Perception, 2007, volume 36, pages 1813 – 1820

doi:10.1068/p5601

Facial averageness and attractiveness in an isolated population of hunter-gatherers

Coren L Apicella, Anthony C Little¶, Frank W Marlowe§
Department of Anthropology, Harvard University, Cambridge, MA 02138, USA; ¶ University of Stirling, Stirling FK9 4LA, Scotland, UK; § Florida State University, Tallahassee, FL 32306, USA;
e-mail: apicella@fas.harvard.edu
Received 28 March 2006, in revised form 31 January 2007

Abstract. Average faces possess traits that are common to a population. Preferences for averageness have been found in several types of study of both real and computer-manipulated faces. Such preferences have been proposed to be biologically based and thus should be found across human populations, though cross-cultural evidence to date has been limited. In this study we examined preferences for averageness in both the West and in an isolated hunter-gatherer society, the Hadza of Northern Tanzania in Africa. We show that averageness is generally preferred across faces and cultures, but there were no significant preferences for averageness in European faces by Hadza judges. The different visual experience of the two cultures may explain the differences in preferences. While Westerners have visual experience of both European and African faces, the Hadza are limited in their experience of European faces, potentially leading to a lack of preference for averageness in this group because of the lack of a representation of the `norm' of European faces.

1 Introduction

Faces which possess traits with mathematically average values for a population are said to be average. Average faces are perceived to be attractive (for review see Rhodes 2006). Computer-generated composite images of both male and female faces are judged as more attractive than almost all individual component faces (Langlois and Roggman 1990; Little and Hancock 2002; Rhodes et al 1999). In addition, individual faces are perceived as more attractive when their configurations are manipulated to be closer to average configurations for the same sex (Rhodes and Tremewan 1996; Rhodes et al 1999, 2001a). Work on unmanipulated faces has shown that average or typical faces are rated as more attractive than distinctive faces (see Light et al 1981; Morris and Wickham 2001; Rhodes et al 1999; Rhodes and Tremewan 1996). In a meta-analysis of 20 studies examining the role of averageness on perceived attractiveness, Rhodes (2006) reported a large effect of averageness on attractiveness.

It has been suggested that the preference for averageness is based on associations with quality making such preferences and adaptation to select mates (Koeslag 1990; Koeslag and Koeslag 1994; Thornhill and Gangestad 1999). Whether this preference is biologically based has yet to be firmly established. It has also been suggested that such preferences may be an artifact of perceptual and/or cognitive information-processing mechanisms (Enquist and Arak 1994; Halberstadt and Rhodes 2000). Both explanations assume that universal biologically based mechanisms are involved. A third explanation of averageness preferences arising from random cultural variability would predict no biological mechanisms that generate preferences for averageness and no agreement on the attractiveness of averageness.

One way to investigate this claim further is by examining whether preferences for averageness emerge early in development. If so, it is less likely that the preference is culturally acquired and more likely to be rooted in our biology (Rhodes et al 2002). Although many researchers have found that infants, even as early as 2-months of age, look longer at attractive faces than at unattractive faces, as judged by adults (for review

see Rhodes et al 2002), few researchers have examined infants' preferences for averageness itself. Rubenstein et al (1999) found that 6-month-old infants look longer at averaged female faces (composite faces made by averaging 32 individual faces) than at an unattractive female face. However, more recently it has been found that, although infants 5^8 months old were able to discriminate between levels of averageness, they actually looked longer at less-average versions of faces (Rhodes et al 2002). Because infants also look longer at stimuli that are unusual or novel (Baillargeon et al 1990; Spelke 1985) it is difficult to interpret these results, since low-average faces are, by definition, unusual (Rhodes et al 2002). Therefore, other evidence is needed to support the notion that the preference for averageness is not dependent on culture. One way to do this is to examine cross-cultural preferences for average faces. Finding that the preference is universal would provide additional support for this notion (Rhodes et al 2001a).

Most work examining the effect of averageness on attractiveness has used both European faces and Western raters. However, there have been some exceptions. Rhodes et al (2001a) found that increasing the averageness of both individual Chinese and Japanese faces increased the attractiveness of those faces to Chinese and Japanese raters, respectively. Jones and Hill (1993) examined preferences for averageness in Americans, Brazilians, Russians, the Ache of Paraguay, and the Hiwi of Venezuela, using faces of Americans, Brazilians, and Ache, and found no clear association between averageness and attractiveness. Only one significant association was found for the Ache. However, age was confounded with averageness in this group and as a result the preference may instead be for youth (Rhodes et al 2001a). Furthermore, it was suggested that the methodology and/or stimuli may have been flawed since they failed to produce significant results in the Western populations where there is a known effect.

In this study, we investigated opposite-sex preferences for averageness in both European and African faces by both Western and African raters in order to assess whether there is cross-cultural agreement. The African population we examined is the Hadza, a relatively isolated society of hunter-gatherers living in Tanzania. This population is ideal for testing attractiveness preferences since their exposure to Western culture and media is negligible. Although same-sex preferences would have also been informative, we decided to limit our data collection to opposite-sex preferences in order to minimize interview time with Hadza participants.

2 Methods

2.1 *Participants*

The Hadza are mobile hunter-gatherers who live in a savanna-woodland habitat in Northern Tanzania. They number about 1000 and live in camps that average about 30 individuals. Camp membership is fluid, with people continually changing residence. Camps are moved to a different location about once every 1^2 months. Women dig for wild tubers, gather berries, and collect baobab fruit, while men hunt mammals and birds with bow and arrow and collect honey.

Marriages are not arranged and both sexes choose whom they want to marry. The median age at first marriage is 21 years for men and 17 years for women (Marlowe 2004). Monogamy is dominant, with about 4% of men having two wives at the same time (Marlowe 2003). Very rarely is there intermarriage with other ethnic groups, but, when it does occur, it is usually Hadza women marrying into a different ethnic group; 4% of Hadza women marry non-Hadza partners (Blurton Jones et al 2005), but many later return to live in Hadza camps.

The Hadza participants were forty-two (twenty-one male) adults from several camps. All were fluent in Swahili as a second language, and Apicella conducted the experiment privately with each individual in Swahili. The Western participants were

recruited over the internet via a mailing list. Western participants included two hundred and fifty-two (one hundred and forty-six male) adults who were tested with an online study. The Hadza men were 21 ^ 56 years old (mean = 35.33 years, SD = 10.40 years) and the Hadza women were 23 ^ 54 years old (mean = 32.14 years, SD = 8.52 years). Western men were 18 ^ 44 years old (mean = 21.99 years, SD = 4 years) and Western women were 18 ^ 60 years old (mean = 21.94 years, SD = 4.67 years).

2.2 *Stimuli*

Four pairs of computer-morphed average and less-average composite faces were made for each race of face. The starting images were 20 male and 20 female young adult white British individuals and 20 male and 20 female young adult Hadza individuals. Given the nature of the study, photographic conditions differed between the image sets. European images were taken under standardised lighting at the same distance from the camera. Hadza images were taken under natural lighting conditions in the field at a variable distance from the camera. All images were standardised on interocular distance before manipulations were made, which normalised head size within all the images used. As all manipulations involved shape only, the differing lighting in the Hadza images is of minimal consequence, influencing the color of the best composites only.

Composite face images become more average the more faces are used to create them. We made our more-average face by blending together all 20 of our starting faces separately for the two sexes and Hadza and European faces (figure 1). Less-average faces were made up of 5 faces each (randomly allocated from the original 20 images). This created four 5-face composite images made up from a unique combination of faces. Test pairs were then each of the four 5-face composites against the 20-face composite for each race and sex. This allowed us to test every face in the 20-face composite against itself to ensure that our test reflected preferences for averageness of shape alone.

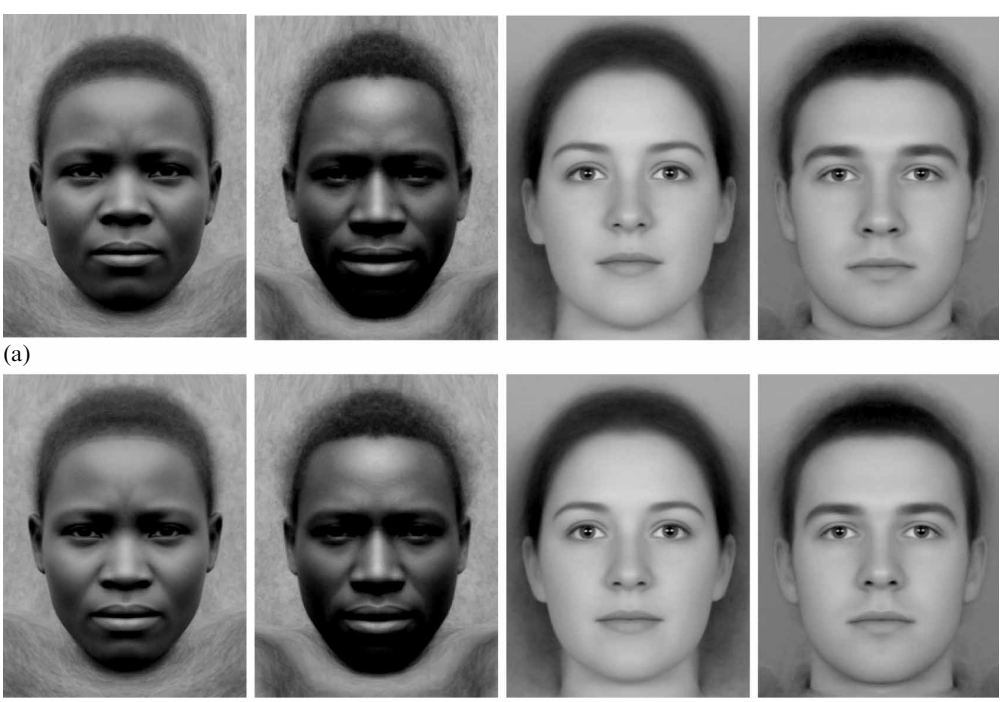

Figure 1. (a) Female and male Hadza and European 5-face composites. (b) Female and male Hadza and European 20-face composites. This figure is available in colour at http://www.perceptionweb.com/misc/p5601/.

For instance, had we used only one 5-face composite, the faces in it may on average have been more attractive or less attractive than the faces in the 20-face composite.

To make each composite, 174 feature points were delineated on each face image in the groups from which the average male and female shapes were calculated. Each individual image was then warped to the relevant average shape, and superimposed one on top of the next to create a photorealistic composite image reflecting both the average shape and average color of its constituent images (see Benson and Perrett 1993; Tiddeman et al 2001). The images were then made perfectly symmetrical by combining them with their mirror images and then the average colour information from the relevant 20-face composite (Hadza or European) was warped into each of the four 5-face composites so that images differed in averageness of shape alone.

2.3 *Procedure*

Each Hadza participated in the experiment individually and privately. Two opposite-sex face pairs were placed randomly in a horizontal row in front of the participant. Each 5-face composite was tested against the 20-face composite. The participant was asked to choose the more attractive face out of the two; 8 face pairs in total were shown to each participant. After each trial the photographs were shuffled and the procedure was repeated until all 4×5 composites were tested against the 1×20 composite for each race. Western participants were shown the same images and given the same instructions as the Hadza; however, the experiment was self-administered via the internet and the image pairs were randomized by using computer generated numbers. In addition, within each trial presentation, whether the more average image was presented on the left or the right side of the computer screen was also randomized.

2.4 *Analyses*

We calculated the proportion of 20-face composite images that were selected as most attractive by each rater over each of the 5-face composite images. These proportions were conducted separately for Hadza images and European images. We then conducted one-sample t-tests to compare the resulting distributions against the random expectation of 0.50. In order to rule out the possibility of a familiarity effect (from repeated viewings of the 20-face composites) driving the preference for the 20-face composites on each subsequent trial in both the European and Hadza face condition, we performed Friedman's test, a non-parametric equivalent of a repeated-measures ANOVA, which tests for significant differences between the four trials. We would expect preferences for the 20-face composite to increase or decrease over the four trials if repeated exposure to the 20-face composite was having an effect. For this test Western and Hadza raters' scores were pooled.

3 Results

Both Western ($t_{251} = 5.087$, $p < 0.001$) and Hadza ($t_{41} = 3.56$, $p = 0.001$) raters selected the 20-face Hadza composite as more attractive than all the 5-face Hadza composites (figure 2). Although Western raters selected the 20-face European composite as significantly more attractive than the 5-face European composite ($t_{251} = 6.35$, $p < 0.001$), the Hadza raters did not ($t_{41} = 0.286$, $p = 0.777$).

The Friedman test revealed significant differences in the distribution of the ranks of each trial for the European faces ($\chi^2_3 = 11.42$, $p = 0.01$) and Hadza faces ($\chi^2_3 = 41.16$, $p < 0.001$). Although there are significant differences in the distributions between the trials in both conditions, no trends emerged suggesting a familiarity effect (see figure 3). In fact, if repeated viewing of the 20-face composites had been driving the preference for the 20-face composite, we should have seen participants' preference for the 20-face image increase over subsequent trials and it did not.

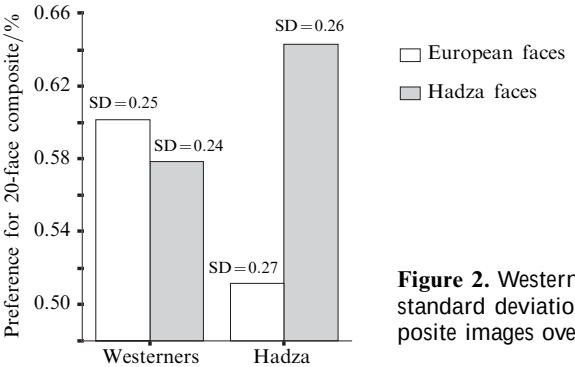

Figure 2. Western and Hadza raters' mean preference and standard deviation for European and Hadza 20-face composite images over respective 5-face composite images.

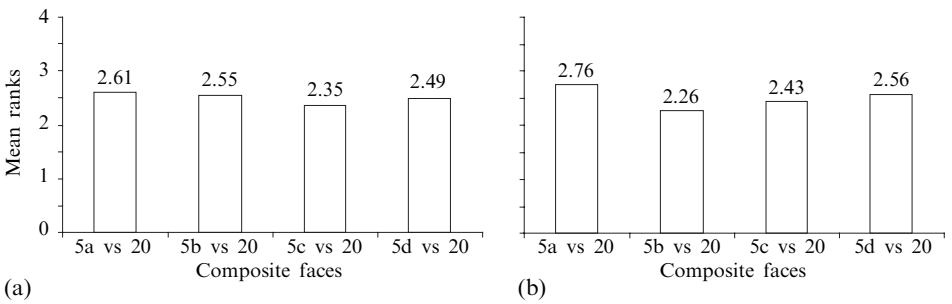

Figure 3. Mean ranks of participants' choice for each trial in (a) the European face condition and (b) the Hadza face condition with 5a representing the first 5-face composite shown to participants and 5d representing the last.

4 Discussion

The current study shows cross-cultural agreement on the attractiveness of averageness, though this was qualified by some differences over the attractiveness of average European faces. Westerners found increased averageness to be attractive in both European and African faces. The Hadza, however, did not find increased averageness in European faces as attractive, although they found it attractive in Hadza faces.

Two main explanations why average faces are preferred have been proposed. It has been suggested that the preference is universal, biologically based (Langlois and Roggman 1990), and may have evolved for selecting optimal mates (eg Koeslag 1990; Koeslag and Koeslag 1994; Thornhill and Gangestad 1999). Alternatively, the preference may be an artifact of a perceptual and/or cognitive information-processing mechanism, which would also be biologically based (Enquist and Arak 1994; Halberstadt and Rhodes 2000). Both explanations can assume that visual experience influences the formation of prototypical faces. Certainly recent exposure to faces can influence perceptions of both normality and attractiveness (see Little et al 2005; Rhodes et al 2003; Webster and MacLin 1999). The third explanation of averageness preferences arising owing to cultural variability would predict no biological mechanisms that generate preferences for averageness and is inconsistent with findings of cross-cultural agreement that averageness is attractive.

Visual experience may explain why averageness is not seen as attractive in European faces by the Hadza. If there exist distinct representations of race, in the same way as there appear to be distinct representations to male and female (Little et al 2005), then differences may be predicted as Westerners are exposed to both European and African faces, whereas the Hadza have virtually no contact with white people, with the exception of the few researchers who study them and, more recently, the occasional tourist. Therefore, it is unlikely that they have enough experience to form a fully applicable mental representation

of what an average European would look like, and so comparison to average representations generates no preference when the experienced `norm' may not necessarily represent average. Thus, differing visual experience is a parsimonious explanation for the observed results of this study, though we cannot rule out other reasons for the absence of preferences for averageness in European faces by the Hadza.

Studying preferences for averageness with the use of composite images has been criticized by some researchers, because facial asymmetry is reduced as more faces are blended together (Alley and Cunningham 1991). This is not true of our stimuli, which were all made perfectly symmetric. Alley and Cunningham (1991) suggest that smoothing of skin texture also contributes to the attractiveness of average faces (also Little and Hancock 2002). Again, this is controlled in our study, as we morphed an average skin texture into each face of our average faces so that the images differed in shape alone. If other incidental properties resulting from combining faces were driving the preference for averageness, then we should have expected the Hadza to also prefer increased averageness in the European faces, but they did not. Our study then provides support that average composites are attractive because they are average in shape.

Rhodes et al (2005) examined whether the `mere exposure effect', that is the phenomenon where individuals prefer stimuli that are familiar to them, could be generalized to preferences for averageness. Although these authors found that seen faces were regarded as more attractive than unseen faces, the findings could not be generalized to averaged composites of seen faces. If we had found that the Hadza prefer average European faces, this would have helped to further rule out the mere exposure effect in driving preferences for averageness, because the Hadza are not familiar with European faces. Instead, we found that the Hadza did not prefer average European faces and, therefore, we cannot rule out whether the preference for averageness is a preference for familiarity, although we are inclined to suspect that the preference is likely for averageness per se. If average faces are preferred independently of familiarity (Halberstadt 2006; Rhodes et al 2005) and independently of other traits associated with averageness, such as symmetry, then why at the proximate level are they viewed as attractive? Our results suggest that the notion that averageness is attractive because of random cultural factors is unlikely, and that such preferences are likely hard-wired.

Ultimately, average traits may be attractive because such preferences are adaptive or because they arise as byproducts of perceptual mechanisms. There are theoretical reasons why averageness would be adaptive, in that stabilizing or normalizing selection would have eliminated deleterious mutations that move traits away from the mean (Koeslag 1990; Koeslag and Koeslag 1994). Therefore, individuals should prefer physical traits in the opposite sex that are close to the mean of the population. If averageness is an indicator of good genes, and therefore mate quality, we should expect individuals who possess this trait to have enhanced health and reproductive success. Rhodes et al (2001b) examined photos and health scores for a sample of 471 male and female adolescents taken from the Intergenerational Studies archive. They found that facial distinctiveness of the participants was negatively associated with childhood health in males and with current health in females. These associations were modest and more work is needed to determine whether averageness is an indicator of health. This relationship may be tenuous in the industrialized world because of rapid advances in both medical and cosmetic technology, and the limited exposure of populations to disease and stressors. To date, the relationship between averageness and reproductive success has yet to be examined. Natural fertility of populations such as the Hadza would be ideal to test whether averageness or other measures of attractiveness are related to health or reproductive success. This type of work is critical in distinguishing between evolutionary psychologists' good-genes account and the cognitive psychologists' perceptual-bias account of preferences for average faces.

Since the Hadza are living in conditions more similar to those of our ancestors and have little exposure to the world outside their area, they provide the strongest support to date for the argument that preference for averageness is universal and not arbitrarily dependent on culture. This finding, coupled with other findings which indicate that preference for averageness may emerge early in development, helps strengthen the claim that this preference is rooted in biology. This study also supports the notion that average faces are attractive because they reflect the mathematical average of faces, rather than other traits associated with averageness. In view of all this, it is likely that the formation of facial prototypes is dependent upon environmental input that is then processed by a hard-wired natural, statistical, calculating mechanism in the brain. Indeed, much evidence suggests that humans are good intuitive statisticians (Cosmides and Tooby 1996). Therefore, we conclude that the preference for averageness itself is biologically based, but it is experience that determines our internal prototypes and influences which faces we find most attractive in the opposite-sex.

Acknowledgments. This paper profited from comments made by three anonymous reviewers. We would also like to thank the Tanzanian Commission on Science and Technology for permission to conduct research, Professor Audax Mabulla of the University of Dar es Salaam, Archeology Unit, for his assistance, and the Hadza for their patience. This research was made possible by a National Science Foundation grant (No 0242455) to Frank Marlowe and the Department of Biological Anthropology at Harvard. Anthony Little is supported by a Royal Society University Research Fellowship.

References

Alley T R, Cunningham M R, 1991 ``Averaged faces are attractive, but very attractive faces are not average'' *Psychological Science* **2** 123 ^ 125

Baillargeon R, Graber M, DeVos J, Black J, 1990 ``Why do young infants fail to search for hidden objects?'' *Cognition* **36** 255 ^ 284

Benson P J, Perrett D I, 1993 ``Extracting prototypical facial images from exemplars'' *Perception* **22** 257 ^ 262

Blurton Jones N G, Hawkes K, O'Connell J F, 2005 ``Older Hadza men and women as helpers: Residence data'', in *Hunter-Gatherer Childhoods: Evolutionary, Development and Cultural Perspectives* Eds B S Hewlett, M E Lamb (New Brunswick, NJ: Transaction) pp 214 ^ 236

Cosmides L, Tooby J, 1996 ``Are humans good intuitive statisticians after all? Rethinking some of the conclusions from the literature on judgment under uncertainty'' *Cognition* **58** 1 ^ 73

Enquist M, Arak A, 1994 ``Symmetry, beauty and evolution'' *Nature* **372** 169 ^ 172

Halberstadt J, 2006 ``The generality and ultimate origins of the attractiveness of prototypes'' *Personality and Social Psychology Review* **10** 166 ^ 183

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

Jones D, Hill K, 1993 ``Criteria of facial attractiveness in five populations'' *Human Nature* **4** 271 ^ 296

Koeslag J H, 1990 ``Koinophilia groups sexual creatures into species, promotes stasis, and stabilizes social behavior'' *Journal of Theoretical Biology* **144** 15 ^ 35

Koeslag J H, Koeslag P D, 1994 ``Koinophilia'' *Journal of Theoretical Biology* **167** 55 ^ 65

Langlois J H, Roggman L A, 1990 ``Attractive faces are only average'' *Psychological Science* **1** 115 ^ 121

Light L L, Hollander S, Kayra-Stuart F, 1981 ``Why attractive people are harder to remember'' *Personality and Social Psychology Bulletin* **7** 269 ^ 276

Little A C, DeBruine L M, Jones B C, 2005 ``Sex-contingent face after-effects suggest distinct neural populations code male and female faces'' *Proceedings of the Royal Society of London, Series B* **272** 2283 ^ 2287

Little A C, Hancock P J B, 2002 ``The role of masculinity and distinctiveness in judgments of human male facial attractiveness'' *British Journal of Psychology* **93** 451 ^ 464

Marlowe F W, 2003 ``A critical period for provisioning by Hadza men: implications for pair bonding'' *Evolution and Human Behavior* **24** 217 ^ 229

Marlowe F W, 2004 ``Mate preferences among Hadza hunter-gatherers'' *Human Nature* **15** 365 ^ 376

Morris P E, Wickham L H V, 2001 ``Typicality and face recognition: a critical re-evaluation of the two factor theory'' *Quarterly Journal of Experimental Psychology A* **54** 863 ^ 877

Rhodes G, 2006 ``The evolutionary psychology of facial beauty'' *Annual Review of Psychology* **57** 199 ^ 236

Rhodes G, Geddes K, Jeffery L, Dziurawiec S, Clark A, 2002 ``Are average and symmetric faces attractive to infants? Discrimination and looking preferences'' *Perception* **31** 315 ^ 321

Rhodes G, Halberstadt J, Jeffery L, Palermo R, 2005 ``The attractiveness of average faces is not a generalized mere exposure effect'' *Social Cognition* **23** 205 ^ 217

Rhodes G, Jeffery L, Watson T L, Clifford C W G, Nakayama K, 2003 ``Fitting the mind to the world: Face adaptation and attractiveness aftereffects'' *Psychological Science* **14** 558 ^ 566

Rhodes G, Sumich A, Byatt G, 1999 ``Are average configurations only attractive because of their symmetry?'' *Psychological Science* **10** 52 ^ 58

Rhodes G, Tremewan T, 1996 ``Averageness, exaggeration and facial attractiveness'' *Psychological Sciences* **7** 105 ^ 110

Rhodes G, Yoshikawa S, Clark A, Lee K, McKay R, Akamatsu S, 2001a ``Attractiveness of facial averageness and symmetry in non-Western cultures: In search of biologically based standards of beauty'' *Perception* **30** 611 ^ 624

Rhodes G, Zebrowitz L A, Clark A, Kalick S M, Hightower A, McKay R, 2001b ``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* **35** 848 ^ 855

Spelke E, 1985 ``Preferential looking methods as tools for the study of cognition in infancy'', in *Measurement of Audition and Vision in the First Year of Postnatal Life* Eds G Gottlieb, N Krasnegor (Norwood, NJ: Ablex) pp 323 ^ 363

Thornhill R, Gangestad S W, 1999 ``Facial attractiveness'' *Trends in Cognitive Sciences* **3** 452 ^ 460

Tiddeman B P, Burt D M, Perrett D I, 2001 ``Prototyping and transforming facial textures for perception research'' *IEEE Computer Graphics and Applications* **21** 42 ^ 50

Webster M A, MacLin O H, 1999 ``Figural after-effects in the perception of faces'' *Psychonomic Bulletin and Review* **6** 647 ^ 653

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION
VOLUME 36 2007

www.perceptionweb.com