Morality is in the eye of the beholder:
Unpacking the neurocognitive basis of the “anomalous-is-bad” stereotype

Clifford I. Workman[1-3]*, Stacey Humphries[1-3], Franziska Hartung[1-3],
Geoffrey K. Aguirre[1], Joseph W. Kable[2-4], & Anjan Chatterjee[1-3]

[1] Department of Neurology, The University of Pennsylvania, Philadelphia, PA

[2] Penn Brain Science Center, The University of Pennsylvania, Philadelphia, PA

[3] Penn Center for Neuroaesthetics, The University of Pennsylvania, Philadelphia, PA

[4] Department of Psychology, The University of Pennsylvania, Philadelphia, PA

***Correspondence:**
University of Pennsylvania
3710 Hamilton Walk
Goddard Laboratories, Room 332
Philadelphia, PA 19104
E-mail: cliffworkmanphd@gmail.com

Abstract: Are people with flawed faces regarded as having flawed moral characters? An “anomalous-is-bad” stereotype is hypothesized to facilitate biases towards people with facial anomalies (e.g., scars), although whether and how these biases affect behavior and brain functioning remain open questions. We replicated previously reported negative character evaluations made about individuals with facial anomalies, and further identified explicit biases directed against them as a group. People with anomalous faces were subjected to less prosociality from those participants highest in socioeconomic status. In amygdala, selective neural responding to anomalous faces—sensitive to beauty and disgust, but not explained by salience or arousal—correlated with stronger just world beliefs (i.e., people get what they deserve), less dispositional empathic concern, and less prosociality towards people with facial anomalies. Characterizing the “anomalous-is-bad” stereotype across and between levels of organization—i.e., attitudes, behavior, and brain–can reveal underappreciated psychological burdens shouldered by people who look different.

Keywords: morality; beauty; disgust; empathy; amygdala; behavioral economics

"You look nervous. Is it the scars? Do you want to know how I got them?"
– The Joker in the 2008 film, *The Dark Knight*

So that moviegoers might understand the nervous feelings wrought by The Joker's arrival, the writers of 2008's *The Dark Knight* instructed that he be shown in "sweaty clown makeup [that obscures] the awful scars which widen his mouth into a permanent, ghoulish smile" (page 26, lines 22-23[1]). He is one in a growing roster of Hollywood bad guys whose flawed moral characters are writ large on their faces [2]. *Harry Potter*, for instance, inspired fear in a reptilian Voldemort, *Nightmare on Elm Street* frightened us awake with Freddy Krueger's marred visage, and–most brazen of all–*The Lion King* taught us you cannot trust a Scar. Facial anomalies (e.g., scars, port-wine stains) may violate expectations about beauty, but why should they bear on judgments of moral character? Addressing this issue is critical for millions of people worldwide whose facial anomalies render them unjust targets of bias and discrimination [3–7].

Groundbreaking research by Dion and colleagues[8] indicated that people are more likely to assign positive traits to attractive people, an effect dubbed the "beauty-is-good" stereotype. Cues of facial beauty, such as averageness and symmetry, have been linked to positive health outcomes, suggesting preferences for beautiful faces are an adaptive way to assess the "quality" of potential romantic pursuits[9]. Brain imaging evidence suggests judgments of facial attractiveness and trustworthiness share a common neural substrate in amygdala, lending neurobiological plausibility to the notion that beauty and morality are intertwined[10]. A meta-analysis of 76 studies found a moderate influence of beauty on subsequent ascriptions of "good" personality traits[11]. Overall, the evidence that attractiveness influences character judgments is compelling. Considerable variation between studies, however, may indicate that evolutionary pressures actually favored an "unattractive-is-bad" stereotype[12], with differences in character attributions more pronounced when comparing *unattractive* to moderately attractive faces than *attractive* to moderately attractive faces. Generally, face perception extracts adaptive information[13] (e.g., detecting facial cues of anger may enable one to avoid life-threatening conflicts). Otherwise adaptive cues, however, may generalize beyond features with known links to evolutionary fitness. Facial anomalies, for instance, can signal poor health even when acquired and unrelated to an underlying illness, and may result in more potent signals than beauty, *per se*.

This study tested the hypothesis that, complementary to the beauty-is-good stereotype, an "anomalous-is-bad" stereotype facilitates biases against people with facial anomalies. We wish to be clear that we are not equating facial anomalies with ugliness. Although people with typical faces tend to rate anomalous faces as less attractive, anybody—regardless of beauty—can acquire anomalies. We recently provided three independent lines of support for the "anomalous-is-bad" hypothesis: First, people judged the character of individuals with facial anomalies more harshly before than after surgical correction[4]. Second, people demonstrated robust negative implicit biases against people with facial anomalies[3], replicating earlier reports[14]. Third, using fMRI, anomalous faces elicited blunted activation in dorsal anterior cingulate cortex (dACC)[3], which previous studies interpreted as a neural biomarker of dehumanization[15]. Dehumanization is defined as "perceiving a person or group as lacking humanness" (page 401)[16]. This tautological definition is symptomatic of theoretical difficulty in conceptualizing "humanness." One theoretical approach allows that dehumanized others can *appear* human while having subhuman essences. Alternatively, dehumanized others may retain their human essences while lacking distinctly human psychological capacities (i.e., not present in non-human animals).

Regardless of the mechanism, dehumanization manifests behaviorally in less prosociality and more antisociality[16]—whether the anomalous-is-bad stereotype facilitates dehumanizing behavior, however, remains an open question. This study used well-characterized paradigms from behavioral economics to examine the impact of facial anomalies on pro- and anti-sociality. These paradigms have been used for over 30 years to study cooperative and competitive behaviors and how they are shaped by social context, including effects of facial beauty (e.g., [17]). In our study, participants played economic games ostensibly with partners whose photographs depicted them with or without facial anomalies. This research examined whether and how anomalies influence moral *behavior* as opposed to moral *beliefs* about hypothetical situations. This distinction is important given mounting evidence that beliefs about hypothetical moral scenarios are poor predictors of actual behavior[18,19].

Although blunted dACC responding to anomalous faces was hypothesized by Hartung and colleagues[3] to reflect a neural marker of dehumanization, their study was not designed to rule out potentially important confounds. For instance, dACC responding could instead track the *beauty* of faces, since people with facial anomalies are regarded as less attractive on average than people without such anomalies [4]. Another possibility is that facial anomalies are particularly *salient*, or that faces with visible anomalies *arouse especially strong emotions* in others. Although visual salience processing does implicate the dACC[20], arousal is more closely linked to the amygdala[21]. Further, if anomalous faces are highly salient and/or arousing, they ought to elicit *increased* activation in dACC, not blunted activation as reported previously. General mechanisms are most assuredly necessary for the anomalous-is-bad stereotype (e.g., for detecting salient facial differences at all). Our focus, however, is on neural responses *selective* for facial anomalies, and on relations between such neural responses and manifestations of the anomalous-is-bad stereotype in attitudes and behavior.

The tendency to react to moral situations with prosocial versus antisocial emotions interacts with social values to shape behavior[22]. An antiegalitarian[23] who is sensitive to disgust will likely react differently to a homeless panhandler, for instance, than someone high in empathic concern[24] who believes the world is unjust[25]. Anomalous faces reportedly trigger disgust responses that are proportionate to disgust sensitivity[26]. When physically interacting with objects that had purportedly been handled by people who were healthy, had the flu, or had facial anomalies, participants in one study showed more facial disgust and avoided touching objects they believed had been touched by people with the flu *and* with facial anomalies[27]. These data are consistent with models suggesting that disgust evolved to facilitate pathogen avoidance[28]. This model suggests a *distal* causal explanation regarding the evolutionary emergence of an anomalous-is-bad stereotype (i.e., pathogen-aversion), and points towards pathogen-related disgust as the *proximate* psychological cause of dehumanizing behavior. A recent meta-analysis, however, reported that effects of incidental disgust on decision-making are small, especially after controlling for publication bias[29]. Disgust alone may be insufficient as a proximate causal explanation for the anomalous-is-bad stereotype, which may instead emerge from diminished prosocial moral emotions like empathic concern[16], and/or from values relating to justice and egalitarianism. Integrating established psychological and behavioral techniques with functional neuroimaging promises to advance our understanding of the causal mechanisms best suited to explain the stereotype.

In the first of two studies, participants completed an online survey to characterize their attitudes towards individuals with facial anomalies. In the second study, a rapid-event-related fMRI implementation of the oddball paradigm probed neural responses to anomalous faces (Fig. 1)[30]. Before scanning, participants learned a set of faces that were independently rated “average” in attractiveness. Inside the scanner, participants saw these familiar faces again but interspersed with novel faces. The novel faces either had visible anomalies, were non-anomalous and average-looking, or were non-anomalous and attractive. In what follows, the design of this task is leveraged along with attitudinal, behavioral, and dispositional measures to address four hypotheses:

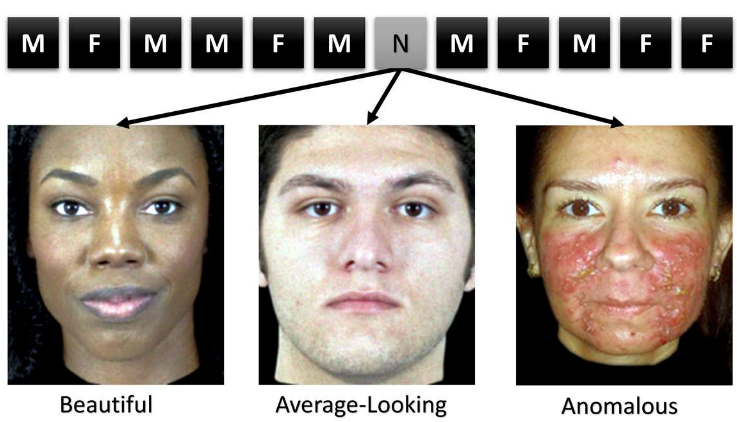
1. Complementary to the beauty-is-good stereotype, an “anomalous-is-bad” stereotype facilitates biases against people with visible facial anomalies. This stereotype operates outside of conscious awareness by prejudicing attitudes towards people with facial anomalies. Participants are predicted to make harsher character judgments about people with facial anomalies relative to typical faces,

replicating Jamrozik andcolleagues[4], and to display implicit biases against anomalous faces, replicating Hartung and colleagues[3].

2. The anomalous-is-bad stereotype facilitates dehumanizing behavior towards people with facial anomalies. Compared to typical faces, people with facial anomalies are expected to elicit more antisociality and less prosociality in behavioral economic games. *Alternatively*, if this hypothesis is incorrect, we would not observe effects of facial anomalies on behavior, or prosociality could increase and/or antisociality could decrease.
3. The dACC and lateral occipital cortices (LOC) respond to faces that are emotionally arousing, reflecting a confound that may explain earlier reports of neural responses to anomalous faces that were hypothesized to reflect dehumanization. The emotional arousal hypothesis predicts stronger activation for anomalous and attractive faces compared to average looking faces in LOC, with the opposite pattern in dACC, replicating Chatterjee et al.[31] and Hartung et al.[3]. *Alternatively,* if LOC and dACC are sensitive to beauty, the activation pattern should be most pronounced for attractive followed by average-looking (novel and learned) followed by anomalous faces. *Alternatively*, if LOC and dACC are sensitive to visual salience, the activation pattern should occur in response to novel (anomalous, beautiful, average-looking) compared to learned faces (average-looking).
4. These levels of organization—attitudes, behavior, brain, and *dispositions*—are interdependent, with information represented on some levels (i.e., in attitudes and dispositions) also represented on others (i.e., in brain activations and behavior). The neural response selective for people with facial anomalies, for instance, is predicted to track implicit biases against them, in line with the view that some neural responses to anomalies are indeed correlates of bias[3]. Such neural responses are predicted to correlate with sensitivity to pathogen-related disgust given the evidence reviewed above. *Alternatively*, participants may struggle to empathize with people in possession of anomalous faces, with selective neural responses to anomalies expected to track trait empathic concern.

Testing these hypotheses promises to expose an under-appreciated social and ethical concern—namely,

Figure 1 | Schematic of the fMRI task design. Participants completed 540 trials of an oddball task (5 repetitions each of 90 learned faces [450 total], 1 repetition each of 90 novel faces [30 anomalous, 30 beautiful, 30 average-looking]. Trials started with a photograph of a male or female face for 500ms, followed by a null event for 1000ms plus jitter lasting an average of 1833ms (3,333 ms per trial). F, female; M, male; N, novel.



that people with facial anomalies are judged and treated adversely because of how they look, which may have no bearing on their actual character.

Results

Hypothesis 1: Attitudinal Correlates of the "Anomalous-is-Bad" Stereotype

Explicit Character Judgments (Study 1)

People with facial anomalies were rated less trustworthy than people with beautiful faces, and less attractive, less content, and more anxious than people with both beautiful and average-looking faces (Table S4). Raters felt more aroused when looking at anomalous compared to average-looking but not beautiful faces. Raters also reported feeling less happy when looking at anomalous relative to typical faces. Ratings of perceiver and face dominance failed to significantly replicate earlier findings. Equivalence testing together with null-hypothesis testing indicated attractiveness, trustworthiness, and anxiousness were statistically not different from zero and statistically equivalent to zero when comparing novel and learned average-looking faces ($p < 0.05$). Contentedness and dominance were equivalent to zero but also statistically different from zero (contentedness: $t(6623) = -2.847$, $p = 0.004$; dominance: $t(6264) = -6.451$, $p < 0.001$). Even though a significant effect was detected, this was likely due to the precision afforded by the sample size, given that the effect was smaller than the *a priori* smallest effect size of interest and, as such, is equivalent in a practical sense. SAM arousal, SAM dominance, and SAM happiness were statistically not different from zero and statistically equivalent to zero when comparing novel and learned average-looking faces ($p < 0.05$). The means and standard deviations for each of these variables is given in Table S2.

Implicit and Explicit Biases (Study 2)

One-tailed non-parametric Wilcoxon signed rank tests were used to detect explicit and implicit biases against anomalous faces in participants. With respect to explicit bias, scores on the EBQ were elevated (expected: 0; Median = 0.071; V(32) = 335, $r = 0.296$, $p = 0.047$, 95% CI [0.03, 0.60]) in contrast to a previous report by Hartung et al.[3]. With respect to implicit bias, elevated scores on the IAT did not obtain statistical significance (expected: 0; Median = 0.200; V(31) = 318, $r = 0.246$, $p = 0.088$, 95% CI [0.01, 0.54]). Nevertheless, IAT d' scores ranging between 0.15 and 0.35 are indicative of slight bias.

Hypothesis 2: Behavioral Correlates of the "Anomalous-is-Bad" Stereotype

Prosocial Behavior in the Dictator Game (Study 2)

To examine *prosocial* behavior towards people with facial anomalies in the DG, a model was constructed that regressed the prosociality of each decision against face type (Anomalous Face Pre-Surgery | Anomalous Face Post-Surgery) by SES by the magnitude difference between each set of presented options, with random intercepts for subject and stimulus. There was a significant interaction between face type, SES, and difference between options ($\beta = -1.002$, SE = 0.464, $z = -2.131$, $p = 0.031$; see Table 1A for remaining fixed effects). Participants with the highest SES—that is, those most likely to have expendable income–were increasingly likely to act as prosocially as possible towards people with "typical" faces (i.e., anomalous faces rendered such by corrective surgery), with the opposite pattern characterizing the treatment of people with facial anomalies (Fig. 3A).

Antisocial Behavior in the Ultimatum Game (Study 2)

To examine *antisociality* towards people with facial anomalies in the UG, a model was constructed that regressed rejections against face type (Anomalous | Beautiful) by SES by the difference in each proposed

Figure 2 | Responses to the online survey of character judgments about people with and without facial anomalies. For semantic differential items (top group of 5 plots), participants responded on scales from 1 to 5. For SAM items (bottom group of 3 plots), participants responded on scales from 1 to 9. Keys positioned atop each group of plots show the colors corresponding to each possible option. Green options were positive, red were negative, and the neutral midpoint is grey. The proportion of participants who chose a given answer is represented in the width of the color corresponding to that option. Anomalous faces were rated less attractive than beautiful, novel average-looking, and learned average-looking faces. They were seen as less trustworthy than beautiful faces but not novel or learned average-looking faces. They were rated less content than beautiful, novel average-looking, and learned average-looking faces. They were also perceived to be less dominant than learned average-looking faces, with no differences for beautiful and novel average-looking faces. Finally, anomalous faces were regarded as more anxious than beautiful, novel average-looking, and learned average-looking faces. With respect to perceiver responses, arousal assessed with the SAM was increased for anomalous compared to novel average-looking faces, with no differences when compared to learned average-looking and beautiful faces. SAM dominance was elevated in response to anomalous compared to beautiful and novel average-looking faces but not learned average-looking faces. The opposite pattern characterized SAM happiness, with anomalous faces evoking less happiness than beautiful, novel average-looking, and learned average-looking faces. SAM, Self-Assessment Manikin.

split, with random intercepts for subject and stimulus. This model examined whether participants were especially likely to reject even the fairest offers ostensibly made by anomalous relative to typical faces. The interaction between face type (Anomalous | Beautiful) and SES approached but did not reach significance (β =1.547, SE = 0.818, z = 1.892, p = 0.059; see Table 1B for remaining fixed effects).

Hypothesis 3: Neural Correlates of the “Anomalous-is-Bad” Stereotype

Neural Sensitivity to Visual Salience (Study 2)

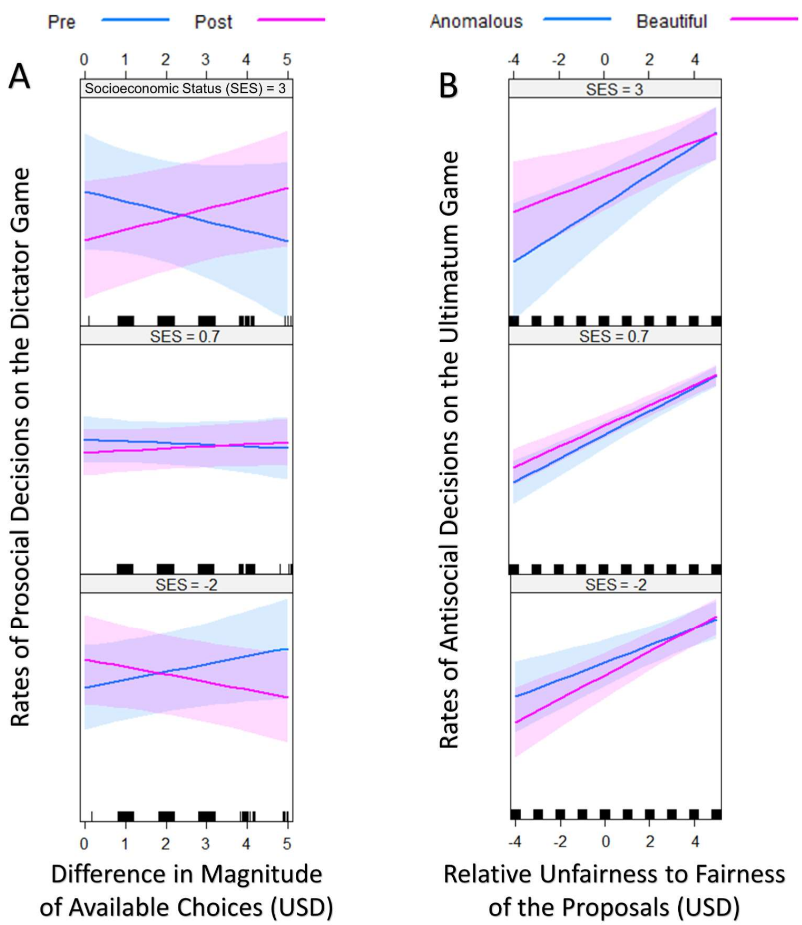
Relative to faces learned outside the scanner, appearances of novel faces elicited stronger activations in bilateral anterior insula (aINS), bilateral supplementary motor area (SMA) / dACC, bilateral ventral striatum (VS), bilateral fusiform gyri, left thalamus, right hippocampus, bilateral inferior parietal lobule (IPL), bilateral dorsolateral prefrontal cortex (dlPFC), left lateral orbitofrontal cortex (latOFC), left inferior frontal gyrus (IFG), right posterior cingulate cortex (PCC), midbrain, and cerebellum (Table S5A; Fig. 4A). Learned faces, by contrast, elicited stronger activation in left hippocampus, bilateral PCC, right precuneus, left ventromedial prefrontal cortex (vmPFC), bilateral VS (non-overlapping with the bilateral VS cluster implicated by the opposite contrast), bilateral middle frontal gyri, right precentral gyrus, right superior temporal gyrus (STG), right posterior insula, left middle temporal gyrus (MTG), and right caudate (Table S5B; Fig. 4A).

Table 1 | Fixed effects from the linear mixed models constructed to examine the Dictator and Ultimatum Games.

A: Dictator Game				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	-8.173	0.969	-8.431	< 0.001
Face type (beautiful)	0.646	0.499	1.294	0.196
Difference of each proposed split	2.120	0.168	12.610	< 0.001
SES	-1.395	1.122	-1.243	0.214
Face type (beautiful) * Difference of each proposed split	-0.066	0.151	-0.434	0.665
Face type (beautiful) * SES	1.547	0.818	1.892	0.059
Difference of each proposed split * SES	0.223	0.196	1.141	0.254
Face type (beautiful) * Difference of each proposed split * SES	-0.345	0.205	-1.679	0.093
B: Ultimatum Game				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	1.971	1.412	1.396	0.163
Face type (pre-surgery)	0.353	0.970	0.364	0.716
Difference between choices	-0.059	0.264	-0.222	0.824
SES	-2.131	1.370	-1.556	0.120
Face type (pre-surgery) * Difference between choices	0.160	0.424	0.377	0.706
Face type (pre-surgery) * SES	2.154	1.052	2.047	0.041
Difference between choices * SES	0.505	0.262	1.925	0.054
Face type (pre-surgery) * Difference between choices * SES	-1.002	0.464	-2.158	0.031

SE, standard error; SES, socioeconomic status.

Figure 3 | A. Regarding prosociality in the Dictator Game (y-axis), there was a significant interaction between face category (anomalous faces before corrective surgery are represented by blue lines in the leftmost plots; faces after corrective surgery are represented by pink lines in the leftmost plot), socioeconomic status (SES; represented across 3 bins: 3.0, 0.7, and -2.0), and difference between options (x-axis; US dollars [USD]). Each slope represents the estimated relationship between prosociality and the difference between options as a function of SES. The error bars represent 95% confidence intervals. Participants with the highest SES were increasingly prosocial towards faces that had undergone corrective surgery (top right), whereas the opposite pattern was observed towards anomalous faces (top left). **B.** Regarding antisociality in the Ultimatum Game (y-axis), the interaction between face category (anomalous faces are represented by blue lines in the rightmost plots; typical faces are represented by pink lines in the rightmost plots) and SES (across the same 3 bins) was shy of reaching significance.



Neural Sensitivity to Emotional Arousal (Study 2)

Stronger activation was observed for highly arousing faces (novel beautiful and anomalous) relative to faces low in arousal (novel average-looking) in a single cluster in right fusiform gyrus (Table S6A; Fig. 4B). Average-looking faces did not elicit stronger activations than beautiful and anomalous faces in any regions. In addition to the oddball task, a functional localizer task identified neural responses sensitive to emotional arousal. Relative to images low in emotional arousal (e.g., a light bulb), highly arousing negative imagery (e.g., a spider) resulted in stronger neural responses in bilateral fusiform, left IFG, and bilateral precentral gyri (Table S6C; Fig. 4B), and weaker responses elsewhere in bilateral fusiform and in left PCC (Table S6D; Fig. 4B).

Neural Sensitivity to Facial Trustworthiness (Study 2)

Across all faces, no significant positive or negative parametric relations were detected between hemodynamic response amplitudes and average facial trustworthiness ratings.

Neural Sensitivity to Disgust (Study 2)

During two blocks of the functional localizer task, participants viewed images rated high on disgust, along with scrambled versions of those same images. Given that neural responses to disgusting images were compared to scrambled versions of those same images, the disgust localizers may have been confounded with simple visual responding (e.g., to objects versus non-objects). In attempting to characterize selective neural responses to facial anomalies below, this hypothesis is tested by *inclusively* masking for disgust while *exclusively* masking for visual salience—the resulting statistical map includes only areas that were sensitive to disgust and insensitive to simple visual responding. In one block, the images depicted typically disgusting animals such as leeches and cockroaches, whereas in the other block the images depicted bodily infections and wounds. Relative to their scrambled counterparts, images of disgusting animals elicited greater activation in bilateral amygdala, bilateral fusiform gyri, bilateral superior occipital cortex, bilateral superior parietal lobule (SPL), and midbrain (Table S7C; Fig. 4C). The reverse contrast identified a single cluster in precuneus (Table S7B; Fig. 4C). Larger neural responses to disgusting injuries, on the other hand, were observed in bilateral fusiform, bilateral precuneus, right IPL, right MTG, and left precentral gyrus (Table S7C; Fig. 4C). The reverse contrast identified clusters in right SPL, left putamen, left MTG, and bilateral STG (Table S7D; Fig. 4C).

Neural Sensitivity to Facial Beauty (Study 2)

No significant positive relations between hemodynamic response amplitudes and average facial beauty ratings were detected. Significant negative relations, however, were observed in bilateral fusiform gyri, bilateral amygdala, bilateral dlPFC, bilateral intraparietal sulcus (IPS), left temporoparietal junction, left IFG, left thalamus, left MTG, and left SMG (Table S8; Fig. 4D).

Selective Neural Responses to Facial Anomalies (Study 2)

Relative to beautiful and average-looking novel faces, anomalous faces increased activation in bilateral fusiform gyri, bilateral amygdala, right hippocampus, right latOFC, right VS, right superior occipital cortex, left IFG, bilateral IPS, and left IPL (Table S9A). Further, decreased activation was detected in bilateral STG and bilateral somatosensory cortices (Table S9B). This analysis was repeated with an inclusive mask representing disgust towards animals (as reported in Table S7A & B) and exclusive masking for visual salience and emotional arousal (as reported in Tables S5A and S5B and in S6A, S6C, and S6D), with increased activation restricted to bilateral fusiform gyri and left amygdala (Table S10A), and no significant areas of decreased activation detected.

The analysis was then repeated with inclusive masking for disgusting injuries (as reported in Table S7C) and exclusive masking for salience and arousal. With masking, increased activation in response to anomalous relative to typical faces was detected in bilateral fusiform gyri (Table S10C), overlapping with the fusiform clusters activated by disgusting animals (Table S7A), and no areas of decreased activation were found. Next, the analysis was repeated with inclusive masking for facial beauty (as reported in Table S8B) and exclusive masking for visual salience and emotional arousal. Increased activation in response to anomalous compared to beautiful and average-looking novel faces was detected in bilateral fusiform gyri, bilateral amygdala, right posterior superior temporal sulcus, left IFG, left IPL, and right IPS

Figure 4 | A. First, the design of the oddball fMRI paradigm was leveraged to characterize the neural response to visually salient faces. Clusters in warm colors showed increased activation to visually salient faces, whereas clusters in cool colors show decreased activation. **B.** To isolate the neural response to emotional arousing faces, images of beautiful and anomalous faces were compared against average-looking faces that are low in arousal. These results are shown together with the results of a functional localizer task to identify neural responses sensitive to arousal. Clusters in warm colors showed increased activation with increasing emotional arousal, whereas clusters in cool colors were negatively related to arousal. **C.** During two blocks of the functional localizer task, participants were presented with images rated high on disgust, along with scrambled versions of those same images—in one block, the images depicted typically disgusting animals such as leeches and cockroaches (top), whereas in the other the images depicted bodily infections and wounds (bottom). Clusters in warm colors were more strongly activated by disgusting images compared to scrambled versions of the same images, whereas clusters in cool colors capture the reverse contrast. **D.** A parametric modulation analysis examined relations between hemodynamic response amplitudes to each face presented during the oddball task and the average attractiveness ratings they received in the online survey study. Clusters in cool colors demonstrated a negative parametric relationship between hemodynamic response amplitudes and the attractiveness ratings for the faces shown to participants during the oddball task. Results were considered significant at a cluster-level FWE-corrected α = 0.01.

1

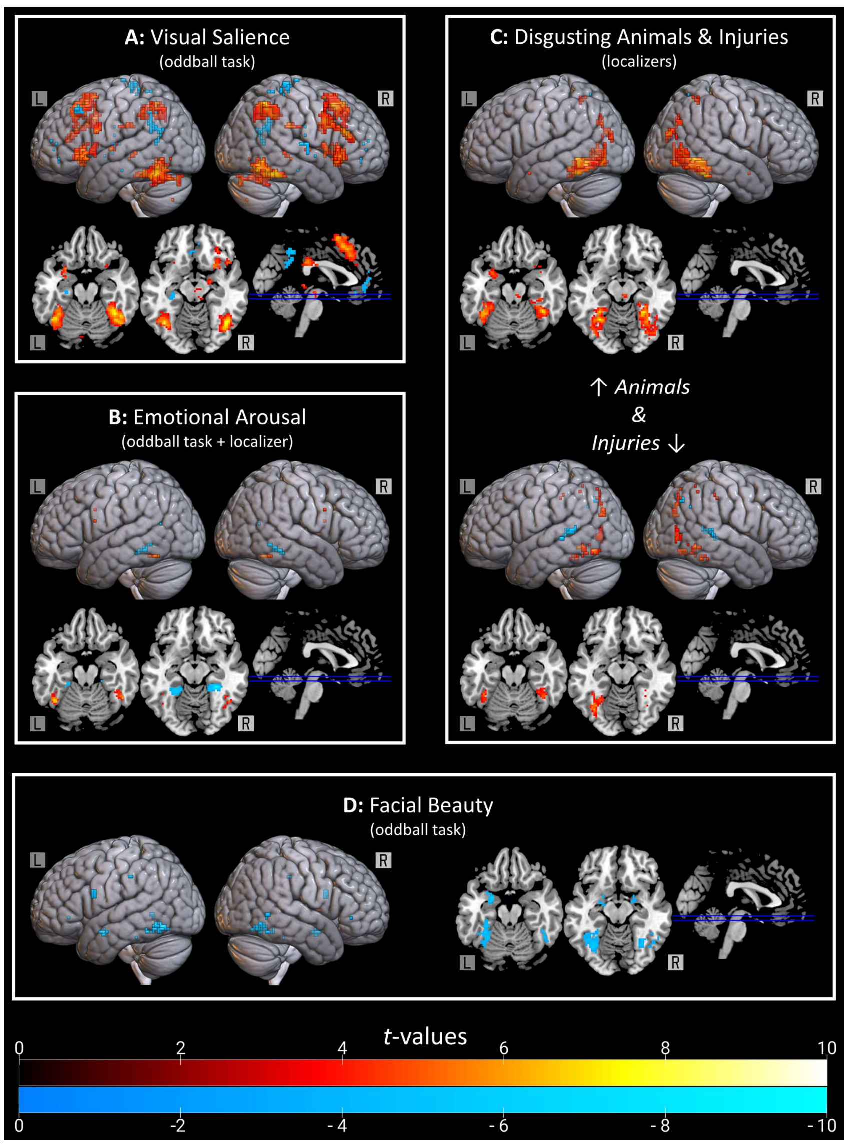


Table 2 | Neural responses selective for faces with visible anomalies.

A: Anomalous > (Beautiful + Average-Looking Novel Faces) Masking: Inclusive Disgusting Animals & Facial Beauty, Exclusive Visual Salience & Emotional Arousal					
Brain Regions by Cluster	**Cluster Size**	**Peak MNI Coordinates**			**Peak *t* Statistic**
		x	*y*	*z*	
L fusiform	92	-35	-62	-13	8.16
L amygdala	13	-25	1	-19	7.01
R fusiform	22	33	-57	-13	6.11
B: (Beautiful + Average-Looking Novel) > Anomalous Faces Masking: Inclusive Disgusting Animals & Facial Beauty, Exclusive Visual Salience & Emotional Arousal					
NS					

L, left; MNI, Montreal Neurological Institute; NS, non-significant; R, right.

(Table S10E). No areas of decreased activation were found.

Hypothesis 4: The “Anomalous-is-Bad” Stereotype Across Levels of Organization

Implicit Attitudes & Neural Responses to Anomalous Faces (Study 2)

When comparing anomalous against beautiful and average-looking novel faces, a positive correlation was detected between IAT d’ scores and activation in bilateral fusiform gyri, bilateral amygdala, right hippocampus, left latOFC, left IFG, left IPL, bilateral IPS (Table S11A). Negative correlations were found in left precuneus, left MTG, bilateral STG, and bilateral somatosensory cortices (Table S11B). Next, this analysis was repeated with inclusive masking for facial beauty (as reported in Table S8B) and disgust towards animals (as reported in Table S7A & B) and exclusive masking for visual salience and emotional arousal (as reported in Tables S5 & S6), with significant positive correlations restricted to bilateral fusiform gyri and left amygdala (Table 3A; Fig. 5A) and no significant negative correlations observed. Of note, the cluster in right fusiform gyrus from Table 2 was negatively associated with implicit biases towards anomalous faces (r_s[27] = -0.515, p = 0.007, 95% CI [-0.752, -0.159]; this non-parametric correlation alone survived correction for multiple comparisons).

Dispositions, Behavior, & Neural Responses to Anomalous Faces (Study 2)

Parameter estimates extracted from an ROI placed in the dACC described by Hartung and colleagues (2019) were negatively related to scores on the empathic concern subscale of the IRI (r_s[27] = -0.431, p = 0.025, 95% CI [-0.697, -0.061]), such that greater empathic concern correlated with less activation in dACC in response to anomalous compared to typical faces. Parameter estimates from the left amygdala cluster reported in Table 2 were positively related to procedural justice beliefs towards others (r_s[27] = 0.463, p = 0.015, 95% CI [0.101, 0.717]; Fig. 6B) and negatively correlated with dispositional empathic concern (r_s[27] = -0.411, p = 0.033, 95% CI [-0.684, -0.037]; Fig. 6C). The same left amygdala cluster demonstrated a negative correlation with the frequency of prosocial decisions towards anomalous (r_s[27] = -0.430, p = 0.025, 95% CI [-0.697, -0.060]) but not typical faces (r_s[27] = -0.302, p = 0.126, 95% CI [-0.612, 0.088]). Of note, similar relations characterized parameter estimates extracted from the fusiform gyri bilaterally (both from Table 2; left: r_s[27] = -0.383, p = 0.048, 95% CI [-0.666, -0.004]; right: r_s[27] = -0.345, p = 0.078, 95% CI [-0.621, 0.074]).

Table 3 | Correlations between implicit biases and neural responses selective for facial anomalies.

A: Positive Correlation with IAT d' Scores, Anomalous > (Beautiful + Average-Looking Novel Faces) Masking: Inclusive Disgusting Animals, Exclusive Visual Salience & Emotional Arousal					
Brain Regions by Cluster	**Cluster Size**	**Peak MNI Coordinates**			**Peak *t* Statistic**
		X	*y*	*z*	
L fusiform	87	-35	-62	-13	8.27
R fusiform	20	33	-57	-13	7.19
L amygdala	13	-25	1	-19	6.61
B: Negative Correlation with IAT d' Scores, Anomalous > (Beautiful + Average-Looking Novel Faces) Masking: Inclusive Disgusting Animals, Exclusive Visual Salience & Emotional Arousal					
NS					

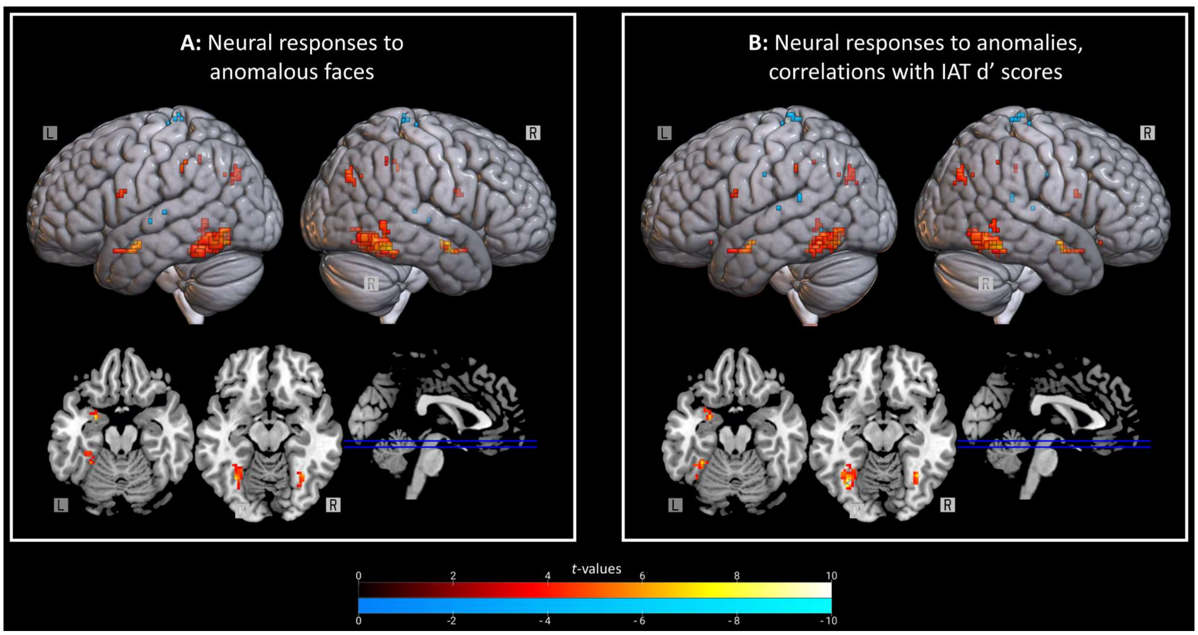
L, left; MNI, Montreal Neurological Institute; NS, non-significant; R, right.

Discussion

Motivated by the well-documented “beauty is good” stereotype, this study tested the hypothesis that there exists a complementary “anomalous-is-bad” stereotype. The proximate causes underpinning the stereotype were characterized by integrating attitudinal, behavioral, neural, and dispositional measures. Regarding attitudes, our findings confirm that people with facial anomalies were seen as less trustworthy than people with beautiful faces. They were also seen as less attractive, less content, and more anxious than people with beautiful and average-looking faces. Participants felt more aroused in response to anomalous compared to average-looking but not beautiful faces. They also felt less happy when looking at anomalous relative to typical faces. Ratings on dominance—both of faces and of rater experiences—failed to replicate earlier findings. The majority of results from the online survey study, however, did replicate the negative character evaluations of people with facial anomalies reported by Jamrozik and colleagues[4]. In the second study, explicit negative evaluations of people with facial anomalies as a group were slightly elevated, in contrast to the absence of explicit bias reported by Hartung and colleagues[3]. Scores on the IAT indicated only slight bias and were not significantly different from zero. Despite deviating from prior studies, we find evidence for the anomalous-is-bad stereotype in explicit negative attitudes about people with facial anomalies as individuals and as a group.

The significance of attitudinal manifestations of the anomalous-is-bad stereotype depends on whether these attitudes facilitate dehumanizing behavior. This point was highlighted by a meta-analysis that did not find evidence that the IAT predicts discrimination[32]. With this caution in mind, the present research tested the hypothesis that the anomalous-is-bad stereotype manifests in behavior as decreased prosocial and/or increased antisocial behavior towards people with anomalous faces. Participants highest in SES—that is, those most likely to have expendable income—were less giving towards individuals with anomalous relative to typical faces on the Dictator Game. Antisociality assessed with the UG, on the other hand, was not significantly elevated against people with facial anomalies. While the DG contrasted prosociality towards otherwise average-looking people with anomalous faces before and after surgical correction, the UG contrasted antisociality towards anomalous and beautiful faces. It is therefore surprising that prosociality was more strongly affected than antisociality, since differences in behavior on the UG should have been exacerbated by simultaneous operation of both the beauty-is-good and anomalous-is-bad stereotypes. These findings suggest evolutionary pressures weigh more heavily on the anomalous-is-bad stereotype, which is corroborated by the observation of only *negative*

Figure 5 | A. Regions in which anomalous faces elicited greater activation compared to beautiful and average-looking novel faces (top; no masking applied). Although bilateral fusiform activation in response to anomalous faces was detected in response to disgusting animals, disgusting injuries, and facial beauty (bottom; selective masking applied), only disgusting animals and facial beauty converged in implicating the left amygdala. Clusters in warm colors showed increased activation in response to anomalous compared to novel beautiful and average-looking faces, whereas cool colors represent decreased activation to the same faces. **B.** Regions where implicit bias scores correlated with activation in regions sensitive to facial anomalies (top; no masking applied). This analysis was then repeated with inclusive masking for facial beauty and disgust towards animals and exclusive masking for visual salience and emotional arousal, with significant positive correlations restricted to bilateral fusiform gyri and left amygdala (bottom; selective masking applied). As before, clusters in warm colors showed increased activation in response to anomalous compared to novel beautiful and average-looking faces—in addition, however, these cluster were further positively correlated with implicit biases. Clusters in cool colors showed negative correlations with implicit biases. Results were considered significant at a cluster-level FWE-corrected $\alpha = 0.01$. IAT, Implicit Association Test.



correlations between unattractiveness and hemodynamic response amplitudes across faces.

Our experimental design allowed us to disambiguate neural activity patterns elicited by visually salient faces, by emotionally arousing faces, and *selectively* by anomalous faces. The neural response to visual salience was characterized by activations in the anterior insula, dACC, ventral striatum, fusiform, lateral prefrontal cortices, and posterior cingulate—consistent with studies describing a salience network anchored in the aINS and dACC and with nodes in amygdala and ventral striatum[20]. The neural response related to emotional arousal implicated the fusiform, precentral, and posterior cingulate cortices, consistent with early functional neuroimaging work that revealed sensitivity to emotional arousal in the fusiform bilaterally[33]. The neural response to facial anomalies captured a fronto-temporo-parietal network with peaks in lateral OFC, fusiform, and amygdala. Next, we tested the possibility that the neural response to facial anomalies implicates brain regions that are not responsive to visual salience or emotional arousal—that is, regions that *selectively* respond to anomalous faces. Selective neural sensitivity to facial anomalies was restricted to the bilateral fusiform gyri and left amygdala.

Does the psychological conflation of beauty and goodness have a biological tether? Our results suggest "yes" and, considered in conjunction with the literature reviewed above, point to four candidate brain areas: dACC, OFC, fusiform, and amygdala. Regarding dACC, we expected anomalous faces to elicit blunted activation compared to typical faces in view of Hartung and colleagues[3]. This expectation is consistent with a previous study that reported transient disruption to the dACC that impaired the integration of aesthetic and moral information[34]. In our study, only the neural response to visual salience implicated the dACC. Additionally, blunted dACC responding to facial anomalies correlated with *elevated* trait empathic concern in contrast to our initial predictions. Although unexpected, similar discrepancies characterize the literatures on the brain basis of moral and aesthetic evaluations[35,36] and on implicit biases following transient and permanent lesions to the dACC. Milne and Grafman[37] found no evidence for an effect of damage to the dACC on implicit biases, suggesting this region is not *necessary* for the "normal" operation of such biases. Although replicated[38], several instances of discrepant findings have also been reported (e.g., [39]). Together, the available evidence suggests blunted dACC responding to anomalous faces is not a consistent neural marker of the anomalous-is-bad stereotype.

Emerging evidence suggests that the lateral and medial orbitofrontal cortices are shared neural substrates of moral and aesthetic evaluation[35,36]. The studies comprising this evidence characterized shared brain activity *during* aesthetic and moral evaluations, often comparing different kinds of judgments (e.g., of beauty and of righteousness[35], or of the beauty *of righteousness*[36]) made about different kinds of stimuli (e.g., of poems and moral beliefs[35], or of faces and moral scenarios[36]). In our study, differences in OFC activity elicited by anomalous compared to typical faces were not detected. Instead, the OFC appeared to track the *salience* of faces, suggesting earlier reports of shared OFC recruitment by aesthetic and moral evaluation may be attributable to the salience of beautiful and morally relevant stimuli. Alternatively, this discrepancy may be attributable to differences in what participants were tasked with doing inside of the scanner. In our study, participants did not evaluate the aesthetic or moral qualities of the faces they saw – even still, neural responses selective for anomalous faces predicted future behavior towards the people ostensibly in possession of those faces.

We provide evidence that neural activity in amygdala and fusiform are candidate neural biomarkers specific to the anomalous-is-bad stereotype. These regions were sensitive both to morality and beauty. This was established with inclusive and exclusive masking to restrict neural responses to facial anomalies to areas that were also sensitive to moral emotion (i.e., disgust) and facial beauty (i.e. parametric modulation of unattractiveness), and which excluded areas involved in salience and arousal. Although we focused on *selective* neural responses to facial anomalies, future research could establish whether and to what extent general psychological processes are constitutive of the anomalous-is-bad stereotype. The inclusive mask for disgusting animals included the amygdala, fusiform, and precuneus, whereas the mask for disgusting bodily injuries included the fusiform and precuneus but not left amygdala. The neural response to disgusting images did not capture the anterior insula, which was unexpected in view of the extant literature on the functional neuroanatomy of disgust[40]. Amygdala responses to disgusting stimuli, however, have been reported in the absence of significant effects in the anterior insula[41]. The mask for facial unattractiveness included the same amygdala region associated with processing disgusting animals (but not injuries), along with bilateral fusiform regions sensitive to both kinds of disgusting imagery. Three clusters responsive to facial anomalies survived masking—one in left fusiform, another in right fusiform, and a final cluster in left amygdala. Activations in all three clusters were associated with implicit biases, suggesting selective neural responses to people with facial anomalies in the fusiform gyri and amygdala underpin negative implicit biases towards them. Notably, though, facial beauty and disgusting animals alone—not injuries—converged in implicating left amygdala.

The amygdala is a key node in the neural circuitry underpinning moral cognition and behavior, with amygdala activation associated with feeling bad when faced with difficult moral choices[19]. For anomalous relative to typical faces, parameter estimates extracted from an ROI in the left amygdala cluster were positively related to beliefs about justice towards others and negatively related to empathic concern. Given that the left amygdala responded to images of disgusting animals that are known disease carriers, we might have anticipated that amygdala activation would correlate with pathogen-related disgust sensitivity, rather than with justice beliefs and emotional empathy. The same amygdala ROI showed a negative correlation with prosociality towards anomalous but not typical faces, with similar relations characterizing the fusiform. These findings are surprising since earlier work suggested that antisocial moral emotions (e.g., disgust) preferentially recruit the amygdala and fusiform, whereas prosocial emotions (e.g., empathic concern) recruit ventromedial prefrontal and mesolimbic regions (e.g., subgenual cingulate cortex[22]). Here, activation in a region typically associated with *antisocial* moral emotions—amygdala—was associated with *prosocial* dispositions and behavior.

We propose that the left amygdala integrates face perception with moral emotions (e.g., empathy) and social values (e.g., pertaining to justice) to guide behavior. A combination of diminished affective empathy interacting with stronger just world beliefs might facilitate dehumanizing people with facial anomalies. At the level of neural circuitry, the amygdala shares afferent and efferent connections with ventromedial prefrontal and mesolimbic cortices, which are implicated in representing prosocial moral emotions like guilt and empathy[22,42,43]. A previous investigation linked the same dispositional empathic concern subscale used here to the representation of guilt in the subgenual cingulate[44]—dysfunctional connectivity between this subgenual region and the amygdala was related to pathological guilt[45]. On the opposite end of the spectrum, fronto-amygdala dysfunction has been reported in psychopathy, in which empathy is impaired[46]. In healthy people, activations in these same regions were related to moral convictions and beliefs about violence against out-group members[47]. These studies suggest that the left amygdala cluster we identified may encode socioemotional information about the value of prospective social partners that is informed by their facial characteristics. We hypothesize that this encoding is achieved through the amygdala's connections with fronto-mesolimbic regions implicated in affective empathy and subjective valuation more broadly[22,48]. Given that the sgACC was captured by the visual salience contrasts in our study, fronto-amygdala connections may tether the selective neural response to facial anomalies to more general psychological processes involved in attention and salience. This perspective is consistent with recent evidence that the amygdala are hubs [49], with specific subnuclei integrating aversive emotional information[50] with representations of social value[51,52].

Two lines of evidence presented here corroborate the proposal that dehumanization facilitates attitudinal, behavioral, and neural manifestations of the anomalous-is-bad stereotype[3,4]. First, masking neural responses to anomalous faces with responses to disgusting animals, but not disgusting bodily injuries, captured the left amygdala cluster reported in Table 2. This finding suggests the left amygdala is preferentially activated, not by disgust towards human body parts, but by disgust towards non-human animals. This interpretation is consistent with work implicating the amygdala in the dehumanization of extreme outgroup members[15]. Second, stronger selective neural responses to facial anomalies were associated with lower dispositional empathic concern and negatively correlated with prosocial behavior towards people with facial anomalies. These findings are consistent with theorizing that suggests the capacity to dehumanize others is associated with empathic functioning[16].

The studies comprising this research integrated attitudinal, behavioral, neuroimaging, and dispositional methods to firmly establish the relationship between beauty and morality. In the first study, we replicated the negative characterological judgments made of people with anomalous faces reported previously. The second study detected explicit bias against people with facial anomalies, while the slight implicit bias that we observed failed to reach significance, thus diverging from evidence presented by Hartung and colleagues[3]. Overall, these findings support the operation of an “anomalous-is-bad” stereotype. Beyond these replications, we also found evidence for a behavioral manifestation of the anomalous-is-bad stereotype, reflected in less prosociality on the Dictator Game. A selective neural response to anomalous relative to typical faces—not explained by visual salience or emotional arousal—was detected in the left amygdala and was sensitive to both a lack of facial beauty and to disgust towards animals. Activation in this cluster was associated with stronger just world beliefs and weaker dispositional empathic concern and correlated negatively with subsequent prosociality towards people with facial anomalies. Understanding how people think about and treat individuals with facial anomalies has the potential to inform interventions aimed at educating the public about the anomalous-is-bad stereotype and the consequent psychological burden this entails for people who look different.

Methods

Participants

For the first study, a sample of N = 451 participants was recruited via Amazon’s Mechanical Turk service to complete a survey delivered online through the Qualtrics platform (178 female; age: 35.53 ± 10.47 years; education: 14.85 ± 2.40 years). Participants were compensated for their time. Using effect sizes calculated from the data reported in Jamrozik et al.[4] (see Table S1 for the effect sizes and settings for the corresponding power analyses), a minimum of 102 responses per dimension was expected to provide sufficient power (80%) to detect differences in character judgments between anomalous and typical faces, and also to be sufficient to reliably estimate face ratings (Cronbach's $\alpha > 0.8$). We targeted 120 responses to each of the face rating dimensions to provide a buffer (around 15% of responses) against exclusions of low-quality data. From the original N = 451 participants, data from 29 participants was excluded for failing more than 2 out of 5 attention checks embedded throughout the survey. Data from 1 participant was excluded because they acknowledged that their responses were of poor quality. Finally, data were also excluded for 18 participants who chose not to report their sex and/or sexual orientation, both of which were expected to interact with ratings of facial attractiveness. The final sample consisted of N = 403 participants (168 female; age: 35.69 ± 10.50 years [range: 19-72]; education: 15.03 ± 2.19 years [range: 4-25]; race/ethnicity: 293 white, 38 black, 21 Asian, 26 Hispanic or Latinx, 2 American Indian / Alaskan Native, 21 multiracial, 1 “other”, and 1 non-response; sexual orientation: 362 heterosexual, 15 homosexual, and 26 bisexual).

A second, independent sample of N = 32 participants was recruited for an fMRI study (13 female; age: 26.75 ± 8.65 years; education: 16.34 ± 2.74 years). This sample size was chosen to maximize power to detect large effects, given that notable improvements in power have been reported after increasing sample sizes from N = 20 to N = 30-40 participants[53,54]. Participants were recruited from the Philadelphia metropolitan area using online advertisements and they received monetary compensation for their time. The inclusion criteria were: 18 years of age or older, no contraindications for MRI scanning (e.g., claustrophobia), no use of psychotropic medications, and no previous head traumas resulting in losses of consciousness for 15 minutes or longer). Two participants completed all study procedures save for the MRI (one was excluded for actively using psychotropic medications and another for having metal in their body that could not be removed before scanning). MRI data for two participants were excluded because of excessive in-scanner head motion in participant (the criteria are elaborated below) and significant morphological abnormalities in the other. The final sample consisted of N = 27 participants

(13 female; age: 25.52 ± 7.12 years [range: 18-46]; education: 16.41 + 2.65 years [range: 12-23]; race/ethnicity: 11 white, 6 black, 6 Asian, 2 Hispanic or Latinx, and 2 multiracial; sexual orientation: 21 heterosexual, 2 homosexual, 1 bisexual, 1 "other", and 2 non-responses). Both studies were approved by the Institutional Review Board at the University of Pennsylvania and all participants gave informed consent prior to starting any study procedures (Protocol 806447).

Online Survey Study: Design and Analysis Plan

Experimental Procedures

Photographs of people *with* facial anomalies were drawn from the ChatLab Facial Anomaly Database (CFAD; https://cfad.clffwrkmn.net/; Fig. 1). Identified with the Face Image Meta-Database (fIMDb; https://fimdb.clffwrkmn.net/) search engine, we chose the Chicago Face Database for images of people *without* facial anomalies (https://chicagofaces.org/default/). The stimuli from these databases were well-matched on sex and race / ethnicity, though people with facial anomalies were generally older than people with typical faces (Table S2). More detail about the stimuli is given in the Supplement.

The online survey in the first study was a truncated version of the survey from Jamrozik et al.[4], with eight questions about emotional reactions to and perceptions of the people in the photographs. The questions were selected to ensure coverage of all four significant principal components described in Jamrozik et al. [4]. Limiting the number of questions reduced the burden placed on participants - with each participant rating 50 (of 180) faces, and 8 questions per face (400 questions total), the survey usually took over 30 minutes to complete. Before starting the survey, participants were instructed they would first rate each photograph according to "how the face made you feel." Next, they were told they would rate their impressions of the person in the photograph. After a practice trial, participants started the survey. As in our previous study, each photograph was presented for about 2.5 seconds before participants were redirected to a separate page to provide their ratings. The nine-point pictorial Self-Assessment Manikin (SAM) scale measured emotional reactions to the stimuli in terms of valence, arousal, and dominance. Five-point semantic differential scales examined perceptions of the people in the photographs in terms of personality characteristics (i.e., anxious), internal attributes (i.e., contentedness), social traits (i.e., confidence, dominance, and trustworthiness), and attractiveness.

Statistical Analyses

The online survey data were submitted to linear mixed effects models (LMEMs) using the lme4 package in RStudio (Version 1.2.5033). These models examined whether attitudinal manifestations of the anomalous-is-bad stereotype are replicable in an independent sample and when comparing anomalous to typical faces with no known history of visible difference. Results are considered significant at an $\alpha = 0.05$ (two-tailed). For each of the eight dimensions on which the face stimuli were rated, three separate LMEMs were constructed. The first regressed a given face rating dimension on face type (i.e., anomalous, beautiful, and two sets of average-looking faces) and included random intercepts for subject and stimulus. The second model was identical to the first but included interactions between face type and face sex, the rater’s biological sex, and the rater’s sexual orientation. The third model was identical to the second but included random slopes for face type. In all cases, however, attempting to fit the third model resulted in failures in model convergence or produced singular model fits. As such, the results of the third models are not considered further. Of the two fitted models, that with the lowest Akaike information criterion (AIC) was selected. Save for the dimension “Content”, the models always performed better after including interacting terms for sex, sexual orientation, and face sex. The models that performed best are described briefly in the Results section, with partial summaries given in Table S4 and complete summaries given in the comments of the corresponding code (https://osf.io/d9sc7/).

In view of the findings reported by Jamrozik et al.[4], we expected anomalous faces to elicit worse ratings than average-looking people on attractiveness, trustworthiness, contentedness, dominance, anxiousness, perceiver arousal (rated with the SAM), perceiver dominance, and perceiver happiness. Building on prior work, which compared anomalous faces to one another before and after corrective treatment, this study examined ratings of anomalous compared to average-looking and beautiful faces. To ensure both sets of average-looking faces were well-matched, equivalence testing was used to demonstrate practical equivalence between sets of faces. Statistical equivalence testing was conducted with the two one-sided tests (TOST) package in RStudio (smallest effect size of interest: Cohen's $d = 0.2$).

fMRI Study: Design and Analysis Plan

Dispositional and Attitudinal Measures

Dispositions: Prior to their in-person study visit, participants in the fMRI study completed an online survey in Qualtrics with the following brief dispositional measures: The Social Dominance Orientation (SDO) scale provided a single measure of attitudes about anti-egalitarianism [23]. The Procedural and Distributive Just World Beliefs Scale (JWBS) measured beliefs about procedural justice (fairness in interpersonal contexts) and distributive justice (fairness in resource allocations) towards oneself and others. The Interpersonal Reactivity Index (IRI) measured cognitive (the perspective taking subscale) and affective (the empathic concern subscale) facets of empathy [24]. The Three-Domain Disgust scale to measure sensitivity to different experiences of disgust, with subscales for pathogen-related, moral, and sexual disgust [55]. We focused solely on pathogen-related disgust sensitivity in what follows. Participants also provided demographic information including age, sex, education, and income.

Explicit Attitudes: The Explicit Bias Questionnaire (EBQ)[3] was used in the fMRI study to examine whether participants harbor explicit biases against people with facial anomalies as a group. This measure consists of 33 questions, with items to assess personal history with disability and facial anomalies, general attitudes about facial anomalies, and evaluations of character (Supplementary Appendix A). The EBQ was administered using a custom interface in Microsoft Excel programmed with a macro in Visual Basic.

Implicit Attitudes: The Implicit Association Test (IAT) described by Hartung and colleagues[3] was again used to quantify implicit biases towards people with facial anomalies. The IAT measures evaluative associations underpinning implicit attitudes[56] and has been used extensively to measure implicit biases towards social out-groups[57]. The logic behind the IAT is that people are quicker to associate negative words with out-group members compared to members of their in-group. If participants are faster at associating anomalous faces with bad words than good, this result would be reflected in a positive d' score and indicative of implicit bias. If implicit bias is not present, however, then the IAT d' score should be close to zero or even negative. A detailed summary of the scoring algorithm used to calculate the IAT d' scores is provided in the Supplement. The IAT, behavioral, and neuroimaging tasks were administered in the open-source experiment platform OpenSesame (https://osdoc.cogsci.nl/).

Behavioral Measures

Dictator Game: In the DG, one player (the dictator) receives an endowment and decides how to split this with another player[58] (the receiver). Since the dictator can keep the entire endowment for themselves without consequence, this game is typically used to assess prosociality. Participants played 15 rounds of the DG as the dictator. In each round, participants decided between different ways of splitting the endowment with the receiver ($5). Before this decision, participants saw a photograph that ostensibly depicted the receiver. The photographs showed people with facial anomalies before (7-8 rounds) or after (7-8 rounds) corrective surgery. These photographs, also from the CFAD, were of people who *did not* appear in the fMRI oddball or UG tasks (see Table S3 for the demographic characteristics of the people whose photographs were used as DG stimuli; for characteristics of the oddball and UG stimuli, see Table S2). Despite being told that they would play both the DG and UG with real people, participants in fact played both games with computer-generated partners. Additional detail about the use of deception is provided in the Supplement. Participants were shown two different options for splitting the endowment (receiver:proposer): \$2.50:\$2.50, \$2.00:\$3.00, … \$0.50:\$4.50, or \$0.00:\$5.00. After deciding, participants saw the outcome for that round before moving onto the next.

Ultimatum Game: The UG is similar to the DG in that one player (the proposer) decided how to split an endowment with another player[59] (the receiver). In contrast to the DG, however, receivers can reject unfair offers so that neither player receives any part of the endowment. Participants played as receivers for 60 rounds. As in the DG, participants saw a photograph of a person with or without a facial anomaly before being presented with an offer. Offers ranged from (receiver:proposer): \$4.50:\$0.50, \$4.00:\$1.00, … \$0.50:\$4.50, to \$0.00:\$5.00), and were ostensibly proposed by the people in the photographs. Participants then decided whether to reject each offer. Although rejections are often interpreted as “altruistic punishment”—costly punishment that discourages future exploitation—we sought to isolate antisociality by contrasting rejections of proposers with and without anomalies.

Scoring and Analysis: For each trial of the DG, the difference in magnitude between the options shown to participants was calculated—the proposal that entailed the smallest disparity between dictator and receiver was categorized as prosocial. In view of mounting evidence that pro- and antisocial behavior are shaped by social class, socioeconomic status (SES; the formula is given in the Supplement) was included in both models [60,61]. A binomial LMEM examined whether prosociality was less frequent towards people with anomalous relative to typical faces. Prosociality was regressed on face type (Pre- | Post-Surgical Intervention), which interacted with the difference in magnitude between options and with SES and included random intercepts for subject and stimulus. For the UG, the disparities in proposed splits were calculated. A binary variable was constructed to index antisociality, reflected in rejection rates of offers from anomalous relative to beautiful faces. A binomial LMEM regressed antisociality on face type (Anomalous | Beautiful), which interacted with disparities in proposed splits and with SES and included random intercepts for subjects and stimulus.

Functional MRI Task

Before scanning, participants were tasked with learning a set of 90 photographs of non-anomalous faces rated average in attractiveness (the Supplement contains a detailed description of the learning task). During fMRI scanning, participants completed 540 trials of an oddball task (Fig. 1; 5 repetitions each of 90 learned faces [450 total], 1 repetition each of 90 novel faces [30 anomalous, 30 beautiful, 30 average-looking]; 83.3% learned, 16.7% novel). Trials started with the presentation of a photograph for 500ms, followed by a null event for 1000ms plus jitter lasting an average of 1833ms (3,333ms per trial). Similar parameters have been used in fMRI studies using the oddball paradigm (e.g., [62,63]). The oddball task was split across five runs lasting around six minutes each (about 30 minutes in total). The optimal ordering and spacing of trials was determined with the optseq2 package in FreeSurfer (http://surfer.nmr.mgh.harvard.edu/optseq/). Participants counted novel faces and, after each run, reported the total number they saw. Relative to e.g. using button presses to track detections of oddball stimuli, counting produces larger sustained neural responses [64].

After the oddball task, another functional scan was acquired while participants passively looked at images as they appeared on the screen. This scan was used to localize neural responses to disgusting animals (the DIsgust-RelaTed-Images [DIRTI] database; https://dx.doi.org/10.5281/zenodo.167037), disgusting injuries (the DIRTI database), and emotional arousal (the Geneva Affective Picture Database; https://www.unige.ch/cisa/research/materials-and-online-research/research-material/). Each functional localizer was comprised of a single block of 16 trials that alternated between subsets of 8 images from the target category and 8 from a control condition (scrambled versions of the DIRTI images for disgusting animals and injuries; images low in rather than high in salience and arousal). Between blocks, participants fixated on a cross until the next block appeared.

MRI Image Acquisition & Pre-Processing

MRI data were acquired on a 3T Siemens Magnetom Prisma (Erlangen, Germany) with a 64-channel head coil using standard anatomical and functional sequences from the Human Connectome Project. Gradient-echo echo planar images (EPIs) were acquired over five runs while participants completed the oddball task (520 volumes per run; 72 axial slices; 2mm slice thickness; interleaved acquisition; repetition time [TR]: 720 ms; echo time: 37 ms; multiband factor: 8; field of view: 208x144mm; reconstructed voxel size: 2mm^3; flip angle: 52°). Identical parameters characterized the EPI acquired during the functional localizers task, with the exception that the entire run consisted of 732 volumes. A 3-dimensional T1-weighted anatomical magnetization-prepared rapid-acquisition gradient-echo image was also be acquired in each participant (160 axial slices; 1mm slice thickness; repetition time: 2400ms; echo time: 2.14ms; field of view: 224x224mm; reconstructed voxel size: 0.7mm^3; flip angle: 8°).

Pre-processing of the MRI data was performed with the SPM12 (http://www.fil.ion.ucl.ac.uk/spm/) MATLAB (MathWorks) toolbox. First, the anatomical and EPI scans were manually realigned to the anterior commissure. Next, frame-by-frame realignment was applied to the EPI scans to correct for head motion, prior to high-pass filtering (128s cutoff). Rigid body motion parameters generated during realignment were eventually entered into first level models as nuisance regressors. Realignment also generated mean EPI images to which the corresponding anatomical images were co-registered. After co-registration, anatomical images were segmented by tissue class to generate nonlinear deformation parameters that were used to warp the anatomical and EPI data into standard space. Finally, the EPI data were smoothed with a 4mm Gaussian kernel. Within each run, volumes characterized by excessive head motion (>0.5mm per TR) were identified with ArtRepair (https://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html), replaced with interpolated volumes, and de-weighted in first-level model estimation. First level contrasts tested the predictions laid out in the Introduction. Two additional sets of first-level models included mean-centered parametric modulators corresponding to average ratings from the online survey study on attractiveness and, separately, trustworthiness. Finally, additional first-levels models were constructed for the functional localizers.

For the oddball task, contrasts at the second level captured neural responses associated with visual salience and emotional arousal. For the functional localizers, second level models identified brain areas implicated in processing disgusting and arousing images. The resulting statistical maps were combined, binarized, and used to mask later analyses. Beyond simply excluding non-selective effects that may be attributable to salience or arousal, we also sought out areas of convergence in processing morality (i.e., trustworthiness, disgust) and beauty (i.e., facial unattractiveness). A mask representing group-level parametric variation in neural responses to facial attractiveness was binarized and combined with masks for disgust towards animals and towards injuries (of note, significant relations between trustworthiness ratings and hemodynamic response amplitudes were not detected). If dehumanization entails the attribution of subhuman essences to dehumanized others, as some have argued, then selective neural responses to anomalous faces should occur in areas responsive to disgusting non-human animals rather than human bodily injuries. It is important to note that masking, in contrast to methods like small volume correction, does not alter significance but only the inclusion and/or exclusion of voxels.

Whole Brain Voxelwise and Region-of-Interest Analyses

At the group-level, voxelwise analyses tested our predictions across the whole brain. In addition, correlational analyses examined whether brain responses to facial anomalies were associated with implicit biases (IAT). To restrict the likelihood of identifying false positives, we used an uncorrected cluster-forming threshold of $\alpha = 0.001$ in conjunction with an extent threshold determined with non-parametric Monte Carlo simulation in the 3dClustSim package in AFNI (k = 13 contiguous voxels; http://afni.nimh.nih.gov/). Results were considered significant at a cluster-level FWE-corrected $\alpha = 0.01$.

We also tested our predictions in the following regions of interest (ROIs): bilateral dorsal anterior cingulate cortex (dACC; Montreal Neurological Institute coordinates: x = -2, y = 36, z = 10), left LOC (-28, -98, 8), right LOC (34, -90, 2), left fusiform gyrus (-32, -60, -13), right fusiform gyrus (33, -57, 13), and left amygdala (-25, 1, -19). Spherical ROIs 6 mm in diameter were placed in each region at the given coordinates, which were either chosen *a priori* from coordinates[3] (dACC and bilateral LOC) or were defined functionally from coordinates reported in Table 2 below (bilateral fusiform gyri and left amygdala). Percent signal change was extracted from each ROI with the MarsBaR MATLAB toolbox (http://marsbar.sourceforge.net/). Attitudinal, behavioral, ROI, and dispositional data were examined with LMEMs and non-parametric correlations to determine whether e.g. signal in the ROIs related to individual differences in psychological dispositions, implicit attitudes, and/or behavior. Results were considered significant at $\alpha = 0.05$ (uncorrected for multiple comparisons).

Acknowledgments: Research reported in this publication was supported by the Penn Center for Human Appearance (AC) and the National Institute of Dental & Craniofacial Research of the National Institutes of Health (F32DE029407 awarded to CIW). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions: All authors contributed to the study design. Data collection was performed by C.I.W. and F.H. Analyses were performed by C.I.W. and S.H. The manuscript was drafted by C.I.W. All authors provided critical revisions and approved the final version of the manuscript.

Competing Interests: The authors declare no competing interests.

Data availability: The study materials and data supporting the findings of this study, including the statistical maps and data used in the figures, are available from https://osf.io/d9sc7/.

Code availability: The code used to run the statistical analyses and to generate the figures reported in the manuscript is available from https://osf.io/d9sc7/.

References

1. Nolan, J. & Nolan, C. *The Dark Knight*. 1–141 (Warner Bros. Pictures, 2008).
2. Croley, J. A., Reese, V. & Wagner, R. F. Dermatologic Features of Classic Movie Villains. *JAMA Dermatology* **153**, 559 (2017).
3. Hartung, F. *et al.* Behavioural and Neural Responses to Facial Disfigurement. *Sci. Rep.* **9**, 8021 (2019).
4. Jamrozik, A., Oraa Ali, M., Sarwer, D. B. & Chatterjee, A. More than skin deep: Judgments of individuals with facial disfigurement. *Psychol. Aesthetics, Creat. Arts* **13**, 117–129 (2019).
5. Houston, V. & Bull, R. Do people avoid sitting next to someone who is facially disfigured? *Eur. J. Soc. Psychol.* **24**, 279–284 (1994).
6. Madera, J. M. & Hebl, M. R. Discrimination against facially stigmatized applicants in interviews: An eye-tracking and face-to-face investigation. *J. Appl. Psychol.* **97**, 317–330 (2012).
7. Mojon-Azzi, S. M., Potnik, W. & Mojon, D. S. Opinions of dating agents about strabismic subjects’ ability to find a partner. *Br. J. Ophthalmol.* **92**, 765–769 (2008).
8. Dion, K., Berscheid, E. & Walster, E. What is beautiful is good. *J. Pers. Soc. Psychol.* **24**, 285–290 (1972).
9. Rhodes, G. The Evolutionary Psychology of Facial Beauty. *Annu. Rev. Psychol.* **57**, 199–226 (2006).
10. Bzdok, D. *et al.* ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Struct. Funct.* **215**, 209–223 (2011).
11. Eagly, A. H., Ashmore, R. D., Makhijani, M. G. & Longo, L. C. What is beautiful is good, but . . .: A meta-analytic review of research on the physical attractiveness stereotype. *Psychol. Bull.* **110**, 109–128 (1991).
12. Griffin, A. M. & Langlois, J. H. Stereotype Directionality and Attractiveness Stereotyping: Is Beauty Good or is Ugly Bad? *Soc. Cogn.* **24**, 187–206 (2006).
13. Zebrowitz, L. A. & Montepare, J. M. Social Psychological Face Perception: Why Appearance Matters. *Soc. Personal. Psychol. Compass* **2**, 1497–1517 (2008).
14. Changing Faces. *Public Attitudes to Disfigurement in 2017*. (2017).
15. Harris, L. T. & Fiske, S. T. Dehumanizing the Lowest of the Low. *Psychol. Sci.* **17**, 847–853 (2006).
16. Haslam, N. & Loughnan, S. Dehumanization and Infrahumanization. *Annu. Rev. Psychol.* **65**, 399–423 (2014).
17. Price, M. E., Brown, S., Dukes, A. & Kang, J. Bodily Attractiveness and Egalitarianism are Negatively Related in Males. *Evol. Psychol.* **13**, 147470491501300 (2015).
18. Bostyn, D. H., Sevenhant, S. & Roets, A. Of Mice, Men, and Trolleys: Hypothetical Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas. *Psychol. Sci.* **29**, 1084–1093 (2018).
19. May, J., Workman, C. I., Haas, J. & Han, H. The Neuroscience of Moral Judgment: Empirical and Philosophical Developments. in *Neuroscience and Philosophy* (eds. Sinnott-Armstrong, W. & De Brigard, F.) 1–21 (MIT Press, 2021).
20. Seeley, W. W. *et al.* Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007).

21. Kragel, P. A. & LaBar, K. S. Decoding the Nature of Emotion in the Brain. *Trends Cogn. Sci.* **20**, 444–455 (2016).

22. Zahn, R., de Oliveira-Souza, R. & Moll, J. Moral Motivation and the Basal Forebrain. *Neurosci. Biobehav. Rev.* **108**, 207–217 (2020).

23. Pratto, F., Sidanius, J., Stallworth, L. M. & Malle, B. F. Social dominance orientation: A personality variable predicting social and political attitudes. *J. Pers. Soc. Psychol.* **67**, 741–763 (1994).

24. Davis, M. H. Measuring individual differences in empathy: Evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* **44**, 113–126 (1983).

25. Lucas, T., Zhdanova, L. & Alexander, S. Procedural and Distributive Justice Beliefs for Self and Others. *J. Individ. Differ.* **32**, 14–25 (2011).

26. Shanmugarajah, K., Gaind, S., Clarke, A. & Butler, P. E. M. The role of disgust emotions in the observer response to facial disfigurement. *Body Image* **9**, 455–461 (2012).

27. Ryan, S., Oaten, M., Stevenson, R. J. & Case, T. I. Facial disfigurement is treated like an infectious disease. *Evol. Hum. Behav.* **33**, 639–646 (2012).

28. Tybur, J. M., Lieberman, D., Kurzban, R. & DeScioli, P. Disgust: Evolved function and structure. *Psychol. Rev.* **120**, 65–84 (2013).

29. Landy, J. F. & Goodwin, G. P. Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence. *Perspect. Psychol. Sci.* **10**, 518–536 (2015).

30. Squires, N. K., Squires, K. C. & Hillyard, S. A. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr. Clin. Neurophysiol.* **38**, 387–401 (1975).

31. Chatterjee, A., Thomas, A., Smith, S. E. & Aguirre, G. K. The neural response to facial attractiveness. *Neuropsychology* **23**, 135–143 (2009).

32. Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J. & Tetlock, P. E. Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *J. Pers. Soc. Psychol.* **105**, 171–192 (2013).

33. Lang, P. J. *et al.* Emotional arousal and activation of the visual cortex: An fMRI analysis. *Psychophysiology* **35**, S0048577298001991 (1998).

34. Ferrari, C. *et al.* The dorsomedial prefrontal cortex mediates the interaction between moral and aesthetic valuation: a TMS study on the beauty-is-good stereotype. *Soc. Cogn. Affect. Neurosci.* **12**, 707–717 (2017).

35. Avram, M. *et al.* Neurofunctional correlates of esthetic and moral judgments. *Neurosci. Lett.* **534**, 128–132 (2013).

36. Wang, T. *et al.* Is moral beauty different from facial beauty? Evidence from an fMRI study. *Soc. Cogn. Affect. Neurosci.* **10**, 814–823 (2015).

37. Milne, E. & Grafman, J. Ventromedial Prefrontal Cortex Lesions in Humans Eliminate Implicit Gender Stereotyping. *J. Neurosci.* **21**, RC150–RC150 (2001).

38. Forbes, C. E. *et al.* Identifying temporal and causal contributions of neural processes underlying the Implicit Association Test (IAT). *Front. Hum. Neurosci.* **6**, 622–645 (2012).

39. Sellaro, R. *et al.* Reducing Prejudice Through Brain Stimulation. *Brain Stimul.* **8**, 891–897 (2015).

40. Calder, A. J., Lawrence, A. D. & Young, A. W. Neuropsychology of fear and loathing. *Nat. Rev. Neurosci.* **2**, 352–363 (2001).

41. Schienle, A. *et al.* fMRI responses to pictures of mutilation and contamination. *Neurosci. Lett.* **393**, 174–178 (2006).

42. Ongür, D., Price, J. L. & Ongur, D. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* **10**, 206–219 (2000).

43. Alheid, G. F. Extended amygdala and basal forebrain. *Ann. N. Y. Acad. Sci.* **985**, 185–205 (2006).

44. Zahn, R., de Oliveira-Souza, R., Bramati, I., Garrido, G. & Moll, J. Subgenual cingulate activity reflects individual differences in empathic concern. *Neurosci. Lett.* **457**, 107–110 (2009).

45. Workman, C. I. *et al.* Subgenual cingulate–amygdala functional disconnection and vulnerability to melancholic depression. *Neuropsychopharmacology* **41**, 2082–2090 (2016).

46. Decety, J., Pape, R. & Workman, C. I. A multilevel social neuroscience perspective on radicalization and terrorism. *Soc. Neurosci.* **13**, 511–529 (2018).

47. Workman, C. I., Yoder, K. J. & Decety, J. The dark side of morality: Neural mechanisms underpinning moral convictions and beliefs about violence. *Under Rev.* 1–20 (2020).

48. Bartra, O., McGuire, J. T. & Kable, J. W. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* **76**, 412–427 (2013).

49. Adolfi, F. *et al.* Convergence of interoception, emotion, and social cognition: A twofold fMRI meta-analysis and lesion approach. *Cortex* **88**, 124–142 (2017).

50. Todorov, A. & Engell, A. D. The role of the amygdala in implicit evaluation of emotionally neutral faces. *Soc. Cogn. Affect. Neurosci.* **3**, 303–312 (2008).

51. Rosenberger, L. A. *et al.* The Human Basolateral Amygdala Is Indispensable for Social Experiential Learning. *Curr. Biol.* 1–6 (2019) doi:10.1016/j.cub.2019.08.078.

52. Liang, X., Zebrowitz, L. A. & Zhang, Y. Neural activation in the “reward circuit” shows a nonlinear response to facial attractiveness. *Soc. Neurosci.* **5**, 320–334 (2010).

53. Yarkoni, T. Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power—Commentary on Vul et al. (2009). *Perspect. Psychol. Sci.* **4**, 294–298 (2009).

54. Geuter, S., Qi, G., Welsh, R. C., Wager, T. D. & Lindquist, M. A. Effect Size and Power in fMRI Group Analysis. *bioRxiv Prepr.* 1–23 (2018) doi:10.1101/295048.

55. Tybur, J. M., Lieberman, D. & Griskevicius, V. Microbes, mating, and morality: Individual differences in three functional domains of disgust. *J. Pers. Soc. Psychol.* **97**, 103–122 (2009).

56. Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. Measuring individual differences in implicit cognition: The implicit association test. *J. Pers. Soc. Psychol.* **74**, 1464–1480 (1998).

57. Kurdi, B., Mann, T. C., Charlesworth, T. E. S. & Banaji, M. R. The relationship between implicit intergroup attitudes and beliefs. *Proc. Natl. Acad. Sci.* **116**, 5862–5871 (2019).

58. Kahneman, D., Knetsch, J. L. & Thaler, R. H. Fairness and the Assumptions of Economics. *J. Bus.* **59**, S285 (1986).

59. Güth, W., Schmittberger, R. & Schwarze, B. An experimental analysis of ultimatum bargaining. *J.*

Econ. Behav. Organ. **3**, 367–388 (1982).

60. Smeets, P., Bauer, R. & Gneezy, U. Giving behavior of millionaires. *Proc. Natl. Acad. Sci.* **112**, 10641–10644 (2015).

61. Korndörfer, M., Egloff, B. & Schmukle, S. C. A Large Scale Test of the Effect of Social Class on Prosocial Behavior. *PLoS One* **10**, e0133193 (2015).

62. Knolle, F. *et al.* Brain responses to different types of salience in antipsychotic naïve first episode psychosis: An fMRI study. *Transl. Psychiatry* **8**, 196 (2018).

63. Warbrick, T., Reske, M. & Shah, N. J. Do EEG paradigms work in fMRI? Varying task demands in the visual oddball paradigm: Implications for task design and results interpretation. *Neuroimage* **77**, 177–185 (2013).

64. Clark, V. P., Fannon, S., Lai, S. & Benson, R. Paradigm-dependent modulation of event-related fMRI activity evoked by the oddball task. *Hum. Brain Mapp.* **14**, 116–127 (2001).

Morality is in the eye of the beholder:

Unpacking the neurocognitive basis of the "anomalous-is-bad" stereotype

Clifford I. Workman, Stacey Humphries, Franziska Hartung,
Geoffrey K. Aguirre, Joseph W. Kable, & Anjan Chatterjee

Supplementary Information

Supplementary Methods

Stimulus Selection

Stimuli depicting people with typical faces and no known history of having had visible facial anomalies were selected from the Chicago Face Database[1] (CFD). Stimuli depicting people with visible facial anomalies were sourced from online and print materials, as described by Workman and colleagues[2]. All of the face stimuli ultimately used in the functional neuroimaging task described herein first underwent the following pre-processing steps: First, images were 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/). Next, images were resized and cropped (width: 345px; height: 407px) before their backgrounds were removed (https://www.remove.bg/) and then replaced with black using the GIMP 2 software package (https://www.gimp.org/). All of the stimuli depicting people with facial anomalies is available at https://cfad.clffwrkmn.net/, and the pre-processed CFD stimuli are available at https://osf.io/d9sc7/.

Scoring the Implicit Association Test (IAT)

The IAT is comprised of seven parts. In the first part, participants used keystrokes to categorize images of faces as anomalous or typical. Second, using the same keys, participants categorized words as good (e.g., happy) or bad (e.g., sickening). Third, participants again used the same keys to categorize both faces and words (e.g., anomalous faces and good words, typical faces and bad words). The fourth part simply replicates the third. In parts 5 through 7, the mapping between faces and keys is swapped (e.g., such that anomalous faces are paired with bad words, and typical faces are paired with good words, or vice versa). Prior to calculating the IAT d' score, the following exclusion criteria were applied[3] (see also http://projectimplicit.net/nosek/papers/scoringalgorithm.sas.txt): error rates could not exceed 30% of trials (i.e., 10 or more), nor could latencies drop below 300ms in over 10% of trials (i.e., 4 or more). Trials with latencies exceeding 10s or falling below 400ms were also removed at this time. Then, the average reaction time when associating anomalous faces with bad words (and typical faces with good words; e.g., part 7) was subtracted from the average reaction time when associating anomalous faces with good words (and typical faces with bad words; e.g. part 4). This difference score was taken over the standard deviation of these reaction times in order to calculate the IAT d' score[4].

Deception in the Economic Games

Participants were led to believe that proposals made by other "players" on the DG and UG were real responses recorded on a different day. Participants were told one proposal would be chosen at random and paid out at the end of the study visit. Ultimately, however, participants always received $50 in compensation for their time, an increase of $10 from what was advertised. Debriefings after the games revealed the deception to participants and gave them an opportunity to withdraw their data from the study (no such requests were made).

Before scanning, participants completed a 1-back working memory task to familiarize them with a set of photographs depicting typical faces of average attractiveness (Table S2) without known histories of visible anomalies. Participants had 1.5 seconds to indicate whether each of 90 different faces was identical to the one immediately preceding it. Participants received feedback about speed when responses were too slow, and about accuracy otherwise. Throughout the 1-back task, participants saw only those 90 photographs that ultimately comprised the learned average-looking faces. After, participants were given a recognition memory test in which they were again presented the 90 face photographs from the working memory task. For each face, participants indicated whether they believed they had seen it earlier in the 1-back task. In fact, participants saw only those 90 faces from the 1-back task, each of which was shown repeatedly until participants achieved 100% recognition accuracy. The recognition memory task did not include distractor faces to limit exposure to novel faces before the oddball task described below.

Calculating Socioeconomic Status

Participants in the fMRI study completed an online survey prior to their study visit. In addition to dispositional measures, this questionnaire included questions about demographic characteristics, including the following about education and income: "What is your yearly income?" (1 = "Less than $25,000", 2 = "$25,000 - $49,999", 3 = "$50,000 - $74,999", 4 = "$75,000 - $99,999", 5 = "$100,000 - $124,9995", 6 = "$125,000 - $149,999", 7 = "$150,000 - $174,999", 8 = "$175,000 - $199,999", 9 = "$200,000 or more"). In order to calculate socioeconomic status (SES), answers to these questions were first rescaled to 5 (i.e., divided by 6 and 9, respectively, and then multiplied by 5) and then summed. These values were then z-scored by subtracting the overall average from each summed value and then taking this over the overall standard deviation. The resulting scores were then used as a proxy for SES.

Supplementary Tables

Table S1 | The minimum number of ratings per photograph required to obtain reliable average ratings and to maximize power for detecting statistical differences in beliefs about anomalous compared to typical faces.

Principle Component (Label): Dimension	**PC1**	**PC2**	**PC4**	**PC5**	**N**	***t***	***df***	***p***	**d_s ≈**	**Min N Reliable**	**Min N/ Group**	**Min N TOTAL**
PC1 (Warmth): Content	**0.22**	0.01	0.03	-0.02	145	-4.27	26.20	0.000	-0.710	15	33	**66**
PC2 (Dominance): Dominant	0.02	**0.38**	0.11	0.05	145	-3.81	25.03	0.001	-0.632	35	41	**82**
PC3 (Competence): NS												
PC4 (Stability): Anxious	-0.13	-0.04	**-0.38**	-0.22	145	4.72	25.02	0.000	0.783	40	27	**54**
PC5 (Objectification): Dominance (viewer)	0.08	0.01	-0.11	**-0.83**	145	-4.64	49.74	0.000	-0.771	20	28	**56**
Attractive	0.18	0.10	-0.05	-0.09	145	-6.04	27.67	0.000	-1.003	10	17	**34**
Trustworthy	0.19	-0.19	-0.02	-0.01	145	-3.38	27.21	0.002	-0.562	25	51	**102**

The first column gives each of the five principle components from [5], a corresponding label, and the face rating dimension selected to represent this component in the online survey study. Columns “PC1” to “PC5” represent relative loadings onto each of the principle components and are provided in the main text, whereas columns “N” through “ds ≈” provide statistical output from models that is reported in the online supplementary material). These data were used to calculate effect sizes (Cohen's d_s) as described by Lakens (2013). For each dimension, raw responses collected by [5] were used to simulate the minimum sample size needed for reliable mean ratings, given in the column “Min Reliable N.” Raw responses were iteratively sub-sampled to different sample sizes and then used to compute Cronbach's α. The sample size at which α exceeded 0.8 therefore reflects the minimum sample size needed to achieve reliability. The last two columns give the minimum numbers of participants—per group and in total—needed to obtain adequate sensitivity. Power analyses were conducted in G*Power [7] with the following settings: Test family - "ttests", Statistical test - "Mean Difference between two independent means (two groups)", Type of power analysis - "A priori: Compute required sample size - given alpha, power, and effect size", Tail(s) - "Two", alpha err prob - "0.05", Power - "0.8", and Allocation ratio - "1." NS, not significant; min, minimum; PC, principle component.

Table S2 | Demographic characteristics of the people whose photographs were used as experimental stimuli in the fMRI and Ultimatum Game (anomalous and beautiful faces only) tasks.

	Anomalous Faces (Novel)				**Beautiful Faces (Novel)**				**Average-Looking Faces (Novel)**				**Average-Looking Faces (Learned)**			
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
Age	50.07	16.39	18	76	27.60	3.23	21	36	27.68	5.59	19	43	29.85	7.32	18	50

Attractive	1.85	0.32	1.40	2.65	3.59	0.45	2.83	4.23	2.68	0.20	2.27	2.97	2.63	0.20	2.25	3.01
Trustworthy	2.72	0.35	2.04	3.39	3.34	0.44	2.30	3.95	2.85	0.32	2.14	3.44	2.83	0.37	2.21	3.51
Content	2.36	0.41	1.71	3.31	3.47	0.55	2.21	4.34	2.84	0.36	2.19	3.43	2.88	0.41	2.15	3.56
Dominant	2.93	0.57	2.10	4.11	3.14	0.38	2.53	4.00	3.07	0.48	2.41	4.30	3.25	0.40	2.28	4.06
Anxious	3.42	0.35	2.71	4.08	2.48	0.44	1.85	3.45	2.95	0.28	2.48	3.63	2.97	0.35	2.20	3.60
SAM Dominance	4.71	0.39	3.87	5.37	5.62	0.40	4.67	6.20	5.32	0.26	4.68	5.72	5.28	0.30	4.56	5.73
SAM Happiness	3.32	0.54	2.46	4.55	5.49	0.65	4.24	6.62	4.59	0.39	3.74	5.31	4.61	0.43	3.84	5.57
SAM Arousal	4.69	0.51	3.58	5.85	3.72	0.35	3.06	4.69	3.55	0.38	2.96	4.60	3.62	0.33	3.09	4.30
Race / Ethnicity *(N)*																
White	22				22				22				33			
Black	2				2				2				6			
Hispanic / Latinx	6				6				6				11			
Facial Anomaly *(N)*																
Scar	15															
Cancer	11															
Paralysis	3															
Swelling	1															

Table S3 | Demographic characteristics of the people whose photographs were used as experimental stimuli in the Dictator Game.

	Anomalous Faces				**Surgically Corrected Faces**			
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
Age	34.87	12.72	17	58	34.87	12.72	17	58
Attractive	2.64	0.40	2.07	3.84	3.22	0.68	2.04	4.44
Trustworthy	3.28	0.26	2.77	3.68	3.44	0.45	2.63	4.18
Content	3.01	0.39	2.10	3.46	3.36	0.73	1.95	4.46
Dominant	2.80	0.43	2.22	3.70	3.04	0.39	2.48	3.72
Anxious	3.15	0.24	2.77	3.61	2.76	0.47	2.06	3.46
SAM Dominance	4.83	0.23	4.47	5.30	5.27	0.48	4.20	5.96
SAM Happiness	4.12	0.60	3.04	5.07	5.33	1.02	3.60	7.31
SAM Arousal	3.89	0.36	3.03	4.57	3.58	0.42	3.11	4.52
Race / Ethnicity								
White	11				11			
Asian	3				3			
Black	0				0			
Hispanic or Latinx	1				1			

	Anomaly	
Scar	3	
Cleft lip	1	
Atrophy	1	
Pigmentation	4	
Swelling	1	
Paralysis	4	
Cancer	1	

Table S4 | Perceivers' judgments of and responses to faces with visible anomalies, beautiful faces, and average-looking faces (learned and average-looking).

A: Attractiveness Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	1.966	0.182	10.818	< 0.001
Face type (beautiful)	2.119	0.196	10.834	< 0.001
Face type (average-looking novel)	0.957	0.192	4.985	< 0.001
Face type (average-looking learned)	0.999	0.182	5.500	< 0.001
B: Trustworthiness Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	2.908	0.210	13.851	< 0.001
Face type (beautiful)	0.786	0.277	2.836	0.005
Face type (average-looking novel)	0.068	0.247	0.277	0.782
Face type (average-looking learned)	0.296	0.233	1.272	0.204
C: Contentedness Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	2.361	0.081	29.135	< 0.001
Face type (beautiful)	1.112	0.112	9.925	< 0.001
Face type (average-looking novel)	0.487	0.112	4.347	< 0.001
Face type (average-looking learned)	0.517	0.100	5.168	< 0.001
D: Dominance Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	2.707	0.186	14.540	< 0.001
Face type (beautiful)	0.449	0.243	1.844	0.065
Face type (average-looking novel)	0.323	0.240	1.347	0.178
Face type (average-looking learned)	0.848	0.229	3.708	< 0.001
E: Anxiousness Ratings				
Fixed Effects	**B**	**SE**	***z* value**	***p* value**

Intercept	3.475	0.184	18.911	< 0.001
Face type (beautiful)	-1.497	0.229	-6.528	< 0.001
Face type (average-looking novel)	-0.499	0.226	-2.210	0.027
Face type (average-looking learned)	-0.867	0.215	-4.041	< 0.001
F: SAM Arousal Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	4.325	0.419	10.321	< 0.001
Face type (beautiful)	-0.110	0.332	-0.330	0.741
Face type (average-looking novel)	-0.861	0.326	-2.639	0.008
Face type (average-looking learned)	-0.599	0.308	-1.945	0.052
G: SAM Dominance Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	4.806	0.398	12.092	< 0.001
Face type (beautiful)	0.986	0.327	3.015	0.003
Face type (average-looking novel)	0.647	0.321	2.018	0.044
Face type (average-looking learned)	0.507	0.302	1.680	0.093
H: SAM Happiness Ratings				
Fixed Effects	**β**	**SE**	***z* value**	***p* value**
Intercept	2.982	0.294	10.131	< 0.001
Face type (beautiful)	3.114	0.322	9.679	< 0.001
Face type (average-looking novel)	1.779	0.316	5.622	< 0.001
Face type (average-looking learned)	1.818	0.300	6.053	< 0.001

SAM, Self-Assessment Manikin.

Table S5 | Neural sensitivity to visual salience.

A: (Anomalous + Beautiful + Average-Looking Novel) > Average-Looking Learned Faces					
Brain Regions by Cluster	**Cluster Size**	**Peak MNI Coordinates**			**Peak *t* Statistic**
		x	*y*	*z*	
L fusiform	515	-40	-62	-13	8.77
R fusiform	500	41	-62	-16	8.40
R IPL	352	38	-55	50	7.98
SMA / dACC	575	8	16	50	7.97
L aINS	198	-27	18	2	7.80
R dlPFC	511	46	11	29	7.33
R aINS	356	36	23	-1	7.04
L VS	87	-12	6	-4	6.66
R hippocampus	75	26	-27	-4	6.63
L thalamus	63	8	-10	5	6.61
R PCC	141	6	-27	26	6.21
L IFG	86	-42	6	23	6.12

Cerebellum	89	-10	-77	-25	6.11
R VS	85	13	8	2	6.04
L latOFC	19	-27	11	-16	5.92
L IPL	196	-32	-52	47	5.72
Cerebellum	34	-30	-72	-49	5.58
Left dlPFC	37	-37	26	23	5.27
Midbrain	13	1	-20	-10	4.67
B: Average-Looking Learned > (Anomalous + Beautiful + Average-Looking Novel Faces)					
L PCC	202	-7	-60	20	5.92
L MFG	83	-22	23	47	5.70
L vmPFC	126	-2	51	14	5.66
R precentral gyrus	198	26	-30	68	5.55
R VS	16	11	21	2	5.49
R STG	104	68	-15	2	5.32
R precuneus	24	16	-52	14	5.12
R pINS	26	41	-15	17	5.11
L MTG	56	-42	-62	23	4.95
L VS	19	-12	23	5	4.94
L hippocampus	20	-27	-22	-16	4.74
R PCC	26	18	-30	44	4.65
R caudate	15	21	-12	32	4.41
R MFG	19	38	-12	50	4.24
R STG	14	48	-32	17	4.19

aINS, anterior insula; dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; L, left; latOFC, lateral orbitofrontal cortex; MFG, middle frontal gyrus; MNI, Montreal Neurological Institute; MTG, middle temporal gyrus; PCC, posterior cingulate cortex; pINS, posterior insula; R, right; SMA, supplementary motor area; STG, superior temporal gyrus; VS, ventral striatum.

Table S6 | Neural sensitivity to emotional arousal.

A: (Anomalous + Beautiful) > Average-Looking Novel Faces					
Brain Regions by Cluster	**Cluster Size**	**Peak MNI Coordinates**			**Peak *t* Statistic**
		x	*y*	*z*	
R fusiform	27	48	-52	-19	6.59
B: (Anomalous + Beautiful) > Average-Looking Novel Faces					
	NS				
C: Functional Localizer High Arousal > Low Arousal					
L fusiform	46	-42	-55	-19	6.79
L IFG	19	-45	6	17	5.85
R precentral gyrus	34	48	8	32	5.44
R fusiform	32	46	-50	-19	5.31
L precentral gyrus	16	-40	-2	35	5.23
D: Functional Localizer Low Arousal > High Arousal					
R fusiform	108	28	-37	-13	9.65

L fusiform	84	-25	-45	-10	6.24
L PCC	15	-15	-60	14	5.17

IFG, inferior frontal gyrus; L, left; MNI, Montreal Neurological Institute; NS, non-significant; PCC, poster cingulate cortex; R, right.

Table S7 | Neural sensitivity to disgust.

Brain Regions by Cluster	Cluster Size	Peak MNI Coordinates			Peak *t* Statistic
		x	*y*	*z*	
A: Functional Localizer Disgusting Animals > Scrambled Images					
L fusiform	551	-25	-57	-10	8.64
R fusiform	676	38	-45	-22	8.61
R superior occipital cortex	132	28	-77	20	5.99
R amygdala	13	31	13	-22	5.80
R SPL	54	26	-52	53	5.55
L amygdala	31	-30	6	-19	5.23
Midbrain	13	6	-27	-10	5.13
L superior occipital cortex	53	-25	-75	29	5.12
L SPL	13	-25	-55	56	4.88
B: Functional Localizer Scrambled Images > Disgusting Animals					
Precuneus	16	0	-2	-62	5.14
C: Functional Localizer Disgusting Injuries > Scrambled Images					
R fusiform	103	31	-50	-10	6.58
L precuneus	129	-25	-75	26	6.05
L fusiform	168	-37	-67	-13	5.89
R precuneus	64	28	-72	32	5.45
R IPL	20	41	-32	44	5.35
R MTG	76	46	-72	-1	4.93
R precuneus	38	23	-60	56	4.58
L precentral gyrus	28	-42	3	32	4.54
D: Functional Localizer Scrambled Images > Disgusting Injuries					
R SPL	60	38	-70	44	6.50
L MTG	97	-52	-37	5	6.10
L putamen	15	-30	-15	8	5.78
R STG	101	63	-32	2	5.32
L STG	13	-55	-30	14	5.00

IPL, inferior parietal lobule; L, left; MNI, Montreal Neurological Institute; MTG, middle temporal gyrus; R, right; SPL, superior parietal lobule; STG, superior temporal gyrus.

Table S8 | Neural sensitivity to facial beauty.

Brain Regions by Cluster	Cluster Size	Peak MNI Coordinates			Peak *t* Statistic
		x	*y*	*z*	
A: Positive Parametric Modulation of Neural Response by Attractiveness Ratings					

	NS				
B: Negative Parametric Modulation of Neural Response by Attractiveness Ratings					
R fusiform	175	46	-45	-16	6.75
L fusiform	237	-35	-65	-13	6.47
L amygdala	56	-22	1	-19	6.46
L IFG	68	-47	8	23	5.34
L dlPFC	34	-45	38	-7	5.28
R IPS	17	31	-65	32	4.96
L thalamus	13	-12	-32	5	4.78
R dlPFC	16	51	36	-1	4.44
L SMG	51	-60	-30	41	4.39
L TPJ	13	-50	-52	23	4.35
R amygdala	15	21	1	-16	4.20
L IPS	13	-30	-42	44	4.19
L MTG	14	-42	-65	-4	3.91

dlPFC, dorsolateral prefrontal cortex; IFG, inferior frontal gyrus IPS, intraparietal sulcus; L, left; MNI, Montreal Neurological Institute; MTG, middle temporal gyrus; NS, non-significant; R, right; SMG, supramarginal gyrus; TPJ, temporoparietal junction.

Table S9 | Neural sensitivity to facial anomalies.

A: Anomalous > (Beautiful + Average-Looking Novel Faces)					
Brain Regions by Cluster	**Cluster Size**	**Peak MNI Coordinates**			**Peak *t* Statistic**
		x	*y*	*z*	
L fusiform	501	-35	-65	-13	8.60
R fusiform	446	38	-65	-10	8.16
L amygdala	108	-22	-7	-13	7.71
R amygdala	76	23	-5	-16	7.62
R hippocampus	22	38	-20	-22	5.86
L IPL	81	-55	-37	41	5.51
R latOFC	16	26	23	-16	5.25
R VS	15	6	-5	2	5.14
R superior occipital cortex	109	33	-77	38	5.03
L IFG	37	-45	6	20	4.80
R IPS	42	38	-45	44	4.36
L IPS	22	-25	-70	29	4.16
R IPS	15	48	-37	47	4.12
B: (Beautiful + Average-Looking Novel) > Anomalous Faces					
L STG	23	-57	-12	-1	5.36
L somatosensory cortex	33	-10	-32	77	5.02
L STG	16	-45	-35	14	4.47
L somatosensory cortex	29	-42	-12	35	4.35
R somatosensory cortex	22	13	-35	74	4.33
L STG	19	-55	-22	8	4.27
R somatosensory cortex	22	11	-30	68	4.27

IFG, inferior frontal gyrus; IPL, inferior parietal lobule; IPS, intraparietal sulcus; L, left; latOFC, lateral orbitofrontal cortex; MNI, Montreal Neurological Institute; R, right; STG, superior temporal gyrus; VS, ventral striatum.

Table S10 | Neural responses to facial anomalies in regions independently implicated in processing disgust and facial beauty.

A: Anomalous > (Beautiful + Average-Looking Novel Faces) Masking: Inclusive Disgusting Animals, Exclusive Visual Salience & Emotional Arousal					
Brain Regions by Cluster	**Cluster Size**	**Peak MNI Coordinates**			**Peak *t***
		x	*y*	*z*	**Statistic**
L fusiform	165	-35	-62	-13	8.16
R fusiform	164	46	-47	-19	7.63
L amygdala	23	-25	1	-19	7.01
L fusiform	22	-42	-72	-7	5.36
B: (Beautiful + Average-Looking Novel) > Anomalous Faces Masking: Inclusive Disgusting Animals, Exclusive Visual Salience & Emotional Arousal					
	NS				
C: Anomalous > (Beautiful + Average-Looking Novel Faces) Masking: Inclusive Disgusting Injuries, Exclusive Visual Salience & Emotional Arousal					
L fusiform	95	-35	-65	-13	8.60
R fusiform	35	43	-47	-19	5.74
D: (Beautiful + Average-Looking Novel) > Anomalous Faces Masking: Inclusive Disgusting Injuries, Exclusive Visual Salience & Emotional Arousal					
	NS				
E: Anomalous > (Beautiful + Average-Looking Novel Faces) Masking: Inclusive Facial Beauty, Exclusive Visual Salience & Emotional Arousal					
L fusiform	169	-35	-65	-13	8.60
R fusiform	122	38	-65	-10	8.16
L amygdala	48	-22	-7	-13	7.71
R amygdala	15	23	-5	-16	7.62
L fusiform	13	-42	-65	-7	5.89
L IPL	32	-55	-37	41	5.51
R pSTS	23	58	-52	-1	5.14
R fusiform	16	-32	-35	-19	4.87
L IFG	28	-45	6	20	4.80
R IPS	13	31	-67	32	4.66
F: (Beautiful + Average-Looking Novel) > Anomalous Faces Masking: Inclusive Facial Beauty, Exclusive Visual Salience & Emotional Arousal					
	NS				

IFG, inferior frontal gyrus; IPL, inferior parietal lobule; IPS, intraparietal sulcus; L, left; MNI, Montreal Neurological Institute; pSTS, posterior superior temporal sulcus; R, right.

Table S11 | Correlations between implicit biases and neural responses to facial anomalies.

Brain Regions by Cluster	Cluster Size	Peak MNI Coordinates			Peak *t* Statistic
		X	*y*	*z*	
A: Positive Correlation with IAT d' Scores, Anomalous > (Beautiful + Average-Looking Novel Faces)					
L fusiform	412	-35	-65	-13	8.79
R fusiform	387	38	-65	-10	7.87
R amygdala	74	23	-5	-16	7.74
L amygdala	99	-22	-7	-13	7.58
R hippocampus	22	38	-20	-22	5.77
L IPL	54	-55	-37	41	5.19
R latOFC	16	26	23	-16	5.08
L IFG	35	-47	8	20	4.97
R IPS	101	33	-77	38	4.91
R IPS	22	31	-47	44	4.39
L IPS	14	-25	-70	29	4.00
B: Negative Correlation with IAT d' Scores, Anomalous > (Beautiful + Average-Looking Novel Faces)					
L precuneus	18	-15	-75	26	5.26
L STG	28	-55	-7	-4	5.06
L somatosensory cortex	32	-10	-32	77	4.94
R somatosensory cortex	26	13	-55	-7	4.86
L STG	26	-55	-22	8	4.56
L STG	20	-45	-37	14	4.54
L MTG	13	-40	-47	2	4.43
R STG	16	58	-17	5	4.42
L somatosensory cortex	23	-42	-15	35	4.4
R somatosensory cortex	27	23	-37	71	4.35
R somatosensory cortex	21	16	-25	71	4.14

IFG, inferior frontal gyrus; IPL, inferior parietal lobule; IPS, intraparietal sulcus; L, left; latOFC, lateral orbitofrontal cortex; MNI, Montreal Neurological Institute; MTG, middle temporal gyrus; R, right; STG, superior temporal gyrus.

Supplementary Appendices

Appendix A | The Explicit Bias Questionnaire[8] (EBQ)

#	Question	Subscale	Options	
1	How often have you encountered or do you interact with a person with a facial disfigurement (marks, rashes, scars, asymmetry, paralysis, etc.)?	EBQ Frequency	Yes	No
2	Do you have a disability?	EBQ Disability	Yes	No
3	Do you have a facial disfigurement?	EBQ Facial Anomaly	Yes	No
4	Do you have a close friend or family member with a disability?	EBQ Family / Friend Disability	Yes	No
5	Do you have a close friend or family member with a facial disfigurement?	EBQ Family / Friend Anomaly	Yes	No
6	How important is your physical appearance on a scale from 1 to 7?	EBQ	1 = Extremely unimportant	7 = Extremely important
7	How warm or cold do you feel towards people with facial disfigurement? (RS)	EBQ	1 = Extremely cold	7 = Extremely warm
8	Which statement best describes you? (RS)	EBQ		
	1 = I strongly prefer people without facial disfigurements t people with facial disfigurements	7 = I strongly prefer people with facial disfigurement to people without facial disfigurement		
	Below is a list of statements about people with facial disfigurements. *Please state how strongly you agree or disagree with each of the following statements.*			
9	They are more happy, confident, assured, and cheerful than others. (RS)	EBQ	1 = Strongly disagree	7 = Strongly agree
10	They are more sad, shy, and miserable than others.	EBQ	1 = Strongly disagree	7 = Strongly agree
11	They are more attractive, desirable, and eligible than others. (RS)	EBQ	1 = Strongly disagree	7 = Strongly agree
12	They are more unattractive, undesirable, ugly, and unsuitable than others.	EBQ	1 = Strongly disagree	7 = Strongly agree
13	They are more easy-going, approachable, likeable, and friendly than others. (RS)	EBQ	1 = Strongly disagree	7 = Strongly agree
14	They are more awkward, unlikeable, unapproachable, and unfriendly than others.	EBQ	1 = Strongly disagree	7 = Strongly agree

15	They are more successful, motivated, accomplished, and more likely to succeed than others. (RS)	EBQ	1 = Strongly disagree	7 = Strongly agree
16	They are more limited and unmotivated and more likely to fail than others.	EBQ	1 = Strongly disagree	7 = Strongly agree

Please indicate how you would describe people with facial disfigurements on the following 7-point scales:

17	Sad to happy	EBQ	1 = Sad	7 = Happy
18	Unconfident to confident	EBQ	1 = Unconfident	7 = Confident
19	Incompetent to competent	EBQ	1 = Incompetent	7 = Competent
20	Shy to assured	EBQ	1 = Shy	7 = Assured
21	Miserable to cheerful	EBQ	1 = Miserable	7 = Cheerful
22	Unattractive to attractive	EBQ	1 = Unattractive	7 = Attractive
23	Undesirable to desirable	EBQ	1 = Undesirable	7 = Desirable
24	Ugly to gorgeous	EBQ	1 = Ugly	7 = Gorgeous
25	Stupid to intelligent	EBQ	1 = Stupid	7 = Intelligent
26	Unsuitable to eligible	EBQ	1 = Unsuitable	7 = Eligible
27	Awkward to easy-going	EBQ	1 = Awkward	7 = Easy-going
28	Untrustworthy to trustworthy	EBQ	1 = Untrustworthy	7 = Trustworthy
29	Unapproachable to approachable	EBQ	1 = Unapproachable	7 = Approachable
30	Unfriendly to friendly	EBQ	1 = Unfriendly	7 = Friendly
31	Non-achiever to achiever	EBQ	1 = Non-achiever	7 = Achiever
32	Ordinary to accomplished	EBQ	1 = Ordinary	7 = Accomplished
33	Unmotivated to motivated	EBQ	1 = Unmotivated	7 = Motivated

1 RS, reverse scored.

References

1. Ma, D. S., Correll, J. & Wittenbrink, B. The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. Methods* **47**, 1122–1135 (2015).

2. Workman, C. . I., Jamrozik, A., Rosen, M. E. & Chatterjee, A. The Face Image Meta-Database and Chatlab Facial Anomaly Database: Tools to Facilitate Neuroscience Research on Face Perception and Social Stigma. in *26th Annual Meeting of Cogntivie Neuroscience Society* (2019).

3. Greenwald, A. G., Nosek, B. A. & Banaji, M. R. Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *J. Pers. Soc. Psychol.* **85**, 197–216 (2003).

4. Greenwald, A. G., Nosek, B. A. & Banaji, M. R. Summary of Improved Scoring Algorithm. *J. Pers. Soc. Psychol.* **85**, 197–216 (2003).

5. Jamrozik, A., Oraa Ali, M., Sarwer, D. B. & Chatterjee, A. More than skin deep: Judgments of individuals with facial disfigurement. *Psychol. Aesthetics, Creat. Arts* **13**, 117–129 (2019).

6. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* **4**, 1–12 (2013).

7. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).

8. Hartung, F. *et al.* Behavioural and Neural Responses to Facial Disfigurement. *Sci. Rep.* **9**, 8021 (2019).