Behavioral Ecology and Sociobiology (2026) 80:7
https://doi.org/10.1007/s00265-025-03680-7

RESEARCH

Light is not always right: peri-iridial lightness reduces attractiveness via perceived sex-typicality across human populations

Vojtěch Fiala[1,2] · Juan Olvido Perea-García[2,3] · Petr Tureček[1,4] · Sławomir Wacewicz[2] · Juan David Leongómez[5] · Šimon Pokorný[1] · Ondřej Pavlovič[1] · S. Adil Saribay[6] · Farid Pazhoohi[7] · Robert Mbe Akoko[8] · Karel Kleisner[1]

Received: 24 June 2025 / Revised: 3 December 2025 / Accepted: 8 December 2025
© The Author(s) 2026

Abstract Evolutionary psychology views the human eye as special. In particular, it claims that the light peri-iridial tissues surrounding a relatively darker iris form a combination that sets us apart from other primates. From this perspective, much less attention has been paid to how eye colouration varies between humans, although evidence indicates that variations in peri-iridial and iridial colouration influence both perceived facial attractiveness and sex-typicality. To determine what aspects of eye colouration influence the perception of faces, we have measured the colour of peri-iridial eye tissues ('the white of the eye') and the iris in nine samples from seven distant cultures ($N = 1033$) across three continents. The faces were rated on facial attractiveness and sex-typicality by raters from the corresponding populations. Accounting for the effects of skin lightness, age, and facial shape, we ran a Bayesian multilevel model to estimate global and sample-specific effects of colouration of the iris and peri-iridial tissues on perceived sex-typicality and facial attractiveness. This exploratory, cross-sectional study revealed an overall preference for slightly darker peri-iridial tissues in women, whereby this association was mediated by perceived sex-typicality. Our findings challenge the notion that the light-eyed phenotype is universally preferred by human raters. Instead, they suggest a preference for a moderate phenotype, perhaps because very light peri-iridial tissues are typical of faces which are generally perceived as less feminine. Women with bluer irises were generally perceived as more attractive but findings related to other colour channels and iris features were inconsistent and varied across samples.

Significance statement The morphological variation of human eyes is an understudied phenomenon. While attention has been paid to the alleged uniqueness of human eyes (compared to other primates), little is known about how variations in eye colouration influence human perception of faces. Our study included over 1000 individuals from seven culturally distinct regions, mapped human eye variation, and tested how eye colouration influences perceived attractiveness and sex-typicality. In humans, variation in eye colouration is relatively large and differs across populations. Our findings suggest that it affects the perception of faces. Paradoxically, darker peri-iridial regions (scleras) slightly enhance the perception of femininity and female attractiveness, which challenges the idea that lighter eyes are universally preferred. Moreover, blue/light irises were in some contexts linked to a higher attractiveness. These results further refine our understanding of the role of eye colouration in perceived attractiveness.

Keywords Facial perception · Eye morphology · Trade-offs · Peri-iridial tissues

Introduction

The eye area and its importance in human facial perception

While some human facial features – such as skin colour (Stephen et al. 2009), facial shape averageness (Kleisner et al. 2024), or sex-typicality (Fiala et al. 2021) – have been extensively studied in the context of attractiveness and its correlates, the influence of the eye morphology remains relatively underexplored (but see Russell et al. 2014; Kramer and Russell 2022; Wacewicz et al. 2022). Given how powerfully the eyes draw our attention, even shortly after birth (Farroni et al. 2002), this lack of scientific interest is surprising. It might be due to some earlier comparative studies which neglected the intraspecific variation in peri-iridial eye tissues (tissues surrounding the iris, sometimes referred to as the 'sclera') in humans (Kobayashi and Kohshima 1997). These earlier

Communicated by T. C.M. Bakker

Extended author information available on the last page of the article

Published online: 02 January 2026

comparative findings had a profound and lasting impact on psychological literature, although clinical evidence had long recognised a substantial variation in both iris and peri-iridial tissue colouration across human populations (e.g., Shields et al. 2007). Moreover, the available data come mostly from Western, educated, industrialised, rich, and democratic populations (WEIRD sensu Henrich et al. 2010), which raises concerns about the generalisability of the findings.

Natural colour of eye tissues as a cue to individual health and status

Previous studies found associations between quantitative measurements of the iris and peri-iridial tissues and certain traits related to human health and condition. For example, peri-iridial tissues – but not irises – become slightly darker, yellower, and redder with age (Russell et al. 2014; Kramer and Russell 2022) and may thus function as a cue to ageing. Red and yellow tones of the peri-iridial tissues moreover tend to indicate various pathological conditions, such as jaundice (Proche and Kobos 2004; Leung et al. 2015), cardiovascular diseases, diabetes (Dąbrowska et al. 2020), or an unhealthy lifestyle (Peragallo et al. 2013). Ageing may also lead to a melanin pigmentation of the conjunctiva (part of the peri-iridial tissues) and that is related to an increased risk of melanoma (Shields et al. 2007). Transient reddening of the peri-iridial tissues (conjunctival hyperaemia) is related to heightened emotions (Provine et al. 2013b), irritation of ocular tissues, or infection (Murphy et al. 2007). To sum up, existing findings suggest that the colouration of peri-iridial tissues may play a role in assessing an individual's current physical condition or emotional state.

Aside from these associations with perceived age, health, and attractiveness, the colour and tone of the iris and peri-iridial tissues are probably also sexually dimorphic. There is some evidence that peri-iridial tissues are redder and yellower in males than in females (Gründl et al. 2012; Kramer and Russell 2022) and, at least in Caucasians, the irises of females seem to be lighter than those of males (Danel et al. 2020). Existing studies thus consistently show that there is an association between the appearance of peri-iridial tissues and age, health, and attractiveness, and that the colour of iridial and peri-iridial tissues is slightly sexually dimorphic.

Light vs dark peri-iridial tissues, blue vs brown iris: Experimental studies

Ample experimental evidence from perceptual studies suggests that, at least in WEIRD populations, the colour of peri-iridial tissues plays a role in ascribing social or psychological characteristics. When Russell et al. (2014) made the visible areas of peri-iridial tissues darker, yellower, and redder, the faces were assessed as older, less healthy, and less attractive. Kramer and Russell (2022) have observed an interaction between the colour of peri-iridial tissues and perceived sex-typicality: the versions of faces with redder and yellower peri-iridial tissues were perceived as more masculine. Provine et al. (2011) found that artificially reddened peri-iridial tissues caused a cropped eye area to be perceived as less youthful, less healthy, and less attractive. In a subsequent study, Provine et al. (2013b) found that reddened peri-iridial tissues were also considered sadder, angrier, more disgusted, and more fearful than the un-reddened eye control group. Wacewicz et al. (2022) used realistic, artificially generated faces of hominins, which were described to participants as 'apemen'. Their results have shown that prosocial and generally positive characteristics (younger, healthier, attractive, more trustworthy, less aggressive) were ascribed to individuals with lighter – as opposed to darker – peri-iridial tissues. Using the same forced-choice alternative design (either light or dark peri-iridial tissues) and purposefully artificial-looking (blue-skinned) faces, Wolf et al. (2023) have reached similar conclusions regarding assessments of trustworthiness.

Such studies indicate that when it comes to the visible areas of the peri-iridial tissues, 'light is right'. Nevertheless, at least two studies (Provine et al. 2013a; Howard 2024) suggest that this is not always the case: facial photographs manipulated to present unnaturally white peri-iridial tissues (whiter than what naturally occurs in the target population) were not rated as more attractive. It seems, therefore, that there might exist a physiological optimum that is close to, but not overlapping with, light peri-iridial tissues, and this colouration elicits the most positive reactions.

Unlike the colour of peri-iridial tissues, iris colouration is strikingly polymorphic across and largely also within human populations. It can range from dark brown to light blue (Mackey et al. 2011) and there is no point on that scale that is considered universally preferred. Nevertheless, iris colour is known to influence both perceived attractiveness and perceived psychological characteristics, such as trustworthiness or dominance. A series of studies by Kleisner et al. (2010, 2013) found that in men, brown iris colour was a significant predictor of higher ratings of trustworthiness and dominance. Moreover, by artificially swapping the eye colour, they demonstrated that this effect was driven by morphological facial features which typically accompany brown eyes. On the other hand, a replication by the same team (Kočnar et al. 2012) found no effect of the iris colour on perceived characteristics (see also Russell et al. 2014). Gründl et al. (2012) showed that iris colour did not affect attractiveness ratings in female eyes but pointed towards the existence of a blue-eyes stereotype, where blue eyes were viewed as an attractive feature by the study's participants.

A problem: lack of diversity of raters and stimuli

A critical weakness of the existing research on human ocular appearance is the insufficient demographic diversity of the studied populations (Perea-García et al. 2025). This applies to both the participants and the stimuli: the participants have been sampled almost exclusively from WEIRD populations and the stimuli showed mostly faces from ethnically uniform WEIRD backgrounds. Although research suggests that living in non-WEIRD societies significantly affects the formation of first impression, a coherent characterisation is lacking. For example, locals may place less (Scott et al. 2014) or more (DeBruine et al. 2011) emphasis on sex-typicality. Moreover, the preferences of non-WEIRD populations are likely to be less shaped by urbanisation, mass media, and the internet (Batres and Perrett 2014), and generally less 'westernised'. Notably, as of 2024, around one third of the global population still does not have access to the internet, whereby in non-WEIRD countries this proportion is even less favourable (International Telecommunication Union 2024).

Non-WEIRD (and specifically non-Western) societies also show a greater variability in the emphasis placed on the eye area than people from WEIRD societies do. For instance, people in South-East Asia and East Asia have been shown to focus more on the eye area when expressing and interpreting emotions (Jack et al. 2012; Park et al. 2013) and during dyadic interaction (Haensel et al. 2022; Hessels et al. 2025). Some evidence also shows that individuals from East Asia (Japan) view a direct gaze as less favourable (e.g., angrier, less pleasant) than people from Europe and North America (Akechi et al. 2013).

Most importantly, non-WEIRD societies remain underrepresented in scientific datasets on ocular morphology. This means that research not only lacks data on how local variations in ocular morphology affect the assigned characteristics but it also lacks data on variation in the ocular morphology as such. Jointly, the differing cultural attitudes to the eye region and the relative scarcity of morphological data on non-WEIRD populations undermine any assumption of a broad cross-cultural consistency in the social effects of the eye phenotype effect during the formation of first impression based on the face.

The studies on the role of eyes in facial perception cited above were conducted mostly in WEIRD societies: In Russell et al. (2014) and Kramer and Russell (2022), both the raters and the stimuli were from WEIRD populations (EU/USA), in Provine et al. (2011, 2013b), the raters were US-based students, making for an ethnically diverse but still WEIRD sample, while in Wolf et al. (2023), the participants were children from the United States. Wacewicz et al. (2022) have crowdsourced participants from across the globe but most raters in that study still came from WEIRD countries. Finally, studies by Kleisner et al. (2010, 2013), Kočnar et al. (2012), Gründl et al. (2012), and Provine et al. (2013a) all worked with samples from the EU and the United States. Two further studies, which nuanced our understanding of the role of the limbal ring and sclera colouration (Howard 2024) and the role of the variability of eye morphology in females (Prantl et al. 2019) in perceived attractiveness were also conducted in WEIRD societies.

Apart from a Chinese-Japanese study by Kuraguchi et al. (2022), which focused on the effects of the iris size, the only non-WEIRD study on the effect of eye colour on perceived characteristics we are aware of is Kočnar et al. (2019), who, however, focused just on the iris colour. In that study, Czech facial stimuli, pre-selected to have either unambiguously brown ($N=60$) or blue eye colour ($N=60$), were rated by an ethnically diverse sample that included European (Czechia, Sweden, Estonia, Romania, Portugal), Asian (India, Turkey), African (Namibia, Cameroon), and South American raters (Brazil). The results showed that only female Portuguese participants rated the portraits of blue-eyed Czech men as more attractive than those of brown-eyed Czech men, and Turkish male raters preferred Czech blue-eyed women. These results are of crucial importance because they demonstrate that populations do vary in their preferences. In order to provide context and nuance to the conclusions of previous research, one should thus explore the perceptions of diverse populations, and that should also involve the use of locally relevant stimuli rather than relying solely on WEIRD samples, as in Kočnar et al. (2019).

The current study

Existing research thus indicates that the colour of peri-iridial tissues functions as a cue to youth and health, including reproductive health and thus also fertility, thus adding to other facial traits perceived as feminine, healthy, and attractive. The perceptual effect of iridial colouration is less straightforward: although we know it plays a role in the rating of attractiveness and other features, its effect probably varies between individuals, populations, and social contexts. As indicated above, a major limitation of the studies which examined the effects of phenotypic variation in human eyes on the ascribed characteristics are the samples on which research was done: they came mostly from European and North American populations. The existing research thus neglects a majority of human populations.

Several remedies could be proposed. Creation of a large, diverse, and standardised facial database would be a good start. Availability of non-WEIRD photosets is currently growing (Saribay Adil et al. 2018; Lakshmi et al. 2021) but, unfortunately, the procedures of stimuli acquisition in the various samples vary, so that cross-sample standardisation remains problematic. Although one can work with unstandardised datasets by analysing each dataset separately (Kočnar et al. 2019; Fiala et al. 2021), it increases the likelihood of false positive results unless the *p*-values are corrected for multiple testing. As a consequence, even large effects can become formally insignificant. While it may be statistically appropriate, this can lead to a dismissal of biologically meaningful associations. One could also apply meta-analytic methods, but their use is more justified when only the effect sizes are available, and not the raw data.

Because our primary aim at present is to explore and estimate the effect sizes and because we have access to all individual data points, we have decided that a Bayesian multi-level model approach should be feasible and allow us to characterise the associations between facial variation and first impressions. Bayesian models can lend support to hypotheses based on the posterior distributions of estimated coefficients instead of *p*-values. Moreover, the trends which appear in analyses can be discussed in reasonable detail, that is, beyond a mere confirmation or rejection of a hypothesis on a selected alpha level. On top of that, the Bayesian approach allows us to specify complex hierarchical models and estimate the magnitude and uncertainty of both the overall effect across all samples ('global' effects) and of particular effects within each sample. The risk of overfitting is lowered by a partial pooling during the parameter estimation.

In our study, we have thus opted for a Bayesian multilevel model. We took a large cross-cultural sample of faces ($N=1033$) and measured a relevant set of ocular and facial features including facial shape, skin lightness, and iridial and peri-iridial lightness and hue. On top of that, we provided estimates of the phenotypic variation of eye tissues in several distinct human populations.

Evidence for a systematic impact of facial morphology and age on perception

The effect of the eye phenotype on first impressions is potentially integrated with the role of previously identified whole-face predictors. Nevertheless, past research – including studies previously published by our team (Fiala et al. 2021; Kleisner et al. 2024) – has not examined eye-related and whole-face predictors together within a single analytical framework.

Given that the eye phenotype seems to exhibit a slight sexual dimorphism (Danel et al. 2020), it is plausible to claim that eye morphology is just one component of the overall sex-typicality of facial morphology. A morphometric study by Kleisner et al. (2021) demonstrates that the eye region is actually amongst the most sexually dimorphic areas of the face, and that this dimorphism cannot be attributed to allometry (Kleisner et al. 2021). In a similar vein, distinctive facial shapes, which are perceived as unattractive across cultures (Kleisner et al. 2024) – potentially due to placing higher demands on cognitive processing (Trujillo et al. 2014) – may be related to a distinctive eye morphology.

Eye colour may be integrated with the overall facial coloration due to a shared genetic background (Katsara and Nothnagel 2019). On top of that, it has been shown that skin lightness predicts perceived attractiveness and sex-typicality but its effects vary across cultures (Coetzee et al. 2014; Kleisner et al. 2017).

Aside from this, the studies cited above also indicate that certain aspects of eye morphology are age-dependent but any age-elicited changes in the eye morphology in our sample will probably be small due to a small age variance (see Table 1). As a result, they are unlikely to affect the results. Nonetheless, it should be noted that age has been identified as a predictor of perceived characteristics, including attractiveness (He et al. 2021), and it predicted both perceived attractiveness and sex-typicality in our previous research (Fiala et al. 2021). All in all, existing evidence supports our decision to include in our analysis as predictors of perceived characteristics – alongside variables related specially to ocular morphology – also whole-face morphometric measures, age, and skin lightness.

Insights from previous research

Our results will be interpreted in the light of the following previously reported findings:

1) Lighter peri-iridial tissues are more attractive (Provine et al. 2011; Gründl et al. 2012; Russell et al. 2014), while redder, yellower, and possibly also darker peri-iridial tissues should have aversive effects. This tendency should be stronger in countries with a lower life expectancy, worse health, and lower accessibility of medical care, such as Cameroon or India (according to the Human Development Index; United Nations 2023).
2) Blue, though not necessarily light, eyes are perceived as more attractive (Kočnar et al. 2019; Gründl et al. 2012). This finding, however, comes from studies where participants rated stimuli collections with significant iris colour variance, which for many populations is not the case. For example, in Vietnamese and Cameroonian samples, no iris could be classified as 'blue' (Mackey et al. 2011). Due to a negative frequency-dependent selection (as proposed by Frost 2006, as well as Forti and Young 2016), blue-iris preference might be more pronounced in countries where blue or very light irises

Table 1 Descriptive statistics of the raters' and stimuli samples

Sample	Sex of the stimuli	No. of Stimuli	Age Stimuli (mean ± SD)	Rated Trait	No. of Raters	Age Raters (mean ± SD)
Cameroon 2013	Women	50	21.24 ± 1.89	Attractiveness	34	22.21 ± 3.14
				Femininity	77	24.00 ± 4.12
	Men	49	22.00 ± 2.24	Attractiveness	28	22,11 ± 3.89
				Masculinity	77	24.00 ± 4.12
Cameroon 2016	Women	49	22.59 ± 4.92	Attractiveness	49	22.96 ± 3.23
				Femininity	94	22.99 ± 3.00
	Men	49	22.00 ± 3.00	Attractiveness	51	23.37 ± 4.25
				Masculinity	94	21.99 ± 3.00
Colombia	Women	66	20.71 ± 2.77	Attractiveness	432	20.71 ± 2.77
				Femininity	432	20.71 ± 2.77
	Men	72	20.38 ± 2.42	Attractiveness	565	21.85 ± 4.81
				Masculinity	565	21.85 ± 4.81
Czechia 2016	Women	50	23.64 ± 4.33	Attractiveness	33	28.18 ± 4.21
				Femininity	44	30.00 ± 9.68
	Men	50	24.04 ± 3.92	Attractiveness	89	27.56 ± 4.23
				Masculinity	231	32.04 ± 7.60
Czechia 2019	Women	56	22.61 ± 3.89	Attractiveness	35	31.89 ± 6.80
				Femininity	60	34.30 ± 9.21
	Men	39	22.54 ± 4.55	Attractiveness	76	32.43 ± 9.61
				Masculinity	63	34.08 ± 8.98
India	Women	52	27.73 ± 9.01	Attractiveness	33.8*	33.51 ± 8.48
				Femininity	33.8*	33.51 ± 8.49
	Men	89	26.96 ± 7.86	Attractiveness	19.5*	33.51 ± 8.50
				Masculinity	19.5*	33.51 ± 8.51
Iran	Women	43	20.63 ± 1.09	Attractiveness	46	37.30 ± 12.30
				Femininity	33	27.73 ± 3.77
	Men	44	20.64 ± 1.60	Attractiveness	41	34.88 ±.9.91
				Masculinity	31	29.35 ± 4.54
Turkey	Women	93	21.2 ± 1.51	Attractiveness	1207	22.09 ± 3.66
				Femininity	1207	22.09 ± 3.66
	Men	91	21.52 ± 1.95	Attractiveness	1207	22.09 ± 3.66
				Masculinity	1207	22.09 ± 3.66
Vietnam	Women	32	21.16 ± 2.68	Attractiveness	86	22.20 ± 3.76
	Men	59	21.1 ± 1.86	Attractiveness	124	22.96 ± 4–26

*Average No. of raters per face. Altogether, 878 raters [238 Females] met the authors' inclusion criteria (Lakshmi et al. 2021)

are present but rare, such as India, Colombia, Turkey, or Iran. All this applies to members of historical local populations and leaves aside the topic of recent immigrants.

3) Lighter, though not necessarily blue, irises are perceived as more attractive (based on Cossu et al. 2024).
4) Redder and yellower peri-iridial tissues contribute to the perception of masculinity in males and are negatively associated with the perception of femininity in females (based on Kramer and Russell 2022).
5) Lighter irises are associated with the perception of femininity in women (Danel et al. 2020), but it should be noted that this result was reported only for Caucasian populations.
6) Literature consistently shows that facial preferences (perceived attractiveness) are related to perceived sex-typicality, especially in women (e.g., Fiala et al. 2021). It is therefore reasonable to expect that the effects of ocular morphology on attractiveness may be mediated by perceived sex-typicality. We will use path analyses to examine this possibility. The analyses will be conducted both on the pooled sample and for each subsample separately. These path analyses are intended to explore the joint effects of perceived sex-typicality and eye morphology on perceived attractiveness.

Methods

Blinded methods were used when appropriate: All facial photos were collected by researchers who were blind to the research hypotheses and all facial measurements were collected independently of the rating data. The participants who rated the facial photographs were blind to the goals of the study.

Acquisition of facial photographs

Details of the procedure of acquisition of facial photographs are presented in Table S1, which also contains the basic descriptive statistics of stimuli included in this study (for the Table, see online Supplementary Materials). The photoset 'Chicago Face Database – Indian Expansion' is described in detail in Lakshmi et al. (2021). The procedure of acquisition of facial photos in Turkey (Bogazici FD) can be found in Saribay Adil et al. (2018). In Cameroon, the Czech Republic, and Vietnam, photographs were taken using a standardised photo acquisition procedure described in Třebický et al. (2016). Acquisition procedure for the Iranian and Colombian dataset can be found in the supplementary materials of Fiala et al. (2021), and the procedure for stimuli acquisition in Vietnam is detailed in Pavlovič et al. (2023).

Participants in each of the nine samples (Table 1) were recruited via online advertising (social media, email), by direct invitation, or through fliers displayed around places where the stimuli acquisition later took place (usually campuses at local universities) and distributed in the country of stimuli collection. The persons who had responded to these calls were then invited to an isolated area where it was possible to control the lighting conditions (e.g., by blocking the windows). They were then asked to remove accessories such as glasses, jewellery, or cosmetics (makeup etc.). The sole exception were Iranian women who were allowed to keep their hijab. In most samples (see column 'neutral clothes' in Table S1 in Supplementary Materials), participants wore a black T-shirt. In most cases, the photographer instructed them to look directly in the camera and adopt a neutral facial expression. In all cases, artificial light was used to light the scene. Photographs were taken either against a neutral grey/white background or in a photographic tent with a white surface. The procedure allowed researchers to maintain stable lighting conditions during each stimuli acquisition session. For more details on the camera settings, see Table S1 in Supplementary materials.

Prior to analyses, photos from the Cameroonian, Czech, and Vietnamese population were colour-corrected in the Adobe Photoshop Lightroom. This procedure adjusted the colour profile and white temperature of the images and converted them from a raw camera format to.jpg (which was then entered into further analyses). We did not participate in stimuli acquisition and processing of the Indian sample; with respect to that sample, we rely on the high standards that had been observed by authors of the Chicago Face Database (see Ma et al. 2015). In the Turkish, Colombian, and Iranian sample, no colour correction was done but at least one of the authors participated in the photo acquisition and confirmed that the camera settings and lighting conditions remained unchanged during the stimuli acquisition and that the photographs were exposed in the same colour profile throughout each sample. In connection with recent findings by Laitly et al. (2021), we should also note that colour differences across-samples seem to affect our conclusions less than previously assumed.

The rating of facial stimuli

The raters (Table 1) were always from the same country as the rated stimuli. For rating, they used a seven-point Likert scale anchored by 1 (very unattractive or not at all masculine/feminine) and 7 (very attractive or very much masculine/feminine). In the case of Colombia, the scale differed (it was 0–10) and individual ratings were afterwards recalculated to the seven-point scale. Then we have averaged the ratings for each face (see Fiala et al. 2021): while this can reduce the variance, it does not affect the per-face (stimulus) estimates of the measured characteristics. The analysis reported below is based on these pooled data. Nevertheless, since individual ratings were available in all samples except one (India), we had also run an ordered logistic regression using individual rating data as the response variable.

Geometric morphometrics

To objectively characterise facial shape, we have used geometric morphometrics based on a paradigm of 36 anatomically or geometrically invariant facial landmarks and 36 semilandmarks, which characterise curved features and outlines. In the Iranian women, the exact position of the semilandmarks that characterise the jawline had to be in some cases estimated because the edge of the jaw was hidden by the hijab. An alternative analysis that omits Iranian faces is reported in the online Supplementary Materials.

Using the gpagen() function of the Geomorph for R package (Collyer and Adams 2018, 2024; Baken et al. 2021; Adams et al. 2024), we have subjected the faces in each set to the Generalised Procrustes Analysis. Then we calculated three kinds of shape metrics based on the aligned Procrustes residuals: Distinctiveness (Dist), Facial Asymmetry (Asym), and Shape Sextypicality (SST). It had been previously shown that these metrics predict both the perceived attractiveness and perceived sex-typicality. In each individual face, the Procrustes distance of its configuration from the sample mean characterises its distinctiveness (Kleisner et al. 2019). Distinctiveness was calculated separately for each sex, whereby the higher the value, the higher the distance between the shape of that face and the sample average.

The position of a face on an axis connecting male and female faces (that had been scaled to zero mean and variance unity, and then multiplied by −1 in women) served as a measure of shape sex-typicality (SST) of that face (Kleisner et al. 2024). Male and female configurations were subjected to the same analysis but computations were run separately for each sample.

Facial asymmetry was calculated by first reflecting the group-specific Procrustes residuals laterally along the midline axis. Then we relabelled the paired landmarks on the left and right side of the face by swapping the numeric indices of the left-side landmarks with those of the right-side landmarks and vice versa. Finally, we calculated the Procrustes distances between the original and the mirrored configuration; higher asymmetry scores indicate greater facial asymmetry (Mardia et al. 2000).

The facial colour and the eye colour

For each stimulus, we have measured facial skin lightness as well as the colour of the iris and the visible area of peri-iridial tissues in the CIELab colour space using the ColorTransformer2 plugin in ImageJ (Schneider et al. 2012). We have selected three squares of the facial skin (one on each cheek and one roughly in the middle of the forehead) and measured the average CIELab L* of these squares to characterise facial skin lightness. This is a standard method for describing surface lightness and colour intensity in a colour space aligned with human perception, often referred to as the 'standard observer' (Hunter 1958; McLaren 1976). For peri-iridial and iris colour measurements, we have recorded the values of the L* (dark–light), a* (red–green), and b* (blue–yellow) channel. We have manually selected the iris, measured its L*a*b* values, and then repeated the process for the peri-iridial tissues. We avoided sampling portions with specular reflections and took care not to accidentally measure the pupil instead of the iris. The distribution of colouration of the iris and peri-iridial tissues is shown in Fig. 1, which was prepared with the vioplot package (Adler et al. 2024).

Sample pooling and variable scaling

Within the samples, the stimuli acquisition procedure was invariant but slight differences between the nine samples were nevertheless present, such as the camera and lens type,

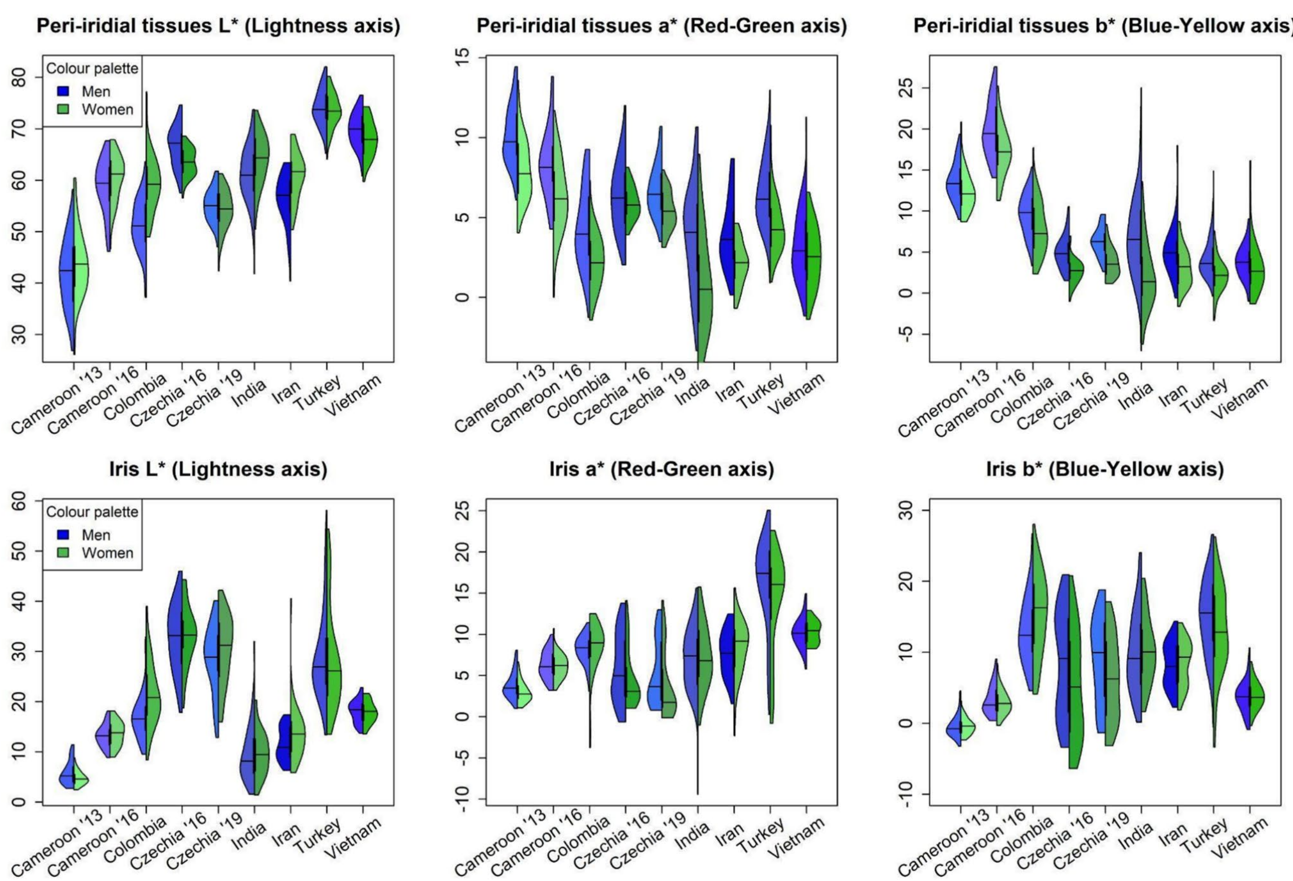

Fig. 1 Violin plots describing (left to right) the distribution of lightness, redness, and yellowness (CIELab L*a*b* axes) of peri-iridial tissues and iris in men (left) and women (right half-violin) in the nine samples. The diagrams show that the distribution spans differ substantially between distant populations rather than within populations or between geographically adjacent samples. Note that the dataset was compiled primarily in order to address the knowledge gap regarding an association between the formation of the first impression and eye morphology. One therefore should not draw any definite conclusions about eye phenotype in the corresponding populations based on our samples

the exact location, standardisation of clothing, or the year (see Table 1). These variations justify our recommendation not to treat the dataset as representative of the colour variance in the sample populations: the abovementioned differences may have led to a slightly different colouration of the stimuli and thereby influence the ratings and ultimately also the modelled estimates.

We have designed the analyses primarily to estimate the effects of predictors across all samples (average effects). To account for differences between the samples, we have included varying terms to estimate the effects of predictors within each sample. In populations for which more than one sample was available (Czech Republic, Cameroon), the effects were estimated separately for each sample because the colour between the samples had slightly varied (Fig. 1).

The studied populations differed in the variance of the predictors, including the eye colour (see Fig. 1). This must be considered because low predictor variance can render the variable irrelevant to the perceived characteristics (see Strom et al. 2012). On the other hand, a standardisation of predictors within each sample would lead to the loss of potentially interesting differences in sample means and variances.

The proposed solution was to standardise across all samples ('overall standardisation') using pooled means and standard deviation, but such overall standardisation can also be problematic because raters always assessed faces from their own population, that is, the population they have the greatest experience with. It is therefore most likely that they are familiar with and accustomed to the extent of within-population variance. As a result, they may be able to incorporate that knowledge in their ratings (Lall and Tanaka 2023) even if that variation is small and cannot be recognised and used by people from distant populations. Moreover, when standardising the variables across pooled samples, the analysis may slightly inflate the overall effects by allowing more variable samples to dominate the analysis. The overall estimates could thus in effect show how the variance would affect human preferences if it were universally present but deliver unreliable predictions for populations where the variance is absent.

In short, standardisation across the entire pooled sample facilitates comparison of parameter estimates between groups but may obscure group-specific variability. On the other hand, standardisation within each subsample preserves group-specific variance but limits direct comparability of parameters between groups.

We have therefore decided to conduct both types of analyses. Here, in the main text, we present the analysis with predictors standardised across all samples (overall standardisation), because this preserves the differences in variance across samples. The alternative analysis, where predictors were standardised across each population, can be found in the Supplementary Materials (Part 3, Figs. S19–S22). While the effect sizes are slightly smaller, probably due to the across-sample variance being reduced by the standardisation procedure, the results consistently point in the same direction. Readers are encouraged to compare the outcomes of both types of analyses. The same document also includes extended and alternative analyses, which are briefly summarised below.

Statistical analyses

Analysis 1: predicting attractiveness and sex-typicality using the measured eye colour

Our research goal was to estimate whether and how the colour (a*b*) and tone (L*) of the iris and peri-iridial tissues affect ascribed sex-typicality and attractiveness in different cultures, and whether the plausible parameters estimated by Bayesian updating based on our data (cf. van Doorn et al. 2021) support conclusions from previous literature (e.g., 'lighter is better' for peri-iridial tissues).

We have conducted two Bayesian analyses: in the first, the dependent variable was attractiveness, while in the second, it was perceived sex-typicality (male masculinity, female femininity). The second analysis used a smaller sample ($N=942$) because masculinity and femininity ratings were not available for the Vietnamese faces. The analyses take into consideration the level of individual data points (faces), higher level units (samples). They also allow for sharing and summarising relevant information across contexts and for estimating random variance, thus adhering to the common rules of multi-level analysis (Leyland and Groenewegen 2020; Newsom 2024).

Our primary focus was on the effects of iris and peri-iridial lightness, redness, and yellowness. Additional independent variables in the analysis were facial skin lightness (L*), age, and facial shape (distinctiveness, corresponding to distance of a face from its sample mean, facial asymmetry, and sexual shape dimorphism).

Following Kleisner et al. (2024), we have estimated the overall ('global') effect of each predictor via fixed terms at two levels: one for women, one for men. This way, the model estimated the effects of eye morphology across the whole sample. Subsequently, we have introduced a varying term that estimated the effect of each predictor ($N=12$) in each sample ($N=9$ [8 in the case of sex-typicality]) and each sex ($N=2$). The varying terms (intercepts and slopes) were estimated as correlated with Cholesky, non-centred decomposition (function compose_noncentered()). Drawing on our previous study (Kleisner et al. 2024), we have used conventional weakly regularising priors centred at zero for the bivariate linear relationship between the ascribed characteristics and the eye phenotype. The priors allowed for a full range of slopes while improving the sampling efficiency by restricting values to a range between app. −1 and 1, because our study used data standardised to zero mean and variance unity. The model structure, further detailed in Part 2 of the Supplementary Materials, allowed us to estimate both the overall effects and effects specific to a particular sample while optimising the sampling performance. Furthermore, due to a partial pooling on the level of sample-specific terms, the model accounted for variable size of individual samples, which improved the precision of such terms' estimates. The results of a sensitivity analysis shows that the model predictions were not affected by the selection of priors.

In all samples except one, we were also granted access to individual ratings. To ensure that the results based on per-face averages are not artifacts of averaging, we have also fitted models using individual ratings as a unit of analysis (for details, see Supplementary materials, Part 5). Additionally, we have conducted analyses with free exponents in order to estimate the shape of the association between colour-related variables and the two perceived scales. We have repeated the main analysis on trimmed data (excluding observations which were over 1.5-times interquartile range below the first quartile or above the third quartile). On top of that, we have fitted a model in which we merged samples from similar geographical background: our aim was to see whether differences in the eye morphology between samples from similar backgrounds (Fig. 1) may have biased the original estimates. All these analyses are reported in the Supplementary Materials.

The models were entered using the rethinking package (McElreath 2024). The posteriors were sampled with the Stan infrastructure using the Hamiltonian Monte Carlo method. We drew the distribution plots based on the posteriors using base R graphics (R Core Team 2024) and edited them in InkScape (ver. 1.2.2). Moreover, since our use of the Bayesian analysis did not require testing for the null hypothesis, we did not run an a priori power analysis for the main analyses, reported in Fig. 2 and Fig. 3.

Analysis 2: associations between peri-iridial lightness, sex-typicality, and attractiveness: a path analysis

The results of Analysis 1 suggested that, in women, attractiveness and sex-typicality are negatively associated with the lightness of peri-iridial tissues. This finding contradicts a substantial body of research. Our previous study moreover suggested a close relation between attractiveness and sex-typicality in women (Fiala et al. 2021). Therefore, we have further inspected the relationship between peri-iridial tissue lightness, measured sex-typicality of facial shape, and perceived femininity and attractiveness.

We have fitted a Bayesian multiple regression where perceived femininity predicted perceived attractiveness, while perceived femininity and attractiveness were both predicted by measured sex-typicality, peri-iridial tissue lightness, and age. Age and peri-iridial tissue lightness also served as predictors of shape sex-typicality (see Fig. 4, the upper row). Subsequently, we inspected whether the estimated strength of these associations changes when the path between perceived femininity and attractiveness is not considered.

In the final version of this analysis (Fig. 4, the lower row), we have omitted from the analysis all variables except for peri-iridial tissue lightness, perceived femininity, and attractiveness. Using the standardised data from above, this analysis inspected the *global* association between iris lightness and the two perceived scales. It also explored whether the strength of the association between these three variables changes if some paths are 'closed' (not fitted in the model). Therefore, this analysis did not consider the varying slopes and intercepts for different samples. As before, we used standard weakly regularising priors centred at zero. Furthermore, we prepared a diagram of the path analyses in InkScape. Path analyses conducted separately for each sample are available in the Supplementary Materials (Part 4).

Results

The effects of variation in eye colouration on the first impressions

Attractiveness

Below, we report both the overall and sample-specific effects of selected variables (CIELab L*, a*, b* colour channels of the iris and peri-iridial tissues) on facial perception. In the following, the first value corresponds to the mean of posterior distribution of the reported bivariate effect, while the numbers in square brackets show the 95% percentile-based compatibility interval. In women, higher attractiveness was associated with darker peri-iridial tissues (β=−0.30 [95% CI: −0.57; 0.00]) and we found a tendency to perceive lighter irises as more attractive (β=0.22 [95% CI: −0.08; 0.48]). In the Colombian (β=0.41 [95% CI: 0.09; 0.83]) and Turkish (β=0.34 [95% CI: 0.09; 0.60]) sample, the positive association between iris lightness and perceived attractiveness was credibly non-zero.

Bluer irises (lower CIELab b* values) predicted higher attractiveness ratings in women (β=−0.20 [95% CI: −0.39; −0.02]) but not in men (β=0.00 [95% CI: −0.22; 0.19]). In

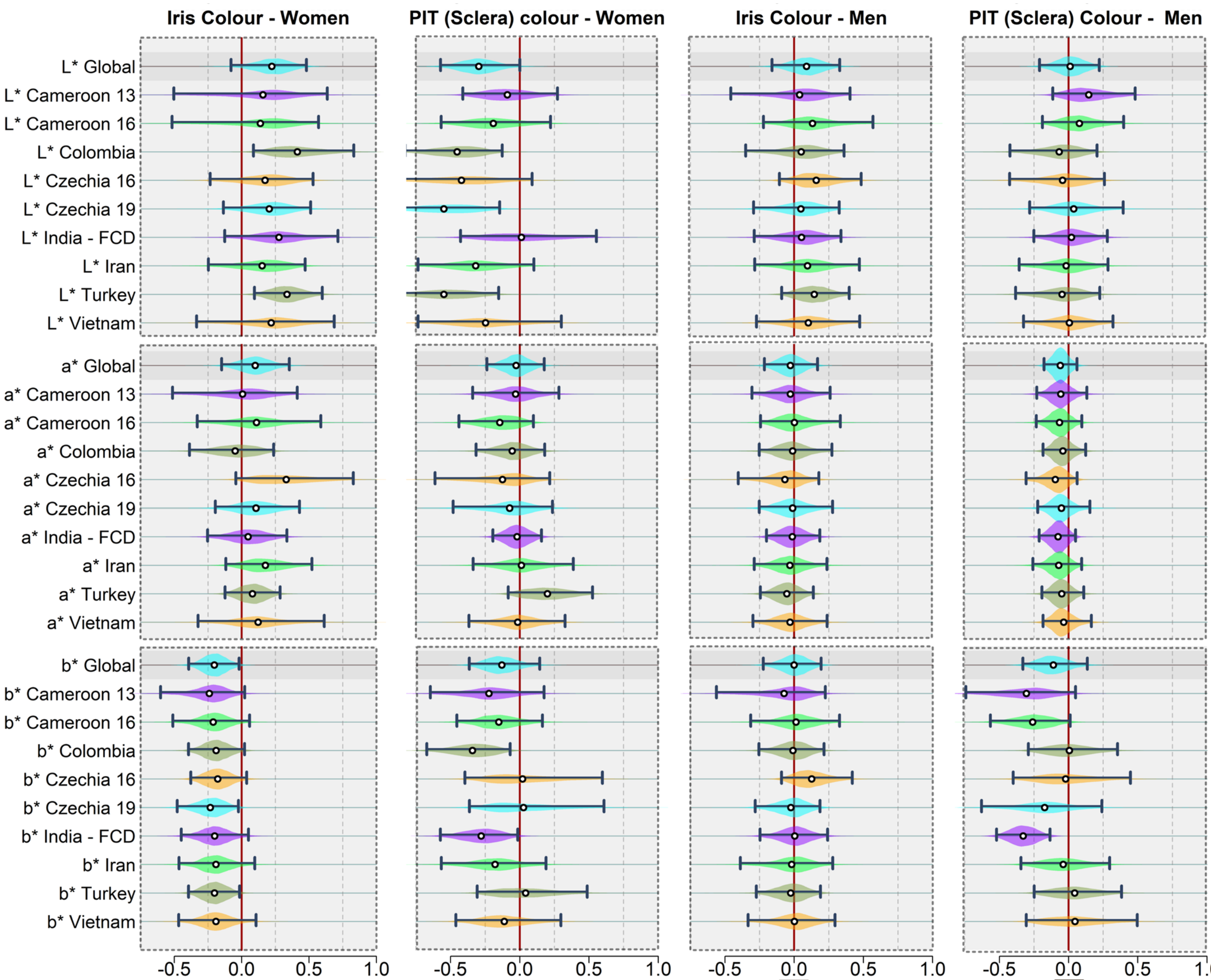

Fig. 2 Associations between perceived attractiveness (N=1033) and selected eye-colour variables related to the iris and peri-iridial tissues. L*a*b* corresponds to CIELab L*a*b* standardised continuous predictors. The horizontal bar with vertical whiskers corresponds to 95% CI. Corresponding coefficients are printed to the left of each posterior distribution. PIT=peri-iridial tissues

women, some of the sample-specific effects of peri-iridial tissue b* were also credibly non-zero: less yellow peri-iridial tissues were found credibly more attractive in the Indian (β=−0.28 [95% CI: −0.57; −0.01]) and Colombian female sample (β=−0.34 [95% CI: −0.67; −0.07]).

The effects were considerably weaker in men, except for the association between peri-iridial b* measures and perceived attractiveness. The overall negative trend of bluer tones being perceived as more attractive than yellowish ones (β=−0.11 [95% CI: −0.33; 0.14]), was credibly non-zero in the Indian male sample (β=−0.33 [95% CI: −0.52; −0.13]).

In the model that predicted attractiveness, the role of the different colour channels was less strictly separated than in the model that predicted perceived sex-typicality (see below). Consequently, the credibility intervals in the model predicting attractiveness were wider, indicating a higher uncertainty of the predictions. Associations between attractiveness, perceived female femininity, and measured predictors are further inspected in Section 3.2 below.

Perceived masculinity and femininity (perceived sex-typicality)

Women were rated for femininity and men were rated for masculinity (these ratings were not available for the Vietnamese sample). In women, the global effect of peri-iridial lightness on perceived femininity was credibly non-zero (β=−0.37 [95% CI: −0.64; −0.10]), meaning that female faces with very light peri-iridial tissues were perceived as less feminine. This effect is unlikely to be an artifact of the different means and ranges of peri-iridial lightness across

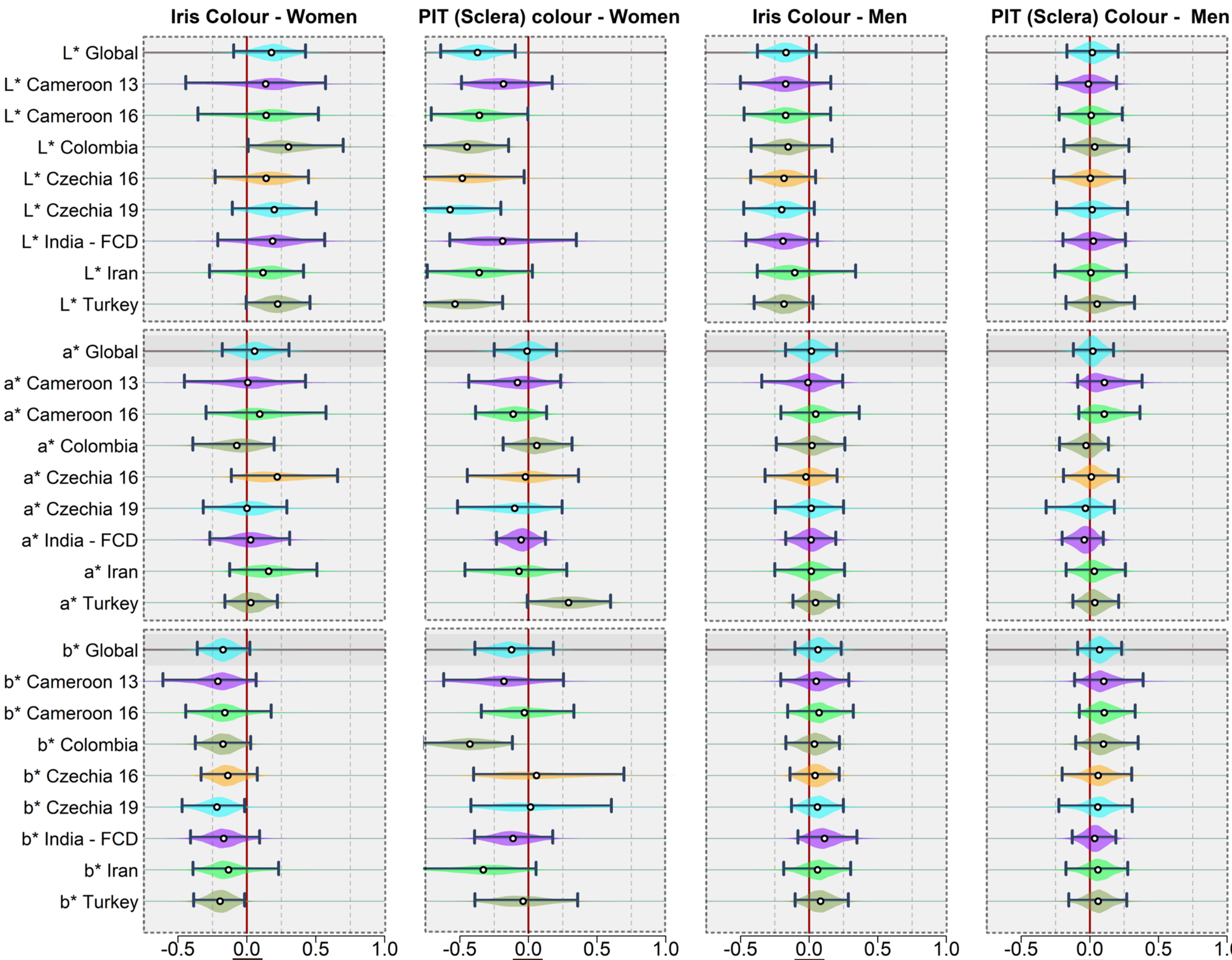

Fig. 3 Associations between perceived sex-typicality (perceived femininity in women, perceived masculinity in men; N=942) and selected eye-colour variables relating to the iris and peri-iridial tissues

cultures. It was credibly non-zero in one-half of the samples, and in the remaining samples we found a trend pointing in the same direction. On top of that, we observed the same trend when the data were standardised for each culture separately (β=−0.18 [95% CI: −0.36; 0.02]), see Fig. S20 in the online Supplementary Materials. This robust pattern suggests that, in women, peri-iridial tissues which appear light beyond a certain point are perceived as less feminine.

For the other two channels of the peri-iridial tissues (a*b*) we have likewise found in the female samples a trending association (β=−0.12 [95% CI: −0.39; 0.18]) between perceived sex-typicality and the blue–yellow colour channel (b*) in the peri-iridial tissues. It was particularly pronounced in the Colombian subsample, where we observed a credibly non-zero effect (β=−0.43 [95% CI: −0.77; −0.12]) showing that neutral or bluer tones of female peri-iridial tissues are perceived as more feminine. In the Turkish women, a reddening of the peri-iridial tissues was found to be conclusively perceived as more feminine (β=−0.29 [95% CI: −0.01; −0.60]). Variance in red–green channel (a*) of the CIELab colour space had no credibly non-zero effect on the ascribed female femininity or male masculinity.

In the Colombian (β=0.30 [95% CI: 0.01; 0.70]) and Turkish (β=0.22 [95% CI: 0.00; 0.46]) sample, lighter female irises were credibly associated with higher femininity ratings, and the global trend pointed in the same direction (β=0.18 [95% CI: −0.10; 0.43]). A trend (β=−0.17 [95% CI: −0.35; 0.02]) had also suggested that bluer irises (b* channel) in women are perceived as more feminine. The effect was credibly non-zero in the Turkish sample (β=−0.19 [95% CI: −0.38; −0.02]) and in one of the two Czech subsamples (2019: β=−0.22 [95% CI: −0.47; −0.02]) but not in the other (β=− 0.14 [95% CI: −0.33; 0.08]).

In men, iris and peri-iridial L*, a*, b* colour variance had only weak effects and neither iridial (β=−0.17 [95%

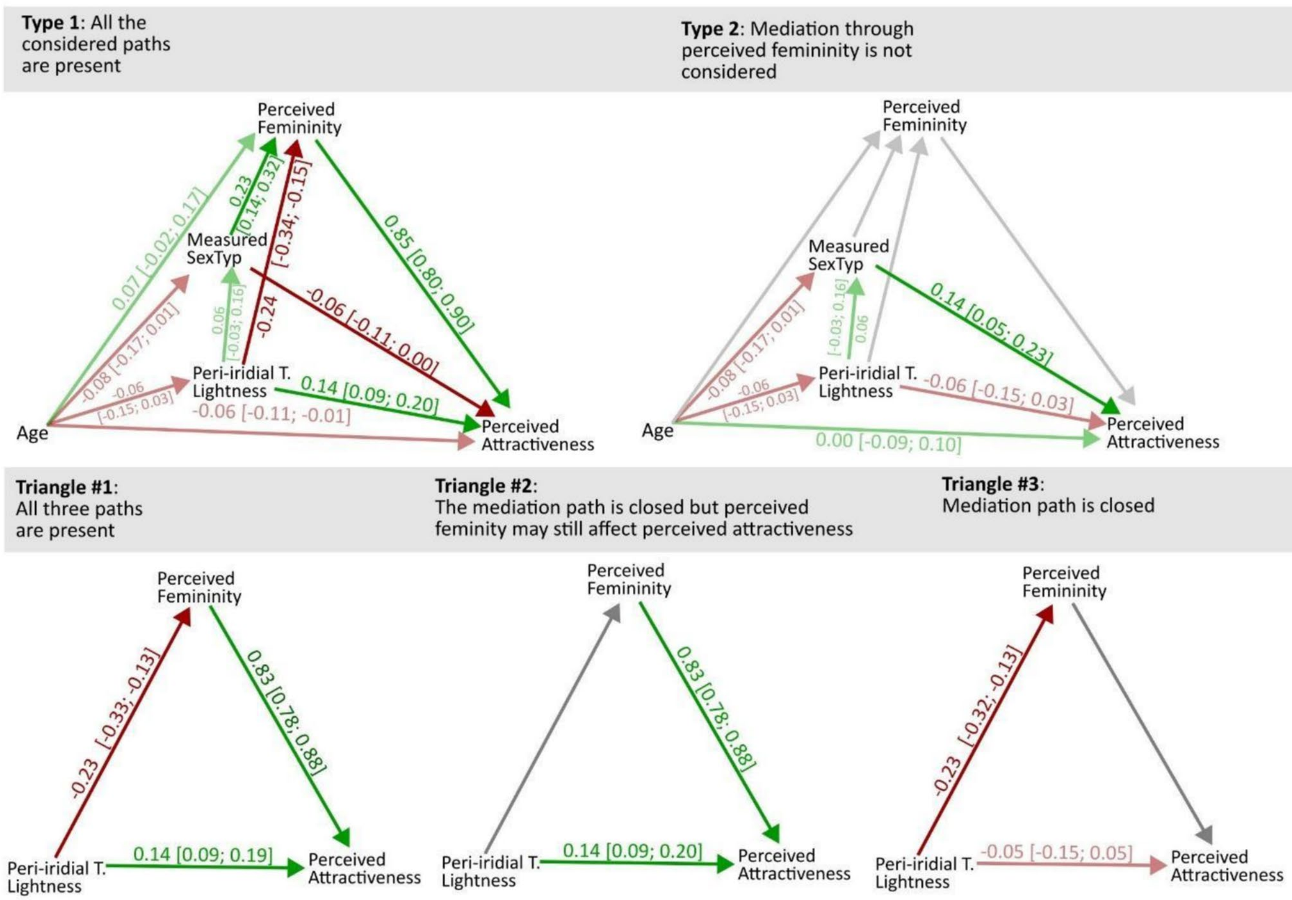

Fig. 4 Bayesian path analysis. In women (N=459), we introduced a path analysis to further inspect the association between attractiveness, perceived sex-typicality (femininity), and measured variables treated as their predictors. Triangles in the lower row depict path analyses where only perceived femininity, attractiveness, and the measured lightness of peri-iridial tissues are included. The paths in *grey* were not estimated in that particular analysis. Change in the magnitude and the prevailing polarity (positive × negative) suggest that the negative association between peri-iridial lightness and perceived attractiveness is mediated by perceived femininity. For a similar diagram for males, see Fig. S26 in the Supplementary materials

CI: −0.38; 0.05]) nor peri-iridial (β=0.02 [95% CI: −0.17; 0.21]) lightness predicted perceived masculinity. Nevertheless, we found one trending association for iris lightness, namely a tendency to perceive darker iris as more masculine (β=−0.17 [95% CI: −0.38; 0.05]). This trend was most pronounced in Turkish men (β=−0.18 [95% CI: −0.40; 0.03]) and in both Czech male samples (2016: β=−0.18 [95% CI: −0.38; 0.05]; 2019: β=−0.20 [95% CI: −0.38; 0.04]).

Facial shape and skin colour: in line with earlier findings

Peri-iridial and iris colour did predict perceived attractiveness and sex-typicality (Figure 2-3) but were not their sole predictors. The models also considered measured facial shape, age, and skin lightness (CIELab L*). In the model fit, attractiveness was credibly negatively associated with distinctiveness in men (β=−0.13 [95% CI: −0.21; −0.05]), while in women we found a trend pointing in the same direction (β=−0.16 [95% CI: −0.34; 0.03]). No effect was observed for asymmetry (men β=0.07 [95% CI: −0.04; 0.18], women β=0.06 [95% CI: −0.02; 0.14]). In women, attractiveness was positively associated with a more female-typical facial shape (β=0.10 [95% CI: 0.01; 0.20]) and negatively associated with age (β=−0.13 [95% CI: −0.24; −0.02]), but the strongest association was observed between skin lightness and female attractiveness (β=0.47 [95% CI: 0.12; 0.77]).

Perceived masculinity and femininity were predicted by measured sex-typicality: the estimated slope of the bivariate association was 0.16 [95% CI: 0.05; 0.25] for women's femininity and 0.17 [95% CI: 0.08; 0.26] for men's masculinity. Relatively younger men were also perceived as more masculine (β=0.31 [95% CI: 0.14; 0.50]). Skin colour was a credible predictor of both perceived masculinity in men (β=−0.35 [95% CI: −0.67; −0.03]) and perceived femininity in women (β=0.55 [95% CI: 0.17; 0.88]), whereby lighter-skinned men were perceived as less masculine and lighter-skinned women as more feminine. The effect of the skin colour on perceived masculinity and femininity was thus moderate to strong, which may have, due to collinearity, inflated the uncertainty of the effect of lightness of peri-iridial tissues on perceived female femininity and attractiveness when both predictors were present.

To further inspect this possibility, we have repeated the analyses from the article without considering skin lightness (reported in Fig. S22 and S23 in Supplementary Materials). The results still showed that perceived femininity and attractiveness tend to be negatively associated with the lightness of peri-iridial tissues, but the credibility intervals spanned from negative to positive values (β=−0.26 [95% CI: −0.56; 0.04] for perceived femininity, β=−0.20 [95% CI: −0.52; 0.12] for attractiveness). These wider compatibility intervals suggest that models without skin colour perform worse when it comes to isolating the effect of peri-iridial tissue colour. This finding supports the hypothesis of independent roles of skin lightness and peri-iridial tissue lightness in attractiveness and sex-typicality ratings, which becomes obvious only if both variables are included in the same model.

Alternative models

We have fitted models using individual ratings as the response variables. The ordered logistic regression followed the structure of the main analysis reported above. The results of this model based on individual ratings support the conclusions regarding the effects of eye morphology on the formation of facial first impressions. We have again observed a weak-to-moderate negative association between attractiveness and sex-typicality and the lightness of peri-iridial tissues. Standard deviations were relatively similar at the level of raters and at the level of stimuli for both rating scales.

First, we want to inspect whether there was evidence of a non-linear association between the lightness of peri-iridial tissues and the two rated scales. To do so, we have fitted models with ‘free exponents’, which estimated the exponent that characterised the shape of the association between the perceived scales and variables related to skin tone and eye morphology (iris and peri-iridial L*a*b* colour values). One of the models also tested a strictly quadratic relationship by fixing the exponent at 2. Based on a WAIC model comparison, linear models showed good out-of-sample predictive accuracy. Also, the models with free exponents suggested that an exponent of approximately 1 was a credible estimate of the association between peri-iridial lightness and the perceived scales in women.

We have fitted counterfactual plots to illustrate the association between the dependent variables and peri-iridial lightness. These plots are available in online Supplementary Materials (Section 6). Jointly, the results and their visualisation suggest that the association between peri-iridial lightness and ascribed characteristics is approximately linear.

Aside from that, we have rerun the linear models after excluding datapoints that fell over 1.5 of the interquartile range units below 25th percentile or above 75th percentile. While the associations became slightly weaker, the overall conclusion remained unchanged (see Section 7 in the Supplementary Materials). Similarly, estimates changed only minimally when we merged the data into country-specific categories (Section 8 in the Supplementary Materials).

The role of peri-iridial lightness and perceived sex-typicality in predicting female attractiveness: a path analysis

A previous study has revealed an association between perceived female femininity and attractiveness that was so close as to render the scales nearly identical (Fiala et al. 2021). To understand the abovementioned unexpected pattern of findings, we have conducted a path analysis where perceived female attractiveness (N=459) was predicted by measured sex-typicality and peri-iridial tissue lightness, but also by perceived femininity. We have also estimated the bivariate effect of measured sex-typicality and peri-iridial tissue lightness on perceived sex-typicality.

Comparing the two alternative paths, the analysis revealed that attractiveness is nearly perfectly predicted by perceived sex-typicality (β=0.85 [95% CI: 0.80; 0.90]). In this layout (Fig. 3), facial attractiveness was actually positively associated with peri-iridial tissue lightness (β=0.14 [95% CI: 0.09; 0.20]). Peri-iridial tissue lightness also predicted perceived femininity but the direction of this association was conclusively negative (β=−0.24 [95% CI: −0.34; −0.15]), just as in the main analysis reported above. These associations held even when only peri-iridial lightness, perceived femininity, and perceived attractiveness were kept in the analysis, which shows that the negative effect of peri-iridial tissue lightness on female attractiveness is due to its close (negative) association with femininity. Once this is accounted for, peri-iridial tissue lightness is positively associated with attractiveness, which is in line with previous investigations.

Discussion

In this multicultural study, we explored how natural variation in iridial and peri-iridial colouration relates to perceived attractiveness and sex-typicality in men and women. Our results regarding a preference for lighter peri-iridial tissues importantly nuance existing understanding of the subject. While lighter peri-iridial tissues can contribute to an attractive appearance, our results show that this is not consistently the case. We found a credibly non-zero negative effect of peri-iridial tissue lightness on perceived female attractiveness in a large, cross-cultural sample. Furthermore, when examining each subsample separately, the association between peri-iridial tissue lightness and attractiveness was credibly negative in half of the female samples. Similar effects were observed for perceived sex-typicality. A path analysis revealed that this pattern is due to a close association between the two rating scales in women, where peri-iridial lightness primarily influences the perceived sex-typicality, which in turn predicts perceived attractiveness.

Furthermore, the expected negative effects of peri-iridial redness and yellowness on perceived attractiveness were quite limited. In particular, we observed credible non-zero adverse effects of yellow tones in peri-iridial tissues only in the Turkish and Indian female sample. Our results thus show that, regardless of the popular heuristic, lighter is *not* necessarily better (i.e., more attractive) and the effects of other colour channels in the peri-iridial tissues are very limited.

Regarding the iris colour, we followed the prediction made by Kočnar et al. (2019) according to whom blue irises may be more attractive than brown ones. Overall effects estimated across the entire dataset indicate a preference for lower b* in female irises. In the CIELab colour space, the b* channel spans from yellow (high values) to blue (low values). The model thus suggests that in populations where blue irises are present, females with bluer irises tend to be perceived as more attractive. In populations where blue irises are absent, raters show a preference for less yellow, though not blue, irises.

When examining individual samples, the trends pointed in the same direction across all female samples, although the effect was credibly non-zero only in two: the Turkish and Czech (2019) samples. These findings contribute to our emerging understanding of how iridial appearance influences attractiveness perceptions across human populations. While there is a tendency to prefer lighter female irises, credible intervals for the relevant coefficients include in this case zero.

No systematic association was found between male iris or peri-iridial colour and the assigned characteristics. Specifically, both redness and yellowness of the peri-iridial tissues seem unimportant for perceived male attractiveness and sex-typicality.

These results are relatively robust and supported by alternative analyses. Regardless of whether we trimmed extreme values, used individual ratings, or modelled non-linear association between the eye morphology and perceived characteristics, the resulting predictions remained consistent with those reported in the main models.

We found a pattern of credibly non-zero associations for perceived female sex-typicality and attractiveness. The path analysis revealed that this is driven by a close relationship between the two rating scales in women; in particular, it suggests that eye morphology influences primarily the perceived sex-typicality, which in turn predicts perceived attractiveness.

Peri-iridial colouration, attraction, and perceived sex-typicality

Darker, but not red or yellow, visible peri-iridial areas may be preferred

As detailed in Section 1, existing research associates lighter peri-iridial tissues with higher perceived healthiness and attractiveness (e.g., Russell et al. 2014). It has also been claimed that peri-iridial tissue colouration is sexually dimorphic: in men, these tissues are redder and yellower (Kramer and Russell 2022). Moreover, redder and yellower peri-iridial tissues affect the judgments of sex-typicality, because they are perceived as more masculine and less feminine. Our exploratory study did not replicate this result. We found that darker female peri-iridial tissues are perceived as more feminine and therefore also as more attractive. This finding is unexpected.

One proposed explanation, namely that our result is an artifact of modelling a linear association between the lightness of peri-iridial tissues and the perceived characteristics instead of an n-shaped association, is not supported. Results from models with non-linear terms (quadratic exponent, free exponent) suggest that a weak linear decline does parsimoniously characterise the association between peri-iridial lightness and the perceived scales. But we do not assume that the relationship is strictly linear across the full range of human peri-iridial lightness. Very dark peri-iridial tissues are likely to have an adverse effect on perception. Similarly, at the other extreme, Provine et al. (2013a) show that peri-iridial tissues that were made artificially lighter were not preferred over natural (slightly darker) peri-iridial tissues. Very light eyes might perhaps appear unnatural or be associated with pathological conditions (see below). Eyes with very dark peri-iridial tissues, on the other hand, may be perceived as atypical for humans, and therefore unattractive. A complete characterisation of this association may not be feasible within a cross-sectional, exploratory study using naturally occurring ranges of peri-iridial lightness. Experimental approaches, such as those used by Wacewicz et al. (2022), are better suited to identifying the lightness or darkness thresholds at which the relationship bends.

Lighter peri-iridial tissues could be actually preferred but they might be associated with facial features which are perceived as less feminine and therefore less attractive. This is a relevant consideration, because eye colour is intricately linked to other aspects of facial morphology. Previous research into associations between facial structures and eye (iris) colour has identified such relationships between ocular and facial features and proposed they might be due to pleiotropy (Kleisner et al. 2010, 2013).

In contrast to previous research, we found no credibly non-zero effects of peri-iridial tissue redness on perceived attractiveness and femininity. On the other hand, the preference for less yellow tones, which was credibly non-zero for the Indian and Colombian sample, pointed in the anticipated direction (Provine et al. 2013b; Kramer and Russell 2022). Still, the fact that preference for less yellow peri-iridial tissues was observed only in the Colombian and Indian sample suggests that, at least in natural unmodified photos, the effects found by Kramer and Russell (2022) with regard to perceived masculinity/femininity are population-dependent.

It has been shown that in countries with vertically structured societies and hierarchies based on castes and ethnicities, more pronounced pigmentation is perceived as a cue to belonging to a lower caste or an ethnic minority, and it lowers female attractiveness (Wagatsuma 1967). In both India and Colombia, more pigmented peri-iridial tissues might be a cue to belonging to a non-privileged group/ethnicity, which would then feed into a stereotypical social classification. An alternative explanation refers to jaundice as a cue to hepatitis and impaired health in general. In that case, however, we would see yellowish tones being likewise non-preferred (at least) in Cameroon, which is a low-income country with relatively inaccessible measures against viral hepatitis (World Health Organization 2024).

Once again, the instances where we failed to find evidence of preference for less yellow peri-iridial tissues might be due to the limited trait variation the raters were exposed to. According to Fig. 1, the Indian sample had a relatively large variation in the b* channel (yellow–blue colour channel) for peri-iridial tissue colouration, but other samples were less variable. Therefore, and this is potentially the case also for the effects of iris colouration, the absence of effects of peri-iridial tissue colouration (a*b*) in most samples could be due to a limited variability of the stimuli.

Adaptive preference for darker peri-iridial tissues

As noted above, it has been shown that reddish tones of the peri-iridial tissues adversely affect attractiveness and other desirable perceived facial characteristics because they function as a potential cue to impaired health (Provine et al. 2011, 2013a). Moreover, bulging eyes appear lighter because a larger part of the eyeball is visible, and such appearance may likewise serve as a cue to impaired health. Exophthalmos, a condition where the eye bulges from its orbit, is related to a wide range of serious health issues (Timmis 1957), including the Graves–Basedow disease, which is an autoimmune disorder of the thyroid. The disease is relatively prevalent in the general population (0.5–2%) but ten-times more frequent in women than in men (Wémenau et al. 2018), which may explain why we found a preference for dark peri-iridial tissues mainly in the female samples. This explanation fits well with our reasoning that light eyes appear feminine and attractive only up to a certain point (Provine et al. 2013a). Moreover, there exist in some cultures certain superstitious beliefs related to 'scleral show' that is smaller or larger than average (Guillaume and Burkat 2025). For instance, in Japan, the term 'Sanpaku eyes' refers to a stereotyped cue to criminal inclinations (superior scleral show) or bad fate (inferior scleral show). Scleral show may also be related to bulging eyes and therefore explained by same physiological reasons as exophthalmos.

Femininity is sometimes confounded with youthfulness, because both are more pronounced in young female faces (Muñoz-Reyes et al. 2015). On the other hand, female facial sex-typicality is a cue to sexual maturity and an important sign of fertility (Pflüger et al. 2012), which is why juvenile and feminine features differ. Smaller and darker peri-iridial tissues may account for the difference: peri-iridial tissue lightness may be a good proxy for the relative eye size especially in female faces, which lack the generally more robust masculine traits, including the brow ridge.

Regarding masculinity, however, our study found no association between peri-iridial lightness and perceived male sex-typicality although male faces are generally considered more masculine when they have a more 'robust' anatomy, including in the eye area, which implies relatively smaller eyes and a more pronounced brow ridge (Kleisner et al. 2021). This anatomy causes relatively less light to fall on the eyes, which should result in relatively darker-looking eye tissues. As a result, one would expect that darker peri-iridial tissues should be related to a higher perceived masculinity. Our null effect may be due to the photo-acquisition methods we have adopted; the light was aimed straight at the face of the stimulus person, while the shadow cast by a pronounced male brow ridge would have played a role if the light were coming from above (as is typically the case for daylight).

Irises, attraction, and perceived sex-typicality

Bluer irises

The only two populations where the effect of female irises' b* channel on perceived attractiveness was credibly non-zero were the Turkish and Czech (2019) sample. In these samples, where less yellow/bluer irises were preferred, a portion of the stimuli can be actually classified as blue-eyed. Moreover, in the case of Turkey, this finding is in good agreement with Kočnar et al. (2019), who found that Turkish male raters preferred blue-eyed Czech women. In the Vietnamese, Iranian, and to a lesser degree also the Cameroonian sample, a large portion of the posterior probability spans over zero; in these populations, the iris colour can only be classified as ranging across various hues of brown (Mackey et al. 2011), while blue, green, and hazel-coloured irises are absent. This was also the case for India, where, although the variability was large, irises were still on average very dark (Low L*, see lower-left panel in Fig. 1).

It seems, therefore, that bluer irises are perceived as more attractive but only in populations where blue eyes are present. The within-sample variance of b* probably also played a role. Nevertheless, neither the presence/absence of blue irises nor the colouration range in a sample explain the data fully. In some highly variable samples (CZ 2016, India), the association between the b* channel of the iris and perceived attractiveness could be null, but the mean slope (based on the posterior probability distribution) was similar in all female samples.

Lighter irises

Blue eyes are usually, though not always, lighter, so the preference for bluer iris tones should go hand in hand with a preference for lighter irises. But while we found an overall tendency to prefer lighter female irises (which is in line with recent literature such as Cossu et al. 2024), on the level of individual samples the preference for lighter irises was credibly non-zero only in two. First of all, we found a moderately strong preference for lighter female irises in the Turkish female sample, which had the relatively largest iris lightness range and some proportion of blue-eyed women. This is not surprising, because the predecessors of the Turkish Republic (the Ottoman and, before that, the Byzantine Empire) were multiethnic and multicultural societies with an immense variation of human phenotypes. Secondly, this preference was also found for the sample of Colombian women, which was not only highly variable in terms of iris colour (Fig. 1) but also ethnically diverse. The Colombian population is a mixture of persons of European, Mestizo, African, and Native American descent (Homburger et al. 2015), and the preference for lighter colours may in fact reflect a preference for traits of the privileged group (in this case, European traits; see Castillo 2009). In the rest of the samples, we observed no credible effects of iris lightness. In men, however, there was a tendency to assign higher masculinity ratings to men with darker irises, which is in line with a previously observed positive association between brown eyes and ratings of dominance, a correlate of masculinity, in men (Kleisner et al. 2010).

To conclude, it seems that a complex interplay of differing within-population variation (Vietnam, Cameroon), multiethnic and stratified population (Colombia, India), and past evolutionary processes (European populations) has generated a pattern where the preferences regarding iris colouration differ across samples. Our data show such a pattern, albeit mainly in women. In men, the effects of iris colouration were most likely null.

The origins of variation in iris colour, which is especially pronounced in Europe, are unknown: the phenomenon may be due to frequency-dependent sexual selection on rare phenotypes, where blue eyes were the desirable rare type (Frost 2006), or it could be an adaptation compensating for the lower amount of light during winters at higher latitudes (Workman et al. 2018). In both cases, the shift could be coupled with preference for a newly introduced phenotype, although that does not necessarily imply that vestiges of such preference are present in the current European (e.g., Czech) populations, especially if various phenotypes are distributed relatively evenly.

Our data also allow us to discuss the potential sexual dimorphism of iris colour. A previous study by Danel et al. (2020) identified sexual dimorphism in the relative iridial illuminance among Caucasians (using one of the Czech samples we used here, namely 2016). Figure 1 suggests that iris colour is slightly sexually dimorphic, although testing it was not a goal of our study. On the other hand, lighter female irises were present mainly in the Czech sample from 2019. We attribute this difference to the measure used in the study: while Danel et al. (2020) used relative illuminance, which measures the relative contrast in the luminance of peri-iridial tissues and the iris, we used the absolute CIELab L* values.

The primary function of the eye is vision

Aside from significantly contributing to the formation of first impressions, facial morphology also serves various other functions (Třebický et al. 2019) – and that applies to the eye morphology as well. The primary function of the eye is vision, which is why we should expect that majority of the variation in eye phenotypes should be explicable by the demands of optimising vision in the local environment. Photo-regulation is an important driver of the external appearance of the eyes in primates (Perea-García et al. 2022) and even mammals in general (Caspar et al. 2023). Photoregulatory functions limit the range of variability which the pressures related to signalling functions and sexual selection can then act upon and maintenance of visual acuity certainly contributes to fitness more than the latter pressures. In short, sexual selection or signalling need not play a major role in facial perception or, specifically, in the communicative or cueing role of human eyes. The small or null effects could be due to this reduced importance.

Limitations

Our study has a number of limitations. Firstly, the lighting conditions slightly differed across the studied datasets. To partly account for the effect of different setups, we ran the analysis both on data standardised within and across the cultures and compared the results of the two. On top of that, we drew only limited conclusions based on the absolute values. An optimal approach, however, would involve datasets created with the same basic settings and under the same lighting conditions.

Even samples from the same country, although collected under standardised conditions, differed in some aspects of facial morphology (Fig. 1). The main challenge for future studies, therefore, is to ensure identical stimuli acquisition procedures, because only so can one make sure that any observed effects are not driven by changes in the photographic setup. In a similar vein, our conclusions regarding the effects of ageing on facial morphology and perceived characteristics are limited due to a relatively small age variance of our stimuli. Nevertheless, studies in evolutionary psychology typically focus on participants in young reproductive age, and so does our study. The data thus fit well into this paradigm and allow for a direct comparison with the results of other studies.

While our goal was to explore the effect of eye morphology on particular whole-face characteristics, the focus on the effect of an isolated facial feature (such as the eye) on whole-face characteristics remains a potential concern. Future studies should, therefore, also collect rating data focused solely on the eye region, and then check for potential associations between the perceived vs. measured eye characteristics and the perceived vs. measured features of the whole face. We should reiterate here that our study was exploratory and used nonmanipulated stimuli. In most of our samples, we are not aware of any previously published data on eye morphology and its effect on facial perception. This lack of data prevented us from predicting what alternation of the eye morphology would elicit detectable effects while remaining within a natural range.

As Danel et al. (2020, 2023) and Pokorný et al. (2024) have shown, there is a broad variety of other facial characteristics, such as feature contrast, eye surface area, and eye fissure shape, which should also be taken into consideration, at least to an extent where the resulting complexity could still be viably handled by available analytic tools. Building on this, future studies should also examine relative (as opposed to absolute) luminance, as well as the effect of eye shape and the relative size of the eye fissure.

Conclusions

We found an overall preference for less light female peri-iridial tissues and less yellow female irises. It seems, however, that when it comes to iris colour preference, the range of colour variation within a population is an important factor, while culture and other locally specific conditions play a role as well. Our results regarding the colouration of peri-iridial tissue nuance the existing research. We found that while the effects of peri-iridial tissue redness and yellowness were limited, the negative association between peri-iridial tissue lightness and both attractiveness and femininity was more stable. A path analysis had shown that women with relatively light peri-iridial tissues are perceived as slightly less sex-typical (feminine), which in turn makes them less attractive. On the other hand, we found no conclusive effects in the male samples. These results highlight the importance of ecologically valid stimuli: experimental studies may find effects which are not necessarily relevant to real life, thus leading to heuristics that are ultimately limited. The population-specificity of many of our effects highlights the desirability of investigating diverse populations using diverse stimuli.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00265-025-03680-7.

Acknowledgments Acknowledgements: We would like to thank Dr. Anna Pilátová for proofreading the manuscript. We thank the reviewers for their thoughtful and constructive comments.

Author contribution VF and JOPG designed the research and drafted the initial version of the manuscript. VF, PT, and KK designed and developed the analyses and VF and PT prepared visualisations for the results. VF, SW, PT and JOPG wrote later versions of the manuscript, and ŠP, OP, KK, JDL, FP, SAS and RMA collected and provided facial stimuli and rating data. All the authors have read and approved the final version of the manuscript.

Funding Open access publishing supported by the institutions participating in the CzechELib Transformative Agreement. This work was supported by the Czech Grant Agency (Project No. 24-11720S). Data collection in Colombia was supported by Universidad El Bosque, Vice-Chancellor of Research (grant number PCI.2017–9444).

Data availability All the data, scripts, and supplementary materials are available at [https://github.com/VojtechFiala/EyeProject_Paper_One].

Declarations

Ethics approval The experimental protocol adhered to all the relevant institutional, national, and international guidelines, including the Helsinki Declaration. This study ensured anonymity of all the participants, with no information or images that could lead to their identification. All procedures were approved by the Institutional Review Board of the Faculty of Science at Charles University (protocol ref. number 09/2023). As this study was based exclusively on human participants and their data, approval from an animal ethics committee was not required.

Informed consent All participants provided informed consent prior to their involvement in the study.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adams DC, Collyer ML, Kaliontzopoulou A, Baken EK (2024) Geomorph: Software for geometric morphometric analyses. R package version 4.0.9, https://doi.org/10.32614/CRAN.package.geomorph

Adler D, Kelly ST, Elliott T, Adamson J (2024) vioplot: violin plot. R package version 0.5.0, https://github.com/TomKellyGenetics/vioplot

Akechi H, Senju A, Uibo H, Kikuchi Y, Hasegawa T, Hietanen JK (2013) Attention to eye contact in the west and east: autonomic responses and evaluative ratings. PLoS One 8:e59312. https://doi.org/10.1371/journal.pone.0059312

Baken EK, Collyer ML, Kaliontzopoulou A, Adams DC (2021) Geomorph v4.0 and gmShiny: enhanced analytics and a new graphical interface for a comprehensive morphometric experience. Methods Ecol Evol 12:2355–2363. https://doi.org/10.1111/2041-210X.13723

Batres C, Perrett DI (2014) The influence of the digital divide on face preferences in El Salvador: people without internet access prefer more feminine men, more masculine women, and women with higher adiposity. PLoS One 9:e100966. https://doi.org/10.1371/journal.pone.0100966

Caspar KR, Hüttner L, Begall S (2023) Scleral appearance is not a correlate of domestication in mammals. Zool Lett 9:12. https://doi.org/10.1186/s40851-023-00210-z

Castillo SS (2009) Discourse and racism in Colombia: five centuries of invisibility and exclusion. In: van Dijk TA (ed) Racism and discourse in Latin America. Lexington Books, New York, pp 131–170

Coetzee V, Greeff JM, Stephen ID, Perrett DI (2014) Cross-cultural agreement in facial attractiveness preferences: the role of ethnicity and gender. PLoS One 9:e99629. https://doi.org/10.1371/journal.pone.0099629

Collyer ML, Adams DC (2018) RRPP: an R package for fitting linear models to high-dimensional data using residual randomization. Methods Ecol Evol 9:1772–1779

Collyer ML, Adams DC (2024) RRPP: Linear Model Evaluation with Randomized Residuals in a Permutation Procedure, R package version 2.0.4, https://doi.org/10.32614/CRAN.package.RRPP

Cossu M, Trupia MG, Estes Z (2024) Beauty is in the iris: constricted pupils (enlarged irises) enhance attractiveness. Cognition 250:105842. https://doi.org/10.1016/j.cognition.2024.105842

Dąbrowska E, Harazny J, Wolf J, Szulc U, Narkiewicz K (2020) The eye — a window to cardiovascular diseases. Artif Intell Hypertens 24:56–60

Danel DP, Perea-Garcia JO, Lewandowski Z, Szala A, Fedurek P, Kleisner K, Wacewicz S (2023) Is there no beauty in sexually dimorphic eyes? Facial attractiveness and white Europeans ocular morphology - brief communication. PLoS One 18:e0284079. https://doi.org/10.1371/journal.pone.0284079

Danel DP, Wacewicz S, Kleisner K, Lewandowski Z, Kret ME, Żywiczyński P, Perea-Garcia JO (2020) Sex differences in ocular morphology in Caucasian people: a dubious role of sexual selection in the evolution of sexual dimorphism of the human eye. Behav Ecol Sociobiol 74:115. https://doi.org/10.1007/s00265-020-02894-1

DeBruine LM, Jones BC, Little AC, Crawford JR, Welling LL (2011) Further evidence for regional variation in women's masculinity preferences. Proc R Soc Lond B Biol Sci 278:813–814. https://doi.org/10.1098/rspb.2010.2009

Farroni T, Csibra G, Simion F, Johnson MH (2002) Eye contact detection in humans from birth. Proc Natl Acad Sci U S A 99:9602–9605. https://doi.org/10.1073/pnas.152159999

Fiala V, Třebický V, Pazhoohi F, Leongómez JD, Tureček P, Saribay SA, Mbe Akoko R, Kleisner K (2021) Facial attractiveness and preference of sexual dimorphism: a comparison across five populations. Evolutionary Human Sciences 3:e38. https://doi.org/10.1017/ehs.2021.33

Forti IRN, Young RJ (2016) Human commercial models' eye colour shows negative frequency-dependent selection. PLoS One 11:e0168458. https://doi.org/10.1371/journal.pone.0168458

Frost P (2006) European hair and eye color. A case of frequency-dependent sexual selection? Evol Hum Behav 27:85–103. https://doi.org/10.1016/j.evolhumbehav.2005.07.002

Gründl M, Knoll S, Eisenmann-Klein M, Prantl L (2012) The blue-eyes stereotype: do eye color, pupil diameter, and scleral color affect attractiveness? Aesth Plast Surg 36:234–240. https://doi.org/10.1007/s00266-011-9793-x

Guillaume G, Burkat CN (2025) Global ophthalmology and superstition: a review on eye-related cultural beliefs, myths, and superstition. In: Burkat CN, Hartstein ME (eds) Global Oculoplastics. Academic Press, London, pp 23–35

Haensel JX, Smith TJ, Senju A (2022) Cultural differences in mutual gaze during face-to-face interactions: a dual head-mounted eye-tracking study. Vis Cogn 30:100–115. https://doi.org/10.1080/13506285.2021.1928354

He D, Workman CI, Kenett YN, He X, Chatterjee A (2021) The effect of aging on facial attractiveness: an empirical and computational investigation. Acta Psychol 219:103385. https://doi.org/10.1016/j.actpsy.2021.103385

Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? Behav Brain Sci 33:61–83

Hessels RS, Iwabuch T, Niehorster DC et al (2025) Gaze behavior in face-to-face interaction: a cross-cultural investigation between Japan and the Netherlands. Cognition 263:106174. https://doi.org/10.1016/j.cognition.2025.106174

Homburger JR, Moreno-Estrada A, Gignoux CR et al (2015) Genomic insights into the ancestry and demographic history of South

America. PLoS Genet 11:e1005602. https://doi.org/10.1371/journal.pgen.1005602
Howard MC (2024) Why do limbal rings cause faces to appear healthier and more attractive? Resolving theoretical and methodological tensions in current research on perceptions of faces with(out) limbal rings. Evol Behav Sci. https://doi.org/10.1037/ebs0000354
Hunter RS (1958) Photoelectric color difference meter. J Opt Soc Am A 48:985–995
International Telecommunication Union (2024) Facts and figures 2024 – internet use. ITU Statistics https://www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-internet-use/
Jack RE, Garrod OGB, Yu H, Caldara R, Schyns PG (2012) Facial expressions of emotion are not culturally universal. P Natl Acad Sci USA 109:7241–7244. https://doi.org/10.1073/pnas.1200155109
Katsara MA, Nothnagel M (2019) True colors: a literature review on the spatial distribution of eye and hair pigmentation. Forensic Sci Int Genet 39:109–118. https://doi.org/10.1016/j.fsigen.2018.12.001
Kleisner K, Kočnar T, Rubešová A, Flegr J (2010) Eye color predicts but does not directly influence perceived dominance in men. Pers Individ Diff 49:59–64. https://doi.org/10.1016/j.paid.2010.03.011
Kleisner K, Kočnar T, Tureček P, Stella D, Mbe Akoko R, Třebický V, Havlíček J (2017) African and European perception of African female attractiveness. Evol Hum Behav 38:744–755. https://doi.org/10.1016/j.evolhumbehav.2017.07.002
Kleisner K, Pokorný Š, Saribay SA (2019) Toward a new approach to cross-cultural distinctiveness and typicality of human faces: the cross-group typicality/ distinctiveness metric. Front Psychol 10:124. https://doi.org/10.3389/fpsyg.2019.00124
Kleisner K, Priplatova L, Frost P, Flegr J (2013) Trustworthy-looking face meets Brown eyes. PLoS One 8:e53285. https://doi.org/10.1371/journal.pone.0053285
Kleisner K, Tureček P, Roberts SC, Havlíček J, Varella Valentová J, Mbe Akoko R, Leongómez JD, Apostol S, Varella MAC, Saribay SA (2021) How and why patterns of sexual dimorphism in human faces vary across the world. Sci Rep 11:5978. https://doi.org/10.1038/s41598-021-85402-3
Kleisner K, Tureček P, Saribay SA et al (2024) Distinctiveness and femininity, rather than symmetry and masculinity, affect facial attractiveness across the world. Evol Hum Behav 45:82–90. https://doi.org/10.1016/j.evolhumbehav.2023.10.001
Kobayashi H, Kohshima S (1997) Unique morphology of the human eye. Nature 387:767–768. https://doi.org/10.1038/42842
Kočnar T, Drbalová KR, Kleisner K (2012) Do dominant-looking males have brown eyes? A further investigation of the role of iris colour for dominance perception. Anthropologie 50:25–32. https://doi.org/10.2307/26272384
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. https://doi.org/10.1371/journal.pone.0225549
Kramer SS, Russell R (2022) A novel human sex difference: male sclera are redder and yellower than female sclera. Arch Sex Behav 51:2733–2740. https://doi.org/10.1007/s10508-022-02304-9
Kuraguchi K, Tanabe-Ishibashi A, Ashida H (2022) Large irides enhance the facial attractiveness of Japanese and Chinese women. Acta Psychol 228:103663. https://doi.org/10.1016/j.actpsy.2022.103663
Laitly A, Callaghan CT, Delhey K, Cornwell WK (2021) Is color data from citizen science photographs reliable for biodiversity research? Ecol Evol 11:4071–4083. https://doi.org/10.1002/ece3.7307
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. https://doi.org/10.3389/fpsyg.2021.627678
Lall MK, Tanaka JW (2023) The culture of perceptual expertise and the other-race effect. Br J Psychol 114:21–23
Leung TS, Kapur K, Guilliam A, Okell J, Lim B, MacDonald LW, Meek J (2015) Screening neonatal jaundice based on the sclera color of the eye using digital photography. Biomed Opt Express 6:4529. https://doi.org/10.1364/boe.6.004529
Leyland AH, Groenewegen PP (2020) What is multilevel modelling? In: Leyland AH, Groenewegen PP (eds) Multilevel modelling for public health and health services research: health in context. Springer, Cham, pp 29–48
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. https://doi.org/10.3758/s13428-014-0532-5
Mackey DA, Wilkinson CH, Kearns LS, Hewitt AW (2011) Classification of iris colour: review and refinement of a classification schema. Clin Exp Ophthalmol 39:462–471
Mardia KV, Bookstein FL, Moreton IJ (2000) Statistical assessment of bilateral symmetry of shapes. Biometrika 87:285–300
McElreath R (2024) rethinking: Statistical Rethinking course and book package, R package version 2.42, commit ac1b3b2cda83f3e14096e2d997a6e30ad109eeee, https://github.com/rmcelreath/rethinking
McLaren K (1976) The development of the CIE 1976 (L*a*b*) uniform colour space and colour-difference formula. J Soc Dye Colour 92:338–341. https://doi.org/10.1111/j.1478-4408.1976.tb03301.x
Muñoz-Reyes JA, Iglesias-Julios M, Pita M, Turiegano E (2015) Facial features: what women perceive as attractive and what men consider attractive. PLoS One 10:e0132979. https://doi.org/10.1371/journal.pone.0132979
Murphy PJ, Lau JSC, Sim MML, Woods RL (2007) How red is a white eye? Clinical grading of normal conjunctival hyperaemia. Eye 21:633–638. https://doi.org/10.1038/sj.eye.6702295
Newsom JT (2024) Fixed and random effects in multilevel models. Class handout, Portland State University https://web.pdx.edu/~newsomj/uws/multilevel/multilevel%20handout%20fixed%20vs%20random.pdf
Park J, Barash V, Fink C, Cha M (2013) Emoticon style: interpreting differences in emoticons across cultures. In: Proceedings of the 7th Int AAAI Conf weblogs Soc media (ICWSM), vol 7, pp 466–475. https://doi.org/10.1609/icwsm.v7i1.14437
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. https://doi.org/10.1038/s41598-023-40458-1
Peragallo J, Biousse V, Newman NJ (2013) Ocular manifestations of drug and alcohol abuse. Curr Opin Ophthalmol 24:566–573
Perea-García JO, Ramarajan K, Kret ME, Hobaiter C, Monteiro A (2022) Ecological factors are likely drivers of eye shape and colour pattern variations across anthropoid primates. Sci Rep 12:17240. https://doi.org/10.1038/s41598-022-20900-6
Perea-García JO, Teuben A, Caspar KR (2025) Look past the cooperative eye hypothesis: reconsidering the evolution of human eye appearance. Biol Rev 100:2038–2054. https://doi.org/10.1111/brv.70033
Pflüger LS, Oberzaucher E, Katina S, Holzleitner IJ, Grammer K (2012) Cues to fertility: perceived attractiveness and facial shape predict reproductive success. Evol Hum Behav 33:708–714. https://doi.org/10.1016/j.evolhumbehav.2012.05.005
Pokorný Š, Pavlovič O, Kleisner K (2024) Sexual dimorphism: the interrelation of shape and color. Arch Sex Behav 53:3255–3265. https://doi.org/10.1007/s10508-024-02918-1
Prantl L, Heidekrueger PI, Broer PN, Knoll S, Thiha A, Gründl M (2019) Female eye attractiveness – where beauty meets science. J Craniomaxillofac Surg 47:73–79. https://doi.org/10.1016/j.jcms.2018.11.011
Proche SP, Kobos R (2004) Jaundice in the adult patient. Am Fam Physician 69:299–304

Provine RR, Cabrera MO, Brocato NW, Krosnowski KA (2011) When the whites of the eyes are red: a uniquely human cue. Ethology 117:395–399. https://doi.org/10.1111/j.1439-0310.2011.01888.x
Provine RR, Cabrera MO, Nave-Blodgett J (2013a) Red, yellow, and super-white sclera: uniquely human cues for healthiness, attractiveness, and age. Hum Nat 24:126–136. https://doi.org/10.1007/s12110-013-9168-x
Provine RR, Nave-Blodgett J, Cabrera MO (2013b) The emotional eye: red sclera as a uniquely human cue of emotion. Ethology 119:993–998. https://doi.org/10.1111/eth.12144
R Core Team (2024) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org
Russell R, Sweda JR, Porcheron A, Mauger E (2014) Sclera color changes with age and is a cue for perceiving age, health, and beauty. Psychol Aging 29:626–635. https://doi.org/10.1037/a0036142
Saribay Adil S, Biten AF, Meral EO, Aldan P, Třebický V, Kleisner K (2018) The bogazici face database: standardized photographs of Turkish faces with supporting materials. PLoS One 13:e0192018. https://doi.org/10.1371/journal.pone.0192018
Schneider CA, Rasband WS, Eliceiri KW (2012) NIH image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675
Scott IM, Clark AP, Josephson SC, Boyette AH, Cuthill IC, Fried RL, Penton-Voak IS (2014) Human preferences for sexually dimorphic faces may be evolutionarily novel. Proc Natl Acad Sci U S A 111:14388–14393. https://doi.org/10.1073/pnas.1409643111
Shields JA, Shields CL, Mashayekhi A, Marr BP, Benavides R, Thangappan A, Phan L, Eagle RC Jr (2007) Primary acquired melanosis of the conjunctiva: experience with 311 eyes. Trans Am Ophthalmol Soc 105:61–72. https://doi.org/10.1016/j.ophtha.2007.07.003
Stephen ID, Law Smith MJ, Stirrat MR, Perrett DI (2009) Facial skin coloration affects perceived health of human faces. Int J Primatol 30:845–857. https://doi.org/10.1007/s10764-009-9380-z
Strom MA, Zebrowitz LA, Zhang S, Bronstad PM, Lee HK (2012) Skin and bones: the contribution of skin tone and facial structure to racial prototypicality ratings. PLoS One 7:e41193. https://doi.org/10.1371/journal.pone.0041193
Timmis P (1957) Exophthalmos. J Laryngol Otol 71:744–753. https://doi.org/10.1017/S0022215100052415
Třebický V, Fialová J, Kleisner K, Havlíček J (2016) Focal length affects depicted shape and perception of facial images. PLoS One 11:e0149313. https://doi.org/10.1371/journal.pone.0149313
Třebický V, Stirrat M, Havlíček J (2019) Fighting assessment. In: Shackelford TK, Weekes-Shackelford VA (eds) Encyclopedia of evolutionary psychological science. Springer International Publishing, Cham, pp 1–11
Trujillo LT, Jankowitsch JM, Langlois JH (2014) Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. Cogn Affect Behav Neurosci 14:1061–1076. https://doi.org/10.3758/s13415-014-0259-5
United Nations (2023) Human Development Index, https://hdr.undp.org/data-center/human-development-index#/indicies/HDI
van Doorn J, van den Bergh D, Böhm U et al (2021) The JASP guidelines for conducting and reporting a Bayesian analysis. Psychon Bull Rev 28:813–826
Wacewicz S, Perea-García JO, Lewandowski Z, Danel DP (2022) The adaptive significance of human scleral brightness: an experimental study. Sci Rep 12:20261. https://doi.org/10.1038/s41598-022-24403-2
Wagatsuma H (1967) The social perception of skin color in Japan. Daedalus 96:407–443
Wémenau J, Klein M, Sadoul J-L, Briet C, Vélayoudom-Céphise F-L (2018) Graves' disease: introduction, epidemiology, endogenous and environmental pathogenic factors. Ann Endocrinol 79:656–664. https://doi.org/10.1016/j.ando.2018.08.005
Wolf W, Thielhelm J, Tomasello M (2023) Five-year-old children show cooperative preferences for faces with white sclera. J Exp Child Psychol 225:105532. https://doi.org/10.1016/j.jecp.2022.105532
Workman L, Akcay N, Reeves M, Taylor S (2018) Blue eyes keep away the winter blues: is blue eye pigmentation an evolved feature to provide resilience to seasonal affective disorder? OA J Behav Sci Psych 1:180002
World Health Organization (2024) Global hepatitis report 2024: action for access in low- and middle-income countries. World Health Organization, Geneva, Switzerland

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Vojtěch Fiala[1,2] · Juan Olvido Perea-García[2,3] · Petr Tureček[1,4] · Sławomir Wacewicz[2] · Juan David Leongómez[5] · Šimon Pokorný[1] · Ondřej Pavlovič[1] · S. Adil Saribay[6] · Farid Pazhoohi[7] · Robert Mbe Akoko[8] · Karel Kleisner[1]

✉ Vojtěch Fiala
fialavoj@natur.cuni.cz

1 Department of Philosophy and History of Science, Faculty of Science, Charles University, Viničná 7, 128 00 Prague, Czech Republic

2 Centre for Language Evolution Studies, Faculty of Humanities, Nicolaus Copernicus University in Toruń, Fosa Staromiejska 3, 87-100 Toruń, Poland

3 Department of Modern Philology, Universidad Las Palmas de Gran Canaria, Juan de Quesada, 30, 35001 Las Palmas de Gran Canaria, Spain

4 Center for Theoretical Study, Charles University and Czech Academy of Sciences, Jilská 1, 110 00 Prague, Czech Republic

5 EvoCo: Human Behaviour and Evolution Lab, Faculty of Psychology, Universidad El Bosque, Bogotá, DC 110121, Colombia

6 Department of Psychology, Kadir Has University, Cibali Mah. Kadir Has Cad., 34083 Fatih-İstanbul, Turkey

7 School of Psychology, University of Plymouth, Portland Square, Plymouth PL4 8AA, UK

8 Department of Sociology and Anthropology, Faculty of Social and Management Sciences, University of Buea, PO Box 63, Buea, South West Region, Cameroon