Perception, 2001, volume 30, pages 611 – 625

DOI:10.1068/p3123

Attractiveness of facial averageness and symmetry in non-Western cultures: In search of biologically based standards of beauty

Gillian Rhodes¶§, Sakiko Yoshikawa§#, Alison Clark¶, Kieran Lee¶, Ryan McKay¶, Shigeru Akamatsu§

¶ Department of Psychology, University of Western Australia, Nedlands, Perth, WA 6907, Australia; e-mail: gill@psy.uwa.edu.au;

§ ATR Human Information Processing Research Laboratories, 2-2 Hikari-dai, Soraku-gun, Kyoto 619-02, Japan;

Department of Cognitive Psychology in Education, Graduate School of Education, Kyoto University, Sakyo, Kyoto 606-8501, Japan

Received 15 August 2000, in revised form 18 December 2000

Abstract. Averageness and symmetry are attractive in Western faces and are good candidates for biologically based standards of beauty. A hallmark of such standards is that they are shared across cultures. We examined whether facial averageness and symmetry are attractive in non-Western cultures. Increasing the averageness of individual faces, by warping those faces towards an averaged composite of the same race and sex, increased the attractiveness of both Chinese (experiment 1) and Japanese (experiment 2) faces, for Chinese and Japanese participants, respectively. Decreasing averageness by moving the faces away from an average shape decreased attractiveness. We also manipulated the symmetry of Japanese faces by blending each original face with its mirror image to create perfectly symmetric versions. Japanese raters preferred the perfectly symmetric versions to the original faces (experiment 2). These findings show that preferences for facial averageness and symmetry are not restricted to Western cultures, consistent with the view that they are biologically based. Interestingly, it made little difference whether averageness was manipulated by using own-race or other-race averaged composites and there was no preference for own-race averaged composites over other-race or mixed-race composites (experiment 1). We discuss the implications of these results for understanding what makes average faces attractive. We also discuss some limitations of our studies, and consider other lines of converging evidence that may help determine whether preferences for average and symmetric faces are biologically based.

1 Introduction

A recent meta-analysis has revealed very high agreement in facial-attractiveness ratings among people from different cultures ($r = 0.94$), as well as among people from the same culture ($r = 0.90$) (Langlois et al 2000). These results challenge the view that standards of beauty are cultural artifacts. So too, does evidence that young infants prefer to look at faces that adults find attractive (for a review see Rubenstein et al 2001). Instead, these findings suggest that some standards of beauty may reflect biologically based preferences that have been shaped by human evolution. Just as we have a language instinct (Pinker 1994), so too may we have an attractiveness instinct that makes us prefer some facial qualities over others.

Two good candidates for biologically based preferences are preferences for average faces and for symmetric faces (Thornhill and Gangestad 1999). In many species, average and symmetric[1] morphological traits reflect an important aspect of mate quality, namely developmental stability—the ability to maintain a stable course of development

[1] Several types of asymmetry can be distinguished by the shape of their distribution in a population. Fluctuating asymmetries, which are randomly distributed (to the left and right), with a mean of zero asymmetry, are thought to reflect developmental stability. Other sorts of asymmetry, such as directional asymmetry, which has a population bias to the left or right, may not reflect developmental stability.

despite environmental and genetic stresses (Møller and Swaddle 1997; Thornhill and Møller 1997). If averageness and symmetry signal mate quality, then individuals with a preference for these traits would have a reproductive advantage (Møller and Swaddle 1997; Thornhill and Gangestad 1993), and the preferences or, more precisely, the information-processing mechanisms that generate them, could have evolved by sexual selection. Alternatively, preferences for averageness and symmetry may be by-products of general perceptual or cognitive mechanisms that evolved by natural selection to process information from the environment, and which respond selectively to average and/or symmetric stimuli (Endler and Basolo 1998; Enquist and Arak 1994; Enquist and Johnstone 1997; Halberstadt and Rhodes 2000; Johnstone 1994).[(2)] In either case, however, the preferences should be universal, and evidence that preferences for average and symmetric faces occur in diverse cultures would support the conjecture that they are biologically based.

The attractiveness of averageness is well established for Western faces and raters. Computer-averaged composite faces are rated as more attractive than almost all the component faces (Langlois and Roggman 1990; Rhodes et al 1999b) and individual faces can be made more (or less) attractive by moving their configurations closer to (or further from) an average same-sex configuration (Rhodes and Tremewan 1996; Rhodes et al 1999b). Faces that naturally lie closer to the population average are also more attractive than less typical (more distinctive) faces (Light et al 1981; Rhodes and Tremewan 1996; Rhodes et al 1999b). A few extreme traits may be more attractive than average ones—eg feminised traits (Perrett et al 1998; Rhodes et al 2000), neotonous traits (Zebrowitz 1997)—but average facial configurations are still more attractive than most faces, and this preference must be explained.

Symmetry is also attractive for Western faces and raters. Natural variations in symmetry covary with attractiveness (Grammer and Thornhill 1994; Mealey et al 1999; Zebrowitz et al 1996) and, when symmetry is increased experimentally, attractiveness generally increases, so long as the manipulation does not produce structural abnormalities (Perrett et al 1999; Rhodes et al 1998; Rhodes et al 1999a). Interestingly, although more average faces are generally more symmetric than less average ones, symmetry and averageness make independent contributions to the attractiveness of Western faces (Rhodes et al 1999b).

Little is known, however, about whether people from non-Western cultures find facial averageness and symmetry attractive. In only two published studies has the attractiveness of average or symmetric faces in non-Western populations been examined. Jones and Hill (1993) examined both preferences in three populations of faces (US Americans, Brazilians, Paraguayan Indians) and five populations of raters (the same three groups plus Russians and Venezuelan Indians). They found a significant correlation between measured averageness and attractiveness in only one group, Paraguayan Indians, but age was confounded with averageness in that group, so that the preference could be for youthful rather than for average faces [a youthful appearance is attractive, at least for Western faces (Zebrowitz et al 1993)]. The fact that they failed to obtain a significant correlation between averageness and attractiveness for either US or Brazilian faces, however, suggests that there may be some problem with the stimuli or methodology used (eg validity of the averageness measurements), because average Western faces are known to be attractive (see above). No significant correlations were found between measured symmetry and attractiveness for the non-Western raters, but again, correlations were not generally found for the Western raters either, suggesting a problem with the stimuli or methodology used.

[(2)] Note that these mechanisms are not mutually exclusive. Both could have shaped the evolution of our current preferences.

Kowner (1996) examined the attractiveness of symmetry in a non-Western culture. She compared the attractiveness of normal and perfectly symmetric versions of Japanese faces. Her Japanese subjects preferred the normal, slightly asymmetric versions (except for elderly faces, where symmetry was preferred). However, the symmetric images were created by reflecting each half of the face about the vertical midline. This method introduces structural abnormalities (eg height : width ratios outside the normal range, abnormally narrow or wide midline features), which make the images unattractive, despite their symmetry, even to Western raters (Rhodes et al 1999a).

These published studies provide little evidence that facial averageness and symmetry are attractive in non-Western cultures. However, all appear to have methodological problems and further tests are clearly needed. The attractiveness of averaged composites in Chinese and Japanese cultures has been examined in one unpublished study (Rhodes et al 2001b). Averaged composites (of own-race faces) were preferred to most individual faces for both Chinese and Japanese raters. These results parallel those found with Western raters (Langlois and Roggman 1990; Rhodes et al 1999b).

In the present study we sought additional evidence whether non-Western populations find facial averageness (experiments 1 and 2) and symmetry (experiment 2) attractive. We examined whether individual faces could be made more (or less) attractive, by increasing (or decreasing) their averageness, as has been found for Western faces and raters. We also examined whether attractiveness correlates negatively with rated distinctiveness, a subjective converse measure of averageness. Chinese raters were tested in experiment 1 and Japanese raters were tested in experiment 2. In experiment 2, we also examined whether symmetry is attractive, by investigating whether perfectly symmetric versions of Japanese faces were preferred to the original versions. Evidence that facial averageness and symmetry are attractive in these non-Western populations would strengthen the claim that they are biologically based standards of beauty.

2 Experiment 1

We manipulated the averageness of Chinese faces by moving the configuration of each face halfway towards an average same-sex, Chinese configuration (an averaged composite of 24 faces), to produce a high-average version of the face. We also moved the configuration away from this average configuration by the same amount, to produce a low-average version of the face (see below for details). This method of manipulating facial averageness alters the spatial configuration or shape, while keeping the texture of the face constant. For Western faces and raters, increasing (decreasing) averageness in this way increases (decreases) attractiveness (Rhodes and Tremewan 1996; Rhodes et al 1999b). We predicted that the same effect would be found for Chinese faces and raters. We also increased averageness of both face shape and skin texture by blending each face 50 : 50 with a same-sex, Chinese averaged composite. We expected these blends to be more attractive than the undistorted faces, and more attractive than the high-average versions, if averageness of skin texture, as well as shape, is attractive (Benson and Perrett 1992).[3]

We also examined the relationship between attractiveness and distinctiveness. Distinctiveness is a useful converse measure of averageness because more distinctive faces differ more from an average configuration than do less distinctive faces (eg Bruce et al 1993, 1994; Johnston and Ellis 1995), and distinctiveness ratings change systematically with manipulations of averageness (Lee et al 2000; Rhodes and Tremewan 1996; Rhodes et al 1998, 1999b). We expected that attractiveness would be negatively correlated

[3] Averaging faces produces smooth complexions, which are attractive (Benson and Perrett 1992). We note, however, that the attractiveness of average faces cannot be attributed solely to these changes in skin texture, because averageness remains attractive when skin texture is controlled (eg Rhodes and Tremewan 1996; Rhodes et al 1999b).

with distinctiveness, as it is for Caucasian faces and raters (eg Rhodes and Tremewan 1996; Rhodes et al 1999b).

We examined the relative attractiveness of averaged composites of own-race, other-race, and mixed-race composites, to determine whether familiarity with the population from which component faces are drawn affects attractiveness of the composites. Langlois et al (1994) suggested that averaged composites may be attractive because of their familiarity, in which case own-race composites should be more attractive than composites of faces from a less familiar race. In this case increasing the similarity of a face to an own-race composite might also enhance its attractiveness more effectively than increasing its similarity to an other-race composite. We tested these predictions by obtaining attractiveness ratings for own-race (Chinese), other-race (Caucasian), and mixed-race composites, and by using both own-race and other-race composites to manipulate the averageness of individual Chinese faces.

2.1 *Method*

2.1.1 *Participants.* The participants were 36 (18 male, 18 female) young, ethnically Chinese adults ($M = 17.0$ years, SD = 1.4 years) who were paid \$10 each. All were born into Chinese communities in Asian countries (Indonesia, $N = 21$; People's Republic of China, $N = 6$; Hong Kong, $N = 4$; Malaysia, $N = 1$; Singapore, $N = 1$; Philippines, $N = 1$; Korea, $N = 1$; Brunei, $N = 1$) and had been living in Perth, Australia, attending local colleges, for 6 months or less ($M = 3.6$ months, range = 0.5–6 months). Despite living in Australia, the participants primarily had Asian friends—14 had no Australian friends, 10 had less than 5, 11 had between 5 and 10, and 1 had more than 10 Australian friends. All but 4 had more than 10 Asian friends.

2.1.2 *Stimuli.* The stimuli were 48 (24 male, 24 female) front-view, black-and-white photographs of young Chinese adults, displaying neutral expressions, which were taken under standard lighting conditions and digitised for use as stimuli. None was familiar to the participants. All images had a standard interpupil distance of 80 pixels, with horizontally aligned pupils. No jewellery or clothing was visible, and blemishes and hair on the forehead were removed with the cloning tool in Adobe Photoshop.

Male and female Chinese averaged composites were created from these faces by using a standard procedure in Gryphon's Morph™ (for details, see Rhodes et al 1999b). Briefly, to create a composite from a set of faces, a set of landmark points ($N = 169$) was found on each face, and used to generate the average locations of each landmark. Each face was then warped into this average configuration, and the grey-level values at corresponding pixel locations in each image were averaged, to generate an averaged composite. Caucasian composites were taken from a previous study (Rhodes et al 1999b). Two mixed-race composites were also created for each sex. One had the same number of component faces as the single-race averages ($N = 24$) and was the average of 12 randomly chosen, same-sex faces from each race (Mixed24). The other was a composite of all 24 Chinese and all 24 Caucasian same-sex faces (Mixed48). The four averaged composites for each sex are shown in figure 1. We will refer to these averaged composites as norms when they are used to manipulate the averageness of individual faces.

We created high-average and low-average versions for each Chinese face, using the procedure described in Rhodes et al (1999b). Briefly, a set of landmark points ($N = 169$) was found on each face and used to alter the shape of the face. A high-average version was created by moving all landmark points on a face 50% closer to their corresponding landmark points on an averaged composite or norm for that sex, and remapping grey-levels from the face into this new configuration. A low-average version was created by moving the landmark points 50% further away from their corresponding points on the norm. This procedure alters the spatial configuration of the face, but leaves the texture unchanged. Analogous low-average and high-average

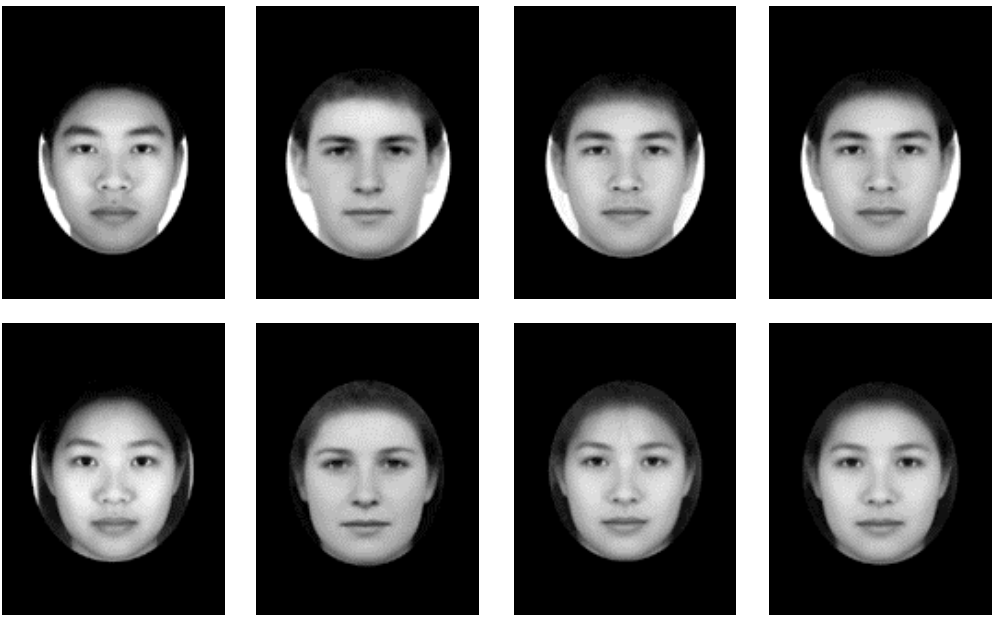

Figure 1. Male (top row) and female (bottom row) averaged composites from experiment 1. Chinese24 is a composite of 24 same-sex Chinese faces, Caucasian24 is a composite of 24 same-sex Caucasian faces. Mixed24 is a mixed-race composite of 12 randomly chosen Chinese and 12 randomly chosen Caucasian same-sex faces. Mixed48 is a composite of all 24 Chinese and 24 Caucasian same-sex faces.

versions were created with Caucasian norms. We also created two blends for each face: a 50 : 50 blend of the face with a same-sex Chinese norm and a 50 : 50 blend with a same-sex Caucasian norm. This blending procedure increases the averageness of both texture and shape information. A full set of images for one face is shown in figure 2.

The averaged composites and blends were sharpened in Adobe Photoshop to reduce any blurring introduced by blending. All images were displayed in oval masks that hid most of the hair but showed the face outline, chin, and ears.

2.1.3 *Procedure.* There were 18 subjects (9 male, 9 female) who rated the images on attractiveness and a different 18 subjects (9 male, 9 female) who rated them on distinctiveness, using 10-point scales. Distinctiveness was explained as the ease with which a face could be picked out of a crowd. Distinctiveness was rated rather than averageness, which can be misinterpreted to mean average-looking as opposed to good-looking. For each task, all seven versions of each face plus the eight averages ($N = 344$ images in total) were presented in random order on a computer screen. Each image remained visible until the rating was made. Subjects chose whether to read the instructions in English, Mandarin, Malay, or Indonesian.

2.2 *Results*

A four-way ANOVA was carried out with sex of subject as a between-subjects variable and sex of face, averageness level (low, normal, high, blend), and race of norm (Chinese, Caucasian) as repeated-measures factors. Planned pairwise comparisons, with Bonferroni correction for multiple comparisons, were used to test for predicted increases in attractiveness from low to normal, normal to high, and high to blend levels (manipulated with the own-race, Chinese norms). A posteriori Tukey tests were used to explore differences that were not predicted or were in the opposite direction to those predicted.

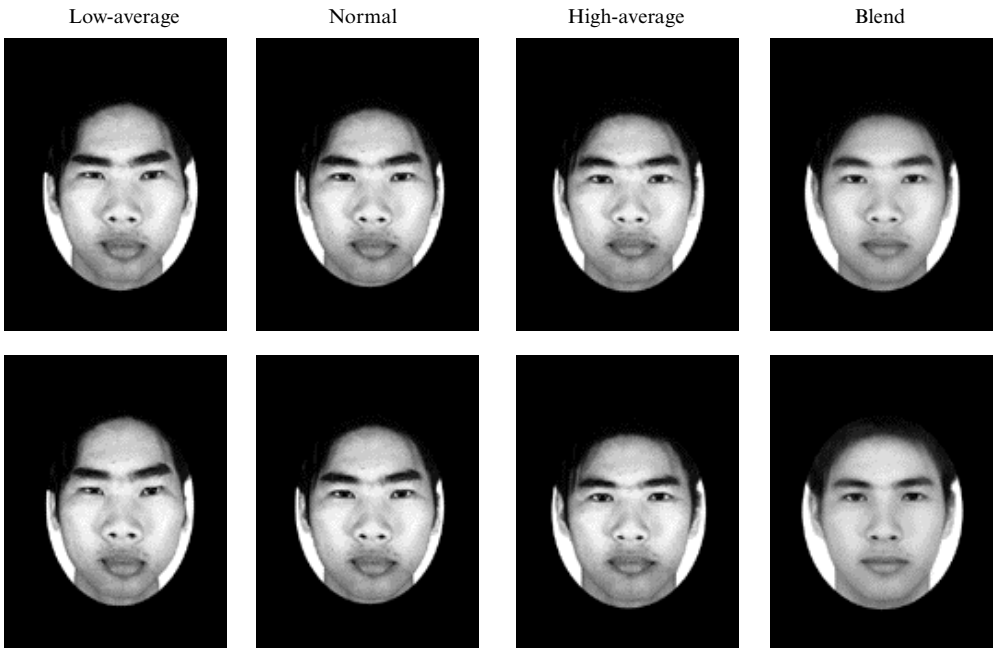

Figure 2. A male face shown at the four averageness levels used in experiment 1. Top row (left to right): low-average, normal, high-average, and blend images made by using the male, Chinese averaged composite (norm). Bottom row (left to right): low-average, normal, high-average, and blend images made by using the male, Caucasian averaged composite (norm). High-average and low-average versions were made by warping the facial configuration half way towards and away from, respectively, the same-sex norm. Averageness of shape is varied with this procedure. The blends are 50 : 50 blends of the original faces with a same-sex averaged composite (norm). Averageness of both texture and shape is varied with this procedure.

There were significant main effects of averageness ($F_{3,48} = 171.62$, $p < 0.0001$) and race of norm ($F_{1,16} = 19.17$, $p < 0.0005$; Chinese—$M = 3.9$, SE $= 0.1$; Caucasian—$M = 4.0$, SE $= 0.1$), and a two-way interaction between these factors ($F_{3,48} = 17.35$, $p < 0.0001$)—see figure 3. As predicted, when the own-race Chinese norm was used, attractiveness increased with averageness, for both male and female faces (for all, $t > 5.19$, $p < 0.0001$, 1-tailed; Bonferroni correction for 6 comparisons, $p = 0.008$). This result parallels previous findings with Caucasian faces (Rhodes and Tremewan

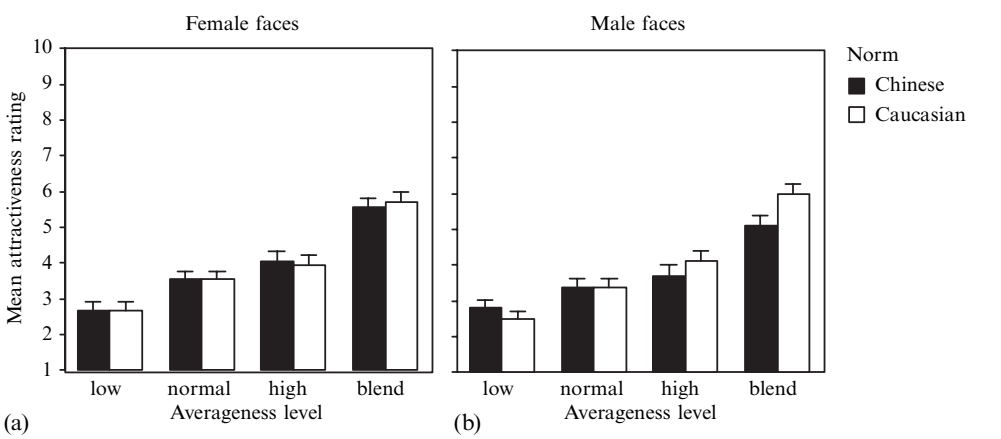

Figure 3. Attractiveness as a function of averageness level and race of norm for female faces (a) and male faces (b) in experiment 1. SE bars are shown.

1996; Rhodes et al 1999b) and clearly shows that the attractiveness of average facial configurations is not restricted to Western culture. Attractiveness also increased when the other-race Caucasian norm was used to manipulate averageness ($p < 0.01$ for all in a Tukey's HSD test). This result suggests that the attractiveness of average facial configurations does not require a high level of familiarity with the population from which the component faces are drawn.

There was a significant three-way interaction between averageness, norm, and sex of face ($F_{3,48} = 28.17$, $p < 0.0001$). For male faces, in contrast to our prediction, blends and high-average versions made using a Caucasian (other-race) norm were more attractive than those made with Chinese norms (for both, $p < 0.01$ in a Tukey's HSD test). For female faces, the race of the norm made little difference to the attractiveness of blends and high-average images. Neither sex preferred blends or high-average images made from own-race norms to those made from other-race norms, even though the former are closer to their own population average. Nor was there a preference for own-race averaged composites over other-race or mixed-race composites (see below). The greater lifetime exposure to Chinese faces did not, therefore, generate a stronger preference for Chinese than Caucasian averages.

There were several other effects, but these did not affect the conclusions drawn above. There was a significant interaction between sex of subject and sex of face ($F_{1,16} = 8.31$, $p < 0.01$), with female subjects rating female faces as more attractive ($M = 4.3$, SE $= 0.2$) than male faces ($M = 3.6$, SE $= 0.2$), but no sex difference in ratings for male faces (for female subjects—$M = 3.9$, SE $= 0.2$; for male subjects—$M = 3.9$, SE $= 0.2$). Sex of face also interacted with norm ($F_{1,16} = 36.63$, $p < 0.0001$; male faces: Chinese norm—$M = 3.8$, SE $= 0.2$; Caucasian norm—$M = 4.0$, SE $= 0.2$; female faces: Chinese norm—$M = 4.0$, SE $= 0.02$; Caucasian norm—$M = 4.0$, SE $= 0.2$).

2.3 *Correlating attractiveness and distinctiveness*

Attractiveness ratings were highly reliable, with a Cronbach coefficient alpha of 0.94 for all the faces. We also examined reliability separately for Chinese and non-Chinese faces, to see whether it was higher for own-race (Chinese) faces. The Chinese category included all the original (undistorted) Chinese faces, blends of those faces with Chinese averaged composites, high-average versions of those faces made by warping the Chinese faces towards a Chinese norm (averaged composite), and the Chinese averaged composites. We excluded Chinese faces that had been warped away from a Chinese average or manipulated using Caucasian norms. These faces, together with the Caucasian averaged composites, formed the non-Chinese category. Alphas were 0.93 and 0.95, respectively, for the Chinese and non-Chinese faces, indicating that attractiveness ratings were equally reliable for own-race and other-race faces. Distinctiveness ratings were less reliable and should be interpreted with caution. Coefficient alphas were 0.54 for all faces, 0.18 for Chinese faces, and 0.55 for non-Chinese faces. It is not clear why distinctiveness ratings should be less reliable for Chinese than non-Chinese faces.

Despite the relatively poor reliability of the distinctiveness ratings, attractiveness and distinctiveness were significantly negatively correlated (for all faces—$r_{342} = -0.488$, $p < 0.0001$; for non-Chinese faces—$r_{196} = -0.491$, $p < 0.0001$; for Chinese faces, all versions—$r_{144} = -0.486$, $p < 0.0001$; and for the original, undistorted Chinese faces—$r_{46} = -0.535$, $p < 0.0001$). These results show that low distinctiveness (high averageness) is attractive for Chinese faces and raters, just as it is for Western faces and raters (Rhodes et al 1999b).

2.4 *Comparing attractiveness of own-race, other-race, and mixed-race averaged composites*

A three-way ANOVA, with sex of subject as a between-subjects factor, and sex of face and type of composite (Chinese24, Caucasian24, Mixed24, Mixed48) as repeated-measures factors was carried out on the mean attractiveness ratings. One missing value was

replaced by the mean for that cell. Planned comparisons with Bonferroni correction for multiple comparisons were used to test the hypothesis that own-race (Chinese) composites would be preferred to other-race and mixed-race composites.

There was no significant effect of type of composite ($F_{3,48} = 1.65$, ns). Attractiveness ratings were high, but below ceiling, for all the composites (figure 4). Type of composite also interacted with sex of face ($F_{3,48} = 3.18$, $p < 0.04$, see figure 4). For male faces, the own-race Chinese composite was not significantly more attractive than the other-race or mixed-race composites (planned comparisons). For female faces the Chinese composite was preferred to the Caucasian (but not the mixed-race) composites, although this difference failed to reach significance with correction for multiple comparisons ($t_{48} = 2.20$, $p < 0.04$; 6 comparisons, corrected $p = 0.008$). A posteriori Tukey HSD tests showed that male Chinese composites were less attractive than either female, mixed-race composite (Mixed24 or Mixed48), $p < 0.05$ in each case, but this result has no obvious theoretical significance and will not be discussed further. In summary, different lifetime exposures to Chinese and Caucasian faces did not produce clear differences in the attractiveness of averaged Chinese and Caucasian composites.

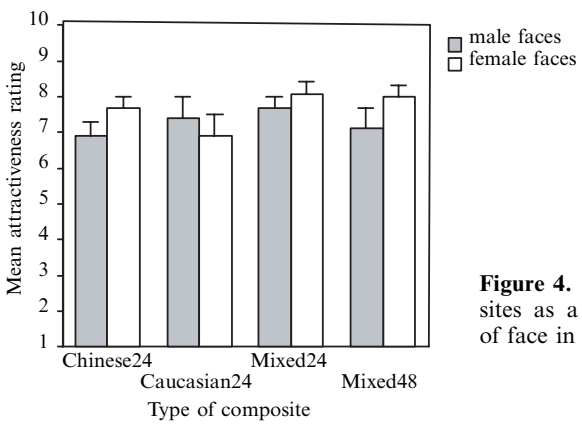

Figure 4. Mean attractiveness of averaged composites as a function of type of composite and sex of face in experiment 1. SE bars are shown.

2.5 *Comparing attractiveness of averaged composites and component faces*

We obtained a mean attractiveness rating for each image by averaging across raters. All female composites received higher ratings (Chinese24—$M = 7.7$; Caucasian24—$M = 7.8$; Mixed24—$M = 8.1$; Mixed48—$M = 8.0$) than any other female image, including the blends and high-average versions (maximum $M = 7.4$). Similarly, all male composites were rated higher (Chinese24—$M = 6.8$; Caucasian24—$M = 7.7$; Mixed24—$M = 7.8$; Mixed48—$M = 7.8$) than any other male image (Maximum $M = 6.6$). Therefore, not only were averaged composites more attractive than their individual component faces, but they were also more attractive than versions of those component faces in which averageness of the shape (high-average) or shape and texture (blends) had been increased by 50% above normal levels.

In summary, these results show that the attractiveness of averageness is not confined to Western faces and raters. Chinese raters clearly found averageness attractive in Chinese faces. Increasing the averageness of face shape (high-average versions) increased attractiveness, and increasing the averageness of skin texture as well (blends) increased attractiveness further. Averaged composites were preferred to the individual component faces, even their high-average versions. Interestingly, the Chinese participants showed little bias towards own-race composites, finding Chinese, Caucasian, and mixed-race composites about equally attractive. We will discuss this result further in section 4.

3 Experiment 2

The Chinese participants in experiment 1 were living in a Western country and clearly had some exposure to Western culture. Their preferences could, therefore, have been influenced by Western standards of beauty. In experiment 2, we tested Japanese participants living in Japan, who had less exposure to Western culture than the Chinese participants of experiment 1.

We investigated whether averageness is attractive for Japanese faces and raters. We manipulated the averageness of individual Japanese faces in the same way as in experiment 1, by moving the faces 50% closer to or further from an averaged same-sex composite of Japanese faces. We predicted that attractiveness would increase with averageness.

We also investigated whether symmetry is attractive for Japanese faces and raters. We manipulated symmetry by blending each face with its mirror image, as in previous studies with Western faces (Rhodes et al 1998, 1999a, 1999b). This blending method produces symmetric faces without structural abnormalities, an advantage over the symmetric chimeras used previously with non-Western individuals (Kowner 1996). Western raters prefer these symmetric blends to the original faces, even when texture differences are controlled, indicating a genuine preference for symmetry, not just for the smoother complexions of blended images (Rhodes et al 1998; see also Perrett et al 1999).

We also created perfectly symmetric blends for the low-average and high-average versions of each face. By including symmetric versions of the low-average, normal, and high-average images, we could determine whether averageness remains attractive when symmetry is controlled, as is the case for Caucasian faces and raters (Rhodes et al 1999b). If it does, then averageness would be attractive in Japanese faces independent of any accompanying increase in symmetry.

3.1 *Method*

3.1.1 *Participants.* The 56 participants (28 male, 28 female) were young adults attending universities in the Kyoto region ($M = 20.2$ years, SD $= 1.8$ years, range $= 18-26$ years). They had little personal contact with Westerners ($M = 2.6$ Westerners known personally, SD $= 4.1$; $M = 1.3$ interactions with Westerners in a month, SD $= 1.9$), but had studied English for an average of 8.0 years (SD $= 1.6$ years) and had some exposure to Western media ($M = 7.2$ h of Western TV per month, SD $= 8.2$; $M = 0.6$ Western magazines read per month, SD $= 0.8$).

3.1.2 *Stimuli.* The stimuli were 48 (24 male, 24 female) front-view, colour photographs of young Japanese adults displaying neutral expressions. The images were from the ATR face database, which contains photographs of Japanese men and women in their twenties who were recruited from the general public in Tokyo in 1998. The photographs showed front-views of faces, displaying neutral expressions, and were taken in a studio with symmetric lighting. They measured approximately 7.5 cm $\times$ 8.0 cm. Blemishes, stray hair, any clothing on the neck, and hair clips were removed using the cloning tool in Photoshop. None of the faces had hair on the forehead and any partings were filled in with hair in Photoshop, so that the symmetric images would not have odd looking hairstyles.

Japanese male and female averaged composites were created from the 24 male and 24 female faces, respectively, with the same procedure as in experiment 1. These were used to create a high-average and low-average version of each face (50% distortions) following the same procedure as in experiment 1. Perfectly symmetric versions of the low-average, normal, and high-average images, and the average composites were created by blending each image with its mirror image, in the following procedure. First, each face was scaled so the pupils were horizontally aligned and in average locations for that sex. Then, a symmetric configuration was obtained by averaging

(separately) the x-deviations and the y-deviations of corresponding (paired) landmark points ($N = 175$), on the left and right side of the face, from a vertical line bisecting the horizontal line joining the pupils. Finally, a symmetric texture was created by averaging the grey levels in corresponding small regions in the symmetric configuration and its mirror image.

A total of 292 images were used in the experiment (48 images $\times$ 3 averageness levels $\times$ 2 symmetry levels, plus normal and symmetric averaged composites for each sex). All images were displayed in oval masks that hid the outer hairline (figure 5).

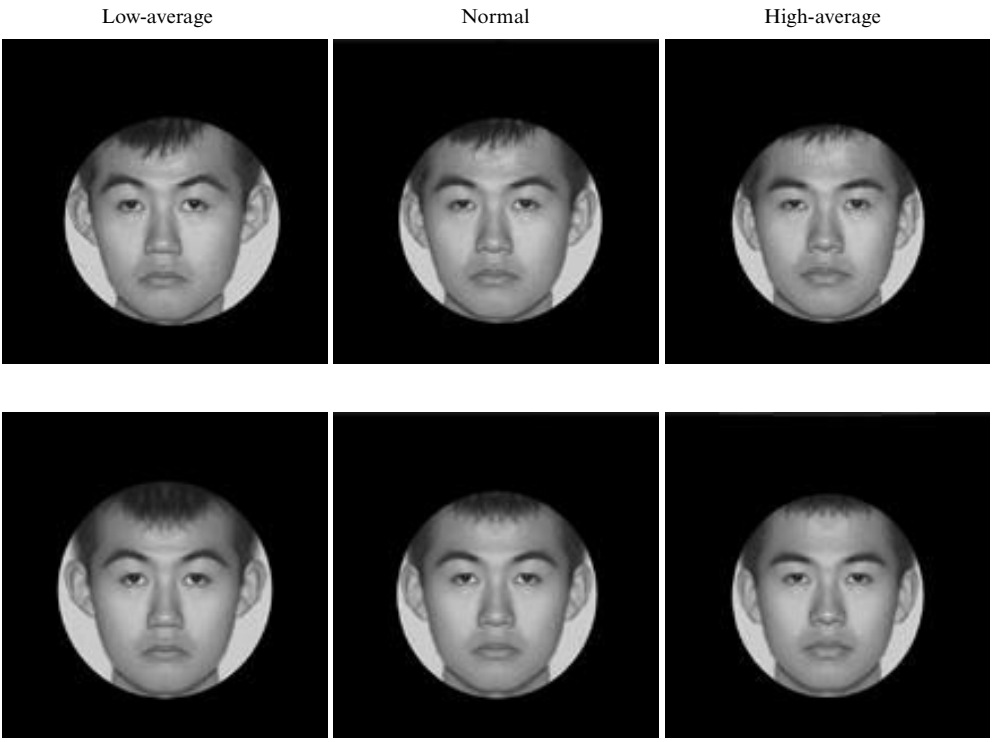

Figure 5. A full set of images for a Japanese face from experiment 2. Top row (left to right): low-average, normal, and high-average versions. Bottom row (left to right): perfectly symmetric versions of the top row images.

3.1.3 *Procedure.* There were 32 (16 male, 16 female) participants who rated all 292 images on attractiveness and a different 24 (12 male, 12 female) participants who rated them on distinctiveness, using 10-point scales. The images were presented in random order (different for each participant) on a computer screen and each image remained visible until rated. Participants were tested individually in a session lasting about 30 min. At the beginning of the session, participants completed a short questionnaire, assessing their amount of contact with Western people and media and their fluency in English. The questionnaire and all instructions were in Japanese.

3.2 *Results*

3.2.1 *Attractiveness.* A four-way ANOVA was carried out on attractiveness ratings, with sex of rater as a between-subjects factor, and sex of face, averageness level (low, normal, high), and symmetry level (normal, perfect), as within-subject factors. Planned t-tests with Bonferroni correction for multiple comparisons were used to test for predicted increases in attractiveness as averageness level increased from low to normal and from normal to high.

Attractiveness increased with averageness level ($F_{2,60} = 207.93$, $p < 0.0001$; Ms = 3.2, 4.3, and 4.8 for low, normal, and high, respectively, all SEs = 0.1; pairwise $t > 6.11$, $p < 0.0001$ for all), and with symmetry level ($F_{1,30} = 32.85$, $p < 0.0001$; normal—$M = 4.0$; perfect symmetry—$M = 4.2$; both SEs = 0.1). These results exactly parallel those for Western faces and raters (Rhodes et al 1999b). The changes in attractiveness with increasing levels of averageness and symmetry are modest, but very consistent across raters, and similar to those found for Western faces and raters. There was no interaction between averageness and symmetry ($F < 1$), see figure 6, indicating that averageness remains attractive when symmetry is controlled. Therefore, although symmetry is attractive, it does not fully account for the attractiveness of average faces.

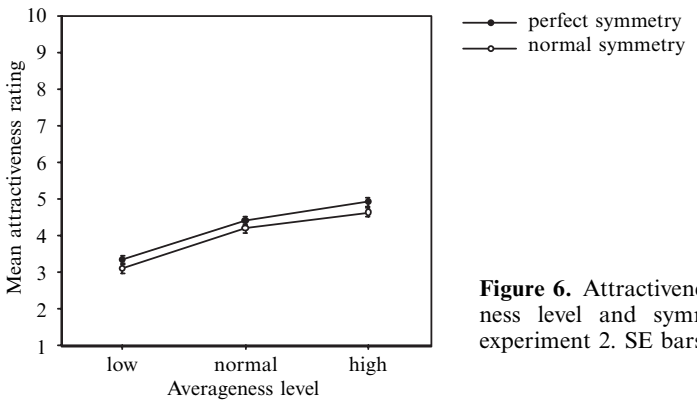

Figure 6. Attractiveness as a function of averageness level and symmetry of Japanese faces in experiment 2. SE bars are shown.

There were several interactions, but they do not alter the conclusions drawn above. Averageness interacted with sex of face ($F_{2,60} = 8.91$, $p < 0.0004$), but as predicted attractiveness increased with averageness for both sexes (female faces—Ms = 3.3, 4.4, 4.7; male faces—Ms = 3.1, 4.2, 4.8; all SEs = 0.1; pairwise $t > 6.92$, $p < 0.0001$ for all). Averageness also interacted with sex of rater ($F_{2,60} = 5.59$, $p < 0.006$), but attractiveness increased with averageness for both female (Ms = 2.8, 4.0, 4.6; SEs = 0.1) and male raters (Ms = 3.6, 4.6, 4.9; SEs = 0.1) ($F > 72.86$, $p < 0.0001$, for all in simple tests of main effects). Symmetry interacted with sex of face ($F_{1,30} = 7.00$, $p < 0.02$), but perfect symmetry was still preferred for both female (normal—$M = 4.0$; perfect—$M = 4.3$; SEs = 0.1) and male faces (normal—$M = 4.0$; perfect—$M = 4.1$; SEs = 0.1) ($t > 3.39$, $p < 0.002$ for all).

Sex of face also interacted with sex of rater ($F_{1,30} = 19.79$, $p < 0.0001$). Female faces received slightly higher attractiveness ratings from male ($M = 4.2$, SE = 0.1) than female raters ($M = 4.1$, SE = 0.1), but male faces received much higher ratings from male ($M = 4.6$, SE = 0.1) than female raters ($M = 3.5$, SE = 0.1).

3.3 *Correlating attractiveness and distinctiveness*

Attractiveness and distinctiveness ratings were highly reliable, with Cronbach coefficient alphas of 0.94 and 0.92, respectively.[4] Attractiveness and distinctiveness ratings were significantly negatively correlated (for all the images—$r_{290} = -0.604$, $p < 0.0001$; and for just the undistorted images—$r_{94} = -0.379$, $p < 0.0001$). Therefore, attractiveness and distinctiveness are negatively correlated for Japanese faces and raters, just as they are for Western and Chinese faces and raters.

[4] Distinctiveness ratings were collected primarily to examine the relationship between attractiveness and distinctiveness, but an ANOVA showed a significant main effect of averageness level ($F_{2,44} = 61.64$, $p < 0.0001$; Ms = 6.2, 4.7, 4.3, for low, normal, and high averageness images, respectively; all SEs = 0.1), confirming the validity of distinctiveness as a converse measure of averageness.

3.4 *Attractiveness of averaged composites compared with individual faces*
All the composites were rated as more attractive (female—$M = 7.5$; perfectly symmetric female—$M = 7.6$; male—$M = 7.9$; perfectly symmetric male—$M = 7.8$), averaging across raters, than any of the other images (maximum $M = 6.6$). Therefore, like Chinese and Western composites, Japanese composites were more attractive than any of their component faces and also more attractive than the high-average versions of those faces.

4 General discussion

Averageness and symmetry are attractive in Western faces and have been conjectured to be biologically based, evolved standards of beauty. A hallmark of such standards is that they show little variation across cultures. We found that average faces were attractive in both Chinese and Japanese cultures. Increasing the averageness of individual own-race faces increased their attractiveness, and decreasing the averageness of individual own-race faces decreased their attractiveness. Averaged composites were rated as more attractive than either the individual component faces or high-average versions of those component faces, whose shape averageness had been increased by 50%. Symmetric Japanese faces were also attractive to Japanese raters (Chinese raters and faces not tested). Perfectly symmetric versions of individual faces were rated as more attractive than the original faces. Furthermore, averageness and symmetry made independent contributions (no interaction) to attractiveness, as they do for Western faces and raters (Rhodes et al 1999b).

These results support the conjecture that averageness and symmetry are biologically based, evolved standards of beauty. We note, however, that our participants certainly had some exposure to Western culture and so could potentially have been influenced by Western standards of beauty. Such influences could plausibly contribute to a symmetry preference, although they seem less plausible as an account of the attractiveness of average own-race faces, because what is average differs for different populations. Therefore, even if our participants had learned that average Western faces are attractive, it is not obvious that this would generate a preference for average faces from their own race. Data from isolated populations could potentially rule out cross-cultural influences as an account for preferences for averageness and symmetry in non-Western cultures, but would be difficult to obtain.

We may, therefore, need to look to other lines of converging evidence to decide whether averageness and symmetry are biologically based standards of beauty. One such line of evidence comes from developmental studies. Preferences that emerge early in development are unlikely to be cultural artifacts and are more likely to have some biological basis.[5] We know that preferences for attractive faces emerge early (Geldart et al 1999; Langlois et al 1987, 1991; Samuels and Ewy 1985; Samuels et al 1994; Slater et al 1998), but we don't yet know whether these are preferences for averageness and symmetry.

There is some evidence that a preference for average faces may emerge early in development. Walton and Bower (1993) found that newborns (8 – 78 h old) preferred to look at averaged composites of faces they had seen than to composites of faces they had not seen. Rubenstein and his colleagues also found that six-month-old infants preferred to look at an averaged composite of 32 female faces than at unattractive faces (Rubenstein et al 1999). An innate ability to abstract the central tendency of a population of seen faces (Walton and Bower 1993), combined with a bias to fixate face-like patterns in the environment (Goren et al 1975; Johnson et al 1991; Valenza

[5] Although early emergence of a preference may indicate a biological basis, biologically based preferences need not necessarily emerge early. For example, a preference that evolved as an adaptation to mate choice might well emerge at puberty, triggered by hormonal changes associated with the onset of mating behaviour.

et al 1996) could generate an early preference for average faces. Experience would determine the actual facial characteristics that are preferred (ie ones that are typical of experienced faces) but the preference for averageness per se would be biologically based, in that it relies on information-processing mechanisms that evolved by either natural or sexual selection.

Much less is known about when a symmetry preference emerges. Samuels et al (1994) found no preference for symmetric faces in infants 4 to 15 months old, but the symmetric faces were made by reflecting each half of the face about the vertical midline, which introduces structural abnormalities and produces images that are unattractive even to adults (Kowner 1996; Langlois et al 1994; Rhodes et al 1999a). More research is needed to determine when a preference for symmetry emerges in development and what information-processing mechanisms are involved.

In the introduction, we suggested that preferences for averageness and symmetry may have evolved because these traits advertise aspects of mate quality, such as developmental stability and health. In this case, individuals with the preferences would enhance their reproductive success relative to individuals without them, and the preferences would be adaptations for identifying high-quality mates (Thornhill and Gangestad 1999). Partial support for this particular evolutionary account comes from evidence that facial averageness is a reliable indicator of health (Rhodes et al 2001c). Facial distinctiveness (a converse measure of averageness) of 17-year-olds was associated with poor childhood health in males, and poor current and adolescent (marginally) health in females, suggesting that averageness may indeed be a sign of health and developmental stability. Facial symmetry, however, did not advertise health in this sample. Shackelford and Larsen (1997) have reported correlations between facial symmetry and self-report measures of psychological, emotional, and physiological health, but most did not replicate across samples and many correlations were examined, raising the possibility of type I statistical errors. Overall, the evidence appears to be stronger for a link between facial averageness and health than between facial symmetry and health.

As noted above, evolved preferences could also reflect more general features of human information processing, such as a prototype-abstraction mechanism that helps us recognise and reason about category members. If a preference for average exemplars does result from such a mechanism, then it should not be restricted to potential mates. Halberstadt and Rhodes (2000) have recently reported that average exemplars are attractive for dogs, birds, and wristwatches, which suggests that humans may have a generic preference for averageness. Thus there is evidence that general information-processing mechanisms may contribute to the preference for average faces (over and above any preference for signs of mate quality), but the precise nature of those mechanisms and why they influence affective responses at all, remains to be determined.

Although not a main aim of the study, our results raise the question of what it is about average faces that makes them attractive. One possibility is that they are attractive because they capture the central tendency of a population of faces (eg Langlois et al 1994). On this account, the preference would be expected to vary with exposure to the relevant population. Yet in experiment 1 there was no advantage for high-average images or blends made with own-race over other-race averages and no preference for own-race (Chinese) averaged composites over other-race (Caucasian) or mixed-race composites. It is possible that the recent exposure of the Chinese participants to Caucasian faces (all had moved to Australia within six months of testing) offset the effect of their greater lifetime exposure to Chinese faces. The fact that very limited exposure to a set of faces can generate a preference for the averaged composite of those faces over a composite of unseen faces (Rhodes et al 2001a; Walton and Bower 1993) adds plausibility to this conjecture.

However, we should consider an alternative interpretation of the failure to find a clear preference for own-race averages. Average faces from any population lack obvious deformities, and look younger (Langlois et al 1994) and healthier than other faces (Rhodes et al 2001c). These are attractive traits (Rhodes et al 2001c; Zebrowitz et al 1993) that can probably be perceived with little or no experience of the specific population for which a face is average. Future studies will be needed to determine how much of the appeal of average faces comes from traits like youthfulness and healthiness and how much comes from displaying typical characteristics of a familiar population.

Acknowledgements. This research was supported by the Australian Research Council, the University of Western Australia, and the Science and Technology Agency of Japan (Neurinformatics Research in Vision, Special Co-ordination Funds for Promoting Science and Technology). We thank Tuart Hill and Beaufort Colleges in Perth for providing testing space and allowing their students to participate in experiment 1 and Yoshi Nagao for testing subjects in experiment 2. We also thank Leigh Simmons and two anonymous reviewers for comments on an earlier version of the manuscript.

References

Benson P, Perrett D, 1992 "Face to face with the perfect image" *New Scientist* **1809** 32 – 35

Bruce V, Burton M A, Dench N, 1994 "What's distinctive about a distinctive face?" *Quarterly Journal of Experimental Psychology A* **47** 119 – 141

Bruce V, Coombes A, Richards R, 1993 "Describing the shapes of faces using surface primitives" *Image and Vision Computing* **11** 353 – 363

Endler J A, Basolo A L, 1998 "Sensory ecology, receiver biases and sexual selection" *Trends in Ecology & Evolution* **13** 415 – 420

Enquist M, Arak A, 1994 "Symmetry, beauty and evolution" *Nature* **372** 169 – 172

Enquist M, Johnstone R A, 1997 "Generalization and the evolution of symmetry preferences" *Proceedings of the Royal Society of London, Series B* **264** 1345 – 1348

Geldart S, Maurer D, Carney K, 1999 "Effects of eye size on adults' aesthetic ratings of faces and 5-month-olds' looking times" *Perception* **28** 361 – 374

Goren C, Sarty M, Wu P, 1975 "Visual following and pattern discrimination of face-like stimuli by newborn infants" *Pediatrics* **56** 544 – 549

Grammer K, Thornhill R, 1994 "Human (Homo sapiens) facial attractiveness and sexual selection: The role of symmetry and averageness" *Journal of Comparative Psychology* **108** 233 – 242

Halberstadt J, Rhodes G, 2000 "The attractiveness of non-face averages: Implications for an evolutionary explanation of the attractiveness of average faces" *Psychological Science* **11** 285 – 289

Johnson M H, Dziurawiec S, Ellis H, Morton J, 1991 "Newborns' preferential tracking of face-like stimuli and its subsequent decline" *Cognition* **40** 1 – 19

Johnston R A, Ellis H D, 1995 "The development of face recognition", in *Cognitive and Computational Aspects of Face Processing: Explorations in Face Space* Ed. T Valentine (London: Routledge) pp 1 – 23

Johnstone R A, 1994 "Female preference for symmetrical males as a by-product of mate recognition" *Nature* **372** 172 – 175

Jones D, Hill K, 1993 "Criteria of facial attractiveness in five populations" *Human Nature* **4** 271 – 296

Kowner R, 1996 "Facial asymmetry and attractiveness judgment in developmental perspective" *Journal of Experimental Psychology: Human Perception and Performance* **22** 662 – 675

Langlois J H, Roggman L A, 1990 "Attractive faces are only average" *Psychological Science* **1** 115 – 121

Langlois J H, Kalakanis L E, Rubenstein A J, Larson A D, Hallam M J, Smoot M T, 2000 "Maxims and myths of beauty: A meta-analytic and theoretical review" *Psychological Bulletin* **126** 390 – 423

Langlois J H, Ritter J M, Roggman L A, Vaughn L S, 1991 "Facial diversity and infant preferences for attractive faces" *Developmental Psychology* **27** 79 – 84

Langlois J H, Roggman L A, Casey R J, Ritter J M, Reiser-Danner L A, Jenkins V Y, 1987 "Infant preferences for attractive faces: Rudiments of a stereotype?" *Developmental Psychology* **23** 363 – 369

Langlois J H, Roggman L A, Musselman L, 1994 "What is average and what is not average about attractive faces?" *Psychological Science* **5** 214 – 220

Lee K, Byatt G, Rhodes G, 2000 "Caricature effects, distinctiveness and identification: Testing the face-space framework" *Psychological Science* **11** 379 – 385
Light L L, Hollander S, Kayra-Stuart F, 1981 "Why attractive people are harder to remember" *Personality and Social Psychology Bulletin* **7** 269 – 276
Mealey L, Bridgestock R, Townsend G C, 1999 "Symmetry and perceived facial attractiveness: A monozygotic co-twin comparison" *Journal of Personality and Social Psychology* **76** 151 – 158
Møller A P, Swaddle J P, 1997 *Asymmetry, Developmental Stability and Evolution* (Oxford: Oxford University Press)
Perrett D I, Burt D M, Penton-Voak I S, Lee K J, Rowland D A, Edwards R, 1999 "Symmetry and human facial attractiveness" *Evolution and Human Behavior* **20** 295 – 307
Perrett D I, Lee K J, Penton-Voak I, Rowland D, Yoshikawa S, Burt D M, Henzi S P, Castles D, Akamatsu S, 1998 "Effects of sexual dimorphism on facial attractiveness" *Nature* **394** 884 – 887
Pinker S, 1994 *The Language Instinct: The New Science of Language and Mind* (London: Penguin Press)
Rhodes G, Tremewan T, 1996 "Averageness, exaggeration, and facial attractiveness" *Psychological Science* **7** 105 – 110
Rhodes G, Halberstadt J, Brajkovich G, 2001a "Generalization of mere exposure effects in social stimuli" *Social Cognition* **19** 369 – 382
Rhodes G, Harwood K, Yoshikawa S, Nishitani M, McLean I, 2001b "The attractiveness of average facial configurations: Cross-cultural evidence and possible biological basis", in *Facial Attractiveness: Evolutionary, Cognitive and Social Perspectives*, volume 1 of *Advances in Visual Cognition* Eds G Rhodes, L A Zebrowitz (Westport, CT: Ablex Publishing) forthcoming
Rhodes G, Hickford C, Jeffery L, 2000c "Sex-typicality and attractiveness: Are supermale and superfemale faces super-attractive?" *British Journal of Psychology* **91** 125 – 140
Rhodes G, Proffitt F, Grady J M, Sumich A, 1998 "Facial symmetry and the perception of beauty" *Psychonomic Bulletin & Review* **5** 659 – 669
Rhodes G, Roberts J, Simmons L, 1999a "Reflections on symmetry and attractiveness" *Psychology, Evolution & Gender* **1** 279 – 295
Rhodes G, Sumich A, Byatt G, 1999b "Are average facial configurations only attractive because of their symmetry?" *Psychological Science* **10** 52 – 58
Rhodes G, Zebrowitz L A, Clark A, Kalick S M, Hightower A, McKay R, 2001c "Do facial averageness and symmetry signal health?" *Evolution and Human Behavior* **22** 31 – 46
Rubenstein A J, Kalakanis L, Langlois J H, 1999 "Infant preferences for attractive faces: A cognitive explanation" *Developmental Psychology* **15** 848 – 855
Rubenstein A J, Langlois J H, Roggman L, 2001 "What makes a face attractive and why?", in *Facial Attractiveness: Evolutionary, Cognitive and Social Perspectives* volume 1 of *Advances in Visual Cognition* Eds G Rhodes, L A Zebrowitz (Westport, CT: Ablex Publishing) forthcoming
Samuels C A, Ewy R, 1985 "Aesthetic perception of faces during infancy" *British Journal of Developmental Psychology* **3** 221 – 228
Samuels C A, Butterworth G, Roberts T, Graupner L, Hole G, 1994 "Facial aesthetics: Babies prefer attractiveness to symmetry" *Perception* **23** 823 – 831
Shackelford T K, Larsen R J, 1997 "Facial asymmetry as an indicator of psychological, emotional, and physiological distress" *Journal of Personality and Social Psychology* **72** 456 – 466
Slater A, Schulenburg C von der, Brown E, Badenoch M, Butterworth G, Parsons S, Samuels C, 1998 "Newborn infants prefer attractive faces" *Infant Behavior and Development* **21** 345 – 354
Symons D, 1979 *The Evolution of Human Sexuality* (Oxford: Oxford University Press)
Thornhill R, Gangestad S W, 1993 "Human facial beauty" *Human Nature* **4** 237 – 269
Thornhill R, Gangestad S W, 1999 "Facial attractiveness" *Trends in Cognitive Sciences* **3** 452 – 460
Thornhill R, Møller A P, 1997 "Developmental stability, disease and medicine" *Biological Reviews* **72** 497 – 548
Valenza E, Simion F, Cassia V M, Umilta C, 1996 "Face preference at birth" *Journal of Experimental Psychology: Human Perception and Performance* **22** 892 – 903
Walton G E, Bower T G R, 1993 "Newborns form 'prototypes' in less than 1 minute" *Psychological Science* **4** 203 – 205
Zebrowitz L A, 1997 *Reading Faces* (Boulder CO: Westview Press)
Zebrowitz L A, Olson K, Hoffman K, 1993 "Stability of babyfacedness and attractiveness across the life span" *Journal of Personality and Social Psychology* **64** 453 – 466
Zebrowitz L A, Voinescu L, Collins M A, 1996 "'Wide-eyed' and 'crooked-faces': Determinants of perceived and real honesty across the life span" *Personality and Social Psychology Bulletin* **22** 1258 – 1269