STABILIZING AND DIRECTIONAL SELECTION ON FACIAL PAEDOMORPHOSIS

Averageness or Juvenilization?

Paul Wehr

University of British Columbia

Kevin MacDonald, Rhoda Lindner, and Grace Yeung California State University, Long Beach

Averageness is purportedly the result of stabilizing selection maintaining the population mean, whereas facial paedomorphosis is a product of directional selection driving the population mean towards an increasingly juvenile appearance. If selection is predominantly stabilizing, intermediate phenotypes reflect high genetic quality and mathematically average faces should be found attractive. If, on the other hand, directional selection is strong enough, extreme phenotypes reflect high genetic quality and juvenilized faces will be found attractive. To compare the effects of stabilizing and directional selection on facial paedomorphosis (juvenilization), graphic morphing and editing techniques were used to alter the appearance of composite faces to make them appear more or less juvenile. Both facial models and judges of attractiveness were from the CSU-Long Beach campus. Although effect sizes for both preferences were large, the effect for averageness was nearly twice that found for juvenilization, an indication that stabilizing selection influences preferences for facial paedomorphosis more so than directional selection in contemporary humans.

KEY WORDS: Averageness; Facial Attractiveness; Paedomorphosis

Received: August 31, 2000; accepted December 7, 2000.

Address all correspondence to Paul Wehr, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, Canada, V6T 1W4. E-mail: pwehr@attglobal.net

Copyright 2001 by Walter de Gruyter, Inc., New York Human Nature, Vol. 12, No. 4, pp. 383–402. Given that genotypes vary in their ability to survive and reproduce successfully in a given environment, sexually reproducing organisms should prefer consorts with high genetic quality to maximize the fitness of their offspring (Andersson 1994). Genetic quality, however, must be inferred from an organism's phenotype since it can't be assessed directly. Attraction, therefore, may be an adaptive preference for phenotypes that imply high genetic quality (Barber 1995; Jones 1996; Symons 1995).

Fluctuating asymmetry (FA), for instance, is positively related to developmental instability and negatively related to attractiveness ratings and sexual activity (Gangestad and Thornhill 1997; Gangestad et al. 1994; Grammar and Thornhill 1994; Thornhill and Gangestad 1994). Likewise, various samples from both the United States and abroad have found female waist-to-hip-ratios (WHR) that are consistent with good health and high fertility to be the most attractive (Singh 1993, 1994; Singh and Luis 1995).¹

While FA and WHR have gained prominence, two variables related to *facial* attractiveness have remained less conspicuous: averageness—facial features reflecting the mathematical average of a population (as opposed to those that are ordinary or common in occurrence), and facial paedo-morphosis—facial features reflecting slowed or truncated development (juvenilization). Although little is known regarding the specific pressures that select for facial averageness and paedomorphosis in humans, the selection forces proposed to account for each are antagonistic. Averageness is purportedly the result of stabilizing selection maintaining the population mean (Langlois and Roggman 1990), whereas facial paedomorphosis is a product of directional selection driving the population mean towards an increasingly juvenile appearance (Alley and Cunningham 1991).

Stabilizing selection occurs in a stable environment (not necessarily a homogeneous one) where organisms are exposed to predictable conditions for long periods of time and become well adapted to their circumstances. Phenotypes close to the population mean are well suited for survival and reproduction whereas phenotypes deviating from the mean, whether because of sexual recombination or mutation, tend to be deleterious. Consequently, natural selection favors intermediate phenotypes over extreme forms. This process, sometimes referred to as centripetal or normalizing selection, is constantly trimming the tails of the population distribution and preserving the population mean (Schmalhausen 1949; Strickberger 1996).

Long-term, uniform changes in the environment, however, may shift the optimal point of a trait's functional range in one direction. Subsequently, extreme phenotypes may be better suited to the new environment than those close to the population mean, which have lost some of their fitnessenhancing qualities. Under these circumstances, recombination and mutation provide the genetic clay necessary for directional selection, extreme phenotypes are favored over intermediate ones, and the population mean shifts towards one tail of the distribution (Strickberger 1996). Although directional selection is more commonly associated with Darwinian evolution, stabilizing selection is actually more commonly occurring (Schmalhausen 1949).

Application of these selection processes to the attractiveness of facial paedomorphosis leads to opposing predictions. If selection is predominantly stabilizing, intermediate phenotypes reflect high genetic quality and mathematically average faces should be found attractive. If, on the other hand, directional selection is strong enough, extreme phenotypes reflect high genetic quality and juvenilized faces will be found attractive. Ultimately, the answer to this question should indicate whether the environment selecting for facial paedomorphosis has stabilized or is in transition.

AVERAGE FACES

Evidence for the attractiveness of average faces comes from studies using composites, which tend to reflect the mathematical average of a group of faces. Attempting to depict assorted social groups graphically, Galton (1878) superimposed photographic exposures over one another to produce a composite appearance. He was unable to discriminate between groups based on their composites, but he did notice that composites were unusually attractive. Building on this observation, Langlois and Roggman (1990) made composite faces using a computer averaging technique. They found composite faces to be significantly more attractive than the individual faces used to construct the composite. Furthermore, there was a significant linear effect demonstrating the influence of their averaging technique: The more faces included in the composite, the more attractive the face. The researchers concluded that "attractiveness is only (mathematically) average" and that physical attractiveness is governed by a preference for the population mean resulting from stabilizing selection (Langlois and Roggman 1990; Langlois et al. 1994).

Grammar and Thornhill (1994) were able to replicate these results for female, but not male, faces. Female composites in their study were judged to be more attractive, less dominant, sexier, and healthier than the corresponding individual faces whereas male analogues were judged to be less attractive, less dominant, less sexy, and less healthy. The most likely reason composite male faces were found less attractive in this study relates to the number of faces used to construct the composites. Langlois and Roggman (1990) found composites using fewer than 16 faces were not significantly more attractive than individual faces. Six of the seven composites included by Grammar and Thornhill contain fewer than 16 faces, which may account for the lack of consensus. Why this would be an issue for male, but not female, faces is unclear. Perhaps there is greater variation in men's faces, resulting from greater facial modification during puberty.

Corroborating evidence for the attractiveness of composites has been provided using a caricature generator, which alters an image to make it appear more average (anti-caricature) or less average (caricature). When this technique was applied to both schematic drawings and black-and-white photos, attractiveness was found to be significantly associated with averageness: anti-caricatures produced more-attractive faces while caricatures produced less-attractive faces (Rhodes and Tremewan 1996; Rhodes et al. 1999).

Critics have attacked the use of averaging techniques in attractiveness studies, claiming that the appeal of such composites may, in part, be an artifact of the process. In particular, they note that composites made from averaging are highly symmetrical, and these faces may be considered attractive because of their symmetry and not necessarily because of any preference for the population mean (Alley and Cunningham 1991; Pittinger 1991). However, averageness accounts for a significant amount of the variance in attractiveness even after symmetry has been partialed out, an indication that averageness contributes to attractiveness independent of symmetry (Rhodes et al. 1999).

While average faces are attractive, they are not necessarily the most attractive. Perrett, May, and Yoshikawa (1994) created a 60-face composite (high averageness) and a 15-face attractive composite (moderate averageness) using a subgroup of the 15 most attractive images. They then produced a third, low-average face by mapping and exaggerating the differences between the attractive and 60-face composites. Ratings indicated that the exaggerated face was the most attractive while the 15- and 60-face composites were progressively less attractive despite greater averageness. Whatever elements separated the attractive from the less attractive individuals in the initial ratings seem to have been preserved in the composites and accentuated in the exaggerated image (Perrett et al. 1994). On the surface, this may seem contradictory to the averageness argument, but it is not. All else being equal, greater averageness is more attractive. All else is not equal in the Perrett study since the mean-deviated images were composed of highly attractive individuals. What the Perrett paper demonstrates so eloquently is the fact that other criteria aside from averageness also influence attractiveness. One such factor is facial juvenilization.

FACIAL PAEDOMORPHOSIS

Paedomorphosis (child formation) describes developmental patterns that "juvenilize" adult forms so that descendant populations are underdeveloped relative to the ancestral population. The converse of paedomorphosis is peramorphosis, where descendant populations are "adultified" or overdeveloped (Gould 1977; McKinney and McNamara 1991). Underdevelopment can result from modification to either the time or the rate of development. The total time of development may be truncated, as is the case in postdisplacement (delayed onset) and progenesis (early termination). Alternatively, the rate of development may be retarded. Deceleration that affects the shape and size of a feature is termed rate-hypomorphosis, while deceleration that affects the shape alone is termed neoteny (McKinney and McNamara 1991; Parker and McKinney 1999; Shea 1989).

Underdevelopment can be global to the entire organism or specific to individual features or even behaviors, and it may involve one or more different modifications. Although paedomorphosis is most often used to describe differences between ancestral and descendant populations, it is also used to describe individual differences within a population (Parker and McKinney 1999). In other words, some individuals are more juvenilized than others. In general, human evolution seems to be marked by peramorphosis, but with regard to facial development specifically, humans are paedomorphic (Parker and McKinney 1999; Shea 1988, 1989).

Social scientists have investigated perceptions of personality and attractiveness in adult faces that have a juvenile appearance. A number of juvenilized features, such as large eyes, small nose, wide eye separation, small eyebrow height, large cranium height, small chin, narrow cheeks, small jaws, and full lips, have been considered (Berry and Zebrowitz-McArthur 1985; Cunningham 1986; Jones 1995; Zebrowitz-McArthur and Montpare 1989). In general, facial paedomorphosis has been found attractive, particularly in women. For instance, Zebrowitz-McArthur and Apatow (1983– 1984) varied the configuration of facial features using police identi-kits and found that juvenile manipulations increased the attractiveness of female faces but decreased the attractiveness of male faces. Also using identi-kits, Keating (1985) discovered that mature traits raise the perceived dominance of faces and ultimately increase the attractiveness of male faces while simultaneously decreasing the attractiveness of female faces, which were judged to be more attractive with immature traits.

Other correlational studies have assessed the relationship between attractiveness and facial morphology in general. Using black-and-white photos of both ordinary women and pageant winners, Cunningham (1986) discovered a dual criterion of attractiveness in which higher juvenilization was preferred in the center of the face while greater maturity was preferred in the periphery. Similar results were found for male faces (Berry and Zebrowitz-McArthur 1985; Cunningham et al. 1990). Furthermore, crosscultural correlations were consistent with North American samples. Asian and Hispanic students who recently arrived in the United States and Taiwanese students living in their home country rated the attractiveness of females of Asian, African, and European descent from around the world. Although Asians were less enthusiastic of mature features in the periphery, all three samples found juvenilized features in the center of the face attractive (Cunningham et al. 1995). Exposure to western culture was ruled out as an explanation for the convergence of attractiveness preferences since there was no difference between high- and low-exposure groups. To explain his results, Cunningham proposes a multiple fitness model wherein juvenile and mature features combine to signal both youth and sexual maturity in a target.

Another cross-cultural study investigated the effect of large eyes, a juvenile feature, on face-perception. Subjects from Brazil, the United States, Russia, and indigenous populations in Venezuela and Paraguay ranked male and female faces from both western and nonwestern populations on attractiveness. Significantly positive correlations were found between female attractiveness and large eye size for all five populations. Male faces, however, were preferred with average eye size (Jones and Hill 1993). Interestingly, the relationship between juvenilization and attractiveness was highest for female faces, and lowest for male faces, in the indigenous populations. The ecology of these hunter-gatherer societies more closely resembles our ancestral environment than ours does and thus may reflect human default mating preferences more accurately. Hence, the preference evident in industrialized populations for a relatively more androgynous face may be partially or entirely due to our inhabiting large urban centers, an environment to which humans are not specifically adapted.

The relationship between paedomorphic features and attractiveness has been confirmed in laboratory experiments as well. Genetic algorithms programmed to imitate the process of natural selection demonstrate a tendency to "evolve" a juvenilized female face. Faces were "crossbred" to create new generations of faces, with faces rated higher in attractiveness being crossbred more often than lower-rated faces. A beautiful face evolves after many such generations. Compared with a composite of ordinary women's faces, this face has a larger forehead, smaller lower face, fuller lips, and a narrower mouth, configurations consistent with a more juvenile female face (Johnston and Franklin 1993).

Jones (1995) measured the impact of neoteny on male and female attractiveness using cardioidal strains to alter the appearance of schematic faces. Positive strains produce a mature appearance by expanding the face outward near the base, and contracting it near the peak. A neotenous appearance is produced using a negative strain that expands the peak and contracts the base. Neotenous strains led to a substantial increase in attractiveness ratings for lower-rated women, whereas the ratings remained steady for medium and highly attractive women, and low- and medium-rated men. Only highly attractive men became less attractive following a neotenous strain. Granted, neoteny did not generally enhance appeal, but neither did it handicap it, whereas mature cardioidal strains decreased attractiveness for men and women of all attractiveness levels (Jones 1995).

In another study, the experimenters used feminizing and masculinizing techniques to alter facial formation. Feminizing a face increases its juvenilization since female morphology (including face morphology) is relatively more paedomorphic. Compared with male faces, female features are smaller and altered less by the time they reach maturity (Montagu 1981; Shea 1988, 1989). Perrett and colleagues (1998) enhanced the gender prototypicality of Japanese and Caucasian composites along a continuum ranging from 50% feminized to 50% masculinized. Both Japanese and European judges preferred feminized faces, regardless of the gender or race of the target. A related experiment, however, discovered that the preferred level of feminization in male faces depended on the ovulatory status of female judges; women preferred a less feminized male face when they were ovulating (Penton-Voak et al. 1999). Although this face was less feminized, it was still more feminized than a typical male face.

Evidence collected thus far indicates a clear relationship between juvenilization and female attractiveness. With regard to male attractiveness, however, the evidence is ambiguous: sometimes juvenilization enhances appearance and sometimes it does not. Perhaps, as Berry (1991) suggests, there is a juvenile/masculine duality in male attractiveness. Berry was able to identify two classes of attractive male faces. The first contained highly juvenile "pretty boys" who aroused impressions of kindness and artistic aptitude; the second contained "rugged" individuals who were perceived as strong and highly masculine. Berry's results alone are insufficient for us to conclude that a dual standard of male attractiveness exists. We wouldn't argue for a dual standard of attractiveness for hair color simply because we can identify individuals with light or dark hair. However, given that female preference for male faces varies with ovulatory status, it seems likely there may be a dual standard to male facial attractiveness.

GOAL OF THE EXPERIMENT

The preceding discussion implicates both averageness and juvenilization in facial attractiveness. However, the deviations associated with juvenilization require a reduction in averageness. Furthermore, the processes attributed to each of these factors are antagonistic; stabilizing selection maintains the population mean while directional selection works to subvert it. In actuality, the case is never either one or the other since stabilizing and directional selections always act jointly (Schmalhausen 1949). The question is: Which factor—and by extension, which selection pressure has relatively greater influence in human facial attractiveness? To answer this question, an array of faces representing varying degrees of juvenilization was constructed. Reduced, average, and enhanced levels of juvenilization were included. Averageness should manifest in a quadratic effect, with deviations from the average leading to lower attractiveness, while juvenilization should manifest in a linear effect, with greater juvenilization leading to higher attractiveness. A comparison of the respective effect sizes should reveal which preference is more important in contemporary humans.

METHOD

Participants

Judges were recruited from undergraduate psychology courses at California State University, Long Beach (CSULB) and received course credit for their participation. Data were collected from a total of 297 participants (218 women and 79 men) over three semesters. Of the total, 114 described their background as European, 70 as Asian, 62 as being of Spanish descent, 18 as African, and 33 others. The average age of the sample was 23.72 years.

Stimuli

Four facial features were manipulated in composite (average) faces: eye size, nose size, lip fullness, and eye separation. Features were manipulated to resemble either greater juvenilization or reduced juvenilization (adultification), and to include both moderate and extreme versions of each. Thus, together with the original composite average, there were five different versions of each feature.

To acquire target faces, individuals were approached on the (CSULB) campus and were asked to pose voluntarily as models for an experiment. Photos were taken using a tripod-mounted camera positioned approximately five feet from the model, who stood in front of a neutral-colored background in natural light. Models were asked to look directly into the camera with a neutral expression. Photos were scanned, cropped to restrict clothing cues, and saved in 24-bit BMP format.

Using a process first introduced by Benson and Perrett (1992), four composite faces (Asian male and female, Caucasian male and female) were generated using Morph 2.0 for Macintosh. The morphing software uses reference points to link the features of two faces and generate a short movie depicting the gradual transformation from one face to the other. The middle frame of this movie is the mathematical average of the two faces. By morphing two morphs together we created a composite face reflecting the mathematical average of four faces.

Each of the four features was manipulated separately. Modifications to composites were accomplished using Adobe Photoshop 3.0 for Macintosh. Features were boxed and saved to a new layer where their location or size could be altered before merging them back with the original image. A smudging tool available in the software was used to eliminate any creasing that resulted from this process. Each level of manipulation reflected a difference in proportion of about one standard deviation-i.e., the moderately juvenilized eye size was about one standard deviation larger than the average, while the extremely juvenilized was about two standard deviations larger. Similarly, the moderately adultified eye size was approximately one standard deviation smaller than the average while the extremely adultified version was about two standard deviations smaller.² Manipulations leading to smaller noses, fuller lips, and wider eye separations were associated with greater juvenilization whereas those leading to larger noses, thinner lips, and narrower eye separations were associated with decreased juvenilization (adultification). See Figure 1 for examples of faces and their manipulations.

Procedure

Judges were seated at individual computer stations in groups ranging from 2 to 16 people. Feature manipulations were extremely subtle (Figure 1), making individual attractiveness ratings of each version impossible. Consequently, a paired comparisons format was chosen wherein each feature version was contrasted with its four alternatives. A program written in Visual Basic 4.0 (32-bit version) presented the judges with comparisons on a computer. For each comparison, judges selected the version they liked best by clicking on it with the mouse. This prompted the program to record their selection and present the next comparisons. The side in which a given feature version appeared and the comparisons themselves were both randomized. After all comparisons had been administered, participants were free to submit anonymous comments regarding any aspect of the experiment and then dismissed. Debriefing occurred in class after all participants had completed the task.





RESULTS

Scoring

The frequency with which a particular feature version was selected over its alternatives determined its attractiveness score, which could range from zero (never selected) to four (always selected). The purpose here was to sample a variety of features representing a more general facial topography, not to study and compare the features themselves. Consequently, attractiveness scores were averaged across features to obtain mean ratings for each of the five versions.

Linear and Quadratic Effects for Juvenilization and Averageness

Figure 2a plots mean attractiveness scores of male and female targets for five levels of juvenilization. Linear and guadratic weights were applied to test for the effects of paedomorphosis and averageness, respectively. In the linear condition, the average face was weighted "0" while juvenilized forms were weighted positively (+1 for moderate and +2 for extreme) and adultified forms were weighted negatively (-1 and -2). Extreme versions were weighted double relative to moderate versions to reflect the fact that manipulations were double in magnitude. If paedomorphosis is attractive, positive values will outnumber negative values and the mean will be greater than zero. If there is no effect for paedomorphosis, positive and negative values will cancel each other out and the mean will approach zero. In the quadratic condition, the average face was weighted positively (+3) while juvenilized and adultified forms were weighted negatively (-1)for extreme and -0.5 for moderate). If averageness is attractive, positive values will outnumber negative values and the mean will be greater than zero. If there is no effect for averageness, positive and negative values will cancel out and the mean will approach zero.

Means for male and female targets are presented in Table 1. Pairedsample *t*-tests indicated there was no difference between male and female targets with regard to either the linear (t(296) = 1.19, ns) or the quadratic (t(196) = 0.08, ns) effects. Male and female data were subsequently pooled to test the linear and quadratic hypotheses. One-sample *t*-tests indicated that both the linear (t(296) = 13.45, p < .001) and quadratic (t(296) = 25.53, p < .001) means were significantly greater than zero, confirming the influences of both paedomorphosis and averageness on attractiveness. Effect sizes were calculated using *d*. In a one-sample *t*-test where the mean is compared to zero, *d* is calculated by dividing the mean by the population standard deviation. Where the population standard deviation is unavailable, it may be estimated from the sample standard deviation (Kirk 1995). Although both the linear and quadratic effects were extremely large, the



Figure 2. (A) Mean attractiveness scores plotted for male and female targets. The average and moderately juvenilized features were highest in attractiveness followed by the moderately adultified and extremely juvenilized features. The extremely adultified features were lowest in attractiveness. (B) Attractiveness ratings broken down by four judge-target combinations: Men-judging-men (MM), women-judging-men (WM), men-judging-women (MW), and women-judging-women (WW). Men-judging-men prefer less juvenilization compared with other judge-target combinations.

Effect	Sex	Mean	SD	N	Effect Size
Linear	Men	2.23	4.01	79	
	Women	2.52	3.43	218	
	Combined	2.37*	3.05	297	0.78
Quadratic	Men	2.98	2.58	79	
	Women	2.99	2.53	218	
	Combined	2.99*	2.01	297	1.48

Table 1.	Means for the Linear and Quadratic Effects in Male and
Fen	ale Targets

Note: Means reflect the degree of juvenilization (linear) and averageness (quadratic) preferred by the sample. Higher means reflect greater preference for juvenilization or averageness. Combined means were compared to zero in one-sample *t*-tests. The effect size for the quadratic is approximately twice that found for the linear. * p < .001

quadratic effect (d = 1.48) was nearly twice as big as the linear effect (d = .78), an indication that, compared to juvenilization, averageness is relatively more important to facial attractiveness.

Interactions with Race and Sex of Judges

Two repeated-measures ANOVAs tested for important interactions: one to test race of target by race of judge effects, and one to test sex of target by sex of judge effects. Sex and race were tested separately because we were only interested in comparing perceptions of own- to opposite-sex and own- to other race. In this analysis there is no theoretical reason to test for three- and four-way interactions, and in any event, statistically significant complex interactions are difficult to interpret and unlikely to be replicable.

To determine the influence of sex in the quadratic and linear effects, target's sex was entered as a two-level within-group variable, judge's sex was entered as a two-level between-group variable, and juvenilization was included as a five-level within-group variable. There was a significant threeway linear interaction, an indication that the increase in attractiveness associated with juvenilization depended on the sex of both the target and the judge (F(1,295) = 5.61, p < .05). Means are plotted in Figure 2b. Linear weights were applied as described above, and pairwise *t*-tests indicated a significant difference between the mean for men-judging-men (0.91, SD = 3.96) and the means for men-judging-women (2.23, SD = 3.35), womenjudging-men (2.63, SD = 3.46), and women-judging-women (2.71, SD = 3.92), even after Bonferroni correction. The results for the latter three groups were not significantly different from one another. The interaction reflects a tendency for men to prefer a lower level of juvenilization in men compared with other target/judge combinations. The size of this interaction was a medium effect (d = .27), or about one-third the size of the main effect for juvenilization.

To determine the influence of race, it was necessary to restrict the analysis to Asian (N = 70), Caucasian (N = 114), and Spanish descent (62) judges since other sample group sizes were too small to include them as separate categories. Judge's race was entered as a three-level between-group variable. Target's race was entered as a two-level (Asian and Caucasian) within-group variable, while juvenilization was included as a five-level within-group variable. There was a significant linear juvenilization-byjudge's race interaction (F(2,243) = 3.65, p < .05), an indication that the attractiveness resulting from juvenilization is tempered by the race of the judge. Linear weights were applied and independent-sample *t*-tests revealed the mean of the Asian group (1.73, SD = 3.47) to be significantly lower than the means of the Caucasian (2.83, SD = 2.66) and Spanish descent (2.89, SD = 2.79) groups. Hence, Asians were relatively less enthusiastic about juvenilization than the other two groups. However, the importance of this interaction is compromised by its small effect size (d = .05).

There were no significant interactions involving the quadratic effect. The preference for averageness was independent of judges' or targets' sex or race; greater averageness was always judged to be more attractive.

DISCUSSION

Stabilizing selection is a maintaining force in evolution whereas directional selection is a driving force. Stabilizing selection should be strong enough to withstand directional selection when the environment is stable, but should capitulate when the environment is unpredictable. Thus, knowing which selection pressure is dominant should inform us as to the nature of the environment. With regard to facial paedomorphosis, stabilizing and directional selection are apparent in preferences for averageness and juvenilization. Although facial attractiveness is subject to both forces, our results indicate that averageness has greater potency by a margin of almost two-to-one, suggesting that present selection pressure(s) are maintaining the population mean and that humans are currently, although unlikely to remain so indefinitely, in a state of stasis.

While our results implicate stabilizing selection as the current dominant force in facial attractiveness, they also evince a substantive preference for juvenilization. Ergo, the directional forces resisting stabilizing selection favor a particular appearance. The selection pressure(s) responsible for this preference are not entirely understood. Juvenilized women may be found attractive because they appear youthful, which in turn is related to fertility and future reproductive capacity (Buss 1987; Johnston and Franklin 1993; Jones 1995; Symons 1979). The same explanation is unlikely to account for the attractiveness of juvenilized features in men for several reasons: (*a*) male reproductive residual decreases at a slower rate compared with that of women, (*b*) status and resources coveted by women are found more abundantly in older men, and (*c*) women typically prefer mates older than themselves (Buss 1994).

Alternatively, facial juvenilization may serve as an accurate indicator of personality and parenting skill; men with lower testosterone levels—as indicated by juvenilized features—make more desirable partners and coparents because they are less aggressive and more amicable. Male faces altered to appear more masculine are perceived as less warm, less emotional, less honest, less cooperative, and less skilled as parents (Penton-Voak et al. 1999), suggesting that these individuals are ill favored as long-term mates. Conversely, baby-faced individuals are commonly perceived as warm, kind, and honest people (Zebrowitz-McArthur and Apatow 1983–1984). These traits are desirable for both men and women and may supplement the youth hypothesis in explaining the attractiveness of juvenilized features.

Some have suggested there are different "types" of attractiveness. For instance, based on a factor analysis of fashion magazines, Ashmore, Solomon, and Longo (1996) believe female attractiveness can be characterized as either sexy, cute, or trendy. Since their project consisted exclusively of professional models, however, it is more than likely that all of their stimuli reflected high levels of averageness, juvenilization, and symmetry, and the "types" they identified could be simulated by a single woman at different points in time. Thus, their results speak more to style and expression than to morphological beauty. Berry (1991) identified two types of male attractiveness: one that was babyish and another that was not. It is important to reiterate, however, that successful categorization falls short of validation.

Assuming there are two classes of male attractiveness, one highly paedomorphic and one not, there is a complementary parallel in female mate selection. Women more often prefer a mate rich in status and resources for the purpose of long-term mating, but prefer a male high in "physical attractiveness" for short-term sexual affairs (Buss 1994). It is possible that the adultified standard of attractiveness may predict a woman's willingness for short-term sexual liaisons, whereas the juvenilized standard may prove useful for predicting her desire to elicit long-term investment. Recent evidence supports this conjecture. Women prefer less feminine (less juvenile) male faces during ovulation, when the probability of conception is high, but otherwise prefer feminized (juvenilized) male faces, when the likelihood of conception is low (Penton-Voak et al. 1999). Given the brevity of its duration, most women in the present project were probably not ovulating at the time, which may explain why their attractiveness ratings are consistent with a highly juvenilized criterion. Male judges don't ovulate, of course, so the same argument cannot explain juvenilization for men judging men. Perhaps men apply the juvenilized criterion to both genders because it is the only criterion relevant to them, or perhaps it was cognitively efficient to do so.

Symmetry or Averageness?

Deviations from the composite average were judged to be less attractive. Symmetry could not account for this effect since the symmetrical integrity of the original composite was maintained by manipulating the right and left sides of the face identically. It might be argued that attractiveness decreased because manipulations led to an odd appearance, but moderate manipulations were only one standard deviation from the mean. Thus, as did the findings of Rhodes et al. (1999), this study supports the contention that average proportions can be attractive independent of symmetry.

Is averageness a necessary component of attractiveness, as Langlois proposes? A paired-sample *t*-test found no significant difference in the mean attractiveness of juvenilized and average faces. This suggests that, while averageness is attractive, it is not the only attractive facial proportion, a conclusion consistent with that arrived at by Perrett et al. (1994), who found exaggerations of an attractive composite to be more appealing than the original average. On the other hand, certainly some minimum level of averageness must be necessary; that conclusion is evident from the size of the quadratic effect, but exactly where that minimum lies has yet to be articulated.

Sex and Race Effects

Baby-faced adults are perceived to be warmer, more honest, submissive, gullible, less intelligent, and younger (Zebrowitz-McArthur and Apatow 1983–1984). While perceptual content is similar for men and women, salience may not be. For instance, submissiveness may be more salient in female than in male faces for two reasons: one, it's more gender prototypical of women than of men (Zebrowitz 1998), and two, it may be more desirable in women than in men (Gowaty 1992; Keating 1985; Sadalla et al. 1987). Furthermore, the perceptions associated with juvenilization may be regulated by the desires of the judge, which differ in men and women (Buss 1994). For instance, when judging men with a juvenilized appearance, submissiveness (a negative perception for male targets) may be relatively more salient to female judges. This may explain the lower juvenilization preferred by men judging men compared with other target-judge combinations.

CONCLUSION

Although our data confirm the joint influences of juvenilization and averageness on facial attractiveness, the effect for averageness was larger than that found for juvenilization. This suggests that stabilizing selection is dominant over directional selection in contemporary humans and that the current environment is stable with regard to facial paedomorphosis. Nevertheless, the effect size for juvenilization was substantial, and future research should strive to elucidate the selection pressure(s) favoring paedomorphosis. In addition, questions remain regarding the relative effect sizes of juvenilization and averageness across the life span. Our targets were composed exclusively of young men and women close to their peak reproductive value. It would be interesting to investigate whether this ratio remains consistent for older faces as well.

The lead author is a third-year Ph.D. student in the psychology department at the University of British Columbia, Canada. The second and third authors are professors of psychology at California State University, Long Beach.

NOTES

1. See Yu and Shepard (1998) for a possible exception.

2. "Extreme" describes the manipulation and not the appearance of the face; extreme features still fall within the normal range of human appearance.

REFERENCES

Alley, T. R. and Cunningham, M. R.

1991 Averaged Faces Are Attractive, But Very Attractive Faces Are Not Average. *Psychological Science* 2:123–125.

Andersson, M. B.

1994 Sexual Selection. Princeton: Princeton University Press.

Ashmore, R. D., M. R. Solomon, and L. C. Longo

1996 Thinking about Fashion Models' Looks: A Multidimensional Approach to the Structure of Perceived Physical Attractiveness. *Personality and Social Psychology Bulletin* 22:1083–1104.

Barber, N.

1995 The Evolutionary Psychology of Physical Attractiveness: Sexual Selection and Human Morphology. *Ethology and Sociobiology* 16:395–424.

Benson, P., and D. Perrett

1992 Face to Face with the Perfect Image. *New Scientist* (Feb.):32–35. Berry, D. S.

1991 Attractive Faces Are Not All Created Equal: Joint Effects of Facial Baby-

ishness and Attractiveness on Social Perception. *Personality and Social Psychology Bulletin* 17:523–531.

Berry, D. S., and L. Zebrowitz-McArthur

1985 Some Components and Consequences of a Babyface. *Journal of Personality and Social Psychology* 48:312–323.

- Buss, D. M.
 - 1987 Sex Differences in Human Mate Selection Criteria: An Evolutionary Perspective. In Sociobiology and Psychology: Ideas, Issues and Applications, C. Crawford, M. Smith, and D. Krebs, Eds. Pp. 335–351. Hillsdale, New Jersey: Lawrence Erlbaum.

1994 The Evolution of Desire: Strategies of Human Mating. New York: Basic Books. Cunningham, M. R.

1986 Measuring the Physical in Physical Attractiveness: Quasi-experiments on the Sociobiology of Female Facial Beauty. *Journal of Personality and Social Psychology* 50:925–935.

Cunningham, M. R., A. P. Barbee, and C. L. Pike

1990 What Do Women Want? Facial-metric Assessment of Multiple Motives in the Perception of Male Facial Physical Attractiveness. *Journal of Personality and Social Psychology* 59:61–72.

Cunningham, M. R., A. R. Roberts, A. P. Barbee, P. B. Druen, and C. H. Wu

1995 "Their Ideas of Beauty Are, on the Whole, the Same as Ours": Consistency and Variability in the Cross-cultural Perception of Female Physical Attractiveness. *Journal of Personality and Social Psychology* 68:261–279.

Galton, F.

1878 Composite Portraits. Journal of the Anthropological Institute of Great Britain and Ireland 8:132–142.

Gangestad, S.W., and R. Thornhill

1997 The Evolutionary Psychology of Extrapair Sex: The Role of Fluctuating Asymmetry. *Evolution and Human Behavior* 18:69–88.

Gangestad, S. W., R. Thornhill, and R. A. Yeo

1994 Facial Attractiveness, Developmental Stability, and Fluctuating Asymmetry. Ethology and Sociobiology 15:73–85.

Gould, S. J.

1977 Ontogeny and Phylogeny. Cambridge: Belknap Press of Harvard University Press.

Gowaty, P. A.

1992 Evolutionary Biology and Feminism. Human Nature 3:217–249.

Grammar, K., and R. Thornhill

1994 Human (*Homo sapiens*) Facial Attractiveness and Sexual Selection: The Role of Symmetry and Averageness. *Journal of Comparative Psychology* 108: 233–242.

Johnston, V. S., and M. Franklin

1993 Is Beauty in the Eye of the Beholder? *Ethology and Sociobiology* 14:183–199.

Jones, D.

1995 Sexual Selection, Physical Attractiveness and Facial Neoteny. Current Anthropology 36:723–748.

- 1996 Physical Attractiveness and the Theory of Sexual Selection. Ann Arbor: Museum of Anthropology, University of Michigan.
- Jones, D., and K. Hill
 - 1993 Criteria of Facial Attractiveness in Five Populations. *Human Nature* 4:271-296.
- Keating, C. F.
 - 1985 Gender and the Physiognomy of Dominance and Attractiveness. Social Psychology Quarterly 48:61-70.
- Kirk, R. E.
 - 1995 Experimental Design: Procedures for the Behavioral Sciences, third ed. Pacific Grove, California: Brooks/Cole.
- Langlois, J. H., and L. A. Roggman
- 1990 Attractive Faces Are Only Average. Psychological Science 1:115–121.
- Langlois, J. H., L. A. Roggman, and L. Musselman
- 1994 What Is Average and What Is Not Average about Attractive Faces? *Psy*chological Science 5:214–220.
- McKinney, M. L., and K. J. McNamara
- 1991 Heterochrony: The Evolution of Ontogeny. New York: Plenum Press.
- Montagu, A.
- 1981 Growing Young. New York: McGraw-Hill.
- Parker, S. T., and M. L. McKinney
 - 1999 Origins of Intelligence: The Evolution of Cognitive Development in Monkeys, Apes, and Humans. Baltimore: Johns Hopkins University Press.
- Penton-Voak, I. S., D. I. and Perrett, D. L. Castles, T. Kobayashi, D. M. Burt, L. K. Murray, and R. Minamisawa
 - 1999 Menstrual Cycle Alters Face Preference. Nature 399:741–742.
- Perrett, D. I., K. J. Lee, I. Penton-Voak, D. Rowland, S. Yoshikawa, D. M. Burt, S. P. Henzi, D. L. Castles, and S. Akamatsu
 - 1998 Effects of Sexual Dimorphism on Facial Attractiveness. *Nature* 394:884-887.
- Perrett, D. I., K. A. May, and A. Yoshikawa
 - 1994 Facial Shape and Judgements of Female Attractiveness. *Nature* 368:239–242.
- Pittinger, J. B.
 - 1991 On the Difficulty of Averaging Faces: Comments on Langlois and Roggman. *Psychological Science* 2:351-353.
- Rhodes, G., and T. Tremewan
 - 1996 Averageness, Exaggeration, and Facial Attractiveness. *Psychological Science* 7:105–110.
- Rhodes, G., A. Sumich, and G. Byatt
 - 1999 Are Average Facial Configurations Attractive Only Because of Their Symmetry? *Psychological Science* 10:52–58.
- Sadalla, E. K., D. T. Kenrick, and B. Vershure
 - 1987 Dominance and Heterosexual Attraction. Journal of Personality and Social Psychology 52:730–738.
- Schmalhausen, I. I.
 - 1949 Factors of Evolution. Philadelphia: Blakiston.

1988 Heterochrony in Primates. In *Heterochrony in Evolution*, M. L. McKinney, ed. Pp. 237–266. New York: Plenum Press.

1989 Heterochrony in Human Evolution : The Case for Neoteny Reconsidered. Yearbook of Physical Anthropology 32:69–101.

Singh, D.

1993 Adaptive Significance of Female Physical Attractiveness: Role of Waistto-hip Ratio. *Journal of Personality and Social Psychology* 65:293–307.

1994 Body Fat Distribution and Perception of Desirable Female Body Shape by Young Black Men and Women. *International Journal of Eating Disorders* 16: 289–294.

Singh, D., and S. Luis

1995 Ethnic and Gender Consensus for the Effect of Waist-to-hip Ratio on Judgment of Women's Attractiveness. Human Nature 6:51-65.

Strickberger, M. W.

1996 Evolution, second ed. London: Jones and Bartlett.

Symons, D.

1979 The Evolution of Human Sexuality. New York: Oxford University Press.

1995 Beauty Is in the Adaptations of the Beholder: The Evolutionary Psychology of Human Female Sexual Attractiveness. In Sexual Nature, Sexual Culture, P. R. Abramson and S. D. Pinkerton, eds. Pp. 80–118. Chicago: University of Chicago Press.

Thornhill, R., and S. W. Gangestad

1994 Human Fluctuating Asymmetry and Sexual Behavior. Psychological Science 5:297–302.

Yu, D. W., and G. H. Shepard Jr.

1998 Is Beauty in the Eye of the Beholder? Nature 396:321-322

Zebrowitz, L.

1998 Reading Faces. Oxford: Westview Press.

Zebrowitz-McArthur, L., and K. Apatow

1983–1984 Impressions of Babyfaced Adults. Social Cognition 2:315–342.

Zebrowitz-McArthur, L., and J. M. Montpare

1989 Contributions of a Babyface and a Neotenous Voice to Impressions of Moving and Talking Faces. *Journal of Nonverbal Behavior* 13:189–203.

Shea, B. T.