

RESEARCH ARTICLE

Cross-cultural perception of female facial appearance: A multi-ethnic and multi-centre study

Rainer Voegeli^{1*}, Rotraut Schoop¹, Elodie Prestat-Marquis², Anthony V. Rawlings³, Todd K. Shackelford⁴, Bernhard Fink^{5,6}

1 DSM Nutritional Products, Kaiseraugst, Switzerland, **2** Newton Technologies, Lyon, France, **3** AVR Consulting Ltd, Northwich, Cheshire, United Kingdom, **4** Department of Psychology, Oakland University, Rochester, Michigan, United States of America, **5** Biosocial Science Information, Biedermannsdorf, Austria, **6** Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria

* rainer.voegeli@dsm.com



Abstract

Humans extract and use information from the face in assessments of physical appearance. Previous research indicates high agreement about facial attractiveness within and between cultures. However, the use of a narrow age range for facial stimuli, limitations due to unidirectional cross-cultural comparisons, and technical challenges have prevented definitive conclusions about the universality of face perception. In the present study, we imaged the faces of women aged 20 to 69 years in five locations (China, France, India, Japan, and South Africa) and secured age, attractiveness, and health assessments on continuous scales (0–100) from female and male raters (20–66 years) within and across ethnicity. In total, 180 images (36 of each ethnicity) were assessed by 600 raters (120 of each ethnicity), recruited in study centres in the five locations. Linear mixed model analysis revealed main and interaction effects of assessor ethnicity, assessor gender, and photographed participant (“face”) ethnicity on age, attractiveness, and health assessments. Thus, differences in judgments of female facial appearance depend on the ethnicity of the photographed person, the ethnicity of the assessor, and whether the assessor is female or male. Facial age assessments correlated negatively with attractiveness and health assessments. Collectively, these findings provide evidence of cross-cultural variation in assessments of age, and even more of attractiveness, and health, indicating plasticity in perception of female facial appearance across cultures, although the decline in attractiveness and health assessments with age is universally found.

OPEN ACCESS

Citation: Voegeli R, Schoop R, Prestat-Marquis E, Rawlings AV, Shackelford TK, Fink B (2021) Cross-cultural perception of female facial appearance: A multi-ethnic and multi-centre study. *PLoS ONE* 16(1): e0245998. <https://doi.org/10.1371/journal.pone.0245998>

Editor: Piotr Sorokowski, University of Wrocław, POLAND

Received: September 23, 2020

Accepted: January 12, 2021

Published: January 22, 2021

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the manuscript. Supporting information on rated samples are available from <https://datadryad.org/stash/dashboard> (doi:10.5061/dryad.t4b8gtj11).

Funding: All funding was provided by DSM Nutritional Products Ltd., Basel, Switzerland. The funders had no role in study design, data collection and analysis, or preparation of the manuscript. The funders approved the manuscript for publication.

Introduction

Research documents influences of facial age, health, and attractiveness on impression formation with regard to human female physical appearance [1–7]. Evolutionary scientists suggest that the interest in and sensitivity to female physical appearance is not culturally arbitrary but reflects evolved cognitive mechanisms that motivated successful ancestral human mate selection [6, 8–12]. Because of the links between female fecundity and youth and health [13–19],

Competing interests: RV and RS are employees of DSM, BF and AVR are consultants to DSM and EP is an employee of Newtone. TKS states no conflict of interest. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

humans universally ascribe importance to attractiveness, health, and youth in women [3, 20–23].

Apparently variable attractiveness standards across populations have been a topic of systematic research since the observations of Darwin in 1871 [24] and Westermarck in 1891 [25], and social and cultural scientists have advocated against communality in attractiveness assessments across cultures [26]. In this view, population-specific attractiveness standards are a cultural product, acquired by social learning (see Jones and Hill for a discussion [3, 27]). In addition to cultural factors, ecological conditions may influence population differences in attractiveness preferences [28], with the typical finding that male preferences vary less cross-culturally than do female preferences [29].

Research consistently finds that certain characteristics are judged attractive across individuals and cultures [23], suggesting an adaptive function of attractiveness, with external features providing information about biological and social qualities that play a role in sexual selection [30–32]. While most evolutionary-based studies acknowledge cross-cultural consistency in attractiveness assessments (with individual differences larger than differences between cultures) [33, 34], other research suggests that the strategies employed to extract information from faces differ across cultures [35]. One reason for the disagreement may be that environmental settings, and not genes, primarily influence face preferences [36]. In addition, it has been suggested that attractiveness preferences may be population-specific, depending on the ecological conditions and population-specific morphology. In this view, the cross-cultural agreement in face assessments would be higher for the assessment of “unattractive” rather than “attractive” faces [37]. Finally, the majority of reports of cross-cultural attractiveness assessments employed a correlational approach, thus comparing correlations between- and within-cultures based on aggregated rating data. These analyses do not consider random effects of facial stimuli and assessors but rely on *p*-level statistics that may be misleading, especially when sample sizes are low.

The generalizability of conclusions about cross-cultural assessments of facial attractiveness has been a matter of concern, in part because many studies secured facial images and panelists’ ratings from individuals that shared an ethnic background. Another concern is the comparability of findings obtained either from pooling information secured in projects with different foci or from studies that used different equipment and/or protocols. Other studies have investigated cross-cultural assessments of facial appearance by presenting images of individuals of one ethnicity to members of several ethnicities (e.g., [38, 39]), or individuals who immigrated to (and live in) a particular country (e.g., [40]). The latter methodology prevents definitive conclusions about the cross-cultural perception of facial appearance, as it is not known whether the immigrants had already adjusted their standards to that of the country to which they immigrated (but see [41] for recent evidence for adaptations to population-specific beauty standards).

Here, we investigate cross-cultural assessments of female facial age, attractiveness, and health in a multi-ethnic and multi-centre study in which female and male individuals identifying with one of five ethnicities (Chinese, French, Indian, Japanese, and South African) judged facial images of women within and across ethnicities. Thus, the current study extends previous research in several ways: i) five ethnicities are considered, concerning both photographed faces and assessors, with stratified random sampling by age (imaged women and assessors) and gender (assessors), ii) imaged women and assessors were selected from an age range of up to ~50 years, iii) the same equipment for imaging and assessment, and the same research protocol, were used in each of five study centres, iv) a mixed-model approach guided analyses of the raw scores (~52,000 judgements per attribute), affording consideration of crossed random effects of facial images and assessors, in addition to fixed effects (face ethnicity, assessor ethnicity,

assessor gender). The present study is the first to capitalize on all these aspects together in an effort to advance understanding of cross-cultural perception of female facial appearance.

Materials and methods

General methodology

Facial images and rating data were secured in five locations—Guangzhou (China), Lyon (France), New Delhi (India), Tokyo (Japan) and Cape Town (South Africa)—using the same experimental equipment and protocol. Data collection occurred from April 14 to September 6 (image recording), 2019 and from October 22, 2019 to February 1, 2020 (image rating). The study was approved by the Reading Independent Ethics Committee (RIEC), Woodley (U.K.), and the ACEAS Independent Ethics Committee, Ahmedabad (India). All participants provided written informed consent before participating. For images of participants shown in this article, the individuals provided written informed consent for publication.

Image recording

Study sample. Five-hundred-twenty-six women (“participants”) were recruited through local agencies and imaged: Chinese ($n = 106$), French ($n = 105$), Indian ($n = 100$), Japanese ($n = 100$), and South African ($n = 115$). Each sample included participants from five age cohorts (20–29, 30–39, 40–49, 50–59, 60–69 years; $n = \sim 20$ per group, equally distributed around the mean age of the respective group) (Table 1, total sample). Facial skin tone of women varied from darkly pigmented to lightly pigmented, as skin pigmentation correlates with latitude and ultraviolet radiation intensity [42]. However, in some countries the variation in skin tone is greater than in others (e.g., S. Africa) due to genetic variants [43]. According to the Fitzpatrick scale (a widely used photo-type classification tool for UV light sensitivity) [44], with type I = lightest pigmentation, and VI = darkest pigmentation, participants corresponded to the following types: Chinese II-IV, French II-III, Indian IV-V, Japanese II-IV, and S. African V-VI (this assessment was made by skin experts of the study centres).

Participants were screened before recruitment and women currently pregnant or lactating, suffering from visible facial pathologies or skin disease, receiving treatment for skin disease, involved in another clinical investigation or having participated in such within the past two months, having facial tattoos or permanent make-up, having topically applied hydroquinone-containing product within the last three months, having a history of facial cosmetic surgery, laser treatment, or application of Botox or hyaluronic acid-based fillers were excluded from participation.

Facial imaging. On the day before imaging, no facial cosmetic or dermatological products (including foundation and/or colour products) were allowed. In the evening before the day of imaging, participants could use their regular facial cleanser or soap. On the morning of the day of imaging, participants washed their face with lukewarm water and patted it dry with a soft

Table 1. Sociodemographic information and skin pigmentation of participants for the total sample and the subsample for the rating study.

Ethnicity	Fitzpatrick Skin Phototype	Gender	Total sample		Rated sample	
			n	Age \pm SD [years]	n	Age \pm SD [years]
Chinese	II-IV	female	107	45.2 \pm 14.1	36	42.6 \pm 13.0
French	II-III	female	105	45.0 \pm 14.6	36	42.6 \pm 13.4
Indian	IV-V	female	100	44.7 \pm 14.4	36	42.7 \pm 13.2
Japanese	II-IV	female	100	44.9 \pm 14.5	36	42.5 \pm 13.3
South African	V-VI	female	115	45.7 \pm 14.2	36	42.5 \pm 13.1

<https://doi.org/10.1371/journal.pone.0245998.t001>

towel. A technician cleaned the participant's face by gentle swabbing with a cotton pad soaked with distilled water of ambient temperature and allowed to dry for 20 min. Facial adornment and glasses were removed for imaging. Before taking photographs, participants were acclimatized for 30 minutes at $21\pm 1^\circ\text{C}$ and $45\pm 10\%$ relative humidity.

Participants wore identical black hairbands and black capes to cover features that might affect facial assessments (e.g., head hair, chest, or clothes) (Fig 1). Their faces were imaged in frontal view, with eyes open, and with a neutral facial expression using the ColorFace system (Newtone Technologies, Lyon, France). ColorFace captures high-resolution (24 MPs, at a maximum image size of 6000 x 4000 pixels, JPEG file format) full-face images without a chin-rest using an in-built single-lens reflex camera (SLR) camera (Nikon D5300; Nikon Inc., Minato, Japan). Earplugs attached to the stand of the device ensured standardized positioning of participants' faces, with fixed distance between lens and face. A horizontal reference line connecting the corners of the mouth was displayed on the facial image visualized in real-time on a remote computer, which served as an additional control before image capture. ColorFace uses LED light sources on the left and right sides of the face. System settings were selected to reduce flash intensity and increase light sensitivity of the camera sensor to avoid disturbance of the participant during imaging. For the presentation of the rating study, earplugs were digitally removed from images, eyes were vertically aligned, and visible area of the neck was standardized across images.

Face ratings

Assessors sample. A sample of 600 volunteers (299 females) ("assessors") participated in the rating study. They were recruited through local agencies in the same locations (and study centres) where the facial images of women were recorded. Participants reported to have lived in the respective location for at least two years. The assessors' skin photo-types (on the Fitzpatrick scale) matched that of imaged women in each of the five study locations (this assessment was made by skin experts of the study centres). Thus, we had male and female participants of five ethnicities ($n = 120$ per location) (Table 2). Each ethnic group included participants from three age groups (20–34, 35–49, and 50–66 years) with 40 assessors (20 females) per age group. The differences in mean ages between adjacent groups were 15 ± 2 years (all $ps < 0.001$).

Procedure. A subset of 180 images (of the initial sample with $n = 526$) was selected for presentation in the rating study (Table 1, rated sample). Before selection of these images, a quality check was performed for suitability of images for inclusion in the rating study. Three raters independently assessed the initial image set on a 4-point scale (1 = *not acceptable*, 4 = *acceptable*) for problems with positioning (e.g., head tilted), visibility of neck, and artefacts due



Fig 1. Samples images of female participants for presentation in the rating study. Chinese a), French b), Indian c), Japanese d), and S. African e).

<https://doi.org/10.1371/journal.pone.0245998.g001>

Table 2. Sociodemographic information and skin pigmentation of assessors in the rating study.

Ethnicity	Fitzpatrick Skin Phototype	Gender	n	Age \pm SD [years]
Chinese	II-IV	female	60	42.4 \pm 12.6
Chinese	II-IV	male	60	43.2 \pm 12.7
French	II-III	female	60	43.7 \pm 13.5
French	II-III	male	60	43.7 \pm 12.9
Indian	IV-V	female	60	43.1 \pm 13.1
Indian	IV-V	male	60	42.4 \pm 13.1
Japanese	II-IV	female	60	42.8 \pm 12.7
Japanese	II-IV	male	60	43.2 \pm 13.2
South African	V-VI	female	59	43.3 \pm 13.7
South African	V-VI	male	61	43.3 \pm 13.7

<https://doi.org/10.1371/journal.pone.0245998.t002>

to digital removal of earplugs. Only images considered “acceptable” by all three raters were considered for subset selection ($n = 382$).

Image selection was randomly stratified for participant/assessor ethnicity, gender, and the three assessor age groups; thus, of the available set of images, 36 images per ethnicity were assigned to female and male assessors of three age groups by considering all possible factor combinations (Fig 2). The images were presented on colour-calibrated, light-shielded, 27-inch LCD monitors (ColorEdge CG277, Eizo, Hakusan, Japan) with faces approximating natural size. Distance of assessor to the monitor during assessment was 50–60 cm. Room conditions during assessment were $21 \pm 1^\circ\text{C}$ and $45 \pm 10\%$ relative humidity with artificial light only.

Participants judged the images for age, health, and attractiveness in monadic presentation design (one after the other). Each assessor judged 90 randomly selected facial images per attribute, balanced across age groups (i.e. 270 images, in total). Thus, each image was assessed ~ 300 times and 10 times per subgroup (age group, ethnicity, gender).

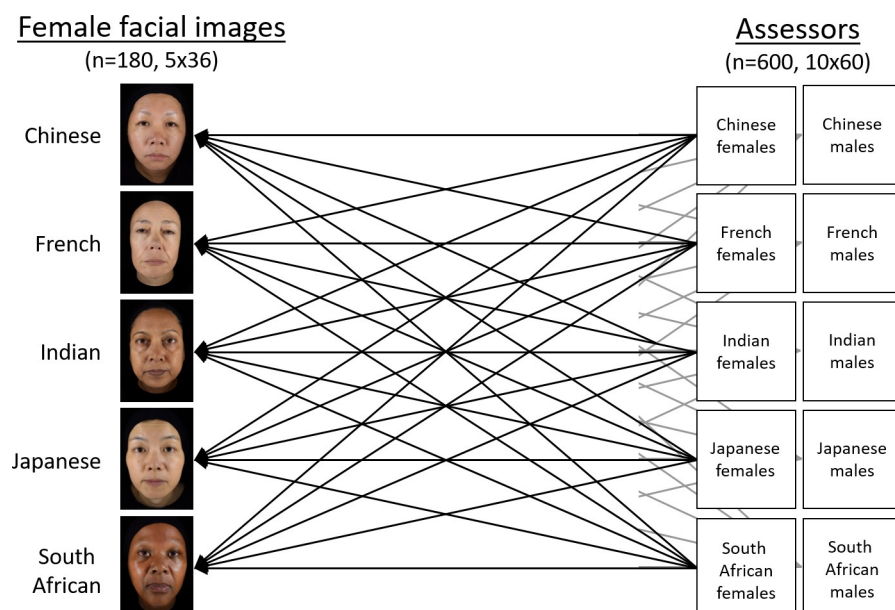


Fig 2. Experimental setup for the face rating. Naïve female and male assessors viewed facial images and provided spontaneous judgments of facial age, attractiveness, and health.

<https://doi.org/10.1371/journal.pone.0245998.g002>

Assessments of the three attributes were made in three separate blocks, using web-based software (PhotoScale; Newton Technologies, Lyon, France). The continuous scales ranged 0–100, with age assessment provided in years, and attractiveness and health assessments ranging “not attractive/not healthy” (0) to “attractive/healthy” (100). Order of blocks was randomized across participants, as was order of images within block. The time for assessment was limited to 3–5 sec. (before the image disappeared) to ensure viewing time was comparable across participants. Breaks of 15 min. were included between blocks to prevent fatigue effects. Statements on the screen and the attributes were created in English and then translated into Mandarin, French, Hindi, Japanese, and Xhosa by native speakers and verified by back-translation.

Statistical analysis

We performed a series of General Linear Mixed Models (GLMMs), separately for age, attractiveness, and health assessment as dependent variables, and with assessor ethnicity and gender, and participant (“face”) ethnicity as fixed effects (including interactions). Participant and assessor were included as crossed, independent random effects (both $ps < 0.001$). p -values of the fixed and interaction effects were corrected for multiplicity using the Benjamini-Hochberg method for control of the false discovery rate [45]. The analysis was performed in R [46], using the packages *lme4* [47] and *lmerTest* [48]. We calculated intra-class-coefficients (ICCs) [49] as variance partition coefficients for the mixed effect models. Finally, we aggregated raw scores for age, attractiveness, and health by participant (“face”) and considering assessor ethnicity, and calculated zero-order correlations (Pearson’s r) among the attributes for all 25 combinations of assessor ethnicity x face ethnicity ($n = 36$ each).

Results

Table 3 reports the main effects and interactions of assessor ethnicity, assessor gender and face ethnicity on perception of facial age and Tables 4–6 present descriptive statistics for age, attractiveness, and health assessments, separately for assessor ethnicity and gender, and participant ethnicity.

Age

Across assessor ethnicities, patterns of facial age perception were similar (Fig 3) across target face ethnicity, with a span of ~5 years (on average) between ethnicities judged youngest (Chinese) and oldest (French). French assessors provided the youngest and S. African assessors provided the oldest age estimations ($p < 0.001$). Neither assessor gender nor face ethnicity showed main effects on age judgements. However, there were interactions of assessor gender with assessor ethnicity and face ethnicity, respectively (both $ps < 0.05$). French women provided younger age estimations than female assessors of other ethnicities, with significant differences between French vs. Chinese and S. African assessors ($p < 0.01$). An interaction was detected for assessor gender with face ethnicity. However, the relevant pairwise comparisons across participants, within females and males, were not significant. The interaction of assessor ethnicity x face ethnicity revealed similar patterns across assessor ethnicities in terms of the rank order of faces. Chinese faces were judged youngest, and French faces oldest, with mean assessments varying across assessor ethnicities (French assessors provided the youngest estimations, and S. African assessors the oldest estimations) (Fig 3). The three-way interaction of assessor ethnicity with assessor gender and face ethnicity on perception of facial age was not significant (Table 3).

Table 3. Main and interaction effects of assessor ethnicity and gender, and participant (“face”) ethnicity on age, attractiveness, and health assessments.

Factor	Attribute	F	DF*	P
Assessor Ethnicity (AE)	Age	10.6	4, 583	<0.001
	Attractiveness	1.58	4, 586	0.18
	Health	19.6	4, 587	<0.001
Assessor Gender (AG)	Age	0.29	1, 583	0.59
	Attractiveness	45.2	1, 586	<0.001
	Health	9.87	1, 587	<0.01
Face Ethnicity (FE)	Age	0.94	4, 175	0.44
	Attractiveness	4.17	4, 175	<0.01
	Health	2.41	4, 175	0.051
AE x AG	Age	2.94	4, 583	<0.05
	Attractiveness	4.86	4, 586	<0.001
	Health	0.20	4, 587	0.94
AG x FE	Age	4.03	4, 51660	<0.01
	Attractiveness	6.92	4, 52137	<0.001
	Health	14.7	4, 52295	<0.001
AE X FE	Age	14.6	16, 51660	<0.001
	Attractiveness	181	16, 52132	<0.001
	Health	74.2	16, 52290	<0.001
AE x AG x FE	Age	1.10	16, 51659	0.35
	Attractiveness	4.26	16, 52127	<0.001
	Health	2.93	16, 52286	<0.001

* Calculated with the Satterthwaite method.

<https://doi.org/10.1371/journal.pone.0245998.t003>

Attractiveness

Attractiveness assessments showed considerable variation across assessor ethnicity and face ethnicity (Fig 4), including the interaction of the two effects and a three-way interaction including assessor gender (Table 3). There were main effects of assessor gender (females > males) and face ethnicity (but not assessor ethnicity), with S. African faces judged less attractive than French and Japanese faces (both $ps < 0.05$). Compared with age assessments, the patterns of facial attractiveness judgements were more diverse across assessor

Table 4. Descriptive statistics of age assessments, separately for assessor ethnicity and gender, and participant (“face”) ethnicity (mean ± SD).

		Face Ethnicity				
		Chinese	French	Indian	Japanese	S. African
Assessor Ethnicity/Gender						
Chinese	F	44.0 ± 13.3	49.6 ± 16.5	47.7 ± 15.9	44.6 ± 14.5	47.0 ± 14.9
Chinese	M	43.4 ± 13.0	49.2 ± 15.5	47.0 ± 13.9	44.4 ± 13.6	46.0 ± 13.1
French	F	40.3 ± 16.0	44.6 ± 17.0	42.9 ± 16.8	41.7 ± 16.6	43.6 ± 16.4
French	M	42.4 ± 15.3	46.0 ± 16.2	45.1 ± 16.3	43.6 ± 16.4	44.9 ± 15.2
Indian	F	42.2 ± 13.3	47.8 ± 15.4	45.3 ± 14.5	43.5 ± 14.3	48.4 ± 14.2
Indian	M	44.3 ± 13.5	49.6 ± 16.0	47.7 ± 14.7	45.4 ± 14.6	49.1 ± 14.9
Japanese	F	42.9 ± 14.8	47.3 ± 17.6	45.1 ± 16.2	43.2 ± 16.5	46.0 ± 15.9
Japanese	M	43.8 ± 14.3	47.7 ± 16.8	45.5 ± 15.4	43.6 ± 15.1	44.9 ± 14.2
S. African	F	45.6 ± 16.1	51.4 ± 17.7	49.8 ± 17.7	46.9 ± 17.1	49.1 ± 16.7
S. African	M	44.3 ± 13.5	49.6 ± 15.2	46.7 ± 15.2	45.1 ± 13.9	46.5 ± 13.7

<https://doi.org/10.1371/journal.pone.0245998.t004>

Table 5. Descriptive statistics of attractiveness, separately for assessor ethnicity and gender, and participant (“face”) ethnicity (mean ± SD).

		Face Ethnicity				
		Chinese	French	Indian	Japanese	S. African
Assessor Ethnicity/Gender						
Chinese	F	49.4 ± 18.9	52.8 ± 21.1	48.4 ± 20.3	50.8 ± 19.5	43.5 ± 20.8
Chinese	M	36.3 ± 21.4	38.5 ± 22.9	34.3 ± 21.5	37.5 ± 21.0	29.3 ± 20.6
French	F	41.5 ± 21.2	45.6 ± 23.0	42.8 ± 21.6	41.6 ± 21.0	43.3 ± 22.6
French	M	31.3 ± 22.0	36.9 ± 25.0	31.7 ± 22.7	31.0 ± 21.5	34.0 ± 23.5
Indian	F	43.6 ± 20.8	46.3 ± 22.6	36.9 ± 18.7	46.3 ± 21.7	25.7 ± 15.9
Indian	M	44.0 ± 19.1	46.2 ± 21.4	37.4 ± 18.1	45.9 ± 19.5	27.2 ± 15.3
Japanese	F	40.4 ± 20.0	51.5 ± 21.2	46.2 ± 21.6	42.8 ± 20.9	44.1 ± 22.0
Japanese	M	31.4 ± 16.3	40.2 ± 20.2	33.6 ± 18.1	33.9 ± 17.6	31.1 ± 18.3
S. African	F	40.0 ± 24.6	37.4 ± 24.7	44.7 ± 26.7	40.4 ± 24.7	39.3 ± 26.0
S. African	M	34.4 ± 22.4	33.9 ± 22.7	37.7 ± 23.7	35.1 ± 23.0	36.3 ± 23.5

<https://doi.org/10.1371/journal.pone.0245998.t005>

ethnicities, with 40 (of 100) significant ($p < 0.05$) pairwise comparisons considering all reasonable comparisons for the interaction of assessor ethnicity x face ethnicity. Chinese, Indian, Japanese, and French assessors judged French women highest in attractiveness (although the pairwise comparisons with those ranked 2nd to 4th were not always significant). S. African assessors, however, provided the highest attractiveness ratings for Indian faces and lowest ratings for French faces, with these two ethnicities showing the only significant difference (Fig 4). Significant gender differences in attractiveness assessments (females > males) were found for all but the Indian and South African assessors. Considering the interaction of assessor ethnicity x assessor gender, eight of 25 pairwise comparisons were significant (all $ps < 0.05$). Chinese women provided the highest attractiveness assessments, followed by Japanese, French, S. African, and Indian assessors (the latter three were significantly different from Chinese assessments at $p < 0.05$). Male attractiveness judgments were highest from Indian assessors, followed by S. African, Chinese, Japanese, and French assessors (with $p < 0.05$ for Indian vs. Japanese and French assessors). The interaction of face ethnicity x assessor gender showed that both female and male assessors judged French faces highest and S. African faces lowest on attractiveness, with females > males ($p < 0.001$) for all pairwise comparisons by face ethnicity. The three-way interaction of assessor ethnicity x face ethnicity x gender indicated that gender

Table 6. Descriptive statistics of health assessments, separately for assessor ethnicity and gender, and participant (“face”) ethnicity (mean ± SD).

		Face Ethnicity				
		Chinese	French	Indian	Japanese	S. African
Assessor Ethnicity/Gender						
Chinese	F	62.5 ± 17.1	63.4 ± 18.5	60.3 ± 17.9	64.1 ± 17.9	58.0 ± 19.5
Chinese	M	59.3 ± 18.4	59.7 ± 20.0	56.7 ± 19.6	59.9 ± 19.1	56.2 ± 20.4
French	F	67.7 ± 20.2	65.5 ± 22.4	62.3 ± 22.2	67.4 ± 20.6	63.1 ± 21.9
French	M	63.5 ± 18.8	61.4 ± 22.5	58.6 ± 20.9	63.1 ± 20.2	62.8 ± 20.5
Indian	F	55.3 ± 19.4	56.0 ± 20.7	48.4 ± 19.6	58.0 ± 19.2	42.7 ± 19.4
Indian	M	54.2 ± 18.4	53.8 ± 20.4	46.9 ± 17.9	56.3 ± 18.6	42.3 ± 18.0
Japanese	F	61.8 ± 16.9	63.7 ± 18.2	58.1 ± 19.3	65.8 ± 18.1	60.7 ± 18.3
Japanese	M	57.1 ± 17.3	58.8 ± 19.1	55.7 ± 18.3	60.2 ± 17.9	57.2 ± 18.4
S. African	F	56.4 ± 27.0	51.9 ± 27.3	58.6 ± 27.4	56.0 ± 26.8	53.3 ± 28.3
S. African	M	51.8 ± 23.4	48.4 ± 23.3	52.1 ± 23.9	52.3 ± 23.7	52.1 ± 24.3

<https://doi.org/10.1371/journal.pone.0245998.t006>

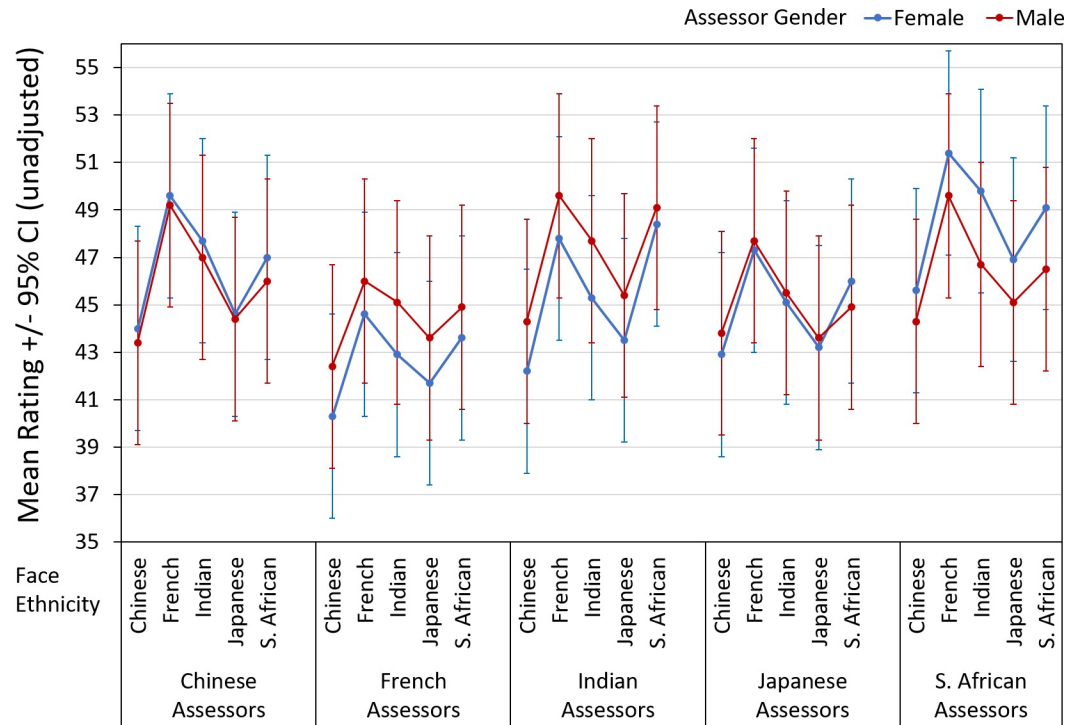


Fig 3. Assessments of facial images, by participant (“face”) ethnicity, assessor ethnicity and gender for age.

<https://doi.org/10.1371/journal.pone.0245998.g003>

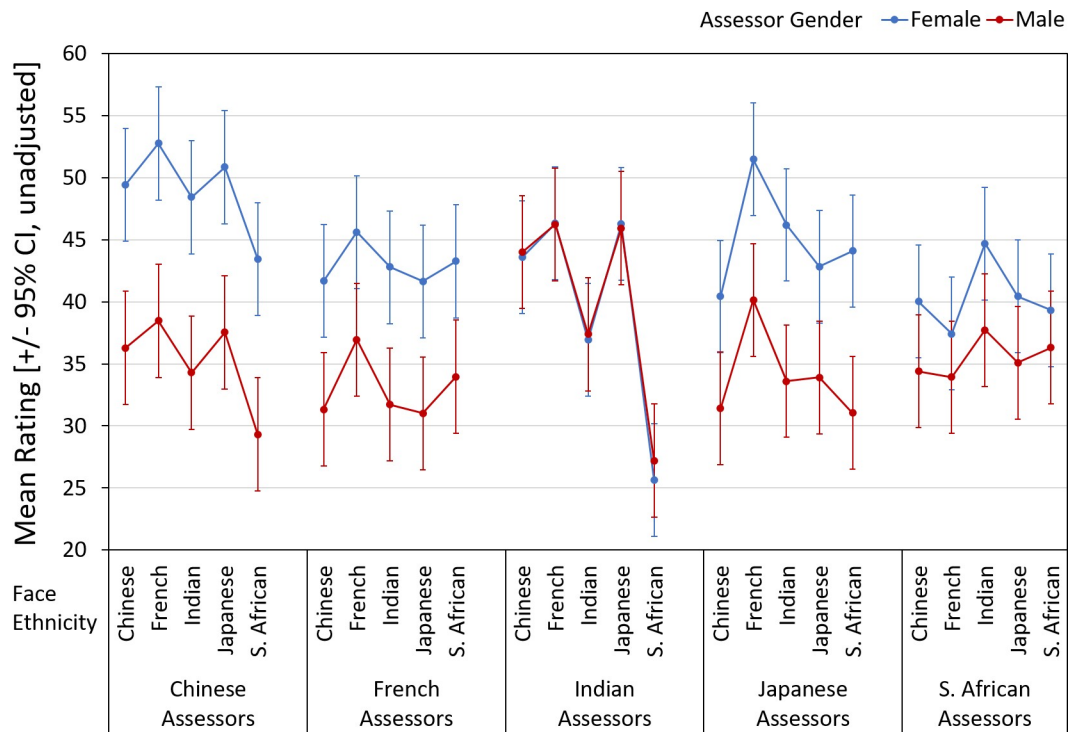


Fig 4. Assessments of facial images, by participant (“face”) ethnicity, assessor ethnicity and gender for attractiveness.

<https://doi.org/10.1371/journal.pone.0245998.g004>

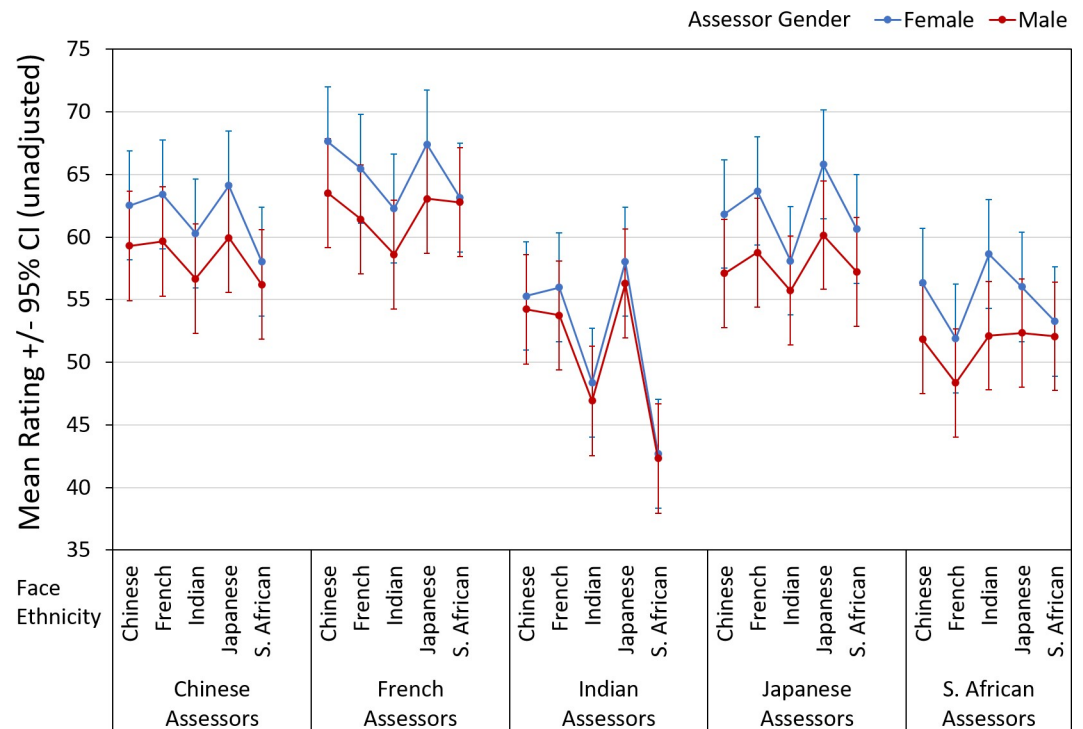


Fig 5. Assessments of facial images, by participant (“face”) ethnicity, assessor ethnicity and gender for health.

<https://doi.org/10.1371/journal.pone.0245998.g005>

differences in mean attractiveness assessments did not generalize across assessor/face ethnicity combinations (with Indian and S. African assessors the exception) (Fig 4).

Health

Health assessments showed main effects of assessor ethnicity and assessor gender but just failed to reach significance for face ethnicity ($p = 0.051$) (Table 3). Female health assessments were higher than male health assessments. Indian health assessments were lowest, followed by those of S. African, Japanese, Chinese, and French assessors (in that order). Indian and S. African assessments were different from those of Chinese, Japanese and French assessors ($p < 0.001$), but no differences were detected among the latter three (Fig 5). There were interactions of assessor ethnicity x face ethnicity, and face ethnicity x assessor gender (Table 3). The lowest health assessments were of S. African faces by Indian assessors (females and males), these being significantly different from the assessments of other ethnicities’ faces. In contrast, S. African assessors judged Indian faces highest and French faces lowest on health (this being the only significant difference in pairwise tests). Japanese assessors provided higher health ratings of faces of their own ethnicity compared with Indian faces ($p < 0.05$). There were no significant differences (across face ethnicities) in health ratings of Chinese and French assessors. The non-significant finding of assessor ethnicity x assessor gender suggests similar female/male judgements across assessor ethnicities. The three-way interaction of assessor ethnicity x face ethnicity x gender suggests some differences depend on face ethnicity. For example, the rank order of face assessments across ethnicities was the same for female and male Chinese and Japanese assessors, but some differences in comparisons of female vs. male ranking were noted for other assessor ethnicities (albeit n.s. in pairwise comparison across face ethnicities) (Fig 5).

The overall ICCs, i.e. the total variation explained by both participant (“face”) and assessor information, were as follows: age ICC = 0.749, attractiveness ICC = 0.584, health ICC = 0.551. At the participant level (correlations between assessments of the same face), the ICCs were 0.629 (age), 0.130 (attractiveness), and 0.162 (health), and at the assessor level (correlation between judgements by the same assessor) were 0.110 (age), 0.400 (attractiveness), and 0.336 (health).

There were negative correlations of age with attractiveness, ranging -0.680 to -0.908 (all $ps < 0.001$), and health, ranging -0.759 to -0.936 (all $ps < 0.001$), and positive correlations of attractiveness with health, ranging 0.792 to 0.983 (all $ps < 0.001$).

Discussion

Previous research suggested strong agreement in attractiveness assessments, both within and across ethnicities [23, 40, 50], especially for female attractiveness [29, 51, 52]. The present study used a simultaneous multi-centre, multi-ethnic approach to secure assessments of female facial age, attractiveness, and health and identified both similarities and differences in assessments across ethnicities. Perhaps most importantly, there were (three-way) interaction effects of assessor ethnicity and gender, and participant (“face”) ethnicity for attractiveness and health (but not for age). This suggests that differences in female facial attractiveness and health judgements depend on who judges the face (i.e. assessor ethnicity), which face is assessed (i.e. target ethnicity), and whether the assessor is female or male. There is stronger agreement in facial age assessments than in attractiveness and health assessments.

Intra-class correlations (ICCs) corroborate the findings of diversity in cross-cultural face assessments; the ICC for age assessments was higher than for attractiveness and health assessments, suggesting greater agreement for the former than the latter assessments. Inter-correlations of female facial age, attractiveness, and health assessments were large and in the direction predicted by evolutionary approaches to female appearance (see for review Grammer *et al.* [9], Rhodes [4], and Thornhill and Gangestad [12]), suggesting a strong relationship of attractiveness with health, and a decline in these qualities with age [3, 17, 53]. Collectively, the findings of the present study suggest greater cross-cultural variation in assessments of female facial appearance than indicated in previous research, especially in attractiveness and health assessments.

Recent research reported disagreement among individual facial attractiveness judgements, highlighting the importance of determining how these preferences vary among individuals [54, 55]. Perhaps most relevant for cross-cultural comparisons is the assumed importance of certain facial characteristics in a given society as derived from the study of another society. Facial characteristics investigated in previous studies (e.g., symmetry, averageness, sex-typical features) may not contribute substantially to judgements of facial attractiveness [56–58] or health [59], but even if they do, the contribution of these features may vary across societies depending on environmental conditions [60, 61] or sociocultural settings [62, 63]. Zhang *et al.* [57] in a data-driven (as opposed to theory-driven) approach detected cross-cultural differences in face preferences not apparent in studies using theory-driven approaches, leading to the conclusion that Chinese and British “White” participants used face information in different ways (i.e. they focused on different features) (see also Kleisner *et al.* [64]). Similar conclusions were derived from the findings of eye-movement patterns of Western and East Asian participants, suggesting that cultural background shapes visual environment affordance [35]. Coetzee *et al.* [65] investigated attractiveness assessments of White Scottish and Black S. African students for own- and other ethnicity faces. Black S. African raters relied more heavily on colour cues in their assessments of Black African female attractiveness, whereas White Scottish

judges relied more heavily on shape cues in their assessments. The researchers concluded that although there was evidence for the universality of facial attractiveness assessments, the ethnicity of the target face moderated this agreement, i.e. agreement on European faces was higher than on African faces (possibly due to a difference in familiarity with other-ethnicity faces).

In the present study, the female participants (imaged women) were recruited in major cities. We might assume that contact with other ethnicities is considerable. Coetzee *et al.* [65] stated for S. Africans, for example, there is variation across samples in terms of familiarity with other ethnicities' facial appearance. However, this alone cannot explain the variation in the facial assessments across ethnicities in our findings. The patterns of age assessments are similar across ethnicities, for both face ethnicity and assessor ethnicity. If assessors of one ethnicity were unable to accurately assess facial appearance of other ethnicities because of unfamiliarity with the variation in morphology, the patterns of age assessments across ethnicities should be more diverse than was the case (although there were differences in mean age assessments). Age-related changes in facial morphology (in terms of shape) and visible skin condition both play a role in age assessments. Yet the relative contribution of these features to age perception may be different across ethnicities depending, for example, on the visibility of skin colouration cues. In lightly pigmented skin, unevenness may be more detectable than in darkly pigmented skin. In the present study, our focus was on the investigation of cross-cultural differences (or similarities) in perceptions of female facial appearance. Thus, we did not quantify facial morphology and/or skin condition. As such, the possibility of cross-cultural variation in the relative importance of these components for age assessments remains to be investigated.

Attractiveness and health assessments showed greater variation across ethnicities, with some large differences associated with face and assessor ethnicity, in addition to gender differences. Perhaps most conspicuous in the pattern of cross-cultural variation in facial attractiveness and health is the low assessments of S. African (and Indian) women (and the absence of a gender difference) made by Indian assessors. This may reflect the influence of socio-cultural factors, namely "colourism" (i.e. a preference for lighter skin colour, possibly dating to colonialism) [66] (but see Wagatsuma, 1967 [67]), on face perception, as darkly pigmented skin in India is perceived negatively, partially due to the hierarchical caste system [68, 69]. Similar "colourism" has been reported for S. Africa where lighter-skinned migrants have been treated more positively than darker-skinned migrants [70]. In the present study, S. African assessors judged French faces lowest and Indian faces highest on attractiveness.

Many additional factors might contribute to cross-cultural differences in attractiveness assessments, including environmental settings [29, 71–73] and measures of national health [28, 39], along with variation within- and between assessors (e.g., hormonal fluctuations), which have been reviewed elsewhere [4, 12, 30, 74, 75] (but see Jones *et al.* and Marcinkowska *et al.* [76, 77]). There is consensus that certain facial cues relate to female age and health, both of which correlate with female fecundity and reproductive potential [9, 13, 15, 78]. From an evolutionary perspective, one might assume that these relationships are found universally, and the evidence from industrialized and pre-industrialized societies suggests that this is the case. However, this universality does not preclude variation in the strength of associations across ethnicities. Our findings of cross-cultural variation in perceptions of female facial appearance do not challenge the evidence that certain facial cues provide information about an individual's mating-related quality. We document negative correlations between age and attractiveness/health, and a positive correlation between attractiveness and health for every combination of face ethnicity and assessor ethnicity. The relative size of effects and the mean assessments may differ across cultures because of differences in environmental conditions, socio-cultural factors, and other variables that contribute to individual differences (see for a review, Pisanski

and Feinberg [79]). Nevertheless, the biological blueprint nature uses to convey certain information about an individual's quality may be the same for all humans [9].

Many studies investigating human physical attractiveness include a statement on the stability of attractiveness ratings across ethnicities ("strong cross-cultural agreement"). However, there is concern about the validity of this statement [54, 55, 80, 81]. The findings of the present study corroborate the presence of differences in the assessment of female facial appearance, depending on the ethnicity of the face and the ethnicity and gender of the assessor. These cross-cultural differences in face assessments are evident especially in attractiveness and health ratings, at least in samples of industrialized and industrializing countries. Previous research reporting differences in face preferences of industrialized vs. pre-industrialized societies [82, 83] suggested that visual experience with facial cues may account for the effect (but see Danel *et al.* [80]). We suggest that visual experience with faces of other ethnicities alone cannot explain our findings. Rather, our findings may be explained through a combination of ethnocentrism [84, 85] and other effects that emerge from different socio-cultural settings. However, the variation in patterns of assessments of female facial appearance may also reflect evolved preferences expressed in response to environmental settings that contributed to the development of plasticity in the perception of female facial appearance across cultures. Future studies should i) quantify cross-cultural variation in facial morphology and visual skin condition, and disentangle the relative impact of these components on face ratings, and ii) consider the influence of ethnocentrism and stereotyping in cross-cultural (facial) assessment, in addition to effects motivated by human sexual psychology. For example, face research has successfully applied geometric morphometrics in the assessment of facial shape variation in samples of industrialized and non-industrialized societies in relation to physical capacity and/or perception (e.g., Butovskaya *et al.*, Fink *et al.*, Schaefer *et al.*, and Kleisner *et al.* [86–89]). Similarly, objective measures of skin color and the evenness of skin tone correlate with assessments of facial age, attractiveness, and health [90, 91]. The application of these technologies in the current multi-ethnic and multi-centre study would take the study findings to the next level by investigating features that predict cross-cultural variation in face assessments.

Although the high level of standardization of facial imaging and assessment protocols is a strength of the current study, we contend that it could be realized only in cooperation with local study centres in major cities. The collection of similar stimuli and information from members of small-scale societies in anthropological fieldwork remains challenging. Therefore, evidence from studies that have investigated face assessments cross-culturally should be considered with caution regarding the comparability of study findings. This includes questions about influences from (Western) media shaped face perception, which can be assumed to be present in all population samples of the current study. France, for example, is a global leader in the cosmetics business, and French cosmetic products are highly regarded especially in China and Japan, possibly leading to stereotypic and higher assessments of French women compared to women of other ethnicities. We suggest that studies investigating cross-cultural agreement in face perception and the reasons for geographical variation need to quantify socio-cultural stereotypes (e.g., Choi *et al.* [92] in inter-population perception in addition to securing objective measures of biological variation in facial appearance.

Acknowledgments

The authors thank the study volunteers who participated in the study.

We also would like to thank the following colleagues, who made the study possible with their enthusiasm and dedication:

The study directors for their indefatigability in conducting the study at the following centers, Etienne Camel and Mathilde De Torsiac at IEC, Lyon, France; Yumiko Seki at IEC Tokyo, Japan; Manoj Choudhry at IEC Cape Town, South Africa; Peggy Ruan at IEC Guangzhou, China and Rashi Nangia and Rajat Gupta at CIDP Biotech, New Delhi, India.

Cyprien Adnet and Adrien Perret of Newton Technologies, Lyon, France, who designed and created the database.

Dominik Imfeld, Rolf Schütz and Remo Campiche of DSM Nutritional Products Ltd, Kaiseraugst, Switzerland for their valuable ideas on the study design.

Author Contributions

Conceptualization: Rainer Voegeli, Elodie Prestat-Marquis, Anthony V. Rawlings, Bernhard Fink.

Data curation: Rainer Voegeli, Rotraut Schoop, Elodie Prestat-Marquis, Bernhard Fink.

Formal analysis: Rainer Voegeli, Bernhard Fink.

Funding acquisition: Rainer Voegeli.

Investigation: Rainer Voegeli, Rotraut Schoop, Elodie Prestat-Marquis, Bernhard Fink.

Methodology: Rainer Voegeli, Rotraut Schoop, Elodie Prestat-Marquis, Anthony V. Rawlings, Bernhard Fink.

Project administration: Rainer Voegeli, Elodie Prestat-Marquis.

Resources: Rainer Voegeli, Rotraut Schoop, Elodie Prestat-Marquis.

Software: Rotraut Schoop, Elodie Prestat-Marquis.

Supervision: Rainer Voegeli, Elodie Prestat-Marquis.

Validation: Rainer Voegeli, Rotraut Schoop, Todd K. Shackelford, Bernhard Fink.

Visualization: Rainer Voegeli, Rotraut Schoop, Bernhard Fink.

Writing – original draft: Rainer Voegeli, Bernhard Fink.

Writing – review & editing: Rainer Voegeli, Rotraut Schoop, Elodie Prestat-Marquis, Anthony V. Rawlings, Todd K. Shackelford, Bernhard Fink.

References

1. Cunningham MR. Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *J Pers Soc Psychol.* 1986; 50(5):925–35.
2. Gray AW, Boothroyd LG. Female facial appearance and health. *Evol Psychol.* 2012; 10(1):66–77. PMID: [22833849](https://pubmed.ncbi.nlm.nih.gov/22833849/)
3. Jones D, Hill K. Criteria of facial attractiveness in five populations. *Hum Nat.* 1993; 4(3):271–96. <https://doi.org/10.1007/BF02692202> PMID: [24214367](https://pubmed.ncbi.nlm.nih.gov/24214367/)
4. Rhodes G. The evolutionary psychology of facial beauty. *Annu Rev Psychol.* 2006; 57:199–226. <https://doi.org/10.1146/annurev.psych.57.102904.190208> PMID: [16318594](https://pubmed.ncbi.nlm.nih.gov/16318594/)
5. Shackelford TK, Larsen RJ. Facial attractiveness and physical health. *Evol Hum Behav.* 1999; 20(1):71–6.
6. Symons D. Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness. *Sexual nature, sexual culture.* Chicago series on sexuality, history, and society. Chicago, IL, US: The University of Chicago Press; 1995. p. 80–119.
7. Zebrowitz LA. First impression from faces. *Curr Dir Psychol Sci.* 2017; 26:237–42. <https://doi.org/10.1177/0963721416683996> PMID: [28630532](https://pubmed.ncbi.nlm.nih.gov/28630532/)
8. Buss DM, Barnes M. Preferences in human mate selection. *J Pers Soc Psychol.* 1986; 50(3):559–70.

9. Grammer K, Fink B, Moller AP, Thornhill R. Darwinian aesthetics: sexual selection and the biology of beauty. *Biol Rev Camb Philos Soc.* 2003; 78(3):385–407. <https://doi.org/10.1017/s1464793102006085> PMID: 14558590
10. Kou H, Xie Q, Bi T. Mechanisms for the cognitive processing of attractiveness in adult and infant faces: From the evolutionary perspective. *Front Psychol.* 2020; 11:436. <https://doi.org/10.3389/fpsyg.2020.00436> PMID: 32218762
11. Symons D. *The evolution of human sexuality.* New York: Oxford University Press; 1979. 358 p.
12. Thornhill R, Gangestad SW. *Facial attractiveness.* Cambridge: Elsevier; 1999.
13. Jones D. Sexual selection, physical attractiveness, and facial neoteny: Cross-cultural evidence and implications. *Curr Anthropol.* 1995; 36(5):723–48.
14. Kenrick DT, Keefe RC. Age preferences in mates reflect sex differences in human reproductive strategies. *Behav Brain Sci.* 1992; 15(1):75–91.
15. Law Smith MJ, Perrett DI, Jones BC, Cornwell RE, Moore FR, Feinberg DR, et al. Facial appearance is a cue to reproductive hormone levels in women. *Proc R Soc B.* 2006; 273(1583):135–40. <https://doi.org/10.1098/rspb.2005.3296> PMID: 16555779
16. Henderson AJ, Holzleitner IJ, Talamas SN, Perrett DI. Perception of health from facial cues. *Philos Trans R Soc Lond B Biol Sci.* 2016; 371:20150380. <https://doi.org/10.1098/rstb.2015.0380> PMID: 27069057
17. Marcinkowska UM, Dixon BJ, Kozlov MV, Prasai K, Rantala MJ. Men's preferences for female facial femininity decline with age. *J Gerontol B Psychol Sci Soc Sci.* 2015; 72(1):180–6. <https://doi.org/10.1093/geronb/gbv077> PMID: 26320118
18. Thornhill R, Gangestad SW. *The evolutionary biology of human female sexuality.* Oxford; New York: Oxford University Press; 2008. 411 p.
19. Kalick SM, Zebrowitz LA, Langlois JH, Johnson RM. Does human facial attractiveness honestly advertise health? Longitudinal data on an evolutionary question. *Psychol Sci.* 1998; 9(1):8–13.
20. Berry DS. Attractiveness, attraction, and sexual selection: Evolutionary perspectives on the form and function of physical attractiveness. *Adv Exp Soc Psychol. Advances in Experimental Social Psychology.* 32: Elsevier; 2000. p. 273–342.
21. Buss DM, Schmitt DP. Mate preferences and their behavioral manifestations. *Annu Rev Psychol.* 2019; 70:77–110. <https://doi.org/10.1146/annurev-psych-010418-103408> PMID: 30230999
22. Etcoff NL. *Survival of the prettiest: The science of beauty.* New York: Doubleday; 1999. 325 p.
23. Langlois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol Bull.* 2000; 126(3):390–423. <https://doi.org/10.1037/0033-2909.126.3.390> PMID: 10825783
24. Darwin C. *The descent of man and selection in relation to sex.* London: John Murray; 1871.
25. Westermarck E. *The history of human marriage.* London: Macmillan & Co.; 1891.
26. Jackson LA. *Physical appearance and gender: Sociobiological and sociocultural perspectives.* Albany: State University of New York Press; 1992. 326 p.
27. Little AC, Jones BC, DeBruine LM, Caldwell CA. Social learning and human mate preferences: a potential mechanism for generating and maintaining between-population diversity in attraction. *Philos Trans R Soc Lond B Biol Sci.* 2011; 366(1563):366–75. <https://doi.org/10.1098/rstb.2010.0192> PMID: 21199841
28. Marcinkowska UM, Kozlov MV, Cai H, Contreras-Garduno J, Dixon BJ, Oana GA, et al. Cross-cultural variation in men's preference for sexual dimorphism in women's faces. *Biol Lett.* 2014; 10(4):20130850. <https://doi.org/10.1098/rsbl.2013.0850> PMID: 24789138
29. Penton-Voak IS, Jacobson A, Trivers R. Populational differences in attractiveness judgments of male and female faces: Comparing British and Jamaican samples. *Evol Hum Behav.* 2004; 25:355–70.
30. Little AC, Jones BC, DeBruine LM. Facial attractiveness: evolutionary based research. *Philos Trans R Soc Lond B Biol Sci.* 2011; 366(1571):1638–59. <https://doi.org/10.1098/rstb.2010.0404> PMID: 21536551
31. Perrett DI, May KA, Yoshikawa S. Facial shape and judgements of female attractiveness. *Nature.* 1994; 368(6468):239–42. <https://doi.org/10.1038/368239a0> PMID: 8145822
32. Perrett DI, Lee KJ, Penton-Voak I, Rowland D, Yoshikawa S, Burt DM, et al. Effects of sexual dimorphism on facial attractiveness. *Nature.* 1998; 394(6696):884–7. <https://doi.org/10.1038/29772> PMID: 9732869
33. Bronstad PM, Russell R. Beauty is in the 'we' of the beholder: greater agreement on facial attractiveness among close relations. *Perception.* 2007; 36(11):1674–81. <https://doi.org/10.1068/p5793> PMID: 18265847

34. Silva AS, Lummaa V, Muller U, Raymond M, Alvergne A. Facial attractiveness and fertility in populations with low levels of modern birth control. *Evol Hum Behav.* 2012; 33:491–8.
35. Blais C, Jack RE, Scheepers C, Fiset D, Caldara R. Culture shapes how we look at faces. *PLoS One.* 2008; 3(8):e3022. <https://doi.org/10.1371/journal.pone.0003022> PMID: 18714387
36. Germine L, Russell R, Bronstad PM, Blokland GAM, Smoller JW, Kwok H, et al. Individual aesthetic preferences for faces are shaped mostly by environment, not genes. *Current biology: CB.* 2015; 25(20):2684–9. <https://doi.org/10.1016/j.cub.2015.08.048> PMID: 26441352
37. Sorokowski P, Kosciński K, Sorokowska A. Is beauty in the eye of the beholder but ugliness culturally universal? Facial preferences of Polish and Yali (Papua) people. *Evol Psychol.* 2013; 11(4):907–25.
38. Kocnar T, Saribay SA, Kleisner K. Perceived attractiveness of Czech faces across 10 cultures: Associations with sexual shape dimorphism, averageness, fluctuating asymmetry, and eye color. *PLoS One.* 2019; 14(11):e0225549. <https://doi.org/10.1371/journal.pone.0225549> PMID: 31751432
39. DeBruine LM, Jones BC, Crawford JR, Welling LL, Little AC. The health of a nation predicts their mate preferences: cross-cultural variation in women's preferences for masculinized male faces. *Proc Biol Sci.* 2010; 277(1692):2405–10. <https://doi.org/10.1098/rspb.2009.2184> PMID: 20236978
40. Cunningham MR, Roberts AR, Barbee AP, Druen PB, Wu C-H. "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *J Pers Soc Psychol.* 1995; 68(2):261–79.
41. Pavlovič O, Fiala V, Kleisner K. Environmental convergence in facial preferences: A cross-group comparison of Asian Vietnamese, Czech Vietnamese, and Czechs. *PsyArXiv.* 2020; March 26.
42. Del Bino S, Bernerd F. Variations in skin colour and the biological consequences of ultraviolet radiation exposure. *Br J Dermatol.* 2013; 169 Suppl 3:33–40. <https://doi.org/10.1111/bjd.12529> PMID: 24098899
43. Crawford NG, Kelly DE, Hansen MEB, Beltrame MH, Fan S, Bowman SL, et al. Loci associated with skin pigmentation identified in African populations. *Science.* 2017; 358(6365). <https://doi.org/10.1126/science.aan8433> PMID: 29025994
44. Fitzpatrick TB. Soleil et peau. *Journal de Médecine Esthétique.* 1975; 2:33–4.
45. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J Royal Stat Soc B.* 1995; 57(1):289–300.
46. Team RC. A language and environment for statistical computing Foundation for Statistical Computing 2020 [Available from: <https://www.R-project.org/>].
47. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models using lme4. *J Stat Softw.* 2015; 67:1–48.
48. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw.* 2017; 82:1–26.
49. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods.* 1996; 1(1):30–46.
50. Zebrowitz LA, Montepare JM, Lee HK. They don't look all look alike: Individual impressions of other racial groups. *J Pers Soc Psychol.* 1993; 65:85–101. <https://doi.org/10.1037//0022-3514.65.1.85> PMID: 8355144
51. Deusch FM, Zalenski CM, Clark ME. Is there a double standard of aging? *J Appl Soc Psychol.* 1986; 16(9):771–85.
52. Gottschall J, Anderson K, Burbank C, Burch J, Byrnes C, Callanan C, et al. The "beauty myth" is no myth: Emphasis on male-female attractiveness in World folktales. *Hum Nat.* 2008; 19(2):174–88. <https://doi.org/10.1007/s12110-008-9035-3> PMID: 26181463
53. Zebrowitz LA, Olson K, Hoffman K. Stability of babyfacedness and attractiveness across the life span. *J Pers Soc Psychol.* 1993; 64(3):453–66. <https://doi.org/10.1037//0022-3514.64.3.453> PMID: 8468672
54. Scott IM, Clark AP, Josephson SC, Boyette AH, Cuthill IC, Fried RL, et al. Human preferences for sexually dimorphic faces may be evolutionarily novel. *Proc Natl Acad Sci USA.* 2014; 111(40):14388–93. <https://doi.org/10.1073/pnas.1409643111> PMID: 25246593
55. Han C, Wang H, Hahn AC, Fisher CI, Kandrik M, Fasolt V, et al. Cultural differences in preferences for facial coloration. *Evol Hum Behav.* 2018; 39(2):154–9.
56. Said CP, Todorov A. A statistical model of facial attractiveness. *Psychol Sci.* 2011; 22(9):1183–90. <https://doi.org/10.1177/0956797611419169> PMID: 21852448
57. Zhang L, Holzleitner JJ, Lee AJ, Wang H, Han C, Fasolt V, et al. A data-driven test for cross-cultural differences in face preferences. *Perception.* 2019; 48(6):487–99. <https://doi.org/10.1177/0301006619849382> PMID: 31068092

58. Stephen ID, Scott IML, Coetzee V, Pound N, Perrett DI, Penton-Voak IS. Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evol Hum Behav.* 2012; 33(4):260–7.
59. Cai Z, Hahn AC, Zhang W, Holzleitner IJ, Lee AJ, DeBruine LM, et al. No evidence that facial attractiveness, femininity, averageness, or coloration are cues to susceptibility to infectious illnesses in a university sample of young adult women. *Evol Hum Behav.* 2019; 40(2):156–9.
60. Little AC, Cohen DL, Jones BC, Belsky J. Human preference for facial masculinity changes with relationship type and environmental harshness. *Behav Ecol Sociobiol.* 2007; 61:967–73.
61. Scott I, Swami V, Josephson SC, Penton-Voak IS. Context-dependent preferences for facial dimorphism in a rural Malaysian population. *Evol Hum Behav.* 2008; 29(4):289–96.
62. Dion KK, Pak AW-P, Dion KL. Stereotyping physical attractiveness: A sociocultural perspective. *J Cross Cult Psychol.* 1990; 21(3):378–98.
63. Zebrowitz LA, Wang R, Bronstad MA, Eisenberg D, Undurraga E, Reyes-Garcia A, et al. First impression from faces among U.S. and culturally isolated Tsimane' people in the Bolivian rainforest. *Journal of Cross-Cultural Psychology.* 2012; 41(1):119–34.
64. Kleisner K, Pokorny S, Saribay SA. Toward a new approach to cross-cultural distinctiveness and typicality of human faces: The cross-group typicality/ distinctiveness metric. *Front Psychol.* 2019; 10:124. <https://doi.org/10.3389/fpsyg.2019.00124> PMID: 30766504
65. Coetzee V, Greeff JM, Stephen ID, Perrett DI. Cross-cultural agreement in facial attractiveness preferences: the role of ethnicity and gender. *PLoS One.* 2014; 9(7):e99629. <https://doi.org/10.1371/journal.pone.0099629> PMID: 24988325
66. Dixon AR, TE E. Skin color and colorism: Global research, concepts, and measurement. *Annu Rev Sociol.* 2017; 43(1):405–24.
67. Wagatsuma H. The social perception of skin color in Japan. *Daedalus.* 1967; 96:407–43.
68. Beal FM. Double jeopardy: To be black and female. *Meridians.* 2008; 8(2):166–76.
69. Nagar I. The unfair selection: A study of skin-color bias in arranged Indian marriage. *SAGE Open.* 2018; Apr-Jun:1–8.
70. Myambo MT. Class identity, xenophobia, and xenophilia. In: Imafidon E, editor. *Handbook of African philosophy of difference*: Springer, Cham; 2019.
71. Dixon BJW, Little AC, Dixon HGW, Brooks RC. Do prevailing environmental factors influence human preferences for facial morphology? *Behav Ecol.* 2017; 28(5):1217–27.
72. Gangestad SW, Buss DM. Pathogen prevalence and human mate preferences. *Ethol and Sociobiol.* 1993; 14(2):89–96.
73. Marcinkowska UM, Rantala MJ, Lee AJ, Kozlov MV, Aavik T, Cai H, et al. Women's preferences for men's facial masculinity are strongest under favorable ecological conditions. *Sci Rep.* 2019; 9(1):3387. <https://doi.org/10.1038/s41598-019-39350-8> PMID: 30833635
74. Little AC. Facial attractiveness. *WIRE Cognitive Sciences.* 2014; 5:621–34. <https://doi.org/10.1002/wcs.1316> PMID: 26308869
75. Penton-Voak IS, Perrett DI. Consistency and individual differences in facial attractiveness judgements: An evolutionary perspective. *Soc Res.* 2000; 67:219–44.
76. Jones BC, Hahn AC, Fisher CI, Wang H, Kandrik M, Han C, et al. No Compelling Evidence that Preferences for Facial Masculinity Track Changes in Women's Hormonal Status. *Psychol Sci.* 2018; 29(6):996–1005. <https://doi.org/10.1177/0956797618760197> PMID: 29708849
77. Marcinkowska UM, Ellison PT, Galbarczyk A, Milkowska K, Pawlowski B, Thune I, et al. Lack of support for relation between woman's masculinity preference, estradiol level and mating context. *Horm Behav.* 2016; 78:1–7. <https://doi.org/10.1016/j.yhbeh.2015.10.012> PMID: 26497247
78. Maestriepieri D, Klimczuk AC, Traficonte DM, Wilson MC. A greater decline in female facial attractiveness during middle age reflects women's loss of reproductive value. *Front Psychol.* 2014; 5:179. <https://doi.org/10.3389/fpsyg.2014.00179> PMID: 24592253
79. Pisanski K, Feinberg DR. Cross-cultural variation in mate preferences for averageness, symmetry, body size, and masculinity. *Cross Cult Res.* 2013; 47(2):162–97.
80. Danel DP, Fedurek P, Coetzee V, Stephen ID, Nowak N, Stirrat M, et al. A cross-cultural comparison of population-specific face shape preferences (Homo sapiens). *Ethology.* 2012; 118(12):1173–81.
81. Fiala V, Třebický V, Leongómez JD, Tureček P, Pazhoohi F, Saribay SA, et al. Cues for facial attractiveness and preference of sexual dimorphism: A comparison across five cultures 2020.
82. Apicella CL, Little AC, Marlowe FW. Facial averageness and attractiveness in an isolated population of hunter-gatherers. *Perception.* 2007; 36(12):1813–20. <https://doi.org/10.1068/p5601> PMID: 18283931

83. Sorokowski P, Sorokowska A, Kras D. Face color and sexual attractiveness: Preferences of Yali people of Papua. *Cross Cult Res.* 2013; 47:415–27.
84. Cashdan E. Ethnocentrism and xenophobia: A cross-cultural study. *Curr Anthropol.* 2001; 42(5):760.
85. Hammond RA, Axelrod R. The evolution of ethnocentrism. *J Conflict Resolut.* 2006; 50(6):926–36.
86. Butovskaya ML, Windhager S, Karelin D, Mezentseva A, Schaefer K, Fink B. Associations of physical strength with facial shape in an African pastoralist society, the Maasai of Northern Tanzania. *PLoS One.* 2018; 13(5):e0197738. <https://doi.org/10.1371/journal.pone.0197738> PMID: 29852024
87. Fink B, Grammer K, Mitteroecker P, Gunz P, Schaefer K, Bookstein FL, et al. Second to fourth digit ratio and face shape. *Proc Biol Sci.* 2005; 272(1576):1995–2001. <https://doi.org/10.1098/rspb.2005.3179> PMID: 16191608
88. Schaefer K, Mitteroecker P, Fink B, Bookstein FL. Psychomorphospace—from biology to perception, and back: towards an integrated quantification of facial form variation. *Biol Theory.* 2009; 4(1):98–106.
89. Kleisner K, Kočnar T, Tureček P, Stella D, Mbe Akoko R, Třebický V, et al. African and European perception of African female attractiveness. *Evol Hum Behav.* 2017 38(6):744–55.
90. Fink B, Grammer K, Thornhill R. Human (*Homo sapiens*) facial attractiveness in relation to skin texture and color. *J Comp Psychol.* 2001; 115(1):92–9. <https://doi.org/10.1037/0735-7036.115.1.92> PMID: 11334223
91. Matts PJ, Fink B, Grammer K, Burquest M. Color homogeneity and visual perception of age, health, and attractiveness of female facial skin. *J Am Acad Dermatol.* 2007; 57(6):977–84. <https://doi.org/10.1016/j.jaad.2007.07.040> PMID: 17719127
92. Choi I, Nisbett RE, Norenzayan A. Causal attribution across cultures: Variation and universality. *Psychol Bull.* 1999; 125:47–63.