol.* **30**, 1436–1442. (doi:10.1093/beheco/arz097)
58. Germine L *et al*. 2015 Individual aesthetic preferences for faces are shaped mostly by environments, not genes. *Curr. Biol.* **25**, 2684–2689. (doi:10.1016/j.cub.2015.08.048)
59. Gouda-Vossos A, Dixson BJ, Brooks RC. 2016 Sexual conflict and gender gap effects: associations between social context and sex on rated attractiveness and economic status. *PLoS One* **11**, e0146269. (doi:10.1371/journal.pone.0146269)
60. Grybinas D, Dobbins IG. 2021 Attempted recall of biographical information influences face attractiveness. *Psychon. Bull. Rev.* **28**, 953–961. (doi:10.3758/s13423-021-01877-1)
61. Gulati A, Martínez-Garcia M, Fernández D, Lozano MA, Lepri B, Oliver N. 2024 What is beautiful is still good: the attractiveness halo effect in the era of beauty filters. *R. Soc. Open Sci.* **11**, 240882. (doi:10.1098/rsos.240882)
62. Gutiérrez-García A, Beltrán D, Calvo MG. 2019 Facial attractiveness impressions precede trustworthiness inferences: lower detection thresholds and faster decision latencies. *Cogn. Emot.* **33**, 378–385. (doi:10.1080/02699931.2018.1444583)
63. Hewer E, Lewis MB. 2024 Unveiling why race does not affect the mask effect on attractiveness: But gender and expression do. *Cogn. Res. Princ. Implic.* **9**, 7. (doi:10.1186/s41235-024-00472-2)
64. Hill AK *et al*. 2017 Are there vocal cues to human developmental stability? Relationships between facial fluctuating asymmetry and voice attractiveness. *Evol. Hum. Behav.* **38**, 249–258. (doi:10.1016/j.evolhumbehav.2016.10.008)
65. Hutcheon TG, McMahon H, Retzloff C, Samat A, Tinker C. 2023 The attractiveness of face cues does not modulate the gaze cuing effect. *Vis. Cogn.* **31**, 655–670. (doi:10.1080/13506285.2024.2335585)
66. Ichimura F, Moriwaki M, Ariga A. 2021 Romantic bias in judging the attractiveness of faces from the back. *J. Nonverbal Behav.* **45**, 339–350. (doi:10.1007/s10919-021-00361-7)
67. Jaeger B, Sleegers WWA, Evans AM, Stel M, van Beest I. 2019 The effects of facial attractiveness and trustworthiness in online peer-to-peer markets. *J. Econ. Psychol.* **75**, 102125. (doi:10.1016/j.joep.2018.12.007)
68. Jones BC *et al*. 2021 To which world regions does the valence–dominance model of social perception apply? *Nat. Hum. Behav.* **5**, 159–169. (doi:10.1038/s41562-020-01007-2)
69. Kočnar T, Saribay SA, Kleisner K. 2019 Perceived attractiveness of Czech faces across 10 cultures: associations with sexual shape dimorphism, averageness, fluctuating asymmetry, and eye color. *PLoS One* **14**, e0225549. (doi:10.1371/journal.pone.0225549)
70. Kramer RSS, Ritchie KL, Flack TR, Mireku MO, Jones AL. 2024 The psychometrics of rating facial attractiveness using different response scales. *Perception* **53**, 645–660. (doi:10.1177/03010066241256221)
71. Kramer RSS, Koca Y, Mireku MO, Oriet C. 2024 Anchoring has little effect when forming first impressions of facial attractiveness. *Perception* **53**, 787–802. (doi:10.1177/03010066241284956)
72. Lakshmi A, Wittenbrink B, Correll J, Ma DS. 2021 The India face set: international and cultural boundaries impact face impressions and perceptions of category membership. *Front. Psychol.* **12**, 627678. (doi:10.3389/fpsyg.2021.627678)
73. Lewis MB. 2012 A facial attractiveness account of gender asymmetries in interracial marriage. *PLoS One* **7**, 31703. (doi:10.1371/journal.pone.0031703)
74. Li W, Zhu H, Zhao K, Zhu H, Wang X, He X. 2024 Good performance-high attractiveness effect: an empirical study on the association between athletes' rankings and their facial attractiveness. *Int. J. Sport Exerc. Psychol.* **22**, 1168–1193. (doi:10.1080/1612197X.2023.2181846)
75. Li J, Lee P, Rafiee Y, Jones BC, Shiramizu VKM. 2025 No evidence that people born to older parents show weaker preferences for younger adult faces. *Adapt. Hum. Behav. Physiol.* **11**, 9. (doi:10.1007/s40750-025-00263-8)
76. Liefooghe B, Oliveira M, Leisten LM, Hoogers E, Aarts H, Hortensius R. 2023 Are natural faces merely labelled as artificial trusted less? *Collabra: Psychol.* **9**. (doi:10.1525/collabra.73066)
77. Lindeberg S, Craig BM, Lipp OV. 2019 You look pretty happy: attractiveness moderates emotion perception. *Emotion* **19**, 1070–1080. (doi:10.1037/emo0000500)

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026

78. Ling L, Luo D, She G. 2019 Judging a book by its cover: the influence of physical attractiveness on the promotion of regional leaders. *J. Econ. Behav. Organ.* **158**, 1–14. (doi:10.1016/j.jebo.2019.01.005)
79. Logemann-Molnár Z, Veres-Székely A, Demetrovics Z, Logemann HA. 2022 Seeing attractive faces challenges inhibitory control, especially when mindful. *PLoS One* **17**, e0273913. (doi:10.1371/journal.pone.0273913)
80. Ma DS, Correll J, Wittenbrink B. 2015 The Chicago face database: a free stimulus set of faces and norming data. *Behav. Res. Methods* **47**, 1122–1135. (doi:10.3758/s13428-014-0532-5)
81. Ma DS, Kantner J, Wittenbrink B. 2021 Chicago face database: multiracial expansion. *Behav. Res. Methods* **53**, 1289–1300. (doi:10.3758/s13428-020-01482-5)
82. Makowski D, Te AS, Neves A, Kirk S, Liang NZ, Mavros P, Chen SHA. 2025 Too beautiful to be fake: attractive faces are less likely to be judged as artificially generated. *Acta Psychol.* **252**, 104670. (doi:10.1016/j.actpsy.2024.104670)
83. Mengelkoch S, Gassen J, Prokosch ML, Boehm GW, Hill SE. 2022 More than just a pretty face? The relationship between immune function and perceived facial attractiveness. *Proc. R. Soc. B* **289**, 20212476. (doi:10.1098/rspb.2021.2476)
84. Mitchem DG. 2015 No relationship between intelligence and facial attractiveness in a large, genetically informative sample. *Evol. Hum. Behav.* **36**, 240–247. (doi:10.1016/j.evolhumbehav.2015.01.004)
85. Pandeirada JNS, Fernandes NL, Vasconcelos M. 2020 Attractiveness of human faces: norms by sex, sexual orientation, age, relationship stability, and own attractiveness judgements. *Front. Psychol.* **11**, 419. (doi:10.3389/fpsyg.2020.00419)
86. Pavlovič O, Fiala V, Kleisner K. 2023 Congruence in European and Asian perception of Vietnamese facial attractiveness, averageness, symmetry, and sexual dimorphism. *Sci. Rep.* **13**, 13320. (doi:10.1038/s41598-023-40649-8)
87. Qi Y, Ying J. 2022 Gender biases in the accuracy of facial judgments: facial attractiveness and perceived socioeconomic status. *Front. Psychol.* **13**, 884888. (doi:10.3389/fpsyg.2022.884888)
88. Roth TS, Du X, Samara I, Kret ME. 2021 Attractiveness modulates attention, but does not enhance gaze cueing. *Evol. Behav. Sci.* **16**, 343–361. (doi:10.1037/ebs0000265)
89. Roth TS, Samara I, Perea-Garcia JO, Kret ME. 2023 Individual attractiveness preferences differentially modulate immediate and voluntary attention. *Sci. Rep.* **13**, 2147. (doi:10.1038/s41598-023-29240-5)
90. Sanders CA, Jenkins AT, King LA. 2023 Pretty, meaningful lives: physical attractiveness and experienced and perceived meaning in life. *J. Posit. Psychol.* **18**, 978–994. (doi:10.1080/17439760.2022.2155222)
91. Sano T, Kawabata H. 2023 A computational approach to investigating facial attractiveness factors using geometric morphometric analysis and deep learning. *Sci. Rep.* **13**, 19797. (doi:10.1038/s41598-023-44907-4)
92. Schirmer A, Franz M, Krismann L, Nöring V, Große M, Mahmut M, Croy I. 2025 Attraction in every sense: how looks, voice, movement and scent draw us to future lovers and friends. *Br. J. Psychol.* **116**, 684–701. (doi:10.1111/bjop.12787)
93. Sugai M, Yonemitsu F, Ariga A. 2024 Romantic bias in judging the attractiveness of faces wearing masks. *i-Perception* **15**, 20416695241287486. (doi:10.1177/20416695241287486)
94. Sutherland CAM, Young AW, Rhodes G. 2017 Facial first impressions from another angle: how social judgements are influenced by changeable and invariant facial properties. *Br. J. Psychol.* **108**, 397–415. (doi:10.1111/bjop.12206)
95. Sutherland CA, Rhodes G, Burton NS, Young AW. 2020 Do facial first impressions reflect a shared social reality? *Br. J. Psychol.* **111**, 215–232. (doi:10.1111/bjop.12406)
96. Talamas SN, Mavor KI, Perrett DI. 2016 Blinded by beauty: attractiveness bias and accurate perceptions of academic performance. *PLoS One* **11**, e0148284. (doi:10.1371/journal.pone.0148284)
97. Wassiliwizky E, Wontorra P, Ullén F. 2023 How being perceived to be an artist boosts feelings of attraction in others. *Sci. Rep.* **13**, 18747. (doi:10.1038/s41598-023-45493-1)
98. Zäske R, Skuk VG, Schweinberger SR. 2020 Attractiveness and distinctiveness between speakers' voices in naturalistic speech and their faces are uncorrelated. *R. Soc. Open Sci.* **7**, 201244. (doi:10.1098/rsos.201244)
99. Bayerl A, Dover Y, Riemer H, Shapira D. 2024 Gender rating gap in online reviews. *Nat. Hum. Behav.* **1**, 1–14. (doi:10.1038/s41562-024-01720-9)
100. Das J, DeLoach SB. 2011 Mirror, mirror on the wall: the effect of time spent grooming on earnings. *J. Socio-Econ.* **40**, 26–34. (doi:10.1016/j.socec.2010.08.018)
101. Kleisner K, Štěrbová Z, Fiala V. 2024 What constitutes the perception of facial typicality? *Personal. Individ. Differ.* **231**, 112838. (doi:10.1016/j.paid.2024.112838)
102. Torrance JS, Wincenciak J, Hahn AC, DeBruine LM, Jones BC. 2014 The relative contributions of facial shape and surface information to perceptions of attractiveness and dominance. *PLoS One* **9**, e104415. (doi:10.1371/journal.pone.0104415)
103. Zhang L. 2019 A data-driven test for cross-cultural differences in face preferences. *Perception* **48**, 487–499. (doi:10.1177/0301006619848992)
104. Sutherland CAM, Burton NS, Wilmer JB, Blokland GAM, Germine L, Palermo R, Collova JR, Rhodes G. 2020 Individual differences in trust evaluations are shaped mostly by environments, not genes. *Proc. Natl Acad. Sci. USA* **117**, 10218–10224. (doi:10.1073/pnas.1920131117)
105. Vásquez-Amézquita M, Castellanos-Chacón A, Medina-Sarmiento W, Cepeda V, Martínez-González MB, Leongómez JD. 2025 Resource availability and experiences of partner violence shape facial masculinity preferences in Colombian women. *Evol. Hum. Behav.* **46**, 106707. (doi:10.1016/j.evolhumbehav.2025.106707)
106. Tureček P, Černý V, Diallo MY, Cissé N, Pokorný Š, Kleisner K. 2025 Violence exposure is associated with preference for masculine faces: evidence from Senegal. *Proc. R. Soc. B* **292**, 20243105. (doi:10.1098/rspb.2024.3105)
107. Kowal M *et al*. 2022 Predictors of enhancing human physical attractiveness: data from 93 countries. *Evol. Hum. Behav.* **43**, 455–474. (doi:10.1016/j.evolhumbehav.2022.08.003)
108. Rice C. 2014 *Becoming women: the embodied self in image culture*. Toronto, Canada: University of Toronto Press.
109. Zweig JS. 2015 Are women happier than men? Evidence from the gallup world poll. *J. Happiness Stud.* **16**, 515–541. (doi:10.1007/s10902-014-9521-8)
110. Winquist LA, Mohr CD, Kenny DA. 1998 The female positivity effect in the perception of others. *J. Res. Personal.* **32**, 370–388. (doi:10.1006/jrpe.1998.2227)
111. Kleisner K *et al*. 2021 How and why patterns of sexual dimorphism in human faces vary across the world. *Sci. Rep.* **11**, 5978. (doi:10.1038/s41598-021-85402-3)
112. Wassiliwizky E, Zietsch BP, Kleisner K, Ullén F. 2026 Supplementary material from: The gender attractiveness gap. Figshare. (doi:10.6084/m9.figshare.c.8476530)

Downloaded from royalsocietypublishing.org/rspb/article-pdf/doi/10.1098/rspb.2026.0362/6140822/rspb.2026.0362.pdf
by guest
on 05 November 2026