AMERICAN PSYCHOLOGICAL ASSOCIATION
DIVISION 10
Psychology of Aesthetics, Creativity, and the Arts

ISSN: 1931-3896
https://doi.org/10.1037/aca0000454

What Is Good Is Beautiful (and What Isn't, Isn't): How Moral Character Affects Perceived Facial Attractiveness

Dexian He[1, 2], Clifford I. Workman[3, 4], Xianyou He[1, 2], and Anjan Chatterjee[3, 4]
[1] Key Laboratory of Brain, Cognition, and Education Sciences, Ministry of Education, South China Normal University
[2] Center for Studies of Psychological Application, South China Normal University
[3] Penn Center for Neuroaesthetics, University of Pennsylvania
[4] Department of Neurology, University of Pennsylvania

A well-documented "beauty is good" stereotype is expressed in the expectation that physically attractive people have more positive characteristics. Recent evidence has also found that unattractive faces are associated with negative character inferences. Is what is good (bad) also beautiful (ugly)? Whether this conflation of aesthetic and moral values is bidirectional is not known. This study tested the hypothesis that complementary "good is beautiful" and "bad is ugly" stereotypes bias aesthetic judgments. Using highly controlled face stimuli, this preregistered study examined whether moral character influences perceptions of attractiveness for different ages and sexes of faces. Compared to faces paired with nonmoral vignettes, those paired with prosocial vignettes were rated significantly more attractive, confident, and friendlier. The opposite pattern characterized faces paired with antisocial vignettes. A significant interaction between vignette type and the age of the face was detected for attractiveness. Moral transgressions affected attractiveness more negatively for younger than older faces. Sex-related differences were not detected. These results suggest information about moral character affects our judgments about facial attractiveness. Better (worse) people are considered more (less) attractive. These findings suggest that beliefs about moral goodness and physical beauty influence each other bidirectionally.

Keywords: attractiveness, morality, age, beauty is good, ugly is bad

Supplemental materials: https://doi.org/10.1037/aca0000454.supp

Although we may be unaware of it, physical attractiveness influences the impressions we form about other people and how we ultimately treat them. Attractive people are expected to have more positive characteristics than unattractive people, an effect known as the "beauty is good" stereotype (Dion et al., 1972). Relative to unattractive people, attractive people are expected to be more intelligent, trustworthy, competent, dominant, and socially skilled and are treated more positively (Eagly et al., 1991; Ferrari et al., 2017; Langlois et al., 2000; Wilson & Eckel, 2006; Zebrowitz et al., 2002; Zhao et al., 2015). A complementary "anomalous is bad" stereotype has also been described (Griffin & Langlois, 2006), which is expressed in negative attitudes about people with facial anomalies (e.g., scars) that may facilitate dehumanizing behavior (e.g., less prosociality; Hartung et al., 2019; Jamrozik et al., 2019; Workman et al., 2021). The attractiveness of faces—whether beautiful or not—affects the inferences we ultimately make about the people harboring those faces.

Attractiveness stereotyping also exerts effects in the opposite direction such that people with desirable personality traits (e.g., ability, honest, and decent) are rated more physically attractive than those without such traits (Gross & Crofton, 1977; Owens & Ford, 1978; Paunonen, 2006; Zhang et al., 2014). Furthermore,

Dexian He https://orcid.org/0000-0003-0087-224X
Clifford I. Workman https://orcid.org/0000-0002-2206-0325
Xianyou He https://orcid.org/0000-0002-3118-4047

The authors declare no potential conflict of interest. This work was supported by the China Scholarship Council (Dexian He), the South China Normal University Study Abroad Program for Elite Postgraduate Students and the Innovation Project of the Graduate School of South China Normal University (Dexian He), the National Institute of Dental and Craniofacial Research of the National Institutes of Health (F32DE029407 awarded to Clifford I. Workman), the National Natural Science Foundation of China (31970984 awarded to Xianyou He), and the Edwin and Fannie Gray Hall Center for Human Appearance (Anjan Chatterjee). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This study was preregistered (https://doi.org/10.17605/OSF.IO/B9FAW). The data, code, and materials reported in this article are publicly available from https://osf.io/aeygb/.

Correspondence concerning this article should be addressed to Xianyou He, School of Psychology, South China Normal University, 55, West of Zhongshan Avenue, Tianhe District, Guangzhou, Guangdong 510631, China, or Anjan Chatterjee, Department of Neurology, University of Pennsylvania, 3710 Hamilton Walk, Goddard Laboratories, Philadelphia, PA 19104, United States. Email: xianyouhe@163.com or anjan@pennmedicine.upenn.edu

links between physical attractiveness and real-world giving behaviors have been reported that cannot simply be attributed to the halo effect (Konrath & Handy, 2021). People who do good things are seen as more attractive than people who do not. Kniffin and Wilson (2004) compared ratings of faces along several dimensions (e.g., attractiveness) made by people personally familiar with the target faces (e.g., classmates) relative to people who never met the target faces. Nonphysical traits affected judgments of physical attractiveness made about familiar faces. An attractive face, for instance, may be seen as ugly by someone familiar with their poor moral character.

Research on the "good is beautiful" stereotype, however, has focused almost exclusively on characterizing attractiveness stereotyping in younger faces, leaving potential interactions with aging underexplored. Older faces are generally perceived as less attractive and are assigned more negative traits than younger faces (He et al., 2021; North & Fiske, 2015), an effect that may be amplified by negative moral character inferences or dampened by beliefs of moral goodness. This study examined whether and how perceived moral character, whether informed by morally good actions or by moral transgressions, influences perceptions of facial beauty among different ages and sexes of faces.

This preregistered study (https://doi.org/10.17605/OSF.IO/B9FAW) tested the hypothesis that a "good is beautiful" stereotype biases facial beauty judgments, with people ostensibly possessing good moral character considered more attractive. We predicted that reading about a person's morally good actions would result in their being found more attractive. We further hypothesized that a complementary "bad is ugly" stereotype operates in the opposite direction. We predicted that reading about a person's moral transgressions would result in their being found less attractive. Alternatively–instead of detecting effects of both moral goodness and moral badness on attractiveness judgments–relations between moral character inferences and attractiveness judgments may be specific for unattractiveness. If detected, this could reflect an evolved disgust response (Klebl et al., 2021).

Face age may interact with moral character inferences to shape judgments of physical attractiveness in one of several ways. First, this interaction could have additive effects (similar to the amplification account; Albrecht & Carbon, 2014; Carr et al., 2017). Positive features would be predicted to be perceived as more positive, and negative features would be predicted to be perceived as more negative. In other words, moral goodness would be predicted to affect attractiveness more positively for younger than older faces. Moral transgressions, on the other hand, would be predicted to affect attractiveness more negatively for older than younger faces. Alternatively, the interaction may result in a selective effect. This account predicts that moral transgressions should exert weaker effects on judgments of attractiveness made in response to older relative to younger faces. People over 50 and people under 21 received less severe criminal sentences compared to other age groups (Steffensmeier et al., 1998). Likewise, Bergeron and McKelvie (2004) found that for murder, 60- and 20-year-old men received more lenient treatment than 40-year-old men in sentencing and parole recommendations. In other words, people were likely to be more tolerant of transgressions committed by older than relatively younger individuals. Thus, this attenuation of moral agency could dampen the negative effect of moral transgressions on judgments of the attractiveness of older faces. A final possibility is that the interaction between moral character inferences and perceived facial beauty will result in equivalent effects (similar to the generalized-positivity-shift account but expanded to incorporate negativity; Carr et al., 2017; Monin, 2003). On this account, moral character inferences would be expected to modulate facial attractiveness ratings by similar magnitudes regardless of face valence. An exploratory aim of this study was therefore to investigate age-related differences in relations between moral character inferences and judgments of facial beauty.

Sex-related differences in aesthetic responses to faces have been reported across a variety of contexts (Leder et al., 2010). In a context conducive to social approach, perceivers spent longer looking at attractive male and female faces than nonattractive male and female faces. In a threat context, however, people spent less time looking at attractive male faces, potentially because men are generally considered more aggressive than women. As such, men may be judged more threatening than women in antisocial contexts. Whyte et al. (2021) conducted an analysis of online survey data from over 7,000 individuals (aged 18 to 65), finding that women care more about resources and personality (e.g., trust) in potential mates than men, whereas men prioritize attractiveness and physical build. In the current research, we assessed whether women are more likely to incorporate moral information into attractiveness judgments than men. We predicted that learning morally relevant information associated with target faces would affect women's ratings more robustly than men's, especially when judging male faces.

Method

Participants

A total of 442 participants were recruited via Amazon's Mechanical Turk service to complete an online survey administered through the Qualtrics platform (249 male; age: 38.11 ± 10.05 years; education: 14.91 ± 2.50 years). Using effect sizes computed from data reported in Paunonen (2006), a minimum sample of n = 322 participants was expected to provide sufficient power (80%) to detect the effects of interest. Data were excluded from 64 participants: 10 due to extreme values for duration (e.g., a duration of 22.07 hr) identified with outlier analysis in SPSS (Curran, 2016), four for reporting that their responses were of poor quality, 14 for failing more than two of five attentional catch trials, one for missing data, and 35 for choosing not to report their sex since participant gender has been reported to play a role in perceived facial attractiveness (He et al., 2021). The final sample consisted of n = 378 participants (age: 38.31 ± 10.06 years; range: 21–72 years; education: 14.96 ± 2.27 years; race/ethnicity: 319 White, 31 Black, 12 Asian, 15 Hispanic or Latinx, one American Indian/Alaskan Native; sexual orientation: 321 heterosexual, 14 homosexual, 40 bisexual, and three other). There were 235 men (age: 37.64 ± 10.06 years; range: 23–72 years; education: 15.05 ± 2.14 years) and 143 women (age: 39.41 ± 9.98 years; range: 21–70 years; education: 14.81 ± 2.46 years). Participants were compensated ($4) for their time and participation in the study. This study was approved by the institutional review board at the University of Pennsylvania. The study—including the sample size rationale—was preregistered (https://doi.org/10.17605/OSF.IO/B9FAW), and the corresponding study materials, code, and data are available from https://osf.io/aeygb/.

Materials

Face Stimuli

The stimuli comprised 12 pairs of images, with each pair depicting the same face but either younger (age: 20–29 years; attractiveness:

4.53 ± 0.84) or older (age: 60 years and older; attractiveness: 3.20 ± 0.46) in appearance. These face images were chosen from a previous study on effects of face age on judgments of different facets of attractiveness (He et al., 2021). See Figure 1 for sample stimuli. The faces were well balanced along the dimensions of sex and race/ethnicity.

Face Stimuli Norming

Face stimuli were selected and generated in the following way: First, we identified 80 middle-aged faces from the Chicago Face Database (Ma et al., 2015; https://www.chicagofaces.org/). These faces were submitted to the FaceApp software package (https://www.faceapp.com/) to generate 80 sets of younger and older faces from the middle-aged faces. Face images were then (a) normalized to interpupillary distance using algorithms provided by the OpenCV computer vision library (https://opencv.org/) and facial landmarks provided by the dlib machine learning toolkit (http://dlib.net/), (b) resized and cropped to 345 pixels (width) × 407 pixels (height), (c) placed onto a plain white background using the GIMP software package (https://www.gimp.org/), and (d) color corrected (Workman & Chatterjee, 2021).

An independent sample of 129 participants—of which 33 were younger (23 male; age: 28.82 ± 3.71 years; range: 20–34 years; education: 14.64 ± 2.56 years), 59 were middle-aged (25 male; age: 47.05 ± 8.14 years; range: 35–59 years; education: 14.41 ± 2.71 years), and 37 were older (11 male; age: 65.00 ± 4.22 years; range: 60–73 years; education: 14.92 ± 2.51 years)—was recruited to rate the computer-generated younger and older faces on expected age, beliefs about realness, and facial attractiveness using 7-point Likert scales. Based on these ratings, 43 sets of faces were identified as potential stimuli. An additional 27 participants (15 male; age: 26.81 ± 3.72 years; range: 22–36 years; education: 18.22 ± 2.64 years) were recruited to judge whether each pair of faces depicted the same person at different ages. These ratings were used to further narrow down the potential stimuli to 30 face pairs. After closely matching the stimuli on attractiveness, ethnicity, and sex, a final set of 12 face pairs was selected for use in the current study (see Table S1 in the online supplemental materials for details).

Moral and Nonmoral Vignettes

Vignettes describing morally good and bad actions were adapted from a previous study (Knutson et al., 2010). These scenarios capture real-world instances of prosociality and antisociality drawn from the experiences of actual people. First, 50 prosocial and 50 antisocial vignettes were selected based on their harm, other-benefit, and moral appropriateness ratings. Second, 100 nonmoral stories were generated by Dexian He and Clifford I. Workman in complement to corresponding prosocial and antisocial vignettes. Next, the vignettes were stripped of demographic details to ensure they could be randomly paired with either young or old faces that were either male or female. A final set of 72 vignettes was selected.

The vignettes were normed by an independent sample of 73 controls (40 male; age: 37.18 ± 11.81 years; range: 20–70 years; education: 15.25 ± 1.82 years) who rated the vignettes along dimensions of harm ("Do the actions of the person in the story you just read harm other people?"), other-benefit ("Do the actions of the person in the story you just read benefit other people?"), and moral relevance ("Are the actions of the person in the story you just read related to morality?"). The actions described in prosocial vignettes were rated higher on other-benefit, $t(23) = 40.92$, $p < .001$, and moral relevance, $t(23) = 55.27$, $p < .001$, than actions described in nonmoral scenarios. The actions described in antisocial vignettes were rated more harmful, $t(23) = 27.06$, $p < .001$, and morally relevant, $t(23) = 57.33$, $p < .001$, than nonmoral actions.

Of the 72 vignettes, 24 described an individual acting prosocially (other-benefit: 6.70 ± 0.39; harm: 1.37 ± 0.22; moral relevance: 91.81% ± 3.79%—in other words, an average of 91.81% participants thought the actions of the person in the story were related to morality), 24 described antisocial actions (other-benefit: 1.57 ± 0.37; harm: 5.83 ± 0.81; moral relevance: 92.05% ± 3.76%), and 24 described nonmoral actions (other-benefit: 1.64 ± 0.35; harm: 1.35 ± 0.28; moral relevance: 26.35% ± 3.03%). The following are samples of the vignettes used in this study:

> Prosocial: During my commute through downtown, I see a lot of homeless people. One day I was driving and saw a homeless woman walking her dog. I pulled over and gave her some money.
>
> Antisocial: When I was younger I worked for my dad in the produce business. I felt that he would underpay me and I deserved more. So I would self-compensate and take money from him.
>
> Nonmoral (neutral): I was in high school and had just finished taking a physiology exam. I didn't have any breakfast before the exam so I had gotten very hungry. I checked my backpack and found a banana to eat.

Figure 1
Sample Stimuli

Note. Middle-aged faces selected from the Chicago Face Database (Ma et al., 2015; middle-aged face not pictured) were morphed to appear either younger (left) or older (right) using the FaceApp software package (https://www.faceapp.com/). See the online article for the color version of this figure.

Procedures

The face rating task was comprised of 72 trials. In each trial, participants saw a face (24 younger and older faces in total) and read a brief story (72 prosocial, antisocial, and nonmoral vignettes in total) ostensibly about the person harboring that face. Each face appeared three times, once with a randomly selected prosocial vignette, once with an antisocial vignette, and once with a nonmoral vignette. Faces remained on the screen while participants rated them. Participants rated each face on facial attractiveness ("How attractive is this face?") using a 7-point scale. Participants also rated the faces on confidence ("How confident is this face?") and friendliness ("How friendly is this face?").

After the face rating task, participants completed a battery of self-report measures assessing psychological dispositions. Specifically,

trait empathy was assessed with the Interpersonal Reactivity Index Scale (Davis, 1980), which comprises four subscales (i.e., empathic concern, perspective taking, personal distress, and fantasy), and sensitivity to disgust was assessed with the Three-Domain Disgust Scale (Tybur et al., 2009), which includes subscales for sensitivities to moral, sexual, and pathogen disgust. Finally, participants completed a short sociodemographic questionnaire. The face images and vignettes, questions, and self-report measures were presented in randomized order. There was no time limit, with ratings proceeding in a self-paced fashion. The experiment lasted approximately 30 min.

Data Analyses

Linear mixed-effects analyses were carried out using the lme4 package (Bates et al., 2015) in RStudio (R Core Team, 2020) to examine whether perceived moral character influences judgments of facial beauty and whether this influence varies as a function of face age, perceiver sex, and face sex. Exploratory analyses investigated whether and how sensitivity to moral disgust and/or trait empathic concern interact with story type (prosocial, antisocial, or neutral) to modulate attractiveness judgments. We obtained p values for the parameter estimates generated by each model using Satterthwaite's approximation as implemented by the lmerTest package (Kuznetsova et al., 2017). Below, regression coefficients (β), standard errors (SE), and t values are reported. Plots were generated with the effects package (Fox & Weisberg, 2018).

Results

Morality and Facial Judgments

To examine the effect of moral information on facial attractiveness, a linear mixed model was constructed with attractiveness as the dependent variable and vignette type (Prosocial | Antisocial | Nonmoral) as a fixed factor. Random intercepts for face stimulus and subject were modeled. Reading about morally good actions and about moral transgressions significantly influenced perceptions of facial attractiveness. Faces paired with prosocial vignettes were rated more attractive than those paired with nonmoral vignettes, $\beta = .138$, $SE = .018$, $t(26813) = 7.712$, $p < .001$, and faces paired with antisocial vignettes were rated less attractive than those paired with nonmoral vignettes, $\beta = -.373$, $SE = .018$, $t(26813) = -20.779$, $p < .001$ (see online Supplemental Tables S2 and S3 for remaining fixed effects and means).

Similar models were constructed to examine effects of vignette type on confidence and friendliness, with significant effects observed in both cases. Faces paired with prosocial vignettes were rated more confident, $\beta = .419$, $SE = .017$, $t(26813) = 24.126$, $p < .001$, and friendly, $\beta = .557$, $SE = .019$, $t(26813) = 29.126$, $p < .001$, than those paired with nonmoral vignettes, whereas faces paired with antisocial vignettes were rated less confident, $\beta = -.201$, $SE = .017$, $t(26813) = -11.540$, $p < .001$, and friendly, $\beta = -1.026$, $SE = .019$, $t(26813) = -53.642$, $p < .001$.

Age-Related Differences

Next, a linear mixed model examined whether an interaction between moral information and face age modulated attractiveness judgments, with attractiveness as the dependent variable and vignette type (Prosocial | Antisocial | Nonmoral) and face age (Younger | Older) as fixed factors. Random intercepts for stimulus and subject were included, and for subject, slopes were allowed to vary according to face age. Significant main effects were detected for face age and vignette type. Relative to nonmoral vignettes, younger and older faces were rated more attractive when paired with prosocial vignettes and were rated less attractive when paired with antisocial vignettes ($p < .001$). Across all contexts, younger faces were rated more attractive than older faces ($p < .001$). There was also a significant interaction between vignette type and face age ($p < .010$; Figure 2A; Table 1). Attractiveness ratings were lower for younger compared to older faces paired with antisocial relative to nonmoral vignettes (see Table 2). No such interaction was detected for confidence and friendliness ratings, however ($p > .050$).

Sex-Related Differences

We then constructed linear mixed models to examine how effects of vignette type on participant ratings varied as functions of perceiver sex and face sex, with facial attractiveness, confidence, and friendliness as the dependent variables (attractiveness in the first model, confidence in the second, and friendliness in the third) and vignette type (Prosocial | Antisocial | Nonmoral) and perceiver sex and face sex (Female | Male) as fixed factors. Random intercepts for stimulus and subject were included. We did not detect a significant interaction between vignette type, perceiver sex, and face sex for attractiveness, confidence, and friendliness judgments ($p > .050$; Figure 2B; see online Supplemental Tables S4–S9 for fixed effects and means).

Sensitivity to Moral Disgust and Empathic Concern

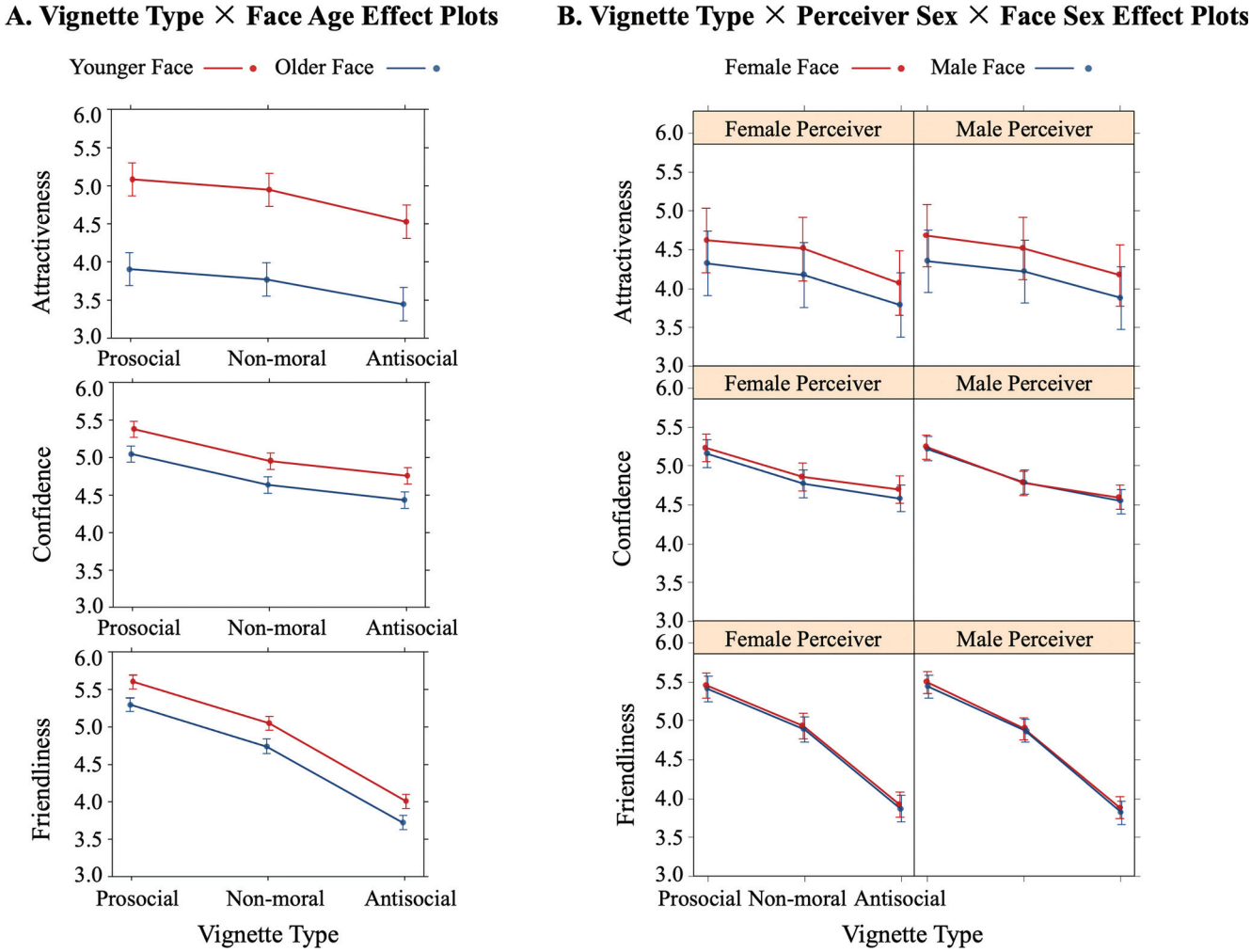
Linear mixed models also examined whether individual differences in propensities for sensitivity to moral disgust and empathic concern bear on attractiveness judgments as a function of vignette type. These models included attractiveness as the dependent variable and fixed factors for vignette type (Prosocial | Antisocial | Nonmoral) and psychological disposition (sensitivity to moral disgust in the first model, empathic concern in the second). Random intercepts for stimulus and subject were modeled. A significant interaction between vignette type and sensitivity to moral disgust was detected ($p < .001$; Figure 3A; Table 3). Participants who were particularly sensitive to moral disgust were also the harshest judges of attractiveness for faces paired with antisocial vignettes compared to prosocial and nonmoral vignettes. A significant interaction was also detected between vignette type and empathic concern for attractiveness judgments ($p < .001$; Figure 3B; Table 4). Similar to sensitivity to moral disgust, those participants who scored highest for trait empathic concern rated faces as less attractive when paired with antisocial vignettes compared to prosocial and nonmoral vignettes.

Discussion

Ample evidence suggests that what is beautiful is also considered good. Are effects of beauty on moral attitudes unidirectional, or might our moral attitudes also shape our judgments of beauty? In the current study, participants evaluated younger- and older-looking versions of the same faces along dimensions of attractiveness, confidence, and friendliness. Prior to making their ratings,

Figure 2

Effects of Age and Sex on the Consequences of Moral Information for Facial Judgments



Note. Panel A: Effects of vignette type on facial attractiveness, confidence, and friendliness ratings as a function of face age. A significant interaction between vignette type and face age was only detected for attractiveness ratings. Panel B: Effects of vignette type on facial attractiveness, confidence, and friendliness ratings for female and male faces as a function of perceiver sex. No significant interactions were detected between vignette type, perceiver sex, and face sex on attractiveness, confidence, or friendliness ratings. The dots represent means. The error bars represent 95% confidence intervals. See the online article for the color version of this figure.

however, each face was paired with a vignette that described a prosocial, antisocial, or nonmoral action. Learning about the morally relevant actions ostensibly carried out by the people whose faces participants saw had consequences for perceptions of attractiveness. Participants rated faces as more attractive, confident, and friendly when they were linked to acts of moral goodness than to moral transgressions and nonmoral actions. In contrast, participants judged faces to be less attractive, confident, and friendly when paired with supposed moral transgressions relative to prosocial and nonmoral actions. A significant interaction was also detected between vignette type and face age, with the attractiveness of older faces showing less sensitivity to moral transgressions than younger faces.

Our results are in line with previous research on the relationship between goodness and beauty (Gross & Crofton, 1977; Owens & Ford, 1978; Paunonen, 2006; Zhang et al., 2014). Evaluations of moral character bear on evaluations of physical attractiveness, which may be underpinned by the engagement of shared neurocognitive mechanisms when making moral and aesthetic judgments. Functional neuroimaging evidence finds that moral and aesthetic judgments implicate overlapping regions of the medial orbitofrontal cortex (Diessner, 2019, p. 186; Luo et al., 2019; Tsukiura & Cabeza, 2011; Wang et al., 2015) and amygdala (Bzdok et al., 2011; Workman et al., 2021). This overlap may have an evolutionary basis. Attractive facial features like symmetry and averageness may signal good health and mate quality (Little et al., 2011; Rhodes, 2006). Similarly, moral behavior has social signaling functions and plays an important role in maintaining social order (Decety et al., 2018). Prosociality (e.g., helping and sharing) may enable social groups to thrive and reproduce (Boyd & Richerson, 2009), while antisocial behavior (e.g., physical aggression and violations of societal rules) may indicate

Table 1
Fixed Effects From the Linear Mixed Models Constructed to Examine the Consequences of Vignette Type and Face Age for Ratings of Facial Attractiveness, Confidence, and Friendliness

Fixed effects	β	*SE*	*t*	*p*
Attractiveness				
Intercept	3.768	.119	31.701	< .001
Vignette type (antisocial)	−.326	.023	−13.981	< .001
Vignette type (prosocial)	.139	.023	5.969	< .001
Face age (younger)	1.177	.146	8.072	< .001
Vignette type (antisocial): Face age (younger)	−.093	.033	−2.829	< .010
Vignette type (prosocial): Face age (younger)	−.002	.033	−.054	.957
Confidence				
Intercept	4.633	.058	80.446	< .001
Vignette type (antisocial)	−.203	.024	−8.455	< .001
Vignette type (prosocial)	.414	.024	17.259	< .001
Face age (younger)	.319	.060	5.296	< .001
Vignette type (antisocial): Face age (younger)	.004	.034	.130	.897
Vignette type (prosocial): Face age (younger)	.011	.034	.318	.750
Friendliness				
Intercept	4.738	.053	90.001	< .001
Vignette type (antisocial)	−1.015	.027	−38.296	< .001
Vignette type (prosocial)	.559	.027	21.086	< .001
Face age (younger)	.306	.053	5.835	< .001
Vignette type (antisocial): Face age (younger)	−.022	.037	−.600	.549
Vignette type (prosocial): Face age (younger)	−.003	.037	−.088	.930

Note. *SE* = standard error.

possible threat and risk of harm (Workman et al., 2020). Together, people who act prosocially to benefit others may be seen as more attractive, more confident, and friendlier than those whose actions are antisocial or nonmoral.

Contextual modulations of hedonic value could also underpin the effects we reported (Skov, 2019). Works of art received significantly higher aesthetic ratings that were more tightly coupled to medial orbitofrontal cortex activation when people believed the artworks were from an art gallery as opposed to being computer generated (Kirk et al., 2009). These findings suggest that different contexts (e.g., art from a gallery vs. from a computer program) induce different expectations about hedonic value. Leder et al. (2010) found that participants looked longer at attractive compared to nonattractive faces, suggesting that individuals are drawn to beauty for its adaptive value. This effect of facial attractiveness on visual attention was influenced by situational demands in the context of experimentally induced threat. Taken together, aesthetic evaluations are shaped by the properties of the aesthetic objects themselves, by individual differences in the psychological dispositions of evaluators, and by contextual demands.

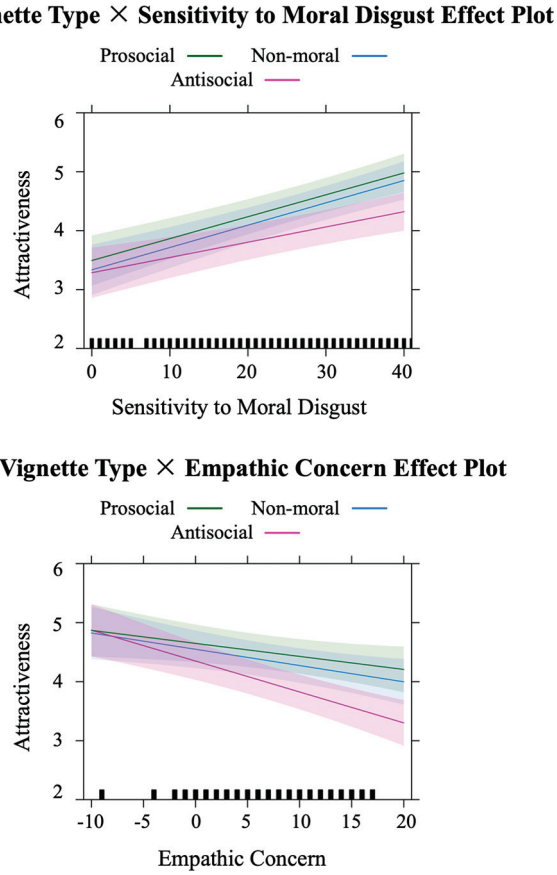
Consistent with prior work, our results suggest that aesthetic evaluations are informed by the properties of aesthetic objects (i.e., whether faces were young or old), by individual differences in psychological dispositions (i.e., sensitivity to moral disgust and empathic concern), and by contextual information (i.e., whether vignettes were prosocial, antisocial, or nonmoral). On the basis of these and earlier findings, we propose a general framework for aesthetic evaluation. Contextual factors modulate hedonic value either by increasing pleasure or by increasing displeasure and disgust. Changes to hedonic value promote approach or avoidance behaviors with consequences for aesthetic evaluation (Skov, 2019). Moral information—whether prosocial or antisocial—is

Table 2
Means and Standard Deviations for Facial Attractiveness, Confidence, and Friendliness Ratings Grouped by Vignette Type and Face Age

Measure	Prosocial vignettes	Antisocial vignettes	Nonmoral vignettes	Overall
Attractiveness				
Younger faces	5.08 (1.43)	4.53 (1.69)	4.95 (1.45)	4.85 (1.54)
Older faces	3.91 (1.68)	3.44 (1.72)	3.77 (1.65)	3.71 (1.69)
Overall	4.49 (1.67)	3.98 (1.79)	4.36 (1.66)	
Confidence				
Younger faces	5.38 (1.23)	4.75 (1.58)	4.95 (1.27)	5.03 (1.39)
Older faces	5.05 (1.36)	4.43 (1.55)	4.63 (1.32)	4.70 (1.44)
Overall	5.21 (1.31)	4.59 (1.57)	4.79 (1.31)	
Friendliness				
Younger faces	5.60 (1.25)	4.01 (1.79)	5.04 (1.24)	4.88 (1.59)
Older faces	5.30 (1.45)	3.72 (1.70)	4.74 (1.34)	4.59 (1.64)
Overall	5.45 (1.36)	3.87 (1.75)	4.89 (1.30)	

Figure 3

Individual Trait Differences in the Effects of Moral Information on Facial Judgments



Note. Panel A: Effects of sensitivity to moral disgust on facial attractiveness as a function of vignette type. Panel B: Effects of empathic concern on facial attractiveness as a function of vignette type. Participants exhibiting greater sensitivity to moral disgust and elevated trait empathic concern were especially prone to rating faces as less attractive when paired with antisocial relative to prosocial and nonmoral scenarios. See the online article for the color version of this figure.

one source for information capable of enhancing pleasure and disgust responses, which then bears on the hedonic valuation of aesthetic objects. Since "bad" people may threaten one's survival, the tendency to prefer "good" people, reflected in elevated attractiveness judgments, may be adaptive. The effect of moral information on attractiveness judgments was mediated by sensitivity to moral disgust and empathic concern. Heightened sensitivities to moral disgust and empathic concern both amplified the negative consequences of antisocial vignettes for facial attractiveness judgments.

We also found that moral transgressions had a stronger impact than prosociality on evaluations of attractiveness. This observation is consistent with theoretical work underscoring the value of allocating attentional resources preferentially for negative compared to positive information (negativity bias; Baumeister et al., 2001).

There was a significant interaction between vignette type and face age on attractiveness ratings. We are cautious in interpreting this finding since younger faces had higher baseline ratings of attractiveness than older faces and scaling effects may have limited decreases in attractiveness for older faces. Antisociality appeared to have selective age-related effects, with older faces judged less harshly for moral transgressions than younger faces. Aging is generally associated with declines in cognitive ability but also with increased wisdom and breadth of knowledge (Lim & Yu, 2015). As described by philosopher Arthur Schopenhauer, "White hair always commands reverence". Schopenhauer suggested that

> the reason . . . respect is paid to age is that old people have necessarily shown in the course of their lives whether or not they have been able to maintain their honor unblemished; while that of young people has not been put to the proof, though they are credited with the possession of it. (Schopenhauer, 1902, p. 39)

The view that elders ought to be accorded respect and honor is one held in many cultures.

According to the stereotype content model, groups of people are judged along two axes: warmth and competence (Fiske, 2018). People high on warmth and competence are admired, and those low on both are denigrated. People high in warmth and low on competence are sometimes pitied, and those low on warmth and high on competence are often feared. One could imagine older people being ascribed high warmth (maintaining social honor) and either low competence (cognitive decline) or high competence (wisdom). One might predict that viewing someone with greater warmth or with less competence might mitigate effects of antisocial information on judgments of their attractiveness. While older faces were perceived as less attractive and were treated more leniently than younger faces when linked to antisocial scenarios, similar interactions were not observed for warmth (friendliness) or competence (confidence). The mechanism giving rise to this effect of antisocial scenarios and age on attractiveness remains to be determined.

Sex differences in the effects of moral information on attractiveness judgments were not detected. It may be that the moral character of potential mates is equally important to both men and women, with both indicating that positive personality traits are an important factor in long-term mates (Buss & Schmitt, 1993; Little et al., 2008). We note, however, that there were differences in the sample sizes of men and women. Specifically, the male sample (n = 235) was larger than the female sample (n = 143).

This study provides evidence for a bidirectional relationship between physical attractiveness and moral character inferences. We also extend prior studies by unpacking the consequences of

Table 3

Fixed Effects From the Linear Mixed Model Constructed to Examine Effects of Vignette Type and Sensitivity to Moral Disgust on Facial Attractiveness

Fixed effects	β	*SE*	*t*	*p*
Intercept	3.336	.218	15.325	< .001
Vignette type (antisocial)	−.049	.059	−.832	.405
Vignette type (prosocial)	.158	.059	2.689	< .010
Moral disgust	.038	.006	6.323	< .001
Vignette type (antisocial): Moral disgust	−.012	.002	−5.774	< .001
Vignette type (prosocial): Moral disgust	−.001	.002	−.355	.722

Note. *SE* = standard error; moral disgust = sensitivity to moral disgust.

Table 4
Fixed Effects From the Linear Mixed Model Constructed to Examine Effects of Vignette Type and Empathic Concern on Facial Attractiveness

Fixed effects	β	*SE*	*t*	*p*
Intercept	4.549	.163	27.957	< .001
Vignette type (antisocial)	−.199	.030	−6.669	< .001
Vignette type (prosocial)	.100	.030	3.340	< .001
Empathic concern	−.028	.010	−2.715	< .010
Vignette type (antisocial): Empathic concern	−.025	.003	−7.232	< .001
Vignette type (prosocial): Empathic concern	.006	.003	1.604	.109

Note. *SE* = standard error.

age and sex (i.e., face age, perceiver sex, and face sex) for judgments of physical attractiveness that are informed by moral information. The present study has several limitations that warrant attention. First, it remains unclear why differential effects of antisocial actions were detected for attractiveness judgments of older and younger faces. Future research should explore these age-related effects in greater detail. Second, this study did not examine middle-aged faces. Given that middle-aged people received more severe sentences compared to other age groups (Bergeron & McKelvie, 2004; Steffensmeier et al., 1998), the effect of moral badness on perceived facial attractiveness may be more pronounced in middle-aged compared to younger and older faces. Finally, the faces shown to participants were neither highly attractive nor highly unattractive, which may have elicited a restricted range of effects. Additional research is therefore needed to establish the generalizability of the effects reported herein.

Conclusion

The present study examined relations between moral character inferences and judgments of facial beauty. The pro- and antisocial actions ostensibly carried out by the people harboring the faces participants saw significantly affected subsequent judgments of physical attractiveness. Individuals were considered more attractive when linked to prosocial acts than to moral transgressions. In addition, acting morally bad had worse consequences for the perceived facial attractiveness of younger relative to older faces. These findings support the notions that what is good is also beautiful and what is bad is also ugly.

References

Albrecht, S., & Carbon, C. C. (2014). The fluency amplification model: Fluent stimuli show more intense but not evidently more positive evaluations. *Acta Psychologica*, *148*, 195–203. https://doi.org/10.1016/j.actpsy.2014.02.002

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Bergeron, C. E., & McKelvie, S. J. (2004). Effects of defendant age on severity of punishment for different crimes. *The Journal of Social Psychology*, *144*(1), 75–90. https://doi.org/10.3200/SOCP.144.1.75-90

Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review*, *100*(2), 204–232. https://doi.org/10.1037/0033-295X.100.2.204

Bzdok, D., Langner, R., Caspers, S., Kurth, F., Habel, U., Zilles, K., Laird, A., & Eickhoff, S. B. (2011). ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Structure & Function*, *215*(3-4), 209–223. https://doi.org/10.1007/s00429-010-0287-4

Carr, E. W., Brady, T. F., & Winkielman, P. (2017). Are you smiling, or have I seen you before? Familiarity makes faces look happier. *Psychological Science*, *28*(8), 1087–1102. https://doi.org/10.1177/0956797617702003

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.

Decety, J., Pape, R., & Workman, C. I. (2018). A multilevel social neuroscience perspective on radicalization and terrorism. *Social Neuroscience*, *13*(5), 511–529. https://doi.org/10.1080/17470919.2017.1400462

Diessner, R. (2019). *Understanding the beauty appreciation trait*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-32333-2_7

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290. https://doi.org/10.1037/h0033731

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., Longo, L. C. (1991). What is beautiful is good, but . . .: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*(1), 109–128. https://doi.org/10.1037/0033-2909.110.1.109

Ferrari, C., Nadal, M., Schiavi, S., Vecchi, T., Cela-Conde, C. J., & Cattaneo, Z. (2017). The dorsomedial prefrontal cortex mediates the interaction between moral and aesthetic valuation: A TMS study on the beauty-is-good stereotype. *Social Cognitive and Affective Neuroscience*, *12*(5), 707–717. https://doi.org/10.1093/scan/nsx002

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, *27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Fox, J., & Weisberg, S. (2018). Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software*, *87*(9), 1–27. https://doi.org/10.18637/jss.v087.i09

Griffin, A. M., & Langlois, J. H. (2006). Stereotype directionality and attractiveness stereotyping: Is beauty good or is ugly bad? *Social Cognition*, *24*(2), 187–206. https://doi.org/10.1521/soco.2006.24.2.187

Gross, A. E., & Crofton, C. (1977). What is good is beautiful. *Sociometry*, *40*(1), 85–90. https://doi.org/10.2307/3033549

Hartung, F., Jamrozik, A., Rosen, M. E., Aguirre, G., Sarwer, D. B., & Chatterjee, A. (2019). Behavioural and neural responses to facial disfigurement. *Scientific Reports*, *9*(1), Article 8021. https://doi.org/10.1038/s41598-019-44408-8

He, D., Workman, C. I., Kenett, Y. N., He, X., & Chatterjee, A. (2021). The effect of aging on facial attractiveness: An empirical and computational investigation. *Acta Psychologica*, *219*, Article 103385. https://doi.org/10.1016/j.actpsy.2021.103385

Jamrozik, A., Oraa Ali, M., Sarwer, D. B., & Chatterjee, A. (2019). More than skin deep: Judgments of individuals with facial disfigurement. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(1), 117–129. https://doi.org/10.1037/aca0000147

Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *NeuroImage*, *44*(3), 1125–1132. https://doi.org/10.1016/j.neuroimage.2008.10.009

Klebl, C., Greenaway, K. H., Rhee, J. J., & Bastian, B. (2021). Ugliness judgments alert us to cues of pathogen presence. *Social Psychological &*

Personality Science, *12*(5), 617–628. https://doi.org/10.1177/1948550620931655

Kniffin, K. M., & Wilson, D. S. (2004). The effect of nonphysical traits on the perception of physical attractiveness: Three naturalistic studies. *Evolution and Human Behavior*, *25*(2), 88–101. https://doi.org/10.1016/S1090-5138(04)00006-6

Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., Adolphs, R., & Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience*, *5*(4), 378–384. https://doi.org/10.1093/scan/nsq005

Konrath, S., & Handy, F. (2021). The good-looking giver effect: The relationship between doing good and looking good. *Nonprofit and Voluntary Sector Quarterly*, *50*(2), 283–311. https://doi.org/10.1177/0899764020950835

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmer Test package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Leder, H., Tinio, P. P. L., Fuchs, I. M., & Bohrn, I. (2010). When attractiveness demands longer looks: The effects of situation and gender. *Quarterly Journal of Experimental Psychology*, *63*(9), 1858–1871. https://doi.org/10.1080/17470211003605142

Lim, K. T. K., & Yu, R. (2015). Aging and wisdom: Age-related changes in economic and social decision making. *Frontiers in Aging Neuroscience*, *7*, Article 120. https://doi.org/10.3389/fnagi.2015.00120

Little, A. C., Burriss, R. P., Jones, B. C., Debruine, L. M., & Caldwell, C. A. (2008). Social influence in human face preference: Men and women are influenced more for long-term than short-term attractiveness decisions. *Evolution and Human Behavior*, *29*(2), 140–146. https://doi.org/10.1016/j.evolhumbehav.2007.11.007

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1638–1659. https://doi.org/10.1098/rstb.2010.0404

Luo, Q., Yu, M., Li, Y., & Mo, L. (2019). The neural correlates of integrated aesthetics between moral and facial beauty. *Scientific Reports*, *9*(1), Article 1980. https://doi.org/10.1038/s41598-019-38553-3

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Monin, B. (2003). The warm glow heuristic: When liking leads to familiarity. *Journal of Personality and Social Psychology*, *85*(6), 1035–1048. https://doi.org/10.1037/0022-3514.85.6.1035

North, M. S., & Fiske, S. T. (2015). Modern attitudes toward older adults in the aging world: A cross-cultural meta-analysis. *Psychological Bulletin*, *141*(5), 993–1021. https://doi.org/10.1037/a0039469

Owens, G., & Ford, J. G. (1978). Further consideration of the "what is good is beautiful" finding. *Social Psychology*, *41*(1), 73–75. https://doi.org/10.2307/3033601

Paunonen, S. V. (2006). You are honest, therefore I like you and find you attractive. *Journal of Research in Personality*, *40*(3), 237–249. https://doi.org/10.1016/j.jrp.2004.12.003

R Core Team. (2020). R: A language and environment for statistical computing (R Version 4.0.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*(1), 199–226. https://doi.org/10.1146/annurev.psych.57.102904.190208

Schopenhauer, A. (1902). *The wisdom of life* (T. B. Saunders, Trans.). https://onemorelibrary.com/index.php/en/?option=com_djclassifieds&format=raw&view=download&task=download&fid=14503

Skov, M. (2019). The neurobiology of sensory valuation. In M. Nadal & O. Vartanian (Eds.), *Oxford handbook of empirical aesthetics* (pp. 1–40). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.7

Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, Black, and male. *Criminology*, *36*(4), 763–798. https://doi.org/10.1111/j.1745-9125.1998.tb01265.x

Tsukiura, T., & Cabeza, R. (2011). Shared brain activity for aesthetic and moral judgments: Implications for the beauty-is-good stereotype. *Social Cognitive and Affective Neuroscience*, *6*(1), 138–148. https://doi.org/10.1093/scan/nsq025

Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: Individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology*, *97*(1), 103–122. https://doi.org/10.1037/a0015474

Wang, T., Mo, L., Mo, C., Tan, L. H., Cant, J. S., Zhong, L., & Cupchik, G. (2015). Is moral beauty different from facial beauty? Evidence from an fMRI study. *Social Cognitive and Affective Neuroscience*, *10*(6), 814–823. https://doi.org/10.1093/scan/nsu123

Whyte, S., Brooks, R. C., Chan, H. F., & Torgler, B. (2021). Sex differences in sexual attraction for aesthetics, resources and personality across age. *PLoS ONE*, *16*(5), Article e0250151. https://doi.org/10.1371/journal.pone.0250151

Wilson, R., & Eckel, C. (2006). Judging a book by its cover: Beauty and expectations in the trust game. *Political Research Quarterly*, *59*(2), 189–202. https://doi.org/10.1177/106591290605900202

Workman, C. I., & Chatterjee, A. (2021). The Face Image Meta-Database (fIMDb) & ChatLab Facial Anomaly Database (CFAD): Tools for research on face perception and social stigma. *Methods in Psychology*, *5*, Article 100063. https://doi.org/10.1016/j.metip.2021.100063

Workman, C. I., Humphries, S., Hartung, F., Aguirre, G. K., Kable, J. W., & Chatterjee, A. (2021). Morality is in the eye of the beholder: The neurocognitive basis of the "anomalous-is-bad" stereotype. *Annals of the New York Academy of Sciences*, *1494*(1), 3–17. https://doi.org/10.1111/nyas.14575

Workman, C. I., Yoder, K. J., & Decety, J. (2020). The dark side of morality – Neural mechanisms underpinning moral convictions and support for violence. *AJOB Neuroscience*, *11*(4), 269–284. https://doi.org/10.1080/21507740.2020.1811798

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, *28*(2), 238–249. https://doi.org/10.1177/0146167202282009

Zhang, Y., Kong, F., Zhong, Y., & Kou, H. (2014). Personality manipulations: Do they modulate facial attractiveness ratings? *Personality and Individual Differences*, *70*, 80–84. https://doi.org/10.1016/j.paid.2014.06.033

Zhao, N., Zhou, M., Shi, Y., & Zhang, J. (2015). Face attractiveness in building trust: Evidence from measurement of implicit and explicit responses. *Social Behavior and Personality*, *43*(5), 855–866. https://doi.org/10.2224/sbp.2015.43.5.855

Received April 19, 2021
Revision received September 17, 2021
Accepted October 5, 2021 ■

What is Good is Beautiful (and What isn't, isn't): How Moral Character Affects Perceived Facial Attractiveness

1. Supplementary Information

Face Stimuli Norming

24 younger and older faces were generated and selected in the following way (also see He et al., 2021):

First, 80 middle-aged faces were selected from the Chicago Face Database (Ma et al., 2015; http://www.chicagofaces.org/), which also provides researchers with information about each face (e.g., race, age, attractiveness). We then used the FaceApp software (https://www.faceapp.com/) to generate 80 sets of younger and older faces based on the middle-aged faces from the CFD.

Second, in order to standardize the stimuli, face images were (a) normalized to inter-pupillary distance using algorithms provided by the OpenCV computer vision library (https://opencv.org/) and facial landmarks provided by the dlib machine learning toolkit (http://dlib.net/), (b) resized and cropped to 345 pixels (width) × 407 pixels (height), (c) placed onto a plain white background using the GIMP 2 software package (https://www.gimp.org/), and (d) color corrected (Workman et al., 2021a, 2021b).

Third, an independent sample of n = 129 participants (race/ ethnicity: 102 white, 14 black, 6 Hispanic or Latinx, 3 Asian, 3 multiracial and 1 chose not to report), of which 33 were young (23 males; age: 28.82 ± 3.71 years; range: 20–34 years; education: 14.64 ± 2.56 years), 59 middle-aged (25 males; age: 47.05 ± 8.14 years; range: 35–59 years; education: 14.41 ± 2.71 years), and 37 older (11 males; age: 65.00 ± 4.22 years; range: 60–73 years; education: 14.92 ± 2.51 years), was recruited via Amazon Mechanical Turk to rate the computer-generated younger and older faces for attractiveness (how attractive do you find the person in the picture?) and realness (does the picture look like a real person?) on a scale from 1 to 7. Participants were also asked to indicate the age range of the faces (how old do you think the person in the picture is? e.g., 20–29 years). 43 sets of faces were selected based on the following criteria: 1) higher rates of being perceived as younger (20–29 years) and older (age 60 or older);

2) highest mean realness ratings.

Next, an independent sample of $n = 27$ participants (15 males; age: 26.81 ± 3.72 years; range: 22–36 years; education: 18.22 ± 2.64 years) was recruited via Amazon Mechanical Turk to judge whether each face from the three different ages belongs to the same person. The 30 sets of faces with the most accurate age group ratings were chosen (accuracy: 0.99 ± 0.005). Finally, after matching the stimuli on attractiveness, ethnicity, and sex, a final set of 12 face pairs was selected for use in the current study (Table S1).

Table S1

Information about the Face Stimuli

	Younger faces	Older faces
N	12	12
M/F	6/6	6/6
Age	20-29 *(67.93%)	60+ *(79.53%)
Attractiveness	4.53 (0.84)	3.20 (0.46)
Realness	5.13 (0.37)	5.57 (0.36)

Note. M - Male; F - Female. Information of younger and older faces derives from the results of face norming tasks in our previous study (He et al., 2021).

*On average, 67.93% participants rated the 12 computer-generated younger faces as 20-29 years; 79.53% participants rated the 12 computer-generated older faces as age 60 or older.

2. Supplementary Tables

Table S2

Fixed Effects from the Linear Mixed Models Constructed to Examine Effects of Vignette Type on Facial Attractiveness, Confidence, and Friendliness Ratings

Fixed Effects	β	*SE*	*t* value	*p* value
a. Attractiveness				
Intercept	4.356	.147	29.687	< .001
Vignette type (antisocial)	-.373	.018	-20.779	< .001
Vignette type (prosocial)	.138	.018	7.712	< .001
b. Confidence				
Intercept	4.793	.058	83.134	< .001
Vignette type (antisocial)	-.201	.017	-11.540	< .001
Vignette type (prosocial)	.419	.017	24.126	< .001
c. Friendliness				
Intercept	4.891	.053	91.771	< .001
Vignette type (antisocial)	-1.026	.019	-53.642	< .001
Vignette type (prosocial)	.557	.019	29.126	< .001

SE = standard error.

Table S3

Means and Standard Deviations for Facial Attractiveness, Confidence, and Friendliness Ratings Grouped According to Vignette Type

	Prosocial vignettes	Antisocial vignettes	Non-moral vignettes
Attractiveness	4.49 (1.67)	3.98 (1.79)	4.36 (1.66)
Confidence	5.21 (1.31)	4.59 (1.57)	4.79 (1.31)
Friendliness	5.45 (1.36)	3.87 (1.75)	4.89 (1.30)

Table S4

Fixed Effects from the Linear Mixed Model Constructed to Examine How the Effects of Moral Character Inferences on Facial Attractiveness Vary as Functions of Perceiver Sex and Face Sex

Fixed effects	β	*SE*	*t* value	*p* value
Intercept	4.506	.212	21.254	< .001
Vignette type (antisocial)	-.438	.041	-10.632	< .001
Vignette type (prosocial)	.108	.041	2.616	< .010
Face sex (male)	-.337	.275	-1.226	.233
Perceiver sex (male)	.011	.114	.093	.926
Vignette type (antisocial) * Face sex (male)	.057	.058	.970	.332
Vignette type (prosocial) * Face sex (male)	.038	.058	.650	.516
Vignette type (antisocial) * Perceiver sex (male)	.085	.052	1.634	.102
Vignette type (prosocial) * Perceiver sex (male)	.048	.052	.909	.363
Face sex (male) * Perceiver sex (male)	.038	.052	.725	.469
Vignette type (antisocial) * Face sex (male): Perceiver sex (male)	-.050	.074	-.678	.498
Vignette type (prosocial) * Face sex (male): Perceiver sex (male)	-.058	.074	-.786	.432

SE = standard error.

Table S5

Means and Standard Deviations for Facial Attractiveness According to Vignette Type and to Perceiver Sex and Face Sex

	Female face			Male face		
	Prosocial	Antisocial	Non-moral	Prosocial	Antisocial	Non-moral
Female perceiver	4.61 (1.63)	4.07 (1.80)	4.51 (1.64)	4.32 (1.67)	3.79 (1.79)	4.17 (1.66)
Male perceiver	4.67 (1.69)	4.16 (1.80)	4.52 (1.69)	4.35 (1.64)	3.87 (1.74)	4.22 (1.62)
Overall	4.65 (1.67)	4.13 (1.80)	4.51 (1.67)	4.34 (1.65)	3.84 (1.76)	4.20 (1.64)

Table S6

Fixed Effects from the Linear Mixed Model Constructed to Examine How the Effects of Moral Character Inferences on Confidence Ratings Vary as Functions of Perceiver Sex and Face Sex

Fixed Effects	β	*SE*	*t* value	*p* value
Intercept	4.859	.091	53.636	< .001
Vignette type (antisocial)	-.174	.040	-4.345	< .001
Vignette type (prosocial)	.365	.040	9.143	< .001
Face Sex (male)	-.089	.093	-.952	.348
Perceiver Sex (male)	-.085	.087	-.982	.327
Vignette type (antisocial) * Face Sex (male)	-.017	.057	-.299	.765
Vignette type (prosocial) * Face Sex (male)	.019	.057	.330	.741
Vignette type (antisocial) * Perceiver Sex (male)	-.005	.051	-.107	.915
Vignette type (prosocial) * Perceiver Sex (male)	.098	.051	1.935	.053
Face Sex (male) * Perceiver Sex (male)	.100	.051	1.978	< .050
Vignette type (antisocial) * Face Sex (male) * Perceiver Sex (male)	-.049	.072	-.679	.497
Vignette type (prosocial) * Face Sex (male) * Perceiver Sex (male)	-.052	.072	-.730	.465

SE = standard error.

Table S7

Means and Standard Deviations for Confidence Ratings According to Vignette Type and to Perceiver Sex and Face Sex

	Female face			Male face		
	Prosocial	Antisocial	Non-moral	Prosocial	Antisocial	Non-moral
Female perceiver	5.22 (1.34)	4.69 (1.61)	4.86 (1.33)	5.15 (1.32)	4.58 (1.60)	4.77 (1.36)
Male perceiver	5.24 (1.29)	4.59 (1.56)	4.77 (1.29)	5.22 (1.29)	4.54 (1.54)	4.79 (1.27)
Overall	5.23 (1.31)	4.63 (1.58)	4.81 (1.31)	5.19 (1.30)	4.56 (1.56)	4.78 (1.30)

Table S8

Fixed Effects from the Linear Mixed Model Constructed to Examine How the Effects of Moral Character Inferences on Friendliness Ratings Vary as Functions of Perceiver Sex and Face Sex

Fixed Effects	β	*SE*	*t* value	*p* value
Intercept	4.927	.085	57.863	< .001
Vignette type (antisocial)	-1.004	.044	-22.826	< .001
Vignette type (prosocial)	.521	.044	11.843	< .001
Face Sex (male)	-.043	.085	-.508	.615
Perceiver Sex (male)	-.034	.086	-.395	.693
Vignette type (antisocial) * Face Sex (male)	-.013	.062	-.215	.829
Vignette type (prosocial) * Face Sex (male)	.001	.062	.019	.985
Vignette type (antisocial) * Perceiver Sex (male)	-.008	.056	-.149	.881
Vignette type (prosocial) * Perceiver Sex (male)	.077	.056	1.372	.170
Face Sex (male) * Perceiver Sex (male)	.024	.056	.430	.667
Vignette type (antisocial) * Face Sex (male) * Perceiver Sex (male)	-.033	.079	-.419	.675
Vignette type (prosocial) * Face Sex (male) * Perceiver Sex (male)	-.038	.079	-.487	.626

SE = standard error.

Table S9

Means and Standard Deviations for Friendliness Ratings According to Vignette Type and to Perceiver Sex and Face Sex

	Female face			Male face		
	Prosocial	Antisocial	Non-moral	Prosocial	Antisocial	Non-moral
Female perceiver	5.45 (1.37)	3.92 (1.82)	4.93 (1.32)	5.41 (1.39)	3.87 (1.79)	4.88 (1.34)
Male perceiver	5.49 (1.35)	3.88 (1.72)	4.89 (1.31)	5.43 (1.36)	3.81 (1.72)	4.87 (1.27)
Overall	5.47 (1.36)	3.90 (1.76)	4.91(1.31)	5.42 (1.37)	3.83 (1.75)	4.88 (1.30)

References

He. D., Workman, C. I., Kenett, Y. N., He. X., & Chatterjee, A. (2021). The effect of aging on facial attractiveness: An empirical and computational investigation. *Acta Psychologica*, *219*, 103385. https://doi.org/10.1016/j.actpsy.2021.103385

Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., …Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience*, *5*(4), 378–384.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122-1135. https://doi.org/10.3758/s13428-014-0532-5

Workman, C. I., & Chatterjee, A. (2021a). The Face Image Meta-Database (fIMDb) & ChatLab Facial Anomaly Database (CFAD): Tools for research on face perception and social stigma. *Methods in Psychology*, *5*, 100063. https://doi.org/10.1016/j.metip.2021.100063

Workman, C. I., Humphries, S., Hartung, F., Aguirre, G. K., Kable, J. W., & Chatterjee, A. (2021b). Morality is in the eye of the beholder: The neurocognitive basis of the "anomalous-is-bad" stereotype. *Annals of the New York Academy of Sciences*, *1494*(1), 3-17. https://doi.org/10.1111/nyas.14575