The dorsomedial prefrontal cortex mediates the interaction between moral and aesthetic valuation: a TMS study on the *beauty-is-good* stereotype

Chiara Ferrari,[1] Marcos Nadal,[2] Susanna Schiavi,[1] Tomaso Vecchi,[3,4] Camilo J. Cela-Conde,[2] and Zaira Cattaneo[1,4]

[1]Department of Psychology, University of Milano-Bicocca, Milan, Italy, [2]Human Evolution and Cognition Group (EvoCog), University of the Balearic Islands and IFISC, Associated Unit to CSIC, Palma de Mallorca, Spain, [3]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy and [4]Brain Connectivity Center, C. Mondino National Neurological Institute, Pavia, Italy

Correspondence should be addressed to Zaira Cattaneo, Department of Psychology, University of Milano-Bicocca, Milan, Italy.
E-mail: zaira.cattaneo@unimib.it.

Abstract

Attractive individuals are perceived as possessing more positive personal traits than unattractive individuals. This reliance on aesthetic features to infer moral character suggests a close link between aesthetic and moral valuation. Here we aimed to investigate the neural underpinnings of the interaction between aesthetic and **moral** valuation by combining transcranial magnetic stimulation (TMS) with a priming paradigm designed to assess the *Beauty-is-Good* stereotype. Participants evaluated the trustworthiness of a series of faces (targets), each of which was preceded by an adjective describing desirable, undesirable, or neutral aesthetic qualities (primes). TMS was applied between prime and target to interfere with activity in two regions known to be involved in aesthetic and moral valuation: the dorsomedial prefrontal cortex (dmPFC, a core region in social cognition) and the dorsolateral prefrontal cortex (dlPFC, critical in decision making). Our results showed that when TMS was applied over vertex (control) and over the dlPFC, participants judged faces as more trustworthy when preceded by positive than by negative aesthetic primes (as also shown in two behavioral experiments). However, when TMS was applied over the dmPFC, primes had no effect on trustworthiness judgments. A second Experiment corroborated this finding. Our results suggest that mPFC plays a causal role linking moral and aesthetic valuation.

Key words: aesthetics; morals; face; attractiveness; trustworthiness; social stereotypes; TMS, prefrontal cortex

Introduction

People's physical appearance has a substantial impact on the way others judge and treat them (Langlois *et al.*, 2000). The evaluation of others' behaviors and personal attributes is influenced by their attractiveness, especially if there is little other information available. Indeed, first impressions are largely based on physical appearance, including attractiveness. The assessment of the attractiveness of others' faces is fast and automatic, and serves as the basis for inferences about their socially relevant personal traits (Locher *et al.*, 1993; Olson and Marshuetz, 2005; Chatterjee *et al.*, 2009; Sui and Liu, 2009). Specifically, attractive individuals are attributed more positive personal and interpersonal qualities (such as trustworthiness,

intelligence and competence) than less attractive individuals (Dion *et al.*, 1972; Eagly *et al.*, 1991; Langlois *et al.*, 2000; Chatterjee *et al.*, 2009)—even when inferring positive traits from an attractive face does not necessarily lead to a correct estimate of the person (Jussim, 1991, 1993; Olivola and Todorov, 2010). This association emerges quite early in development (Langlois *et al.*, 2000; Griffin and Langlois, 2006), and influences the way people treat others. Attractive children and adults are treated better, and given better opportunities, than unattractive people (Langlois *et al.*, 2000). This is the case even in contexts in which such differential treatment is explicitly prohibited or discouraged, including schools, job interviews, salary negotiations, and court sentences (e.g., Hamermesh and Parker, 2005; Frevert and Walker, 2014).

This evaluation heuristic, known as the *Beautiful-is-Good* stereotype (e.g., Eagly *et al.*, 1991), has been discussed in Western thinking at least since Classical Greece, where the notion of beauty included what was morally good (intellectual beauty) and what was pleasant (sensuous beauty). Thus, it was used both in reference to the human physique, buildings and artworks, and in reference to personal character, actions and beliefs (Tatarkiewicz, 1970a). In his *Tusculan Disputations*, Cicero went even further, highlighting the parallels between the specific features of physical and moral rightness in humans: "And as in the body a certain symmetrical shape of the limbs combined with a certain charm of colouring is described as beauty; so in the soul the name of beauty is given to an equipoise and consistency of beliefs and judgments, combined with a certain steadiness and stability following upon virtue or comprising the true essence of virtue." (cited in Tatarkiewicz, 1970b, p. 206).

From the perspective of contemporary neuroscience, the intersection between morals and aesthetics suggests the possibility that the valuation of aesthetic and moral attributes may rely on partially overlapping neural and cognitive mechanisms (see Zaidel and Nadal, 2011, for a review). Does beauty appreciation of a painting, for instance, elicit similar brain responses to admiration for an altruistic gesture? This possibility is in line with the notion of a common network of brain regions, including the medial prefrontal cortex and ventral striatum, that computes the valuation of items of different sorts (Levy and Glimcher, 2012). This network represents the reward value of diverse objects, situations, or events, in a common neural currency, enabling the straightforward assessment and comparison of the value and motivational relevance of options of different kinds (Levy and Glimcher, 2012; Phelps *et al.*, 2014; Berridge and Kringelbach, 2015). Neural value computation plays an important role in social cognition too (Ruff & Fehr, 2014). Social features and preferences are also encoded in terms of a common neural currency that assigns value and motivational relevance to them (Zaki *et al.*, 2014). But although the activity of this circuit represents the integrated value of all relevant factors, different brain networks provide the perceptual and cognitive information relative to the different social features (Levy and Glimcher, 2012; Phelps *et al.*, 2014). It is therefore conceivable that attractive faces lead to the attribution of positive personal qualities because, although facial attractiveness and personal qualities *per se* are processed by different brain networks, the values of both physical and personal qualities are coded along a single dimension by a common valuation brain network.

Thus, understanding the neural mechanisms that link moral and aesthetic valuation has a double interest. First, it affords the opportunity to explain a prevalent and socially relevant stereotype that leads people to be assessed and treated differently based solely on their attractiveness. Second, it offers a novel domain in which to explore the mechanisms by which the human brain computes the value of socially relevant physical and personal features. In this

regard, the *Beauty-is-Good* association is interesting for our purposes not just as a case of social stereotyping (e.g. race or gender), but inasmuch as this association represents a "window" into the overlap between moral and aesthetic valuation.

But what is known about the neural correlates of this phenomenon? Although most research has selectively focused on facial attractiveness judgments or social evaluation, some neuroimaging studies have directly investigated the connection between moral and aesthetic valuation (Zaidel and Nadal, 2011, for review). These studies have revealed an extended cortical and subcortical network mediating the evaluation of both aesthetic and moral value including the amygdala, insula, nucleus accumbens, and also the orbitofrontal cortex (OFC) and medial and lateral sectors of the prefrontal cortex (Tsukiura and Cabeza, 2011; Bzdok *et al.*, 2012a; Avram *et al.*, 2013; Mende-Siedlecki *et al.*, 2013; Wang *et al.*, 2014). For instance, Bzdok *et al.* (2012a) found that explicit face trustworthiness judgments and face attractiveness judgments both induced activation in the dorsomedial prefrontal cortex (dmPFC) and in the inferior frontal gyrus. Common responses in the insula and in the medial OFC were also reported in Tsukiura and Cabeza's (2011) fMRI study when participants evaluated attractiveness of faces and when they decided about the morality of behavioral statements. Similarly, evaluating beauty in faces and morality in vignettes representing positively-valenced or neutral behaviors resulted in the activation of a common network comprising the OFC, the inferior temporal gyrus and the medial superior frontal gyrus (Wang *et al.*, 2014). Interestingly, a similar neural circuit (encompassing the OFC and mPFC) was observed when participants judged morality and aesthetics in poems (Avram *et al.*, 2013).

This evidence shows that medial sectors of the prefrontal cortex are involved in both aesthetic and moral valuation. Indeed, the mPFC is a core region of the social brain (Amodio and Frith, 2006; Van Overwalle, 2009): neuroimaging evidence suggests that it is involved in several aspects of social cognition, mediating self-representation (e.g., Gusnard *et al.*, 2001; D'Argembeau *et al.*, 2007; Jenkins and Mitchell, 2011), first impression formation (e.g., Mitchell *et al.*, 2005a; Baron *et al.*, 2011), personality trait inference (e.g., Ma *et al.*, 2011, 2013a), attribution of mental states (Mitchell *et al.*, 2005b), and social categorization, including stereotyping (Knutson *et al.*, 2007; Quadflieg *et al.*, 2009; Gilbert *et al.*, 2012). Studies in the aesthetic domain found also consistent activation in the mPFC in response to preferred stimuli, whether faces or artworks (Jacobsen *et al.*, 2006; Chatterjee *et al.*, 2009; Chatterjee and Vartanian, 2016). Lesion studies confirm the central role of the mPFC in social cognition. Indeed, damage to the ventromedial prefrontal cortex may lead to impaired theory of mind abilities (Jenkins *et al.*, 2014), abnormal social functioning, and limited attention to moral rules (e.g. Anderson *et al.*, 1999). Furthermore, patients with mPFC lesions not only are more inclined to approve moral violations compared to healthy participants (Ciaramelli *et al.*, 2007), but also show less or more pronounced stereotypical attitudes depending on the damaged portion of the mPFC (Gozzi, Raymont *et al.*, 2009), and abnormal trustworthiness perception in trust-games (Krajbich *et al.*, 2009). However, whether damage to medial sectors of the prefrontal cortex also biases aesthetic evaluations (of faces), or whether it impacts on how face attractiveness affects social (moral) evaluation, is not known.

Another region that deserves attention when investigating the link between moral and aesthetic evaluation is the dorsolateral prefrontal cortex (dlPFC). Activity in the dlPFC is also related to face attractiveness ratings (Nakamura *et al.*, 1998; Winston *et al.*, 2007; Chatterjee *et al.*, 2009; Ferrari *et al.*, 2015) and moral reasoning (Greene *et al.*, 2001; Greene *et al.*, 2004; Tassy *et al.*, 2011; Jeurissen

et al., 2014), and it has been found to respond to both moral and aesthetic evaluation within the same participants (Bzdok *et al.*, 2012a). However, the dlPFC is not part of the core social brain (Van Overwalle, 2009), and its involvement in social decisions may reflect a general role of this structure in decision-making and conflict regulation (e.g., Fleck *et al.*, 2006; Kim *et al.*, 2014). Nonetheless, we were interested in studying whether the dlPFC regulates flow of information from one system of value (i.e., aesthetics) towards another system of value (i.e., morals), as it does for instance when controlling emotional responses in social evaluation (Knutson *et al.*, 2007; Ito and Bartholow, 2009; Quadflieg *et al.*, 2011; Cattaneo *et al.*, 2011; Kubota *et al.*, 2012).

In this study, we combined a paradigm designed to assess the *Beauty-is-Good* stereotype with transcranial magnetic stimulation (TMS) to investigate the causal role of the mPFC and of the dlPFC in bridging moral and aesthetic valuation (note here that the OFC would also be an interesting area to study in this context, but unfortunately it is very difficult to be reliably reached by TMS). Brain stimulation allows interfering with the neural activity in a targeted region in a controlled and reversible manner. It is thus able to shed light on the *causal* role of different brain areas mediating a particular function/behavior, adding to the correlation evidence provided by neuroimaging studies. Importantly, participants in TMS experiments act as their own controls, overcoming some of the limitations intrinsic in patient studies, such as potential differences in pre-morbid ability, and variability depending on high heterogeneity of lesion extent and severity. We used a priming paradigm to elicit the *Beauty-is-Good* stereotype because priming/adaptation effects are well suited to the effects of TMS (e.g., Cattaneo *et al.*, 2008; Cattaneo and Silvanto, 2008; Cattaneo *et al.*, 2012). Participants were asked to judge the trustworthiness of faces that were preceded by an adjective conveying desirable aesthetic qualities (e.g., *attractive*), undesirable aesthetic qualities (e.g., *ugly*) or neutral qualities (e.g. *horizontal*). In two behavioral studies (see Supplementary Material) we showed that faces appeared more trustworthy when preceded by aesthetically positive adjectives, in line with the *Beauty-is-Good* stereotype, and suggesting that prime and target stimuli were being aligned on a common valuation scale. In fact, other prime cues unrelated to physical appearance but still representing the poles of negative/positive continua (e.g. *less/more*; *little/a lot*) did not affect trustworthiness evaluation (see Supplementary Material), ruling out unspecific *halo* effects.

If the mPFC mediates the link between moral and aesthetic valuation, as suggested by prior fMRI evidence (Bzdok *et al.*, 2012a; Avram *et al.*, 2013; Wang *et al.*, 2014), interfering with activity in this region should interfere with the effect of the aesthetic prime on the trustworthiness evaluation, possibly attenuating the behavioral expression of the *Beauty-is-Good* stereotype. Predicting the effects of stimulation of the dlPFC are less straightforward. Although this region plays a role in controlling inappropriate emotional responses in social contexts (i.e., stereotyping) (Knutson *et al.*, 2007; Quadflieg *et al.*, 2009; Gilbert *et al.*, 2012), in our task participants are unlikely to be aware of the priming effect and/or to consider it as socially 'inappropriate' and hence as a response to inhibit. Nonetheless, it may be that TMS over the dlPFC interferes with the way a general evaluative system allows information coming from different domains (aesthetics, moral) to interact, thus also affecting priming effects.

Experiment 1

Method

Participants. Twenty Italian participants (5 males, mean age = 22.4 years, SD = 2.0) volunteered to participate in the study. They were all right handed as assessed by a standard questionnaire (Oldfield, 1971) and all had normal or corrected to normal vision. Prior to the TMS experiment, each participant filled in a questionnaire (translated from Rossi *et al.*, 2011) to evaluate TMS safety. An additional 34 participants were tested in two behavioral experiments (for details, see Supplementary Material). The experiment was approved by the local ethical committee and participants were treated in accordance with the Declaration of Helsinki.

Stimuli. Experimental stimuli consisted of 32 young Caucasian faces displayed in frontal pose and with a neutral expression and of 6 adjectives. Face stimuli (7×7 deg of visual angle) were selected from a larger set of computer-generated faces (cf. http://tlab.princeton.edu/databases/randomfaces/) for which rating scores (on a 9-point Likert scale) on different trait dimensions (including trustworthiness) are available (for details, see Oosterhof and Todorov, 2008). From this set, we selected 16 unambiguously males and 16 unambiguously females of medium trustworthiness (within ± 1 SD from the mean of the whole sample, mean = 4.8, $SD = 0.7$). Medium trustworthiness faces were intentionally chosen, allowing for the possible influence of the prime-adjective on participants' evaluations. All the adjectives were selected from the Corpus CODIS of written Italian (http://corpora.dslo.unibo.it/coris_ita.html) and referred either to desirable human aesthetic attributes (we used two adjectives: *attractive* and *beautiful*), or to undesirable aesthetic features (we used two adjectives: *horrid* and *ugly*), or described neutral traits not related to human qualities (we used two adjectives: *horizontal* and *diagonal*).

Procedure. Participants were seated in front of a 15.5'' PC (1280×800 pixels) screen at an approximate distance of 57 cm, in a **normally lightened** and silent room, and performed a computerized task. Before starting the experiment, participants were informed that they would be viewing a set of faces and that their task was to indicate whether each face appeared trustworthy to them or not. Figure 1 shows the timeline of an experimental trial. Each trial started with a central black fixation point (1200 ms). Next, the adjective-prime appeared for 300 ms, followed by a blank screen (150 ms) and by the target face stimulus immediately after, which remained on the screen until participants responded. Participants were instructed to (silently) read the prime adjective and to judge the face as trustworthy or not by left/right key pressing using their right hand (response key assignment was counterbalanced across participants). Participants were instructed to be as accurate and fast as possible. TMS was delivered between the appearance of the adjective-prime and the face to be judged (see below for TMS details). Each participant performed three experimental blocks, one for each TMS targeted site. In each block, each face was presented three times, once for each prime-adjective type (*beauty*-prime, *ugliness*-prime, or *neutral*-prime), for a total of 96 trials in each block. Faces were presented in random order within each block, with the only constraint that the same face never appeared in two consecutive trials. The three experimental blocks were performed within the same session (participants were given a few minutes break after the first and second block); the order of the TMS targeted sites was counterbalanced across participants. Participants performed six practice trials at the beginning of the experimental session to familiarize with the task. The software E-prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, USA) was used for stimuli presentation, TMS triggering and data recording.
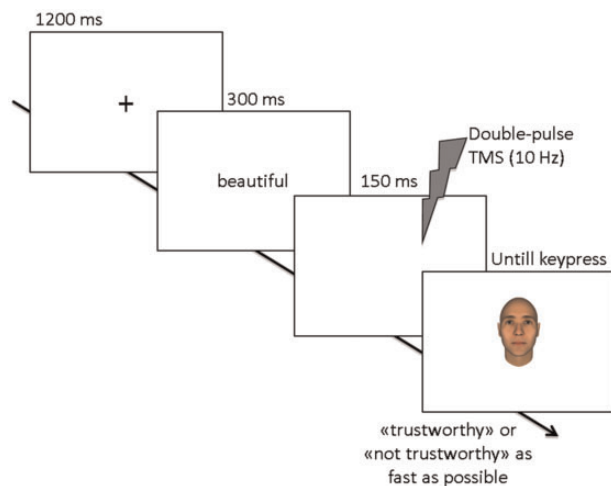
Fig. 1. The timeline of an experimental trial. Participants had to classify a face as trustworthy or not trustworthy. Each face was preceded by an adjective-prime that was either neutral; related to beauty (e.g., *beautiful, attractive*) or related to ugliness (e.g., *ugly, horrid*). 10 Hz double-pulse TMS was applied over the dmPFC, the right dlPFC, or over the vertex (control site) between the presentation of the prime and the target face.

Transcranial magnetic stimulation (TMS). Online neuronavigated TMS was performed with a Magstim Rapid[2] stimulator (Magstim Co Ltd, Whitland, UK) connected to a 70 mm butterfly coil at a fixed intensity of 60% of the maximum stimulator output (e.g., Campana *et al.*, 2007, 2014a, 2015; Bona *et al.*, 2014). Double-pulse TMS (10 Hz) was delivered 50 ms after the offset of the adjective-prime. Accordingly, the first TMS pulse was given 100 msec before the onset of the face, and the second pulse upon onset of the face. Targeted sites were the dmPFC, the right dlPFC, and the vertex (control site). We targeted the dlPFC in the right hemisphere in light of converging evidence indicating that the right dlPFC is involved more than the left in social decisions, including face attractiveness evaluation (e.g., Ferrari *et al.*, 2015), social categorization (e.g., Mitchell *et al.*, 2009; for a review, Amodio, 2014), implementation of fairness-related behaviours (Knoch *et al.*, 2006), and moral reasoning (Green *et al.*, 2004; Tassy *et al.*, 2011). The vertex was localized as the point falling half the distance between the nasion and the inion on the same midline. The dmPFC and the right dlPFC were localized by means of stereotaxic navigation on individual estimated magnetic resonance images (MRI) obtained through a 3D warping procedure fitting a high-resolution MRI template with the participant's scalp model and craniometric points (Softaxic, EMS, Bologna, Italy) (see Figure 2). This procedure has been proven to ensure a global localization accuracy of roughly 5 mm, a level of precision closer to that obtained using individual MRIs than to what can be achieved using other localization methods (Carducci and Brusco, 2012). Anatomical MNI coordinates were obtained from previous neuroimaging studies on traits perception and stereotypes (Mitchell, Cloutier, Banaji, & Macrae, 2006; Mitchell et al., 2009) and were $x = -3$, $y = 48$, $z = 48$ for the dmPFC, and $x = 32$, $y = 22$, $z = 38$ for the right dlPFC. MNI coordinates were then converted into the Talairach space (Talairach and Tournoux, 1988) to be suitable for the stereotaxic navigation. The coil was placed tangentially to the scalp with the handle pointing backward and held parallel to the midsagittal line in the vertex and mPFC stimulation conditions, and pointing backward and rightward at a 45° angle from the mid-sagittal line in the right dlPFC condition.
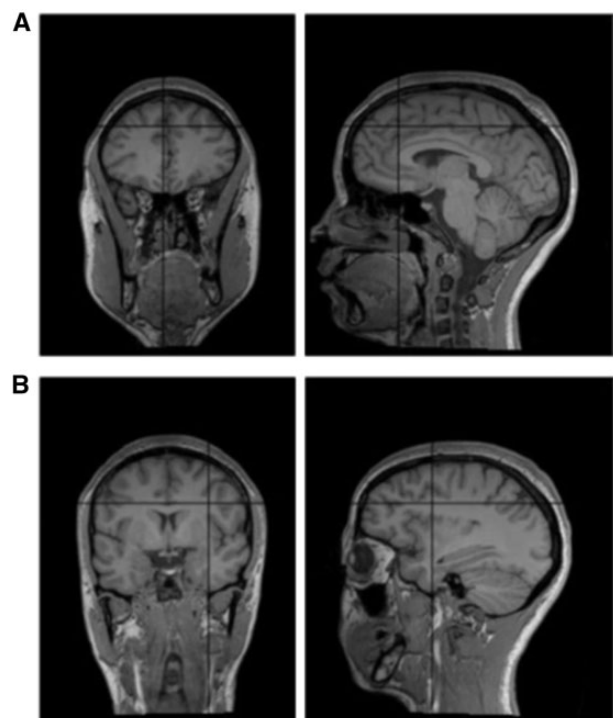

Fig. 2. The coronal (left) and sagittal (right) section of the estimated MRI of a representative participant showing the targeted site in the (**A**) dorsomedial prefrontal cortex (dmPFC, MNI $x = -3$, $y = 48$, $z = 48$) and (**B**) dorsolateral prefrontal cortex (dlPFC, MNI $x = 32$, $y = 22$, $z = 38$).

Results

The number of positive (i.e., 'this face is trustworthy') responses was calculated for each participant in each block and converted into a percentage score. Similarly, mean response latencies (RT) were calculated for each participant in each TMS condition. Trials in which participants' RT were ±3SD above or below their own average response time were excluded from the analyses (0.99% of trials were excluded following this criterion). The dependent variables were analyzed via repeated-measures ANOVAs with prime (*beauty, ugliness, neutral*) and TMS (dlPFC, dmPFC, vertex) as within-subjects factors. The Bonferroni-Holm correction was applied to all post-hoc comparisons.

The analysis revealed a significant main effect of TMS on the percentage of faces judged as trustworthy (Figure 3a) [$F(2,38) = 5.15$, $P = 0.010$, $\eta_p^2 = 0.21$], a significant effect of prime [$F(2,38) = 8.39$, $P = 0.001$, $\eta_p^2 = 0.31$], and a significant interaction TMS by prime [$F(4,76) = 2.67$, $P = 0.039$, $\eta_p^2 = 0.12$]. TMS over dlPFC lowered the percentage of faces judged as trustworthy, compared to both vertex [$t(19) = 3.18$, $P = 0.015$] and dmPFC stimulation [$t(19) = 2.99$, $P = 0.014$]. In turn, the percentage of 'trustworthy' responses did not differ significantly in the dmPFC and vertex TMS conditions [$t(19) < 1$, $P = 0.98$]. The effect of prime was modulated by the TMS condition. In the baseline (vertex) condition, the effect of prime was significant [$F(2,38) = 7.78$, $P = 0.001$, $\eta_p^2 = 0.29$]. Specifically, faces were judged as trustworthy significantly more frequently when preceded by *beauty*-related primes than when preceded by *ugliness*-related primes [$t(19) = 3.73$, $P = 0.003$] (Figure 3b). A similar trend emerged also for the *beauty*-related vs. neutral comparison [$t(19) = 2.07$, $P = 0.11$ (without correction, $P = 0.053$)]. Also, faces tended to be judged as trustworthy less frequently when preceded by the *ugliness*-related than neutral primes [$t(19) = 2.03$, $P = 0.057$]. Overall, this pattern
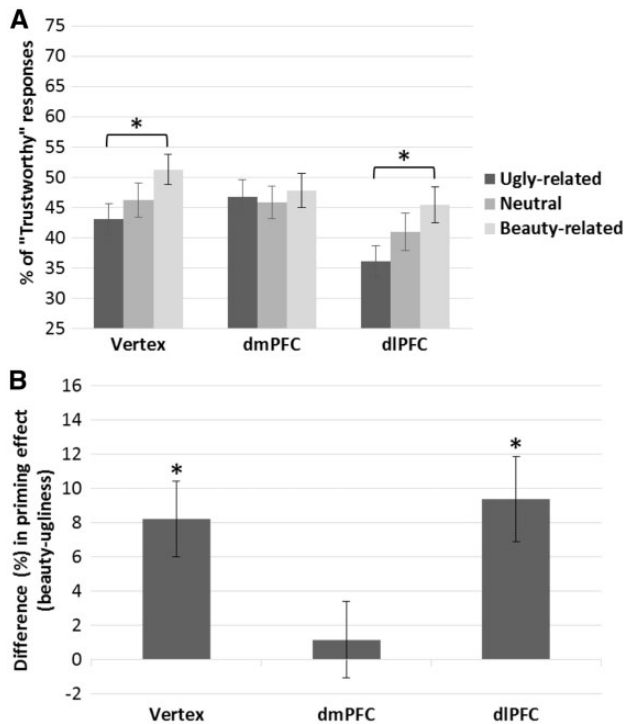
Fig. 3. (**A**) Percentage of positive responses (i.e., *The face is trustworthy*) as a function of prime (ugliness, neutral, beauty) and TMS condition (vertex, dmPFC, dlPFC). In the baseline (vertex) and in the dlPFC TMS conditions, faces were classified as trustworthy significantly more frequently following *beauty* primes than *ugliness* primes. Although participants evaluated faces as overall less trustworthy when TMS was applied over the dlPFC, stimulation over this region did not impact on the *Beauty-is-Good* stereotype. In turn, TMS over the dmPFC abolished the effect of priming. Error bars indicate ± 1 SEM. Asterisks indicate significant differences in priming effects within each TMS condition. (**B**) Difference in the percentage of faces classified as trustworthy when faces were preceded by *beauty*-primes vs. *ugliness*-primes (i.e., *beauty* minus *ugliness*). Asterisks indicate a significant difference compared to zero (i.e., no priming effect). Error bars indicate ± 1 SEM.

resembled the one found in the pilot behavioral experiment (see Supplementary Material).

In the dlPFC condition, similar priming effects were observed [$F(2,38)$ = 6.86, P = 0.003, η_p^2 = 0.27]. In particular, faces preceded by *ugliness*-related primes were judged as trustworthy significantly less frequently than faces preceded by *beauty*-related primes [$t(19)$ = 3.77, P = 0.003] and neutral primes [$t(19)$ = 2.49, P = 0.044]. The priming effect for beauty vs. neutral failed to reach statistical significance [$t(19)$ = 1.47, P = 0.16], although the pattern was similar to the one observed in the baseline Vertex condition (Figure 3a). In the dmPFC condition, critically, the main effect of prime was not significant [$F(2,38)$ < 1, P = 0.60]. It seems, thus, that TMS over this region prevented the emergence of the *Beauty-is-Good* stereotype.

Mean RT for positive and negative responses are reported in Figure 4. The ANOVA on the mean RT for positive responses revealed a significant main effect of prime [$F(2,38)$ = 9.07, P = 0.001], indicating that responses were faster overall when following *beauty*-related primes than when following *ugliness*-related primes [$t(19)$ = 4.20, P < 0.001], and neutral primes [$t(19)$ = 2.27, P = 0.070]. Furthermore, RT were slightly slower following *ugliness*-related primes than neutral primes [$t(19)$ = 2.05, P = 0.06]. Neither the main effect of TMS [$F(2,38)$ < 1, P = 0.63], nor the interaction prime by TMS [$F(4,76)$ = 1.18, P = .33], reached significance. The ANOVA on the mean RT for negative

responses revealed neither a significant effect of prime [$F(2,38)$ < 1, P = 0.42], TMS [$F(2,38)$ < 1, P = 0.40], or their interaction [$F(4,76)$ = 2.18, P = 0.08].

Experiment 2

Experiment 1 showed that interfering with dmPFC activity abolishes the effect of aesthetic primes over face trustworthiness decisions. However, the effects (also at baseline) were overall of small size. In order to rule out the possibility that our findings possibly reflected a false positive, we decided to carry out a second experiment to verify whether the pattern of results obtained in Experiment 1 could be replicated in a new sample of participants. The experimental procedure was identical to that of Experiment 1 except for the fact that only positive and negative aesthetic primes were used. Neutral primes were not used in this second Experiment given that our interest was mainly in the differential effect of the two poles of the aesthetic dimension (ugliness vs. beauty) on trustworthiness valuation.

Methods

Participants. Twenty Italian participants (3 males, mean age = 22.6 years, SD = 1.4) volunteered to participate in the study. None of them had participated in Experiment 1. Inclusion criteria were the same as for Experiment 1.

Stimuli and procedure. The experimental paradigm was identical to Experiment 1, with the exception that the neutral adjectives were not included. TMS sites, parameters and timing were the same as those of Experiment 1.

Results

Analyses were carried out as in Experiment 1. Trials in which participants' RT were ±3SD above or below their own average response time were excluded from the analyses (0.87% of trials were excluded following this criterion). A repeated-measures ANOVA with prime (*beauty* vs. *ugliness*) and TMS (dlPFC, dmPFC, and vertex) as within-subjects factors was carried out on percentage scores and mean RT. The analysis on the percentage of faces judged as trustworthy revealed a significant main effect of TMS [$F(2,38)$ = 3.80, P = 0.031, η_p^2 = 0.17]. As shown in Figure 5a, TMS over the dlPFC lowered the percentage of faces judged as trustworthy, compared to vertex stimulation [$t(19)$ = 2.97, P =
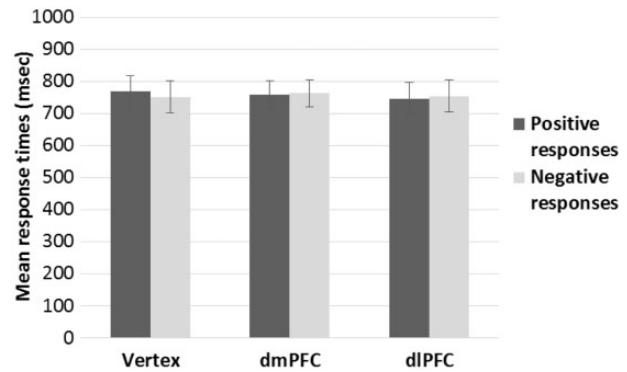
Fig. 4. Mean reaction times as a function of participants' positive (i.e., *The face is trustworthy*) or negative (i.e., *The face is not trustworthy*) responses and TMS condition in Experiment 1. TMS did not affect response times. Error bars indicate ± 1 SEM.
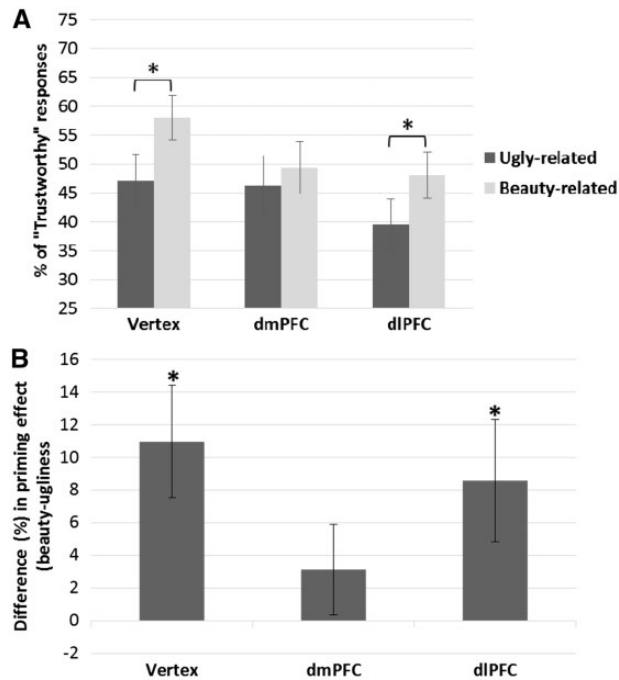
Fig. 5. (**A**) Percentage of positive responses (i.e., *The face is trustworthy*) of Experiment 2 as a function of prime (ugliness vs. beauty) and TMS condition (vertex, dmPFC, dlPFC). Similarly to Experiment 1, the *Beauty-is-Good* stereotype was observed when TMS was delivered over the vertex and over the dlPFC, but not when TMS was delivered over the dmPFC. TMS stimulation of the dlPFC decreased the overall number of positive responses. Error bars indicate ± 1 SEM. Asterisks indicate significant differences in priming effects within each TMS condition. (**B**) Difference in the percentage of faces classified as trustworthy when faces were preceded by *beauty*-primes vs. *ugliness*-primes (i.e., *beauty* minus *ugliness*) in Experiment 2. Asterisks indicate a significant difference compared to zero (i.e., no priming effect). Error bars indicate ± 1 SEM.

0.024] (Bonferroni-Holm correction applied), whereas there were no main differences between dlPFC and dmPFC stimulation [$t(19)=1.22$, $P = 0.24$], and between dmPFC and vertex stimulation [$t(19) = 1.44$, $P = 0.17$]. The main effect of prime [$F(1,19) = 7.30$, $P = 0.014$, $\eta_p^2 =.28$], and the interaction TMS by prime [$F(2,38) = 3.26$, $P = 0.049$, $\eta_p^2 = 0.15$] were also significant. In the baseline (vertex) condition, faces were judged as trustworthy significantly more frequently when preceded by *beauty*-related primes than when preceded by *ugliness*-related primes [$t(19) = 3.20$, $P = 0.005$] (Figure 5b). In the dlPFC TMS condition, a similar priming effect was observed [$t(19) = 2.30$, $P = 0.033$] (Figure 5b). In turn, when TMS was delivered over the dmPFC condition, no difference was observed between positive and negative primes in biasing 'trustworthy' responses [$t(19) = 1.14$, $P = 0.27$].

Mean RT for positive and negative responses are reported in Figure 6. The ANOVA on mean RT for positive responses revealed no significant effect of either TMS [$F(2,38) < 1$, $P = 0.84$], or prime [$F(1,19) = 1.46$, $P = 0.24$]. The interaction prime by TMS was not significant [$F(2,38) = 2.02$, $P = 0.15$]. The ANOVA on mean RT for negative responses did not reveal any significant effect: TMS [$F(2,38) < 1$, $P = 0.41$], prime [$F(1,19) = 2.42$, $P = 0.14$], prime by TMS interaction [$F(2,38) < 1$, $P = 0.73$].

Discussion

In two different experiments, participants had to evaluate the trustworthiness of computer-generated faces that were preceded by prime adjectives denoting desirable (beauty),
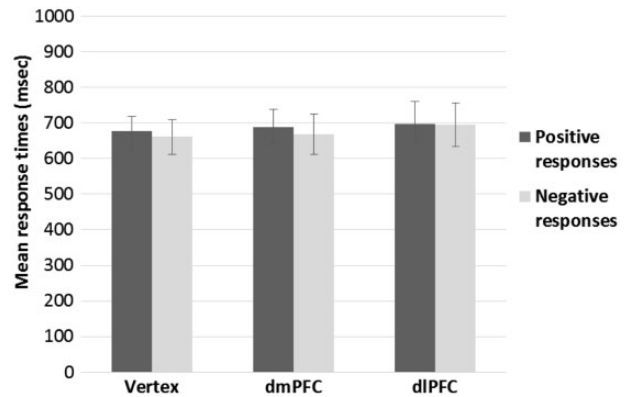

Fig. 6. Mean reaction times as a function of participants' positive (i.e., *The face is trustworthy*) or negative (i.e., *The face is not trustworthy*) responses and TMS condition in Experiment 2. Response times were not affected by TMS. Error bars indicate ± 1 SEM.

undesirable (ugliness), or neutral aesthetic qualities (in Experiment 2, neutral primes were not used). Participants in both experiments rated faces as more trustworthy when preceded by *beauty*-related primes than when preceded by *ugliness*-related primes in the baseline control condition (Vertex stimulation). Results in this condition replicated the same pattern obtained in a pilot behavioral study (see Supplementary Material), and reflect the *Beauty-is-Good* stereotype (Eagly *et al.*, 1991; Langlois *et al.*, 2000). Critically, when TMS was applied over the dmPFC, the stereotypical association between attractiveness and trustworthiness disappeared. In turn, following dlPFC stimulation faces tended to appear overall as less trustworthy, but the *Beauty-is-Good* stereotype was still observed. Overall, thus, our data suggest that the dmPFC (but not the dlPFC) plays a key role in linking aesthetic and moral valuation.

Interfering with neural activity in the dmPFC did not affect face trustworthiness evaluation *per se*: when faces were preceded by neutral primes (Experiment 1), participants' responses did not differ between the dmPFC and the control condition. This is in line with prior TMS evidence showing that interfering with dmPFC activity did not impact perceived face trustworthiness when the judgment was exclusively based upon facial appearance (Ferrari *et al.*, 2016, supplementary data), and with prior neuroimaging evidence indicating that evaluations uniquely based on face appearance are likely to elicit responses in subcortical (e.g., amygdala) more than in cortical regions (Said *et al.*, 2009; Baron *et al.*, 2011; de Gelder *et al.*, 2012, Fouragnan *et al.*, 2013; Mende-Siedlecki *et al.*, 2013; but see Bzdock *et al.*, 2012a). In turn, TMS over the dmPFC reduced the effect of priming (more consistently so across the two experiments for the positive primes), such that trustworthiness responses in this TMS condition were similar regardless the prime type (Experiments 1 and 2). This is in line with reports of (anterior) mPFC critical involvement in social priming in prior fMRI research (Wang and Hamilton, 2015).

The lack of priming effects following dmPFC TMS is unlikely to reflect a general role of this region in mediating semantic priming *per se*. Indeed, semantic priming tasks unrelated to a social dimension do not recruit the dmPFC (e.g., Copland *et al.*, 2007; Kircher *et al.*, 2009). Accordingly, neuroimaging evidence suggests that person knowledge is functionally dissociable within the brain from other classes of semantic knowledge (for instance, related to objects features) (Mitchell *et al.*, 2002; see also Ma *et al.*, 2013b). In line with this, damage to medial sectors of

the PFC tends to elicit specific deficits in social reasoning and cognition (e.g., Anderson *et al.*, 1999; Gozzi *et al.*, 2009; Jenkins *et al.*, 2014), but does not typically affect semantic knowledge in general, mainly mediated by temporal lobe regions (e.g., Gainotti, 2000; Campanella *et al.*, 2010; Piretti *et al.*, 2015). Moreover, it is unlikely that TMS over the dmPFC acted by disrupting maintenance of the verbal cue in memory. Indeed, interfering with dmPFC activity with TMS in prior studies did not affect maintenance of verbal primes (e.g., Mattavelli *et al.*, 2011; Ferrari *et al.*, 2016). In turn, short–term memory for visually presented words is usually affected by stimulation of visual (e.g., Amassian *et al.*, 1989; van de Ven *et al.*, 2012) or language-related areas (e.g., Deschamps *et al.*, 2014).

If, on the one hand, our results are unlikely to depend on an unspecific role of the dmPFC in semantic priming or short term memory (see above), on the other hand, we do not argue for a selective role of the dmPFC in mediating aesthetic-to-moral (priming) associations. In fact, although in our study we focused on the *Beauty-is-Good* stereotype as a "window" onto the intersection of **moral** and aesthetic evaluation, previous neuroimaging studies have shown preferential activation in the medial PFC when responses matched other stereotypical social beliefs, as those concerning gender or race (Mitchell *et al.*, 2006; Knutson *et al.*, 2007; Ito and Bartholow, 2009; Quadflieg *et al.*, 2009; Gilbert *et al.*, 2012). Accordingly, interfering with mPFC activity via brain stimulation was found to affect implicit measures of stereotypical beliefs about gender and in-group/out-group (positive vs. negative) attributes (Cattaneo *et al.*, 2011; Sellaro *et al.*, 2015). Brain-lesion evidence also supports the involvement of the (ventro-) medial PFC in stereotypical beliefs (see Gozzi *et al.*, 2009). Still, although activity in the dmPFC is certainly modulated by the stereotypicality of the information available about another agent (Van der Cruyssen *et al.*, 2015), the dmPFC is also involved in social evaluation beyond stereotypical categorizing. Indeed, converging evidence points to an involvement of the mPFC in different aspects of social evaluation such as first impression formation, personality traits inference, and attribution of mental states (Baron *et al.*, 2011; Contreras *et al.*, 2012; Fouragnan *et al.*, 2013; Ma *et al.*, 2013a,b; Mitchell *et al.*, 2005; Van den Stock *et al.*, 2014; for a review, Van Overwalle, 2009). Our study critically adds to this prior evidence by showing that the mPFC is also a key region in mediating the 'transfer' from the domain of aesthetics to the domain of morality (in the form of a stereotypical *Beauty-is-Good* association). This is also in agreement with prior evidence pointing to a critical role of the mPFC not just in selectively mediating moral judgments (e.g., Greene and Haidt, 2002; Beer and Ochsner, 2006; Bzdok *et al.*, 2012b; Englander *et al.*, 2012; Yoder and Decety, 2014) and aesthetic judgments (Jacobsen *et al.*, 2006; Kirsch *et al.*, 2015; Pegors *et al.*, 2015; Vessel *et al.*, 2015), but also in linking aesthetic and moral valuations (Bzdok *et al.*, 2012a; Avram *et al.*, 2013; Wang *et al.*, 2014).

Following dlPFC stimulation, faces tended to be generally judged as less trustworthy, but the effect of *beauty*-related primes on trustworthiness decisions was still observed. We were interested in verifying whether the dlPFC plays a role in regulating flow of information between the aesthetic and moral dimensions in impression formation, in light of its regulatory role in controlling emotional responses in social categorization tasks (e.g., Cattaneo *et al.*, 2011; Knutson *et al.*, 2007; Ito and Bartholow, 2009; Kubota *et al.*, 2012; Quadflieg *et al.*, 2011). Our data suggest that this was not the case. However, it is important to note that dlPFC activity is related to social decision making when a conflict is detected, such as when a stereotypical

representation is violated (e.g., a woman depicted in a male-stereotypical occupation, such as a 'chef', Quadflieg *et al.*, 2011). In our paradigm, there was no conflict between the aesthetic cues and the faces, because faces were all of average-trustworthy. It may be that using other paradigms eliciting a conflict between the aesthetic and moral dimension (for instance, a very beautiful male face associated with the description of a very bad act) may then recruit the dlPFC. Future research may address this issue.

In turn, the overall decrease in the number of faces perceived as trustworthy following dlPFC stimulation (irrespective of the prime) compared to the control condition is in line with neuroimaging evidence on the role of this region in the evaluation of trustworthiness in faces (Bzodck et al., 2012a) and behaviors (Watabe *et al.*, 2011), in addition to evaluations of moral appropriateness and moral reasoning (Greene *et al.*, 2001, 2004; Tassy *et al.*, 2011; Jeurissen *et al.*, 2014). Furthermore, the dlPFC may regulate subjective evaluations of positive traits in general: for instance, increasing excitability in the dlPFC resulted into higher attractiveness judgments for faces (Ferrari *et al.*, 2015), and interfering with its activity also affected appreciation of visual artworks (Cupchik *et al.*, 2009; Cattaneo *et al.*, 2014b; 2015; Chatterjee and Vartanian, 2016).

The priming effect we reported in our baseline condition most probably did not depend on unspecific halo effects, since other verbal cues unrelated to physical appearance but still evoking a negative/positive continuum (e.g. *less/more*; *little/a lot*) did not affect trustworthiness ratings (see Supplementary Material). Hence, one may question whether priming occurred because faces were perceived as less/more beautiful and hence less/more trustworthy following *ugliness-* vs. *beauty*-related primes. We think that this possibility is unlikely, and that the aesthetic adjectives directly biased trustworthiness decisions (possibly relying on a common valuation system), without going through an intermediate visual step in which faces also appeared less/more attractive. In fact, deciding about attractiveness of computer-generated faces such as the ones we used (that were specifically created to vary along the trustworthiness dimension, Oosterhof and Todorov, 2008), especially if a yes/no decision response is required, feels unnatural because these faces lack important features that are typically used to determine attractiveness (e.g., hair, skin texture, eye color, variation in symmetry, masculine/feminine traits). Accordingly, several studies focusing on mechanisms implied in face attractiveness evaluation employed real faces (e.g., Jones *et al.*, 2004; Little *et al.*, 2008; Mitrovic *et al.*, 2016), whereas the use of computer generated-faces may be suboptimal for such a purpose (e.g., Komori *et al.*, 2009; Sutherland *et al.*, 2013).

In our study, TMS affected the decision output but not response latencies. Dissociation of TMS effects on responses bias/accuracy and reaction times are not uncommon (Devlin and Watkins, 2007; Robertson *et al.*, 2003), and largely depend on the specific paradigm used. As we mentioned above, we used average trustworthy faces so that decisions were uncertain and could be modulated by the primes we used: TMS is more effective in affecting responses when uncertainty is higher (Robertson *et al.*, 2003). In turn, when there are clear correct vs. incorrect responses (with accuracy being high), behavioral effects induced by TMS tend to manifest more in terms of differences in reaction times (Devlin and Watkins, 2007). In the case of our paradigm, priming effects became manifest essentially in the response given and only marginally so in the response latencies (with no priming effects on RT in the purely behavioral studies, see Supplementary Material, and in the baseline

condition of Experiment 2); it is thus less surprising that TMS mainly modulated the response bias than RT.

The effects of TMS over the dmPFC were overall of small size. In this regard, it is important to consider that other cortical and subcortical regions have been found to respond to both moral and aesthetic evaluation, such as the orbitofrontal cortex (involved in reward processing, common to both aesthetic and moral valuation, see Tsukiura and Cabeza, 2011) and the insular cortex (critical in processing negative emotions and social negative signals, see Tsukiura and Cabeza, 2011). The temporal lobe may also be involved given its role in mediating emotional memories and social knowledge (see Zaidel and Nadal, 2011, for a review). The relative small size of the effect of TMS over the dmPFC may also partially reflect the work of these other regions in mediating the association between moral and aesthetic valuation. Nonetheless, it is worth noting that TMS can modulate activity not only in the neurons under the coil but also in interconnected regions (e.g., Avenanti *et al.*, 2013; Siebner *et al.*, 2009). For instance, the OFC and the mPFC are known to be inherently connected (Öngür and Price, 2000); it is thus possible that the TMS effects we reported did not solely reflect direct interference with the mPFC activity but also indirect modulation of a larger network including the OFC.

In our study, we did not consider whether positive personality traits would also prime a face to appear more attractive, and whether this would in case rely on similar neural mechanisms. Accordingly, literature has mainly focused on the 'what is beautiful is good' rather than the reverse inference, possibly reflecting the precedence of the aesthetic attribute over other personal attributes in first impression formation (see Eagly *et al.*, 1991). Nonetheless, available evidence suggests that attractiveness judgments can also be permeable to the influence of 'goodness' evaluation. In particular, Little and colleagues (Little *et al.*, 2006) found that individuals positively valuing particular personality traits found faces displaying those traits to be more attractive. Similarly, Zhang *et al.* (2014, see also Eagly *et al.*, 1977) reported that faces presented simultaneously with positive personality traits were rated as more attractive than faces presented with negative personality traits or no-information. Although we are not aware of any study that directly looked at the neural underpinnings of the influence of perceived goodness of a person over her/his face attractiveness, it is reasonable to speculate that the dmPFC would be involved, given its role in encoding personality traits (Ma *et al.*, 2013a,b; Van Overwalle *et al.*, 2015). Future neuroimaging and brain stimulation research may shed light on this interesting issue.

In sum, our study provides evidence for a *causal* role of the dmPFC in mediating the link between aesthetic and moral valuation, critically stepping beyond prior evidence based on correlational techniques supporting the existence of a common brain network mediating aesthetic and moral evaluation (e.g., Tsukiura and Cabeza, 2011; Bzdok *et al.*, 2012a; Avram *et al.*, 2013; Mende-Siedlecki *et al.*, 2013; Wang *et al.*, 2014). This network is believed to encode value in terms of a common neural currency and assign value and motivational relevance to social and non-social stimuli alike (Ruff *et al.*, 2013; Zaki *et al.*, 2014). From this perspective, therefore, aesthetics and morals are linked in terms of a common valuation neural system that assigns congruent values to beauty and goodness, and common motivational dispositions to attraction and trustworthiness. Morals and aesthetics are likely to be distinctively human traits, and have been systematically associated in the history of Western philosophy. Our data, together with prior neuroimaging findings (e.g., Tsukiura and Cabeza, 2011), suggest the possibility that the association we experience between aesthetics and morals may actually be due to the two systems of value exploiting a common neural network, at least in as much they apply to the valuation of other's socially relevant features (Zaidel and Nadal, 2011).

References

Amassian, V.E., Cracco, R.Q., Maccabee, P.J., Cracco, J.B., Rudell, A., Eberle, L. (1989). Suppression of visual perception by magnetic coil stimulation of human occipital cortex. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, **74**, 458–62.

Amodio, D.M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, **15**, 670–82.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, **7**, 268–77.

Anderson, S.W., Bechara, A., Damasio, H., Tranel, D., Damasio, A.R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, **2**, 1032–7.

Avenanti, A., Annella, L., Candidi, M., Urgesi, C., Aglioti, S.M. (2013). Compensatory plasticity in the action observation network: virtual lesions of STS enhance anticipatory simulation of seen actions. *Cerebral Cortex*, **23**, 570–80.

Avram, M., Gutyrchik, E., Bao, Y., Pöppel, E., Reiser, M., Blautzik, J. (2013). Neurofunctional correlates of esthetic and moral judgments. *Neuroscience Letters*, **534**, 128–32.

Baron, S.G., Gobbini, M.I., Engell, A.D., Todorov, A. (2011). Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Social Cognitive and Affective Neuroscience*, **6**, 572–81.

Beer, J.S., Ochsner, K.N. (2006). Social cognition: A multi level analysis. *Brain Research*, **1079**, 98–105.

Berridge, K.C., Kringelbach, M.L. (2015). Pleasure systems in the brain. *Neuron*, **86**, 646–64.

Bona, S., Herbert, A., Toneatto, C., Silvanto, J., Cattaneo, Z. (2014). The causal role of the lateral occipital complex in visual mirror symmetry detection and grouping: An fMRI-guided TMS study. *Cortex*, **51**, 46–55.

Bzdok, D., Langner, R., Hoffstaedter, F., Turetsky, B.I., Zilles, K., Eickhoff, S.B. (2012a). The modular neuroarchitecture of social judgments on faces. *Cerebral Cortex*, **22**, 951–61.

Bzdok, D., Schilbach, L., Vogeley, K., *et al.* (2012b). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, **217**, 783–96.

Campana, G., Cowey, A., Casco, C., Ousen, I., Walsh, V. (2007). Left frontal eye field remembers "where" but not "what". *Neuropsychologia*, **45**, 2340–5.

Campanella, F., D'Agostini, S., Skrap, M., Shallice, T. (2010). Naming manipulable objects: anatomy of a category specific effect in left temporal tumours. *Neuropsychologia*, **48**, 1583–97.

Carducci, F., Brusco, R. (2012). Accuracy of an individualized MR-based head model for navigated brain stimulation. *Psychiatry Research*, **203**, 105–8.

Cattaneo, Z., Bona, S., Silvanto, J. (2012). Cross-adaptation combined with TMS reveals a functional overlap between vision and imagery in the early visual cortex. *Neuroimage*, **59**, 3015–20.

Cattaneo, Z., Lega, C., Ferrari, C., *et al.* (2015). The role of the lateral occipital cortex in aesthetic appreciation of representational and abstract paintings: A TMS study. *Brain and Cognition*, **95**, 44–53.

Cattaneo, Z., Lega, C., Gardelli, C., Merabet, L.B., Cela-Conde, C.J., Nadal, M. (2014a). The role of prefrontal and parietal cortices in esthetic appreciation of representational and abstract art: a TMS study. *NeuroImage*, **99**, 443–50.

Cattaneo, Z., Lega, C., Flexas, A., Nadal, M., Munar, E., Cela-Conde, C.J. (2014b). The world can look better: enhancing beauty experience with brain stimulation. *Social Cognitive and Affective Neuroscience*, **9**, 1713–21.

Cattaneo, Z., Mattavelli, G., Platania, E., Papagno, C. (2011). The role of the prefrontal cortex in controlling gender-stereotypical associations: a TMS investigation. *NeuroImage*, **56**, 1839–46.

Cattaneo, Z., Rota, F., Vecchi, T., Silvanto, J. (2008). Using state-dependency of transcranial magnetic stimulation (TMS) to investigate letter selectivity in the left posterior parietal cortex: A comparison of TMS-priming and TMS-adaptation paradigms. *European Journal of Neuroscience*, **28**, 1924–9.

Cattaneo, Z., Silvanto, J. (2008). Time course of the state-dependent effect of transcranial magnetic stimulation in the TMS-adaptation paradigm. *Neuroscience Letters*, **443**, 82–5.

Chatterjee, A., Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New York Academy of Sciences*, **1369**, 172–94.

Chatterjee, A., Thomas, A., Smith, S.E., Aguirre, G.K. (2009). The neural response to facial attractiveness. *Neuropsychology*, **23**, 135–43.

Ciaramelli, E., Muccioli, M., Ladavas, E., di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, **2**, 84–92.

Contreras, J.M., Banaji, M.R., Mitchell, J.P. (2012). Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Social Cognitive and Affective Neuroscience*, **7**, 764–70.

Copland, D.A., de Zubicaray, G.I., McMahon, K., Eastburn, M. (2007). Neural correlates of semantic priming for ambiguous words: an event-related fMRI study. *Brain Research*, **1131**, 163–72.

Cupchik, G.C., Vartanian, O., Crawley, A., Mikulis, D.J. (2009). Viewing artworks: contributions of cognitive control and perceptual facilitation to aesthetic experience. *Brain and Cognition*, **70**, 84–91.

D'Argembeau, A., Ruby, P., Collette, F., *et al.* (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *Journal of Cognitive Neuroscience*, **19**, 935–44.

de Gelder, B., Hortensius, R., Tamietto, M. (2012). Attention and awareness each influence amygdala activity for dynamic bodily expressions-a short review. *Frontiers in Integrative Neuroscience*, **6**, 54.

Deschamps, I., Baum, S.R., Gracco, V.L. (2014). On the role of the supramarginal gyrus in phonological processing and verbal working memory: evidence from rTMS studies. *Neuropsychologia*, **53**, 39–46.

Devlin, J.T., Watkins, K.E. (2007). Stimulating language: insights from TMS. *Brain*, **130**, 610–22.

Dion, K., Berscheid, E., Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, **24**, 285–90.

Eagly, A.H., Ashmore, R.D., Makhijani, M.G., Longo, L.C. (1991). What is beautiful is good, but...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, **110**, 109–28.

Eagly, A.H., Gross, A.E., Crofton, C. (1977). What is good is beautiful. *Sociometry*, **40**, 85–90.

Englander, Z.A., Haidt, J., Morris, J.P. (2012). Neural basis of moral elevation demonstrated through inter-subject synchronization of cortical activity during free-viewing. *PloS One*, **7**, e39384.

Ferrari, C., Lega, C., Tamietto, M., Nadal, M., Cattaneo, Z. (2015). I find you more attractive... after (prefrontal cortex) stimulation. *Neuropsychologia*, **72**, 87–93.

Ferrari, C., Lega, C., Vernice, M., *et al.* (2016). The dorsomedial prefrontal dortex plays a causal role in integrating social impressions from faces and verbal descriptions. *Cerebral Cortex*, **26**, 156–65.

Fleck, M.S., Daselaar, S.M., Dobbins, I.G., Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex*, **16**, 1623–30.

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., Coricelli, G. (2013). Reputational priors magnify striatal responses to violations of trust. *Journal of Neuroscience*, **33**, 3602–11.

Frevert, T.K., Walker, L.S. (2014). Physical attractiveness and social status. *Sociology Compass*, **8**, 313–23.

Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: a review. *Cortex*, **36**, 539–59.

Gilbert, S.J., Swencionis, J.K., Amodio, D.M. (2012). Evaluative vs. trait representation in intergroup social judgments: Distinct roles of anterior temporal lobe and prefrontal cortex. *Neuropsychologia*, **50**, 3600–11.

Gozzi, M., Raymont, V., Solomon, J., Koenigs, M., Grafman, J. (2009). Dissociable effects of prefrontal and anterior temporal cortical lesions on stereotypical gender attitudes. *Neuropsychologia*, **47**, 2125–32.

Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, **44**, 389–400.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, **293**, 2105–8.

Greene, J., Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, **6**, 517–23.

Griffin, A.M., Langlois, J.H. (2006). Stereotype directionality and attractiveness stereotyping: Is beauty good or is ugly bad? *Social Cognition*, **24**, 187–206.

Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, **98**, 4259–64.

Hamermesh, D.S., Parker, A. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review*, **24**, 369–76.

Ito, T.A., Bartholow, B.D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, **13**, 524–31.

Jacobsen, T., Schubotz, R.I., Höfel, L., Cramon, D.Y.V. (2006). Brain correlates of aesthetic judgment of beauty. *Neuroimage*, **29**, 276–85.

Jenkins, A.C., Mitchell, J.P. (2011). Medial prefrontal cortex subserves diverse forms of self-reflection. *Social Neuroscience*, **6**, 211–8.

Jenkins, L.M., Andrewes, D.G., Nicholas, C.L., *et al.* (2014). Social cognition in patients following surgery to the prefrontal cortex. *Psychiatry Research: Neuroimaging*, **224**, 192–203.

Jeurissen, D., Sack, A.T., Roebroeck, A., Russ, B.E., Pascual-Leone, A. (2014). TMS affects moral judgment, showing the role of DLPFC and TPJ in cognitive and emotional processing. *Frontiers in Neuroscience*, **8**, 1–9.

Jones, B.C., Little, A.C., Feinberg, D.R., Penton-Voak, I.S., Tiddeman, B.P., Perrett, D.I. (2004). The relationship between shape symmetry and perceived skin condition in male facial attractiveness. *Evolution and Human Behavior*, **25**, 24–30.

Jussim, L. (1991). Social perception and social reality: A reflection–construction model. *Psychological Review*, **98**, 54–73.

Jussim, L. (1993). Accuracy in interpersonal expectations: A reflection–construction analysis of current and classic research. *Journal of Personality*, **61**, 637–68.

Kim, C., Johnson, N.F., Gold, B.T. (2014). Conflict adaptation in prefrontal cortex: now you see it, now you don't. *Cortex*, **50**, 76–85.

Kirsch, L.P., Urgesi, C., Cross, E.S. (2015). The Shaping and Reshaping of the Aesthetic Brain: Emerging Perspectives on the Neurobiology of Embodied Aesthetics. *Neuroscience & Biobehavioral Reviews*, **62**, 56–68.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, **314**, 829–32.

Knutson, K.M., Mah, L., Manly, C.F., Grafman, J. (2007). Neural correlates of automatic beliefs about gender and race. *Human Brain Mapping*, **28**, 915–30.

Komori, M., Kawamura, S., Ishihara, S. (2009). Averageness or symmetry: which is more important for facial attractiveness? *Acta Psychologica*, **131**, 136–42.

Krajbich, I., Adolphs, R., Tranel, D., Denburg, N.L., Camerer, C.F. (2009). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *Journal of Neuroscience*, **29**, 2188–92.

Kircher, T., Sass, K., Sachs, O., Krach, S. (2009). Priming words with pictures: Neural correlates of semantic associations in a cross-modal priming task using fMRI. *Human Brain Mapping*, **30**, 4116–28.

Kubota, J.T., Banaji, M.R., Phelps, E.A. (2012). The neuroscience of race. *Nature Neuroscience*, **15**, 940–8.

Langlois, J.H., Kalakanis, L., Rubenstein, A.J., Larson, A., Hallam, M., Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, **126**, 390–423.

Levy, D.J., Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, **22**, 1027–38.

Little, A.C., Burt, D.M., Perrett, D.I. (2006). What is good is beautiful: Face preference reflects desired personality. *Personality and Individual Differences*, **41**, 1107–18.

Little, A.C., Jones, B.C., DeBruine, L.M. (2008). Preferences for variation in masculinity in real male faces change across the menstrual cycle: Women prefer more masculine faces when they are more fertile. *Personality and Individual Differences*, **45**, 478–82.

Locher, P., Unger, R., Sociedade, P., Wahl, J. (1993). At first glance: Accessibility of the physical attractiveness stereotype. *Sex Roles*, **28**, 729–43.

Ma, N., Baetens, K., Vandekerckhove, M., Kestemont, J., Fias, W., Van Overwalle, F. (2013a). Traits are represented in the medial prefrontal cortex: an fMRI adaptation study. *Social Cognitive and Affective Neuroscience*, **2**, 123–38.

Ma, N., Baetens, K., Vandekerckhove, M., Van der Cruyssen, L., Van Overwalle, F. (2013b). Dissociation of a trait and a valence representation in the mPFC. *Social Cognitive and Affective Neuroscience*, **9**, 1506–14.

Ma, N., Vandekerckhove, M., Van Overwalle, F., Seurinck, R., Fias, W. (2011). Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: spontaneous inferences activate only its core areas. *Social Neuroscience*, **6**, 123–38.

Mattavelli, G., Cattaneo, Z., Papagno, C. (2011). Transcranial magnetic stimulation of medial prefrontal cortex modulates face expressions processing in a priming task. *Neuropsychologia*, **49**, 992–8.

Mende-Siedlecki, P., Said, C.P., Todorov, A. (2013). The social evaluation of faces: a meta-analysis of functional neuroimaging studies. *Social Cognitive and Affective Neuroscience*, **8**, 285–99.

Mitchell, J.P., Ames, D.L., Jenkins, A.C., Banaji, M.R. (2009). Neural correlates of stereotype application. *Journal of Cognitive Neuroscience*, **21**, 594–604.

Mitchell, J.P., Banaji, M.R., Macrae, C.N. (2005b). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage*, **28**, 757–62.

Mitchell, J.P., Cloutier, J., Banaji, M.R., Macrae, C.N. (2006). Medial prefrontal dissociations during processing of trait diagnostic and non-diagnostic person information. *Social Cognitive and Affective Neuroscience*, **1**, 49–55.

Mitchell, J.P., Heatherton, T.F., Macrae, C.N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, **99**, 15238–43.

Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2005a). Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. *Neuroimage*, **26**, 251–7.

Mitrovic, A., Tinio, P.P., Leder, H. (2016). Consequences of beauty: effects of rater sex and sexual orientation on the visual exploration and evaluation of attractiveness in real world scenes. *Frontiers in Human Neuroscience*, **10**, 122.

Nakamura, K., Kawashima, R., Nagumo, S., *et al.* (1998). Neuroanatomical correlates of the assessment of facial attractiveness. *Neuroreport*, **9**, 753–7.

Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, **9**, 97–113.

Olivola, C.Y., Todorov, A. (2010). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, **46**, 315–24.

Olson, I.R., Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, **5**, 498–502.

Öngür, D., Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, **10**, 206–19.

Oosterhof, N.N., Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, **105**, 11087–92.

Pegors, T.K., Kable, J.W., Chatterjee, A., Epstein, R.A. (2015). Common and unique representations in pFC for face and place attractiveness. *Journal of Cognitive Neuroscience*, **27**, 959–73.

Phelps, E.A., Lempert, K.M., Sokol-Hessner, P. (2014). Emotion and decision making: multiple modulatory neural circuits. *Annual Review of Neuroscience*, **37**, 263–87.

Piretti, L., Carnaghi, A., Campanella, F., Ambron, E., Skrap, M., Rumiati, R.I. (2015). The neural network associated with lexical-semantic knowledge about social groups. *Cortex*, **70**, 155–68.

Quadflieg, S., Flannigan, N., Waiter, G.D., *et al.* (2011). Stereotype-based modulation of person perception. *Neuroimage*, **57**, 549–57.

Quadflieg, S., Turk, D.J., Waiter, G.D., Mitchell, J.P., Jenkins, A.C., Macrae, C.N. (2009). Exploring the neural correlates of social stereotyping. *Journal of Cognitive Neuroscience*, **21**, 1560–70.

Robertson, E., Theoret, H., Pascual-Leone, A. (2003). Studies in cognition: the problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, **15**, 948–60.

Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A. (2011). Screening questionnaire before TMS: an update. *Clinical Neurophysiology*, **122**, 1686.

Ruff, C.C., Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, **15**, 549–62.

Ruff, C.C., Ugazio, G., Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, **342**, 482–4.

Said, C.P., Baron, S.G., Todorov, A. (2009). Nonlinear amygdala response to face trustworthiness: contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience*, **21**, 519–28.

Sellaro, R., Derks, B., Nitsche, M.A., *et al.* (2015). Reducing prejudice through brain stimulation. *Brain Stimulation*, **8**, 891–7.

Siebner, H.R., Hartwigsen, G., Kassuba, T., Rothwell, J.C. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. *Cortex*, **45**, 1035–42.

Sui, J., Liu, C.H. (2009). Can beauty be ignored? Effects of facial attractiveness on covert attention. *Psychonomic Bulletin & Review*, **16**, 276–81.

Sutherland, C.A., Oldmeadow, J.A., Santos, I.M., Towler, J., Burt, D.M., Young, A.W. (2013). Social inferences from faces: ambient images generate a three-dimensional model. *Cognition*, **127**, 105–18.

Talairach, J., Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme Medical, New York.

Tassy, S., Oullier, O., Duclos, Y., *et al.* (2011). Disrupting the right prefrontal cortex alters moral judgement. *Social Cognitive and Affective Neuroscience*, **7**, 282–8.

Tatarkiewicz, W. (1970a). Did Aesthetics Progress?. *Philosophy and Phenomenological Research*, **31**, 47–59.

Tatarkiewicz, W. (1970b). *History of Aesthetics (Vol. I. Ancient Aesthetics)*. Warsaw: Mouton.

Tsukiura, T., Cabeza, R. (2011). Shared brain activity for aesthetic and moral judgments: implications for the Beauty-is-Good stereotype. *Social Cognitive and Affective Neuroscience*, **6**, 138–48.

van de Ven, V., Jacobs, C., Sack, A.T. (2012). Topographic contribution of early visual cortex to short-term memory consolidation: a transcranial magnetic stimulation study. *Journal of Neuroscience*, **32**, 4–11.

Van den Stock, J., Tamietto, M., Zhan, M., *et al.* (2014). Neural correlates of body and face perception following bilateral destruction of the primary visual cortices. *Frontiers in Behavioral Neuroscience*, **8**, 30.

Van der Cruyssen, L., Heleven, E., Ma, N., Vandekerckhove, M., Van Overwalle, F. (2015). Distinct neural correlates of social categories and personality traits. *NeuroImage*, **104**, 336–46.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, **30**, 829–58.

Van Overwalle, F., Ma, N., Baetens, K. (2015). Nice or nerdy? The neural representation of social and competence traits. *Social Neuroscience*, **11**, 567–78.

Vessel, E., Stahl, J., Purton, I., Starr, G. (2015). Domain-general representation of visual aesthetic appreciation in the medial prefrontal cortex. *Journal of Vision*, **15**, 124-124.

Wang, T., Mo, L., Mo, C., *et al.* (2014). Is moral beauty different from facial beauty? Evidence from an fMRI study. *Social Cognitive and Affective Neuroscience*, **10**, 814–23.

Wang, Y., Hamilton, A.F.D.C. (2015). Anterior medial prefrontal cortex implements social priming of mimicry. *Social Cognitive and Affective Neuroscience*, **10**, 486–93.

Watabe, M., Ban, H., Yamamoto, H. (2011). Judgments about others' trustworthiness: An fMRI study. *Letters on Evolutionary Behavioral Science*, **2**, 28–32.

Winston, J.S., O'Doherty, J., Kilner, J.M., Perrett, D.I., Dolan, R.J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, **45**, 195–206.

Yoder, K.J., Decety, J. (2014). Spatiotemporal neural dynamics of moral judgment: A high-density ERP study. *Neuropsychologia*, **60**, 39–45.

Zaidel, D.W., Nadal, M. (2011). Brain intersections of aesthetics and morals: perspectives from biology, neuroscience, and evolution. *Perspectives in Biology and Medicine*, **54**, 367–80.

Zaki, J., López, G., Mitchell, J.P. (2014). Activity in ventromedial prefrontal cortex co-varies with revealed social preferences: evidence for person-invariant value. *Social Cognitive and Affective Neuroscience*, **9**, 464–9.

Zhang, Y., Kong, F., Zhong, Y., Kou, H. (2014). Personality manipulations: Do they modulate facial attractiveness ratings? *Personality and Individual Differences*, **70**, 80–4.