Psychology of Aesthetics, Creativity, and the Arts

1931-3896/13/$12.00 DOI: 10.1037/a0031591

Middle Temporal Gyrus Encodes Individual Differences in Perceived Facial Attractiveness

AQ: au

Oshin Vartanian
York University and University of Toronto–Scarborough

Vinod Goel
York University and University of Hull

Elaine Lam
York University

Maryanne Fisher
St. Mary's University

Josipa Granic
York University

Given the far-reaching implications of facial attractiveness for human behavior, its neural correlates have been the focus of much recent interest. However, whereas the focus of previous studies has been on highlighting a common network that underlies attractiveness judgments for all participants, we were also interested in examining individual differences in perceived facial attractiveness. We scanned 29 participants with functional MRI as they evaluated the attractiveness of faces. Activation in left anterior frontal cortex and right middle occipital gyrus covaried as a function of attractiveness ratings, which we attribute to making evaluative judgments involving the rewarding properties of faces. In addition, comparing participants who on average gave higher versus lower ratings to faces revealed activation in right middle temporal gyrus exclusively. We suggest that the activation in middle temporal gyrus reflects an aspect of individual differences in perceived facial attractiveness, possibly driven by the convergence of information from a variety of sources that extend beyond the domain of faces exclusively.

Keywords: attractiveness, beauty, faces, mate selection, reward

Physical attractiveness is a key interpersonal characteristic, with consequences in a variety of domains. For example, it has been shown to play an important role in the formation of interpersonal evaluations, in turn influencing evaluations of other characteristics such as personality, marital satisfaction, and employment success (Berscheid & Walster, 1974; Dion, Berscheid, & Walster, 1972). From an evolutionary perspective, attractiveness is significant because it may signal fertility, gene quality, and health (Buss, 1989; Dixson, Halliwell, East, Wignarajah, & Anderson, 2003; Shackelford & Larsen, 1999; Thornhill & Gangestad, 1993), as reflected in the preference both sexes express for attractive people as partners (Berscheid, Dion, Walster, & Walster, 1971; Berscheid & Walster, 1974; see also Fletcher, Simpson, Thomas, & Giles, 1999; Hatfield, Aronson, Abrahams & Rottman, 1966; Stelzer, Desmond, & Price, 1987; Urbaniak & Kilmann, 2003). AQ: 1

Given its importance, researchers have been investigating the cortical systems that underlie evaluations of facial attractiveness. Nakamura et al. (1998) demonstrated that rating faces on attractiveness increased regional cerebral blood flow in a network including the frontal cortex, the fronto-temporal junction, the orbitofrontal cortex (OFC), the caudate nucleus, and the visual cortex. They argued that the involvement of these regions is a specific example of their more general involvement in evaluative judgments involving an affective component. In other words, attractive faces were considered rewarding stimuli that elicit emotional responses. Subsequent functional MRI (fMRI) studies have essentially confirmed this interpretation, linking attractiveness judgments to a number of structures implicated in evaluative judgment or reward processing, such as the nucleus accumbens (Aharon et al., 2001), the dorsal striatum (Kampe, Frith, Dolan, & Frith, 2001), and the OFC (O'Doherty et al., 2003; Winston, O'Doherty,

Oshin Vartanian, Department of Psychology, York University, Toronto, Ontario, Canada; and Department of Psychology, University of Toronto–Scarborough, Toronto, Ontario, Canada; Vinod Goel, Department of Psychology, York University; and Department of Psychology, University of Hull, Hull, United Kingdom; Elaine Lam, Department of Psychology, York University; Maryanne Fisher, Department of Psychology, St. Mary's University, Halifax, Nova Scotia, Canada; Josipa Granic, Department of Psychology, York University.

This research was supported by grants from the National Sciences and Engineering Research Council and the Canadian Institutes of Health Research and by a Premier's Research Excellence Award to Vinod Goel. We thank Kevin S. Weiner for the anatomical localization of the middle temporal gyrus.

Correspondence concerning this article should be addressed to Oshin Vartanian, Department of Psychology, University of Toronto–Scarborough, 1265 Military Trail, Toronto, ON, Canada, M1C 1A4. E-mail: oshinv1@mac.com

Kilner, Perrett, & Dolan, 2007). In fact, it has been suggested that a system involving the orbitofrontal and striatal neurons may underlie valuation of rewards irrespective of the modality giving rise to the rewarding stimuli (Montague & Berns, 2002).

However, much of the work in this area is characterized by one of two themes. First, there is a strong tendency to focus on how variations in features of the faces (e.g., gaze direction, symmetry, averageness, etc.) affect attractiveness ratings or brain activations. For example, Kampe et al. (2001) reported that there was a positive correlation between attractiveness ratings and activation in the dorsal striatum when the gaze of the face was directed at the participant. In contrast, when the gaze of the face was averted, the correlation reversed to negative. This emphasis on how variations in features of faces affect judgment and its neural correlates treats the former as sources of information about the reproductive fitness (i.e., mate value) of the target (Thornhill & Gangestad, 1999). Thus, the human brain is considered to have undergone evolutionary adaptations enabling it to solve the problem of mate selection using those relevant cues. Furthermore, it is assumed that variations in the relevant cues trigger certain responses essentially *universally* among raters. Indeed, such signals would be ineffective if they did not elicit the desired cognitive or emotional response in most members of the opposite sex (Johnston, 2006). AQ: 2

We argue that focusing exclusively on such shared responses among raters may miss the unique or private component of subjective responses to attractiveness. For example, a literature review by Thornhill and Gangestad (1999) demonstrated that facial attractiveness judgments between two raters correlate in the 0.3–0.5 range, accounting for only 9%–25% of the observed variance in ratings. This finding suggests that there are individual differences in what people perceive as attractive, acting against a backdrop of shared preferences. Indeed, Hönekopp (2006) determined that shared and private taste contribute approximately equally (≈50%) to the variance in attractiveness ratings, where s*hared taste* was defined as encompassing all attractiveness standards that enable two judges to agree about the attractiveness of faces, and *private taste* as encompassing all attractiveness standards of a single judge that give rise to replicable disagreement with shared taste.

Individual differences in attractiveness judgments are important because they can have a direct bearing on behavior—ultimately the variable of most interest to behavioral scientists. Hönekopp (2006) asks us to imagine a situation where the average attractiveness rating given by a group of female judges to the faces of John, Paul, George, and Ringo are 1, 2, 3, and 4, respectively (where 1 reflects *extremely unattractive* and 9 reflects *extremely attractive*). Now suppose that Alexandra rated them 1, 2, 3 and 4, respectively, whereas Natasha's ratings were 6, 7, 8 and 9, respectively. One could argue that the absolute differences between Alexandra's and Natasha's ratings are unimportant because they reflect trivial differences in scale use and the same preference order (i.e., ranking) is preserved. For example, both preferred Paul over John. However, one could also argue that the differences do matter because they could culminate in differences in behavior. For example, whereas Natasha may opt to date John (because she rated him 6 out of 9 and therefore above average), Alexandra may not (because she rated him 1 out of 9 and therefore by definition extremely unattractive).

In fact, it has recently been argued that such individual differences in scale use are not trivial and reflect an important aspect of preference (Hönekopp, 2006). Furthermore, the average attractiveness rating generated by each participant can be used as a proxy measure to determine an internal reference point based on which judgments are made. In other words, a participant who on average gives higher ratings (than average) to a group of faces may have a different internal metric for what constitutes an attractive face than does a participant who on average gives lower ratings to the same group of faces. Thus, differences in average ratings may reflect fine-tuning or parameter setting relative to a group of stimuli.[1] Fn1

The second theme that characterizes much of the research on facial attractiveness is an exclusive focus on targets located in the age range of maximal reproductive fitness. Typically, the studies rely upon heterosexual participants viewing members of the opposite sex. This is not surprising, given that most studies have been motivated by evolutionary hypotheses geared toward mate selection. However, the use of the term "attractive" is not limited to our assessments of people exclusively in their reproductive prime. The term is also used to characterize elderly persons, as well as children. For example, attractive elderly individuals and children are frequently used in advertising to market products. Hence, to reconcile this observation with the evolutionary argument that those in their reproductive prime are maximally attractive, we argue that judgments of attractiveness are triggered by different cues (and related strategies) as a function of the age of the target. For example, attractiveness may tap "cuteness" when assessing babies, "reproductive fitness" when assessing young adults, and "character" or "worldliness" when assessing older adults. This variability also implies that the ability to find faces in various age ranges attractive may be indicative of one's flexibility to respond appropriately to different cues. Therefore, an additional feature of our study was to extend the age of target faces from infancy to old adulthood, in order to highlight a system that responds to attractiveness across the age range, rather than exclusively the age range of maximal reproductive fitness. Although there is a strong possibility that differential responsiveness to various cues of attractiveness across the age range may be built on distinct neural systems (and associated psychological metrics), we were interested in exploring the possibility of a *common* evaluative system operating across the age range.

We tested two neural hypotheses. First, we predicted that much as in earlier studies, a distributed network involving structures shown to underlie evaluative judgments of rewarding stimuli would underlie judgments of attractiveness, despite the extension of the age range from infancy to old adulthood. This network would represent the shared neural system responsive to attractiveness across participants. Furthermore, given conflicting results in the literature about sex differences in the neural system underlying attractiveness ratings (Cloutier, Heatherton, Whalen, & Kelley, 2008; Ishai, 2007; Kranz & Ishai, 2006), we also explored the possibility that there may be sex differences in the activation of this system while making judgments.

Second, given that people have different internal metrics for facial beauty (see Hönekopp, 2006), we explored the possibility that certain cortical structures might be sensitive to individual

[1] There are of course other methods for exploring individual differences in scale use. For example, the variance of ratings might correspond to sensitivity to variations in attractiveness.

differences in rating faces on attractiveness. Specifically, we predicted that activations in structures that encode individual differences would vary as a function of deviations from the average ratings given to the faces—revealed by comparing brain activation between participants who on average give higher versus lower ratings to the same faces. This would reflect an aspect of individual differences in perceived facial attractiveness related to the internal reference point based on which attractiveness judgments are made.

AQ: 3

Method

To test the aforementioned two hypotheses, we conducted two separate studies. First, we conducted a behavioral study to evaluate the relevant features of our stimulus set. Second, we conducted an fMRI study to test our neural hypotheses.

Experiment 1

Participants

The sample consisted of 30 men (M = 21.4, SD = 1.48) and 30 women (M = 20.9, SD = 1.35), recruited from the university community of southern Ontario.

Materials

Previous investigations have used stimulus sets where faces were prerated as high or low on attractiveness (e.g., Aharon et al., 2001). In addition, when the average age of the stimuli (faces) was provided, it tended to be youthful (e.g., mean age of 24 years, Nakamura et al., 1998; and between 20 and 35 years, O'Doherty et al., 2003). Our stimuli were composed of a wide assortment of faces, varying in terms of age and attractiveness. This design feature enabled us to examine attractiveness beyond the age range of maximal reproductive fitness, as well as adding a measure of ecological validity by not focusing exclusively on extremely attractive or unattractive faces. The stimuli consisted of 110 black-and-white photographs of male (55) and female (55) faces (see Figure 1). The models were a convenience sample from the Toronto community who ranged in age from children (1-year-old) to the elderly (82-year-old), M = 29.56, SD = 22.10 (note the large standard deviation). Specifically, there were five male and five female models in each of 11 age categories (0–2, 3–6, 7–11, 12–16, 17–22, 23–27, 28–35, 36–45, 46–55, 56–65, and >65) (see Berk, 2000). Four male and four female models in each age category were Caucasian, and one male and one female model were non-Caucasians of the same ethnicity (based on self-report). Models looked straight into the camera, with an upright head facing forward and with eye gaze directed toward the viewer, and displayed a neutral emotional expression. Their photographs, initially captured in color, were converted to grayscale, shadows were removed, and backgrounds were converted into white using Photoshop. Furthermore, the resolutions were equalized at 150 dpi, and the canvas size was converted to 5.03 × 3.74 in.

AQ: 4

F1

Procedure. The participants were instructed to rate each face on attractiveness, using a 6-point scale ranging from *extremely unattractive* (0) to *extremely attractive* (5). All ratings were collected on an individual basis.

Results. The results of our study demonstrated that overall, female faces (M = 2.93, SD = .45) were rated as more attractive than male faces (M = 2.65, SD = .46). Male participants rated female faces (M = 2.87, SD = .43) as more attractive than male faces (M = 2.56, SD = .44), $t(29) = 7.09$, $p < .001$, and the same was true for female participants who also rated female faces (M = 2.98, SD = .46) as more attractive than male faces (M = 2.72, SD = .47), $t(29) = 6.83$, $p < .001$. There was no main effect for rater, and no interaction between sex of rater and sex of target face.

For female faces, the results demonstrated a main effect for age range (11 levels), $F(10, 580) = 158.65$, $p < .001$. As expected, attractiveness ratings decreased as a function of increasing age. The main effect for sex of rater was not significant. There was also a significant interaction between sex of rater and age, $F(10, 580) = 4.73$, $p < .001$. To probe the interaction further, we investigated the simple main effect of sex of rater at each age range. The results demonstrated that female participants gave significantly higher

Figure 1. Sample stimuli (faces) from the study. The stimuli represent a sample of female and male faces around 5, 20, 30, 40, and 50 years of age (±2 years) used in the study (see Experiments 1 and 2, Materials and Procedures for details). All participants granted written permission to have their likenesses reprinted.

ratings than male participants to female faces in the 0–2 and 3–6 levels ($p < .05$). In addition, there was a trend toward higher ratings given by male participants than female participants to female faces in the 17–22 age level, $p < .10$. The difference between male and female participants was not significant in the remaining eight age levels.

For male faces, the results demonstrated a main effect for age range (11 levels), $F(10, 580) = 74.67$, $p < .001$. As expected, attractiveness ratings decreased as a function of increasing age. The main effect for sex of rater was not significant. There was also a significant interaction between sex of rater and age, $F(10, 580) = 6.62$, $p < .001$. To probe the interaction further, we investigated the simple main effect of sex of rater at each age range. Again, the results demonstrated that female participants gave significantly higher ratings than male participants to male faces in the 0–2 and 3–6 levels ($p < .05$). The difference between male and female participants was not significant in any of the remaining nine age levels.

Conclusion

The results of our first study demonstrated that (a) for both male and female faces, perceived attractiveness decreased as a function of increasing age, (b) male and female raters assigned higher attractiveness ratings to female faces, (c) female raters assigned higher attractiveness ratings to male and female faces in the 0–2 and 3–6 age levels, and (d) there was a trend toward higher ratings assigned by male participants to female faces in the 17–22 age level.

Experiment 2

Participants

Twenty-nine right-handed participants (14 female and 15 male) with no history of neurological or psychiatric disorders participated in this study. Participants were recruited from the university community of southern Ontario. The mean age of the sample was 25.1 years ($SD = 5.4$) and the mean education level was 17.1 years ($SD = 2.3$). Participants were financially compensated for their travel and time.

Materials

The stimulus set was identical to the one used in Experiment 1.

Procedure

In the fMRI scanner the ratings were obtained based on a 7-point rating scale ranging from *extremely unattractive* (0) to *extremely attractive* (6). We added an additional level to the rating scale to increase sensitivity. A Kolmogorov–Smirnov test demonstrated that the age distribution of faces did not vary from normality, $z = 1.25$, *ns.* The faces were presented in an event-related design. There were no gaps in presentation between successive stimuli. There were no practice trials. Each face was presented for 4 s.

The specifications for the fMRI scanning and analysis were as follows. A 4-Tesla Oxford Magnet Technologies magnet with a Siemens Sonata gradient coil was used to acquire T1 anatomical volume images (1 × 1 × 2-mm voxels) and 22 $T2^*$-weighted interleaved multishot contiguous echoplanar images (3 × 3 × 5-mm voxels), sensitive to blood oxygenation level-dependent (BOLD) contrast. The images were acquired axially and positioned to cover the whole brain. A total of 146 volumes were recorded during a single session, acquired with a repetition time (TR) of 3.0 s/vol. The first six volumes were discarded to allow for T1 equilibration effects (leaving 140 volumes for analysis). The stimuli were presented to the participants using an LCD projector (NEC MultiSync MT800) with a video resolution of 640 × 480 pixels and a light output of 370 lumens.

Data were analyzed using Statistical Parametric Mapping (SPM2) (Friston et al., 1994). All functional volumes were spatially realigned to the first volume. Head movement was less than 2 mm in all cases. A mean image created from the realigned volumes was spatially normalized to the Montreal Neurological Institute EPI brain template (Evans et al., 1993), using nonlinear basis functions (Ashburner & Friston, 1999). The derived spatial transformation was then applied to the realigned $T2^*$ volumes, which were finally spatially smoothed with a 12-mm full width at half maximum isotropic Gaussian kernel in order to make comparisons across participants and to permit application of random field theory for corrected statistical inference (Worsley & Friston, 1995). The resulting time series across each voxel were high-pass filtered with a cutoff of 128 s, using cosine functions to remove section-specific low-frequency drifts in the BOLD signal. Global means were normalized by proportional scaling to a Grand Mean of 100.

We conducted whole brain analyses. Condition effects at each voxel were estimated according to the general linear model, and regionally specific effects were compared using linear contrasts. Each contrast produced a statistical parametric map of the t-statistic for each voxel, which was subsequently transformed to a unit normal z-distribution. For the parametric analysis, the activations reported survived a voxel-level intensity threshold of $p < .05$ using a random effects model, corrected for multiple comparisons using the False Discovery Rate (Genovese, Lazar, & Nichols, 2002) and a minimum of 40 contiguous voxels. The BOLD signal was modeled as a canonical hemodynamic response function. In addition, motor response was entered into the analysis but modeled as an event of no interest (see below, Functional MRI results). Given the exploratory nature of our individual-differences hypothesis involving the contrast between participants who on average give higher versus lower ratings to faces, we used a more liberal threshold by reporting any activation that survived a voxel-level intensity threshold of $p < .005$ (uncorrected for multiple comparisons) using a random-effects model.

Results

Behavioral results. As expected, although Cronbach's alpha for the ratings was high (.97), average interrater correlation was modest (.42). This pattern indicates that although agreement about relative ranking of faces is high among raters, attractiveness ratings are also influenced by individual differences among raters (see Hönekopp, 2006). We ran a mixed-model ANOVA, with the sex of the rater as the between-subjects variable, sex of the stimulus (i.e., face) as the within-subjects variable, and attractive-

ness ratings as the dependent variable. On average, male (M = 2.66, SD = .37) and female (M = 2.56, SD = .48) participants did not differ in the attractiveness ratings they assigned to the faces, T1 $F(1, 27) = .11$, *ns* (see Table 1). However, there was a main effect for sex of stimuli such that participants rated female faces (M = 2.76, SD = .41) as significantly more attractive than male faces (M = 2.46, SD = .45), $F(1, 27) = 41.68$, $p < .001$. There was no interaction between sex of rater and sex of stimulus.

Average response latency across all stimuli was 1912 ms (SD = 620). The correlation between response latency and attractiveness ratings was not significant, $r = .14$, *ns*. We investigated the possibility that participants might have spent a longer duration on making the potentially more difficult intermediate judgments, by running a one-way ANOVA with rating as a three-level factor (0–1 range, 2–3 range, 4–6 range) and reaction time (RT) as the Fn2 dependent variable.[2] The results demonstrated that rating had no effect on RT, $F(2, 52) = 2.40$, *ns*. Then, we ran a mixed-model ANOVA, again with the sex of the rater as the between-subjects variable, sex of the stimulus as the within-subjects variable, and RT as the dependent variable. None of the main effects or the interaction reached significance.

The correlation between attractiveness ratings and the age of target faces was significant, $r(108) = -.66$, $p < .001$. As expected, confirming the results of our first study reported above, attractiveness ratings decreased as a function of increasing age. There was no difference in the magnitude of this negative correlation between male ($r = -.61$) and female ($r = -.67$) participants, indicating that both groups found younger faces to be more F2 attractive (see Figure 2).

However, follow-up analyses revealed distinct differences between male and female participants in rating patterns as a function F3 of the sex of the target face (see Figure 3). Specifically, for female participants, a linear trend accounted for approximately 37% of the observed variance in the attractiveness ratings assigned to male faces across the age range and approximately 66% of the observed variance in the attractiveness ratings assigned to female faces across the age range. In contrast, for male participants, although a linear trend accounted for approximately 45% of the observed variance in the attractiveness ratings assigned to male faces across the age range, a cubic trend accounted for approximately 75% of the observed variance in the attractiveness ratings assigned to female faces across the age range. These results demonstrate that when it comes to perceiving attractiveness in the faces of the *opposite* sex, male and female participants assign higher ratings to different segments of the age range.

The average attractiveness rating across all participants was 2.66 (SD = .43). We computed a median split, discarding the sole participant whose average score for all faces fell exactly at the median (*Median* = 2.74). Thus, two groups of 14 participants who were placed in the "high" and "low" rating groups were created. There was no difference in the distribution of male and female participants between the two groups, $\chi^2(1) = 2.29$, *ns*. As expected, on average, the 14 participants who were placed in the "high" group generated significantly higher ratings (M = 3, SD = .13) than the 14 participants who were placed in the "low" group (M = 2.38, SD = .35), $t(26) = 6.09$, $p < .001$.

Table 1
Average Attractiveness Ratings and Response Latencies for Each Rater and Target (Face) Condition

	Male faces	Female faces
Male raters	2.51 (0.93)	2.81 (1.06)
Respone latencies (ms)	1849	1829
Female raters	2.44 (1.13)	2.70 (1.12)
Respone latencies (ms)	2001	1972

Note. Standard deviations appear in parentheses.

fMRI results. We conducted three sets of analyses to test our predictions. First, participants' attractiveness ratings were analyzed using a parametric analysis of fMRI data. Specifically, presentations of faces were treated as events (coupled with attractiveness ratings as the parameter of interest), and the motor response was entered as an event of no interest. Consistent with previous studies, the results revealed that activity in the left anterior frontal cortex (BA 10) (−36, 58, 8, $z = 4.26$) and right middle occipital gyrus (BA 19) (10, −82, 30, $z = 3.71$) covaried as a function of attractiveness ratings (see Figure 4). Second, given the F4,AQ:5 observed behavioral difference in rating members of the opposite sex, we investigated but failed to find any difference in the neural correlates of this parametric response for male and female participants rating faces of the opposite sex.

Third, we conducted a direct contrast between the 14 participants who on average gave higher ratings to the faces and the 14 participants who on average gave lower ratings to the faces. This contrast revealed significantly higher activation exclusively in right middle temporal gyrus (MTG) (54, −50, −8, $z = 2.75$) in the former group (see Figure 5). The reverse contrast did not reveal F5 any significant area of activation.

General Discussion

Experiment 2 was conducted to determine the neural correlates of evaluating facial attractiveness. Importantly, the key behavioral results of Experiment 2 replicated all the results from Experiment 1: (a) perceived attractiveness decreased as a function of increasing age, (b) female faces were assigned higher attractiveness ratings, (c) female raters assigned higher attractiveness ratings to male and female faces in the 0–2 and 3–6 age levels, and (d) male participants assigned higher ratings to female faces in the age range of maximal reproductive fitness. This suggests that the behavioral results observed in Experiment 1 and Experiment 2 were reliable.

Our results indicated that activation in left anterior frontal cortex (BA 10) covaried parametrically with facial attractiveness ratings. This activation (−36, 58, 8) corresponds closely to the activation reported by Nakamura et al. (1998) in which male participants rated female facial attractiveness (−37, 50, 2). We argue that this activation is linked to the act of making self-referential evaluative judgments, akin to what Chatterjee has referred to as the decision step of aesthetic judgment (Chatterjee, 2003; Chatterjee, Thomas, Smith, & Aguirre, 2009). For example, a number of studies have linked the anterior frontomedian cortex (BA 10/9) to engagement in evaluative judgments involving faces (O'Doherty et al., 2003), abstract geometric figures (Jacobsen, Schubotz, Höfel, & von

[2] Only three responses in the entire data set were 6 s (.1%), explaining why the highest level includes a range of three ratings (i.e., 4–6).

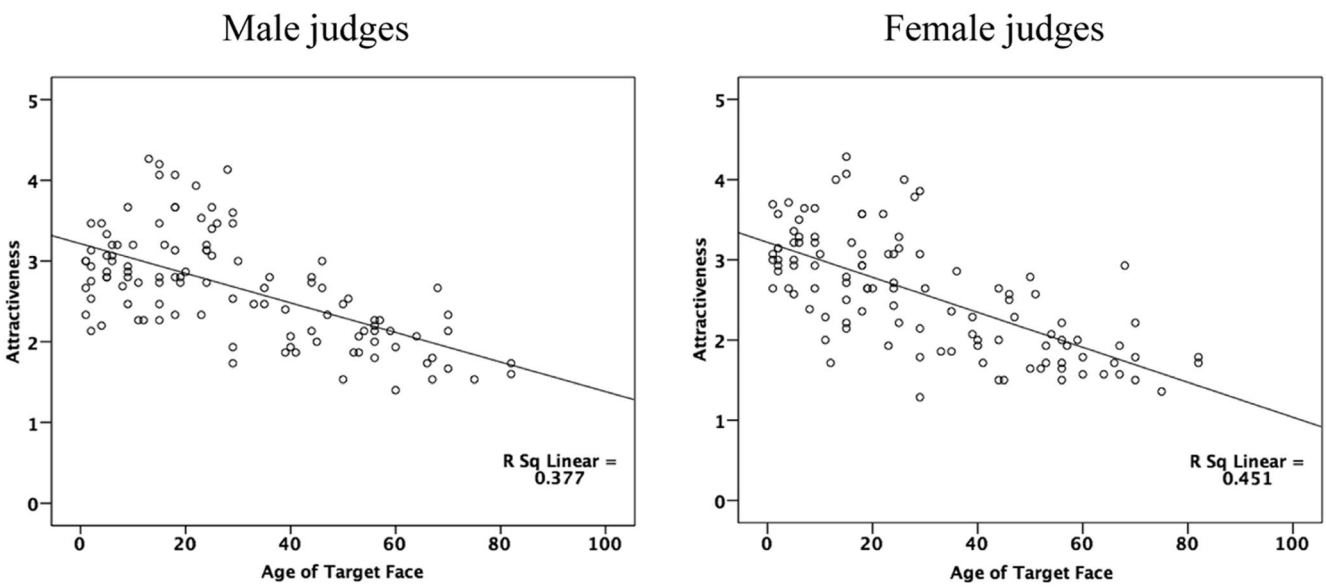

Figure 2. The effect of age of target (face) on attractiveness ratings.

Cramon, 2006), and verbal statements of the form "I like George W. Bush: yes/no" (Zysset et al., 2002). Although the activation observed in our study is somewhat more lateral compared with those reported by the aforementioned studies, our interpretation is nevertheless consistent with the broader role of BA 10 in self-referential evaluative judgment (Burgess, Dumontheil, & Gilbert, 2007; Christoff, Ream, Geddes, & Gabrieli, 2003; Cupchik, Vartanian, Crawly, & Mikulis, 2009; Northoff et al., 2006).

AQ: 6

Our results also indicated that activation in right middle occipital gyrus (BA 19) covaried parametrically with attractiveness ratings for faces, which we attribute to processing faces that vary in reward properties. Attractive faces are rewarding and can motivate behavior (Quinsey, Ketsetzis, Earls, & Karamanoukian, 1996). For example, people seek interactions and try to form relationships with individuals they consider attractive (Berscheid & Walster, 1974; Berscheid et al., 1971; Buss, 1989). Thus, attractiveness may be considered a reward that motivates or elicits adaptive behaviors, such as the selection of healthy and fertile mates. In support of this inference, the right middle occipital gyrus (BA 19) has been activated in studies that involved viewing or rating visual stimuli varying in emotional or rewarding properties such as faces, paintings, and stimuli from the International Affective Picture System (Dolan et al., 1996; Nakamura et al., 1998; Paradiso et al., 1999; Vartanian & Goel, 2004).[3]

Fn3

Behaviorally, we observed similar patterns for male and female participants when rating faces of the same sex (see Figure 3). Specifically, participants gave the highest ratings to younger faces of the same sex. Konrad Lorenz (1942) argued that infantile features may trigger "innate releasing mechanisms" in the viewer. In turn, such mechanisms serve to motivate caring behaviors, enhancing reproductive success when the infant and the viewer are related (Darwin, 1872). It is possible that when rating faces of the same sex on attractiveness, both sexes respond strongly to their "babyishness" (Lorenz, 1942). However, female subjects assigned higher ratings than male subjects to younger faces of the same sex, consistent with evidence indicating that females respond more strongly to infants than do males, as measured by pupil size (Hess & Polt, 1960).

In contrast, we observed certain sex differences when rating faces of the opposite sex (see Figure 3). Specifically, whereas female subjects continued to assign higher ratings to younger male faces, males assigned the highest ratings to female faces centered around 20 years of age, coupled with a dramatic drop in ratings following the mid-thirties. These differences could reflect differences in rating strategy. It is known that for men, attractiveness of potential mates is linked strongly to reproductive capacity (see Buss, 1989). As such, it follows that men would assign the highest attractiveness ratings to women in the age range associated most strongly with reproductive fitness and assign lower ratings beginning in the mid-thirties when reproductive fitness is perceived as starting to decline. This result also implies that when viewing female faces, men may not be responsive to cues that define attractiveness in children and in elderly persons, such as cuteness or character, respectively. Our results are also consistent with evidence suggesting that females' evaluations of male attractiveness is partially linked to his ability to provide resources (see Buss, 1989), a capacity that should be operative in older males. This explains the fact that older male faces continued to receive ratings within one standard deviation of the rating that females assigned to the average-aged male face ($M = 2.44$, $SD = 1.13$).

At the neural level, no corresponding sex difference was observed while rating faces of the opposite sex. This finding is

[3] Contrary to expectation, we did not observe any activation in the striatum or the OFC, two structures frequently activated in relation to observing rewarding stimuli, including attractive faces. This is likely due to the large signal dropout in OFC in the 4 T scanner, verified after scanning. Data acquisition at higher magnet strengths is frequently plagued by this problem, which may also explain why OFC has not been activated in several prior studies of facial attractiveness.

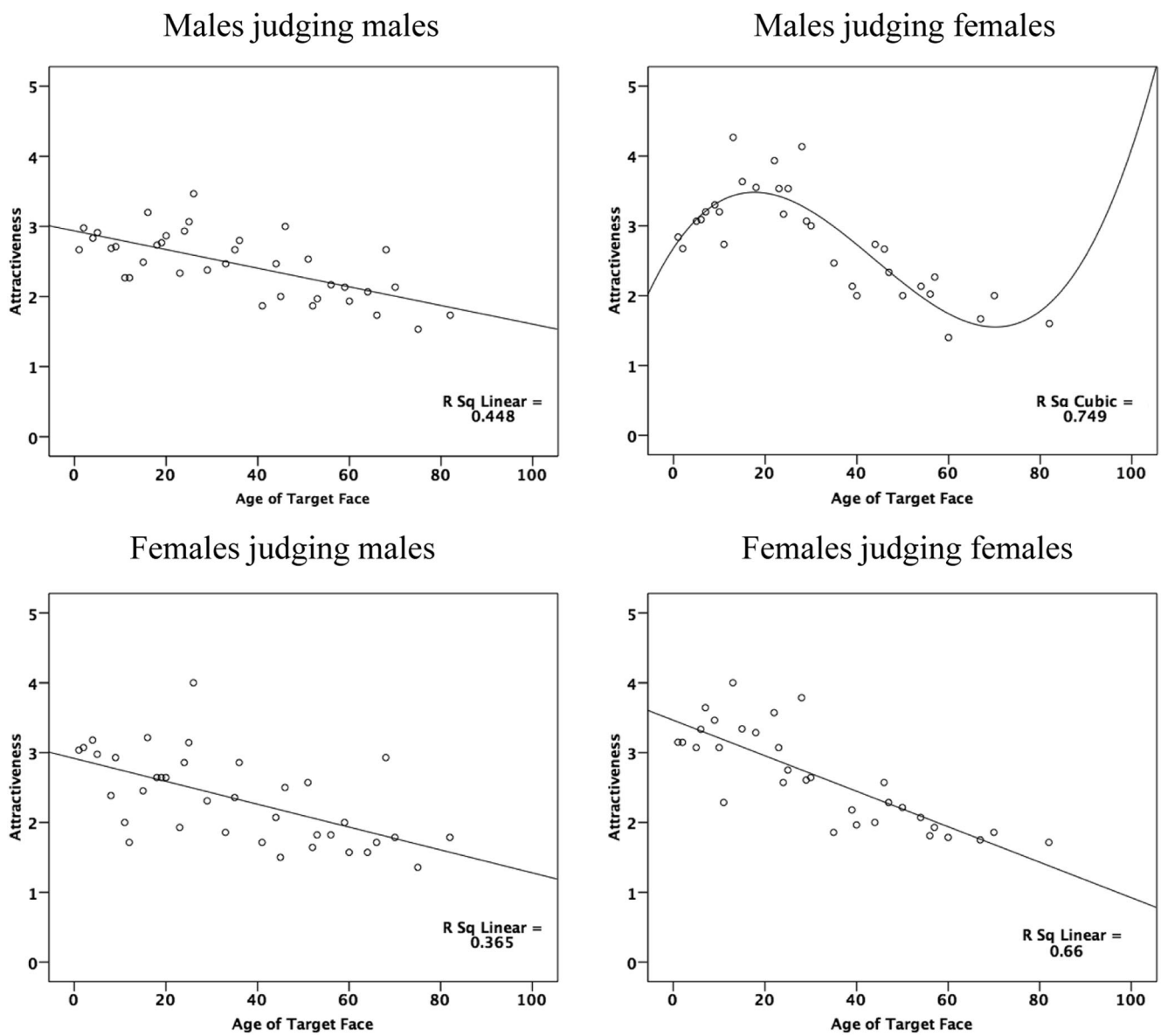

Figure 3. Differences between male and female participants in rating faces of the same and opposite sexes.

similar to the results reported by Kranz and Ishai (2006; see also Ishai, 2007), but contrary to those reported by Cloutier et al. (2008). In the latter study where sex differences were observed, participants exclusively viewed faces of the opposite sex. This may have contributed to the ability to detect differences; the focus on opposite sex faces may have facilitated the expectation to engage in strategies deployed for viewing the opposite sex, without the requirement for switching between strategies. Our results suggest that despite responding to different cues (e.g., reproductive fitness, etc.) when judging attractiveness in the opposite sex (see Figure 3), male and female subjects may nevertheless rely on the same neural system for computing it.

However, the novel contribution of the current study to the facial attractiveness literature involves our focus on the contribution of individual differences to attractiveness judgment. Our exploratory analysis revealed higher activation in MTG in participants who on average gave higher ratings to the faces (see Figure 5). Interestingly, MTG does not form one of the three main cortical regions involved in face perception in humans, namely the fusiform face area, the superior temporal sulcus, and the occipital face area (Kanwisher & Yovel, 2009). This suggests that its involvement here is likely not related to the perception of the physical features of the faces per se. Rather, recent high-resolution fMRI scans of the occipitotemporal cortex have suggested a new functional organization wherein MTG appears to play an important role in the convergence of information across modalities. Specifically, the available evidence implicates MTG in visual, haptic, tactile, and motor processing, with potential roles in language and social communication as well (Weiner & Grill-Spector, 2011). Its activation here suggests that the judgment of facial attractiveness

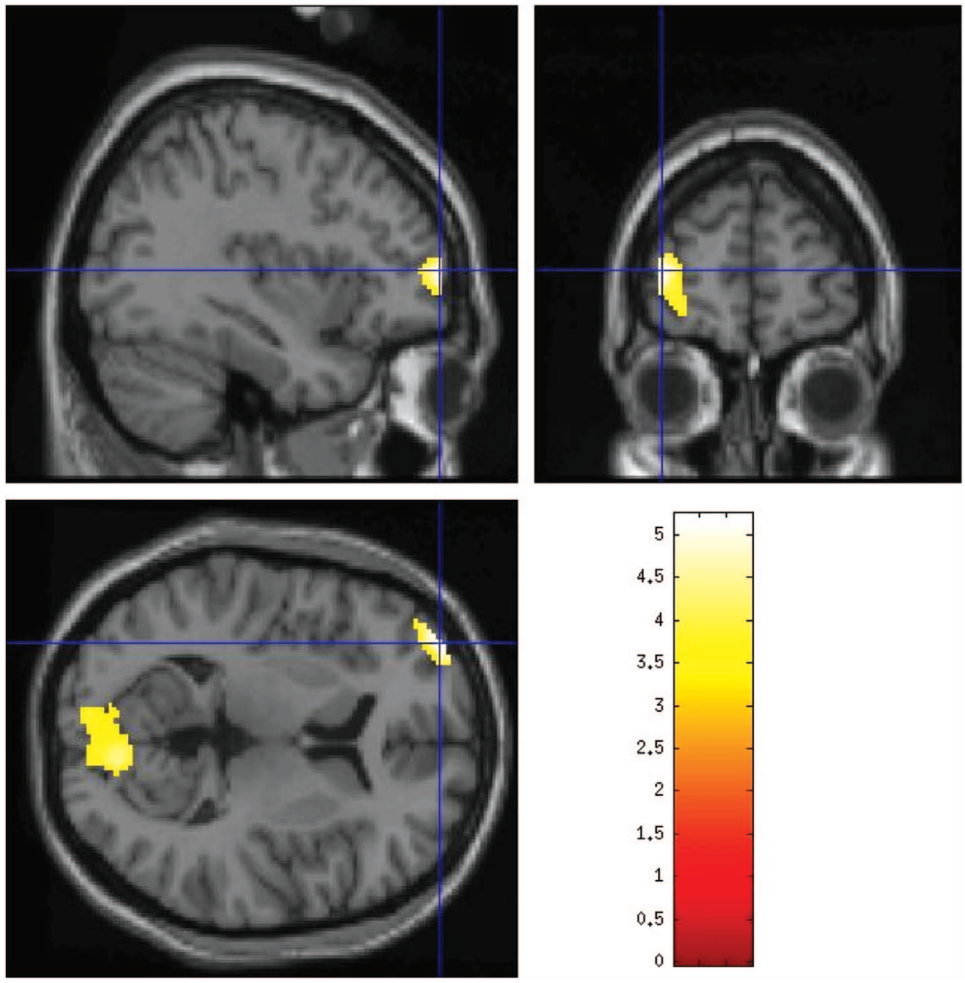

Figure 4. Activation in left anterior frontal cortex and middle occipital gyrus covaried as a function of attractiveness ratings. Statistical Parametric Mapping rendered into standard stereotactic space and superimposed on to transverse, coronal, and saggital MRI in standard space. Regions are designated using the MNI (Montreal Neurological Institute) coordinates. Bar = magnitude of *T*-score.

might rely on the convergence of information from a variety of sources that extend beyond the domain of faces exclusively. For example, relevant semantic, emotional, social, and cultural factors may contribute to individual differences in perceived facial attractiveness, in turn reflected in MTG activation.

This brings us to acknowledge some important caveats that remain for interpreting our results. First, for adult males in western societies, there are strong cultural taboos against finding young males and females of *either* sex attractive. In fact, strong legislations exist for punishing action motivated by such attitudes. It is possible that such societal values may have contributed to lower attractiveness ratings assigned by our male participants to target faces in the young age range. However, we believe that among male participants such taboos should suppress ratings for young male *and* female faces. As it turns out, this was not the case. In fact, for target faces that fell within the first four age levels (i.e., 0–16 years of age), male participants gave significantly higher ratings to female than male faces, $t(14) = 6.40$, $p < .001$. This suggests that cultural taboos alone cannot account for the lower ratings assigned by male participants to young target faces, given that these ratings were moderated by the sex of the target face.

Second, Senior (2003) has argued that for heterosexual males, it may be difficult to dissociate between "liking" and "wanting" when rating female faces on beauty. Our results suggest that to the extent that perceptions of beauty in female faces are linked to perceptions of attractiveness, those ratings will be strongly tied to perceptions of reproductive fitness among males (see Figure 3). Given that we measured perceived attractiveness and not beauty, we are not able to assess Senior's hypothesis directly. Indirectly, however, our results are consistent with a strong binding of liking and wanting in female face perception among males.

Conclusion

In summary, we demonstrated that there is a distributed system for assessing facial attractiveness involving structures geared toward evaluation and processing of reward (Aharon et al., 2001; Chatterjee et al., 2009; Nakamura et al., 1998; Kampe

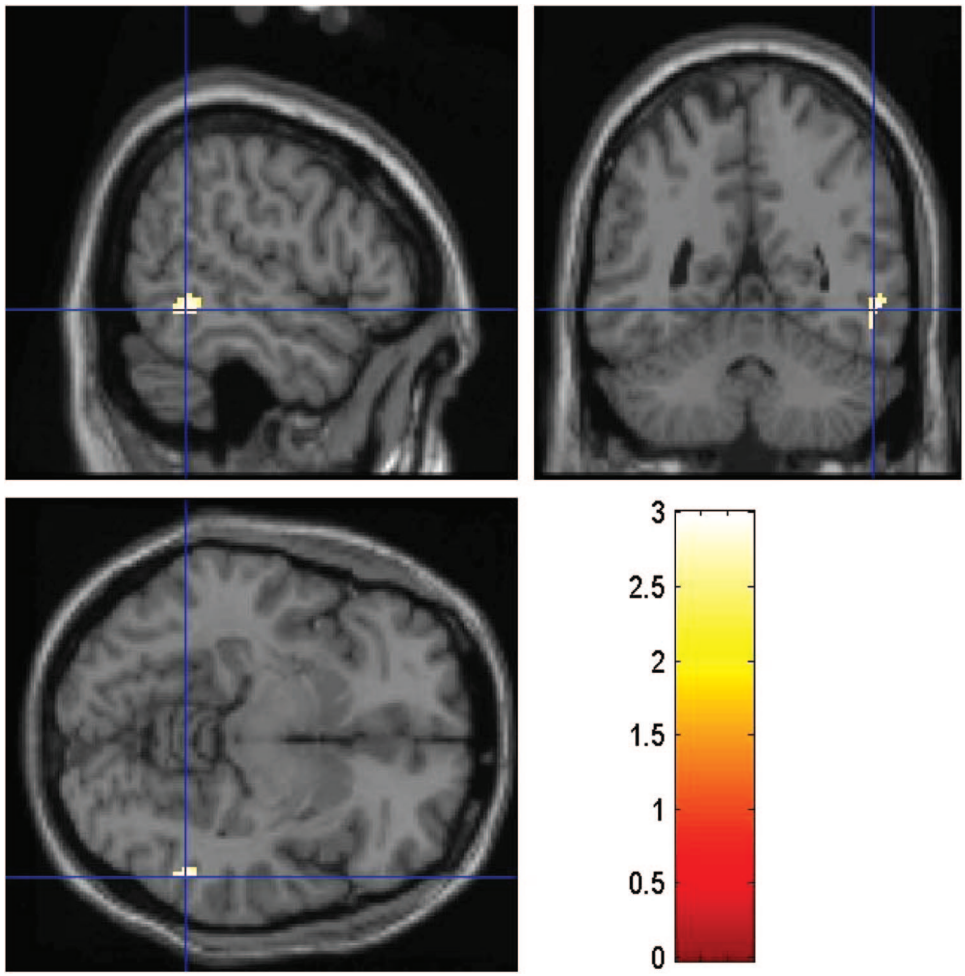

Figure 5. Activation in middle temporal gyrus was higher in participants who on average gave higher rather than lower ratings to faces. Statistical Parametric Mapping rendered into standard stereotactic space and superimposed on to transverse, coronal, and saggital MRI in standard space. Region is designated using the MNI (Montreal Neurological Institute) coordinates. Bar = magnitude of *T*-score.

et al., 2001; O'Doherty et al., 2003; Winston et al., 2007). Additionally we explored individual differences in the neural correlates of attractiveness ratings. We operationalized this difference using average facial attractiveness ratings and provided evidence that MTG responds to variations among participants in the average attractiveness rating ascribed to faces. This reflects a role for MTG in encoding individual differences in perceived facial attractiveness.

References

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron, 32,* 537–551. doi:10.1016/S0896-6273(01)00491-3

Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping, 7,* 254–266. doi:10.1002/(SICI)1097-0193(1999)7:4<254::AID-HBM4>3.0.CO;2-G

Berk, L. E. (2000). *Child development* (5th ed.). Boston, MA: Allyn & Bacon.

Berscheid, E., Dion, K., Walster, E., & Walster, G. W. (1971). Physical attractiveness and dating choice: A test of the matching hypothesis. *Journal of Experimental Social Psychology, 7,* 173–189. doi:10.1016/0022-1031(71)90065-5

Berscheid, E., & Walster, E. (1974). Physical attractiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 157–213). New York, NY: Academic Press.

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences, 11,* 290–298. doi:10.1016/j.tics.2007.05.004

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12,* 1–49. doi:10.1017/S0140525X00023992

Chatterjee, A. (2003). Prospects for a cognitive neuroscience of visual aesthetics. *Bulletin of Psychology of the Arts, 4,* 55–60.

Chatterjee, A., Thomas, A., Smith, S. E., & Aguirre, G. K. (2009). The neural response to facial attractiveness. *Neuropsychology, 23,* 135–143. doi:10.1037/a0014430

Christoff, K., Ream, J. M., Geddes, L. P., & Gabrieli, J. D. (2003). Evaluating self-generated information: Anterior prefrontal contributions

to human cognition. *Behavioral Neuroscience, 117,* 1161–1168. doi:10.1037/0735-7044.117.6.1161

Cloutier, J., Heatherton, T. F., Whalen, P. J., & Kelley, W. M. (2008). Are attractive people rewarding? Sex differences in the neural substrates of facial attractiveness. *Journal of Cognitive Neuroscience, 20,* 941–951. doi:10.1162/jocn.2008.20062

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. doi:10.1016/j.bandc.2009.01.003

Darwin, C. (1872). *The expression of the emotions in man and animals*. Chicago, IL: University of Chicago Press. doi:10.1037/10001-000

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

Dixson, A. F., Halliwell, G., East, R., Wignarajah, P., & Anderson, M. J. (2003). Masculine somatotype and hirsuteness as determinants of sexual attractiveness to women. *Archives of Sexual Behavior, 32,* 29–39. doi:10.1023/A:1021889228469

Dolan, R. J., Fletcher, P., Morris, J., Kapur, N., Deakin, J. F. W., & Frith, C. D. (1996). Neural activation during covert processing of positive emotional facial expressions. *NeuroImage, 4,* 194–200. doi:10.1006/nimg.1996.0070

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *Proceedings of the IEEE-Nuclear Science Symposium and Medical Imaging Conference* (Vol. 95, pp.1813–1817). London, UK: MTP Press.

Fletcher, G. J., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology, 76,* 72–89. doi:10.1037/0022-3514.76.1.72

Friston, K., Holmes, A., Worsley, K., Polinr, J.-B., Frith, C., & Frackowiak, T. (1994). Statistical parametric maps in functional imaging: A general approach. *Human Brain Mapping, 2,* 189–210. doi:10.1002/hbm.460020402

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage, 15,* 870–878. doi:10.1006/nimg.2001.1037

Hatfield, E., Aronson, V., Abrahams, D., & Rottman, L. (1966). The importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology, 4,* 508–516.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science, 132,* 349–350. doi:10.1126/science.132.3423.349

Hönekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 199–209. doi:10.1037/0096-1523.32.2.199

Ishai, A. (2007). Sex, beauty and the orbitofrontal cortex. *International Journal of Psychophysiology, 63,* 181–185. doi:10.1016/j.ijpsycho.2006.03.010

Jacobsen, T., Schubotz, R., Höfel, L., & von Cramon, D. (2006). Brain correlates of aesthetic judgments of beauty. *NeuroImage, 29,* 276–285. doi:10.1016/j.neuroimage.2005.07.010

Johnston, V. S. (2006). Mate choice decisions: the role of facial beauty. *Trends in Cognitive Sciences, 10,* 9–13.

Kampe, K. K. W., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Reward value of attractiveness and gaze. *Nature, 413,* 589. doi: 10.1038/35098149

Kanwisher, N., & Yovel, G. (2009). Cortical specialization for face perception in humans. In J. T. Cacioppo, & G. G. Berntson (Eds.), *Handbook of neuroscience for the behavioral sciences* (pp. 841–858). Hoboken, NJ: J. Wiley and Sons.

Kranz, F., & Ishai, A. (2006). Face perception is modulated by sexual preference. *Current Biology, 16,* 63–68. doi:10.1016/j.cub.2005.10.070

Lorenz, K. (1942). Die angeborenen Formen Möglicher Erfahrung [Innate forms of potential experience]. *Zeitschrift für Tierpsychologie, 5,* 235–409. doi:10.1111/j.1439-0310.1943.tb00655.x

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36,* 265–284. doi:10.1016/S0896-6273(02)00974-1

Nakamura, K., Kawashima, R., Nagumo, S., Ito, K., Sugiura, M., Kato, T., . . . Kojima, S. (1998). Neuroanatomical correlates of the assessment of facial attractiveness. *Neuroreport, 9,* 753–757. doi:10.1097/00001756-199803090-00035

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage, 31,* 440–457. doi:10.1016/j.neuroimage.2005.12.002

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41,* 147–155. doi:10.1016/S0028-3932(02)00145-8

Paradiso, S., Johnson, D. L., Andreasen, N. C., O'Leary, D. S., Watkins, G. L., Ponto, L. L. B., & Hichwa, R. D. (1999). Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. *The American Journal of Psychiatry, 156,* 1618–1629.

Quinsey, V. L., Ketsetzis, M., Earls, C., & Karamanoukian, A. (1996). Viewing time as a measure of sexual interest. *Ethology & Sociobiology, 17,* 341–354. doi:10.1016/S0162-3095(96)00060-X

Senior, C. (2003). Beauty in the brain of the beholder. *Neuron, 38,* 525–528. doi:10.1016/S0896-6273(03)00293-9

Shackelford, T. K., & Larsen, R. J. (1999). Facial attractiveness and physical health. *Evolution and Human Behavior, 20,* 71–76. doi:10.1016/S1090-5138(98)00036-1

Stelzer, C., Desmond, S. M., & Price, J. H. (1987). Physical attractiveness and sexual activity of college students. *Psychological Reports, 60,* 567–573. doi:10.2466/pr0.1987.60.2.567

Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3,* 452–460. doi:10.1016/S1364-6613(99)01403-5

Urbaniak, G. C., & Kilmann, P. R. (2003). Physical attractiveness and the "Nice Guy Paradox": Do nice guys really finish last? *Sex Roles, 49,* 413–426. doi:10.1023/A:1025894203368

Vartanian, O., & Goel, V. (2004). Neuroanatomical correlates of aesthetic preference for paintings. *NeuroReport, 15,* 893–897. doi:10.1097/00001756-200404090-00032

Weiner, K. S., & Grill-Spector, K. (2011). Neural representations of faces and limbs neighbor in human high-level visual cortex: Evidence for a new organization principle. *Psychological Research*. Epub ahead of print. doi:10.1007/s00426-011-0392-x

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. doi:10.1016/j.neuropsychologia.2006.05.009

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—Again. *NeuroImage, 2,* 173–181. doi:10.1006/nimg.1995.1023

Zysset, S., Huber, O., Ferstl, O., & von Cramon, D. Y. (2002). The anterior frontomedian cortex and evaluative judgment: an fMRI study. *NeuroImage, 15,* 983–991.

Received February 6, 2012
Revision received July 26, 2012
Accepted July 30, 2012 ■