ELSEVIER

Evolution and Human Behavior 22 (2001) 251–267

Evolution and Human Behavior

Male facial attractiveness
Evidence for hormone-mediated adaptive design

Victor S. Johnston[a,*], Rebecca Hagel[a], Melissa Franklin[b], Bernhard Fink[c], Karl Grammer[c]

[a]*Department of Psychology, New Mexico State University, Las Cruces, NM 88003, USA*
[b]*University of New Mexico, Albuquerque, NM, USA*
[c]*Ludwig-Boltzmann Institute for Urban Ethology Vienna, Austria*

Received 11 February 2000; accepted 4 April 2001

Abstract

Experimenters examining male facial attractiveness have concluded that the attractive male face is (1) an average male face, (2) a masculinized male face, or (3) a feminized male face. Others have proposed that symmetry, hormone markers, and the menstrual phase of the observer are important variables that influence male attractiveness. This study was designed to resolve these issues by examining the facial preferences of 42 female volunteers at two different phases of their menstrual cycle. Preferences were measured using a 40-s QuickTime movie (1200 frames) that was designed to systematically modify a facial image from an extreme male to an extreme female configuration. The results indicate that females exhibit (1) a preference for a male face on the masculine side of average, (2) a shift toward a more masculine male face preference during the high-risk phase of their menstrual cycle, and (3) no shift in other facial preferences. An examination of individual differences revealed that women who scored low on a "masculinity" test (1) showed a larger menstrual shift, (2) had lower self-esteem, and (3) differed in their choice of male faces for dominance and short-term mates. The results are interpreted as support for a hormonal theory of facial attractiveness whereby perceived beauty depends on an interaction between displayed hormone markers and the hormonal state of the viewer.

Keywords: Facial attractiveness; Sexual selection; Mate choice; Menstrual cycle; Hormones

* Corresponding author. Tel.: +1-505-646-2800; fax: +1-505-646-6218.
E-mail address: vic@crl.nmsu.edu (V.S. Johnston).

PII: S1090-5138(01)00066-6

1. Introduction

In a fraction of a second, the brain of a human male or female can ascertain the physical attractiveness of another person's face (Johnston & Oliver-Rodriguez, 1997; Oliver-Rodriguez, Guan, & Johnston, 1999). This remarkable feat appears to depend on the delicate interplay between physical markers on the face, perhaps fitness indicators (Miller, 2000), and exquisitely sensitive brains that generate feelings, perhaps fitness monitors (Johnston, 1999), in response to such signals. Support for this model comes from studies that have used various image-processing techniques to systematically manipulate the features and proportions of female facial images and observe the behavioral and/or emotional responses of men and women exposed to such images. The current experiment attempts to evaluate and refine this model by examining how the attractiveness of male and female faces varies with both their displayed hormone markers, and the hormonal state of female viewers, as indicated by their menstrual phase.

Early studies suggested that the most attractive female face was the average face in a population (Langlois & Roggman 1990; Langlois, Roggman, & Musselman, 1994; Langlois, Roggman, Musselman, & Acton, 1991). Several experimenters, however, have concluded that the image-processing technique used in these studies may be flawed, and proposed that although the average face is attractive, it is not the most attractive face in a population (Alley & Cunningham, 1991; Johnston, 2000). Strong support for the "non-average" hypothesis comes from Perrett, May, and Yoshikawa (1994), who demonstrated that an average face made by combining random faces is judged to be less attractive than the average of attractive faces drawn from the same sample of faces. Indeed, there is now substantial evidence indicating that attractive female faces are not average, but differ from the average in a systematic manner. More specifically, they possess a shorter, narrower, lower jaw, fuller lips, and larger eyes than an average face (Cunningham, Roberts, Barbec, Druen, & Wu, 1995; Johnston & Franklin, 1993; Perrett et al., 1994). These specific markers have been shown to be effective across cultures (Cunningham et al., 1995; Perrett et al., 1994), and electrophysiological studies have revealed that they elicit emotional responses in male, but not female, viewers of female faces (Johnston & Oliver-Rodriguez, 1997). Because pubertal bone growth (brow ridges and lower jaw) is stimulated by androgens (Tanner, 1978), and lip fullness parallels estrogen-dependent fat deposits elsewhere on the female body (Farkas, 1981), Johnston and Franklin (1993) have hypothesized that an attractive female face may be displaying hormone markers (high estrogen/low androgen) that serve as reliable indicators of fecundity.

In contrast to the research on female facial attractiveness, studies examining the importance of hormone markers on male faces have produced apparently incompatible results. For example, although a number of experimenters have demonstrated that women favor a "masculinized" male face possessing a large jaw and prominent brow ridges and cheekbones (Grammer & Thornhill, 1994; Scheib, Gangestad, & Thornhill, 1999), other studies have reported that both British and Japanese females prefer a more "feminized" male face with a shorter-than-average lower jaw (Penton-Voak et al., 1999; Perrett et al., 1998). Still others have found that a mixture of mature features (large lower jaw, prominent cheekbones, and thick eyebrows) and neotenous features (large eyes and small nose) is the most desirable configuration of male faces (Cunningham, Barbee, & Pike, 1990). Some of these inconsistencies across studies may result from the different procedures used to generate the male facial stimuli. For example, both studies reporting a preference for feminized male faces used a caricaturing technique to create their masculinized and feminized facial images. Producing a masculinized male face using this procedure involves an algorithm that further exaggerates the differences between an average female and an average male face. This mathematical procedure is based on the assumption that extreme masculine and feminine faces can be produced by a linear extrapolation of the differences between these average faces. However, given that the majority of consistent male–female differences are the result of bone growth, which depends on a complex interaction between androgens, estrogens, and growth hormone, this linear growth assumption may not be valid (Grumbach, 2000; Tanner, 1978).

A second consequence of caricaturing is its potential effects on modifying facial symmetry. Average male and female faces are inevitably low in fluctuating asymmetries (traits that differ among individuals but are symmetrical at a population level), so morphs generated between such faces (feminized males and masculinized females) will invariably be quite symmetrical. It is also inevitable, however, that any small deviation from perfect symmetry that exists in either the average male or average female face will be substantially exaggerated in caricatured faces (masculinized males and feminized females). Caricaturing, therefore, is likely to generate feminized male facial images that are more symmetrical than masculinized male images. Because facial symmetry may influence attractiveness (Gangestad, Thornhill, & Yeo, 1994; Grammer & Thornhill, 1994; Mealey, Bridgstock, & Townsend, 1999; Perrett et al., 1999), any observed preference for a feminized or a masculinized male face may be confounded by differences in the symmetry of these faces. The current study attempts to improve on this research methodology by examining preferences for male (and female) faces that have been masculinized (or feminized) on the basis of perceived masculinity (or femininity), rather than employing the caricaturing technique.

Low fluctuating asymmetry (FA) is thought to reflect an ability to resist the harmful effects of mutations, parasites, and/or toxins during development (Møller & Swaddle, 1997). Because such resistance is partially heritable (Møller & Thornhill, 1997), there may be important fitness benefits for females who exhibit a preference for such mates. A variety of experimental findings support this hypothesis. Across species, symmetrical males have significantly greater mating success (Møller & Thornhill, 1998), and symmetrical males have been shown to be more desirable and have more sexual opportunities than asymmetrical men (Gangestad & Thornhill, 1997; Thornhill & Gangestad, 1994). In populations lacking widespread contraception and modern medicine, these enhanced sexual opportunities can be translated into more offspring and better health (Waynforth, 1998). However, despite the correlation between symmetry and attractiveness, it appears that human females may not use or even perceive fluctuating asymmetries when judging the attractiveness of male faces (Scheib et al., 1999).

Scheib et al. (1999) found that the measured FA of male faces was not only correlated with their attractive ratings, but was also correlated with the attractiveness ratings of a right or left half-face, circumstances where all direct cues to bilateral symmetry are absent. Clearly, their experimental participants did not require explicit cues of bilateral symmetry to make accurate judgments of attractiveness. It appears that the salient cues for attractiveness are apparent on each half of a male's face and their presence is correlated with computed FA, but not perceived symmetry. An examination of the attractive faces led the authors to conclude that the specific features correlated with attractiveness and FA were a longer lower jaw and prominent cheekbones. Keating (1985) also found that the shape of the lower jaw was an important attribute of male facial attractiveness. Using an Identi−Kit methodology, she examined the effects of eye size, lip fullness, brow thickness, and jaw shape on both dominance and attractiveness ratings. As she had predicted, the combination of masculine features (square jaw, narrow eyes, thick eyebrows, and thin lips) enhanced the dominance ratings of male faces, but only a subset of these attributes (square jaw and thin lips) resulted in significantly higher attractiveness ratings. It appears that some high testosterone markers (square jaw) and low estrogen markers (thin lips) influence both the dominance and attractiveness of male faces, but dominance and attractiveness are not identical attributes. Given these prior results, the current study attempts to clarify the role of hormone markers in the perception of dominance and attractiveness, in both male and female facial images.

Some of the discrepancy in findings among male attractiveness studies may be a consequence of differences in the participant populations. One potential source of variance is the hormonal status of female participants. Penton-Voak et al. (1999) have shown that females' preferences for male faces changed as a function of the viewer's menstrual phase at the time of testing. Specifically, females tested during the 9 days prior to ovulation (high conception risk group) preferred a less feminized male face than females tested outside of this window (low conception risk group). In a more recent study, using 139 participants who responded to a magazine survey, Penton-Voak and Perrett (2000) reported that women in the high conception risk group were significantly more likely to prefer a masculine face than those in the low conception risk group. The authors interpret their findings as evidence for a conditional mate choice strategy whereby females in the high conception risk group are exhibiting a preference for male facial cues that signal adaptive heritable genetic characteristics, such as immunocompetence. However, these menstrual studies have not shown that the observed change in preference over the menstrual cycle is specific to attractive male faces. Do females' preferences for female faces also change? If so, then the observed effect may simply reflect a general change in mood over the menstrual cycle (Dalton, 1982), rather than a specific adaptation. To explore the generality of the menstrual effect, the current study examines how a variety of different facial preferences (attractive male, attractive female, dominant male, dominant female, etc.) are, or are not, modified by the hormonal state of female viewers. An effect on male facial attractiveness alone would provide strong evidence for complex adaptive design (Williams, 1966). This is the hypothesis that is evaluated in the current study.

2. Methods

2.1. Participants

The participants were 42 female volunteers between 18 and 35 years of age ($M = 22$). These women were recruited from the undergraduate population at New Mexico State

University in Las Cruces (USA) and the Ludwig−Boltzmann Institute for Urban Ethology in Vienna (Austria) ($n = 40$ and $n = 2$, respectively). Participation in the experiment was limited to volunteers who stated that they were (1) heterosexual, (2) not currently pregnant or breastfeeding a child, and (3) not currently taking any steroid medications or birth control pills. All participants signed an informed consent document indicating that they were volunteering for an experiment on facial preferences that would be conducted over two experimental sessions. They were also informed that they would be asked to provide relevant personal information, all such data would be confidential, and they could withdraw from the experiment at any time.

2.2. Apparatus

In order to measure facial preferences, the method developed by Perrett et al. (1998) was modified so that average male and female facial images could be masculinized or feminized to varying degrees on the basis of perceived masculinity and femininity, rather than the caricature procedure used by these authors. The final experimental tool was in the form of a 1200-frame QuickTime movie that gradually morphed an extremely masculine male facial image into to an extremely feminine face, at a rate of 30 frames/s. The movie was created in four steps (Fig. 1).

First, 16 random male and 16 random female facial images were separately morphed to produce a composite average male and a composite average female image. The male photographs were taken from college students between 18 and 26 years of age. The female pictures were digital photographs of Caucasians from a CD-ROM by the Japanese artist, Gomi (1998), and ranged in age from 18 to 30. All of the photographs were taken under constant light conditions and showed faces with neutral expression and with no apparent make-up, facial hair, or adornments (e.g., earrings). Prior to morphing, all pictures were

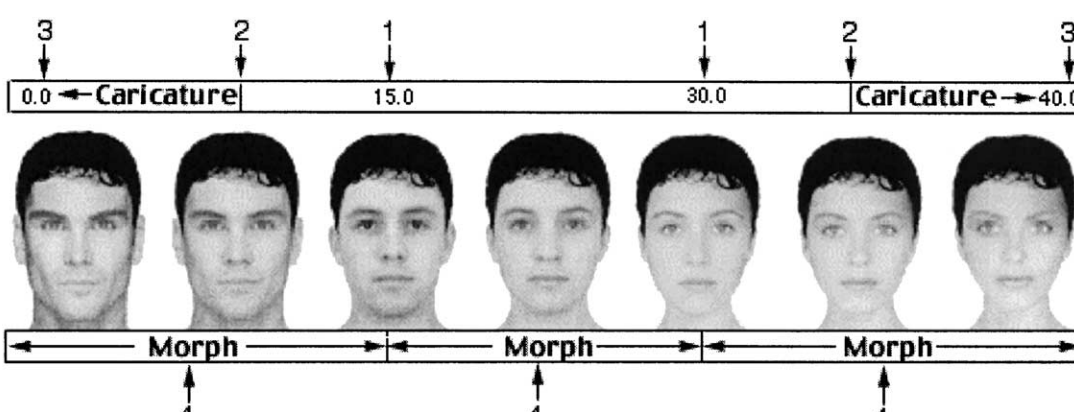

Fig. 1. Schematic diagram of the steps used to produce the facial morph movie. (1) Simultaneous morphs produce average male and female faces. (2) FacePrints program defines features and proportions of masculine and feminine faces. (3) Caricatures produce extreme male and extreme female faces. (4) Three morphs produce movie clips that are combined into a single movie. The originals were color images.

standardized to the same orientation using the procedure described by Rikowski and Grammer (1999). Using the "Facial Explorer" program (Grammer, Fieder, & Fink, 1998), the composite average male and the composite average female image were produced in a single step (for details on the morphing algorithm itself, see Beier & Neely, 1992; Gomes, Darsa, Costa, & Velho, 1999; Wolberg, 1990).

Next, the features and proportions of a perceived masculine and a perceived feminine face were "evolved" using the FacePrints software program (Johnston, 1994). The details of this program have been described elsewhere (Johnston & Franklin, 1993). In essence, FacePrints employs a genetic algorithm that allows participants to search a multidimensional face space of more than 34 billion possibilities and find their most masculine or most feminine facial image. Together with the average male and female faces, these masculine male and feminine female faces defined the features and proportions of the four key data points needed to construct a movie that slowly morphed a masculine male face to a feminine female face. However, in order to conceal the locations of the perceived masculine and feminine faces, both ends of the movie were extrapolated using a 5-s caricature. That is, the average male to masculine male difference was extrapolated to produce an extremely masculine male facial image and the average female to feminine female was extrapolated to produce an extremely feminine face. In the final step, all faces were fitted with the same androgynous hairstyle and the movie clips were combined into a single QuickTime movie using Adobe Premier. In essence, the movie permitted systematic modification of the major features and proportions that differentiate human male from female faces following puberty (Farkas, 1981). Using a slider control and single frame buttons, an experimental participant could move back and forward through the 1200-frame movie in order to find the frame number corresponding to any desired target face (e.g., an attractive male face).

Facial symmetry measures were obtained following the procedure developed by Grammer and Thornhill (1994). Based on the intercorrelations between naive users, this procedure has been shown to produce a highly reliable index of bilateral symmetry (Grammer & Thornhill, 1994; Rikowski & Grammer, 1999). Twelve standard landmarks were identified on the average and extreme male and female images used in the morph movie. These landmarks included the innermost and outermost corners of the eyes and the leftmost and rightmost edges of the nose. The points for measuring the cheekbones were defined as the *x*-coordinates of the widest points of face below the eyes. The *x*-coordinates of the jaw and mouth were identified as points on a horizontal line passing through the corners of the mouth. The "Facial Explorer" program (Grammer et al., 1998) then measured the midpoints of the six resulting horizontal lines and computed an index of horizontal symmetry by summing the *x*-axis differences between the midpoints. This analysis revealed that (1) the average male face was very symmetrical (FA = 16) and (2) there was a small systematic decrease in symmetry toward the extreme masculine end of the movie (FA = 16.5).

2.3. Procedure

Each woman was evaluated during two experimental sessions that were exactly 2 weeks apart. During session one, she viewed the morph movie and was told how to use the slider and single step controls in order to find the frame number corresponding to the facial image that was closest to an androgynous face. Following this practice trial, she was then required to use these controls to find the frame number of the facial image that most closely resembled a set of 15 target faces. The targets were an average male face (AvM), an average female face (AvF), an attractive male face (AtM), an attractive female face (AtF), a dominant-looking male face (DoM), a dominant-looking female face DoF), a healthy-looking male face (HtM), a healthy-looking female face (HtF), a masculine-looking male face (MaM), a feminine-looking female face (FmF), an intelligent-looking male face (ItM), an intelligent-looking female face (ItF), a good-father male face (GfM), a good-mother female face (GmF), and an androgynous face (Pat). In each case, a descriptive phrase was used to clarify the desired target face. For example, an average face was described as "a typical male (female) on the street," a dominant face as that of a male (female) who was "more likely to give than take orders," an androgynous face as "a face that could be either male or female," and an attractive face as "the male (female) that you like best from this range of possibilities." The order of these target faces was counterbalanced among participants and between sessions.

As each target face was located, the experimenter noted the corresponding frame number. When complete, the male faces (including the androgynous face) that had been selected by the participant were re-displayed in a random order. While each face was visible, the participant was asked to rate the face on 20 different attributes, using a seven-point Likert scale for each rating. These attributes were: physically attractive, sexually exciting, protective, intelligent, coercive, sensitive, impulsive, selfish, trustworthy, good parent for child, dominant, healthy, masculine, wealthy, volatile, threatening, cooperative, manipulative, helpful, and controlling. Using the same scale, participants also evaluated the desirability of each face as a short-term mate (STM) and a long-term mate (LTM).

In session two, each returning woman was asked to select and rate the same set of faces, using the procedures employed during session one. When these tasks had been completed, she was required to respond to the 60-item Bem Sex Role Inventory (BSRI) and the Rosenberg Self-Esteem (SES) questionnaire (Bem, 1974; Rosengerg, 1965). Each woman then completed a personal history form that included questions concerning the exact date of her last menses (first day of bleeding), the typical length of her menstrual cycle, regularity of her cycle, her age, her prior hormone use, and her pregnancy history. Finally, each participant was given a debriefing statement that explained the purpose of the experiment, and she was requested to telephone or e-mail the date of the onset of her next menses using an assigned identification number. The requirement for a post-experiment menses report was waived for those participants who, on the second session, reported a menses onset between the two experimental sessions.

For each participant, the date of their last ovulation was computed using either their post-experimental menses report or the reported menses between their experimental sessions. Although the duration of the menstrual cycle varies among females, this variance is almost exclusively confined to the follicular (pre-ovulatory) phase. Thus, ovulation is almost exactly 14 days prior to the onset of the next menses, irrespective of the cycle length (Fluhmann, 1957; Lein, 1979; Matsumoto, Nogami, & Okhuri, 1962). In a 28-day cycle, ovulation occurs on about the 14th day; but in a 34-day cycle, ovulation occurs on about the 20th day

(Katchadourian, 1980). In the absence of direct hormone measures, this counting backwards procedure offered the most accurate method for determining the date of ovulation that was closest to these participant's test dates. Knowing the date of ovulation allowed each subject's menstrual state on the days of testing to be defined in terms of the number of days before or after ovulation.

3. Results

For each target face, the participants' selections made during the follicular phase (pre-ovulation) were compared to those made during the luteal phase (post-ovulatory) of their menstrual cycle. A series of paired t tests revealed that that there was no significant difference in any face choice as a function of whether the woman was in her follicular or luteal menstrual phase on the day of testing. A second analysis examined only those participants who had been tested during the time of highest conception risk. Based on the probability of conception over the menstrual cycle, the time of highest conception risk was defined as the 9 days prior to ovulation (Barrett & Marshall, 1969). Twenty-nine of the women had been tested within this high-risk window. A within-subject analysis of their data revealed that within the high-risk window, these women selected an attractive male face that was significantly more masculine than their choice outside of this window [$t(28)=2.20$, $P=.02$]. Furthermore, of all the male and female facial preferences that were collected from these women, the attractive male face was the only preference that changed significantly with conception risk. These findings replicate and extend the menstrual shift effect first reported by Penton-Voak et al. (1999). The observed shift in preference toward a more masculine male face was a change in preference from a mean frame number of 299 during the low-risk phase, to frame 270 during the high-risk phase — a mean shift of 29 frames of the morph movie.

There was considerable variance in the size of the menstrual shift among the female participants. These individual differences were clarified by examining the size of the menstrual shift as a function of participants' scores on the BSRI. On the BSRI, individuals are classified as androgynous if they score above the median value of 4.9 on both the masculinity (Bem-M) and femininity (Bem-F) scales (Bem, 1974). When female participants were classified in this manner, the size of the menstrual shift was significantly larger for the non-androgynous women (56 frames) compared to the androgynous women (-5 frames); [$t(27)=2.55$, $P=.017$]. A more detailed analysis revealed that the shift in preference toward a more masculine male during the high-risk days of the menstrual cycle was most closely related to a participant's score on the masculinity scale. That is, there was a significant inverse correlation between the size of participants' menstrual shifts and their score on the Bem-M scale ($r=-.40$, $P=.03$). Also, the size of the menstrual shift was significantly different for participants classified as above or below the median on the Bem-M scale [$t(27)=2.79$, $P=.009$]. The low masculinity group's average attractive male face changed from frame 298 to frame 245 (a 53-frame shift), whereas the high masculinity group changed from 299 to 315, 16 frames in the reverse direction. Fig. 2 shows the preference shift for the low masculinity group.

Fig. 2. From left to right: the perceived average male face (frame 394) and the attractive male faces selected by the low masculinity female participants when conception risk was low (frame 298) and when conception risk was high (frame 245). The originals were color images.

Participants in the high Bem-M group also scored higher on the Rosenberg SES questionnaire [$t(27)=2.12$, $P=.04$]. Finally, irrespective of menstrual phase, high Bem-M and low Bem-M women differed on the frame they selected as exemplifying a dominant male face [$t(27)=2.46$, $P=.02$]. The high Bem-M group selected an extreme masculine face (frame 6), whereas the low Bem-M women perceived a much less masculinized male face (frame 63) as depicting a dominant male. This was the only facial preference that differed between these two groups of participants.

For all 29 participants, the selected mean frame number, over sessions, was computed for each target face. For male faces, the means were as follows: DoM (43), MaM (115), AtM (284), HtM (275), GfM (341), ItM (385), AvM (394), Pat (699). The attractive male face was significantly more masculine than the average male face [$t(28)=4.39$, $P<.0001$] and significantly different from all other target faces with the exception of the healthy male face [$t(28)=0.64$, $P=.74$]. A similar analysis of the female target faces revealed the following mean frame numbers: Pat (699), AvF (925), GmF (959), ItF (995), HtF (1021), FmF (1053), AtF (1074), DoF (1195). In agreement with prior studies, the attractive female face was significantly more feminized than the average female face [$t(28)=6.93$, $P<.0001$] and differed from all others except the feminine female face [$t(28)=1.17$, $P=.88$].

Each of the 29 high-risk participants had selected eight male face types (including Pat) during each experimental session, and subsequently rated each of these faces on 20 different attribute scales. Averaged across sessions, this procedure yielded 232 male faces with mean ratings for all 20 attributes. A Principal Components Analysis was performed on the correlation matrix between the 20 attribute ratings that participants had assigned to these faces. Three factors accounted for 75% of the variance: 46%, 22%, and 7%, respectively. All three factors were rotated using the Varimax method. The rotated factor pattern indicated that the first factor was heavily loaded with attributes that are considered socially negative. In order of importance, these were: threatening, volatile, controlling, manipulative, coercive, selfish, dominant, and impulsive. On this basis, factor one (F1) was labeled the "Enemy" factor. F2 was named the "Friend" factor because, in order of magnitude, it was loaded with the following positive attributes: helpful, cooperative, trustworthy, good father, wealthy, and intelligent, all characteristics of a desirable friend. F3 was descriptively called the "Lover" factor. In order of importance, F3 factor loadings were as follows: physically attractive, sexually exciting, masculine, healthy, and protective. Fig. 3 shows third-order polynomial curves fitted to the mean factor scores of all three factors over the first 700 frames of the morph movie, from the extreme masculine face to the androgynous face. The curve for F1 reveals a progressive increase in perceived threat with increasing masculinity and a pronounced increase in F1 scores with extreme masculinity. Perceived friendliness declines rapidly with increasing masculinity, and the Lover factor, F3, peaks close to the location where masculinized faces received higher scores on F1 than F2.

Fig. 4 shows third-order polynomial curves fitted to the mean ratings of male faces with respect to their suitability as a STM and as a LTM. These preference curves are plotted separately for the high Bem-M and low Bem-M groups. For the high Bem-M participants, the LTM and STM rating curves are similar. In both cases, the preference rating initially increases with increasing masculinization, reaches a maximum value, and then declines with further masculinization. This same preference pattern is reflected in the LTM curve of the low-Bem-M participants, but the STM curve for this group shows a remarkably different pattern. In this case, the desirability of a male as a STM continues to increase with facial masculinization. That is, women who shift their preference toward a more masculine male face during the high-risk phase of their menstrual cycle (low Bem-M group) also prefer STMs with very

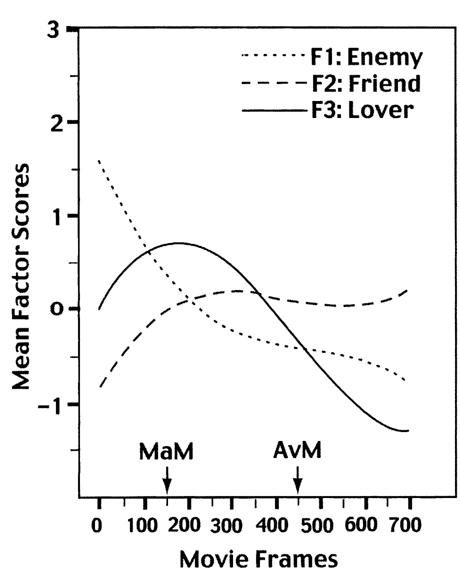

Fig. 3. Third-order polynomial curves fitted to the mean factor scores of all three factors (Enemy, Friend, and Lover) over the first 700 frames (masculine images) of the morph movie. For reference, the positions of the masculine male (MaM) and average male (AvM) are marked.

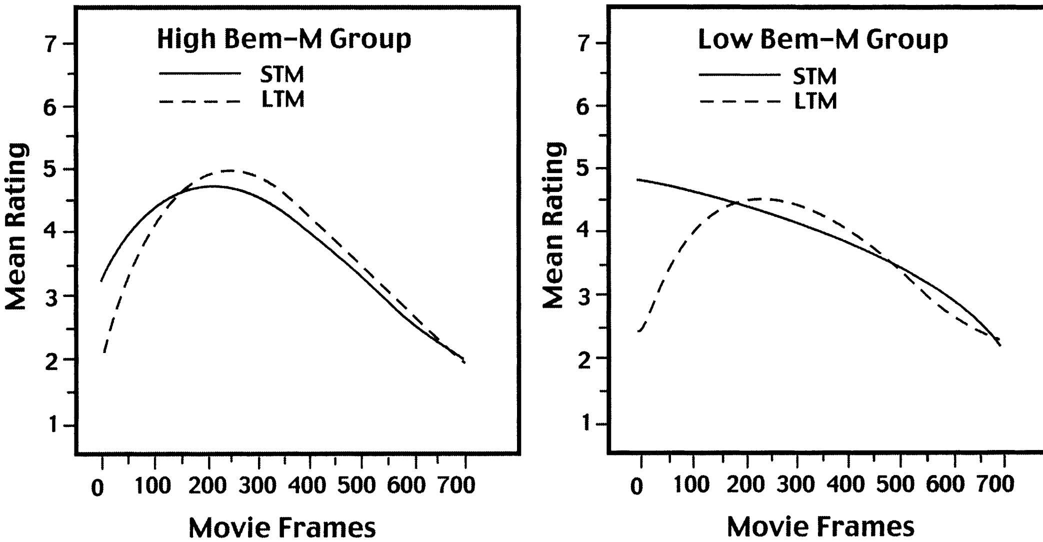

Fig. 4. Third-order polynomial curves fitted to the mean ratings of faces for desirability as a STM and a LTM, for participants who scored high or low on the Bem masculinity scale.

masculine male features. To examine the relationship between these facial preferences and the personality traits assigned to faces, LTM and STM ratings were correlated with the F1, F2, and F3 factor scores of the same faces. For high Bem-M participants, their LTM and STM mean ratings were significantly correlated with both the Friend ($r=.57$, $P<.001$ and $r=.44$, $P<.001$, respectively) and Lover ($r=.62$, $P<.001$ and $r=.74$, $P<.001$, respectively) factors, but not the Enemy factor. Similarly, the low Bem-M group's LTM ratings were significantly correlated with the Friend ($r=.41$, $P<.001$) and Lover ($r=.54$, $P<.001$) factors, but their STM ratings were correlated with the Enemy factor ($r=.45$, $P<.001$) and not significantly correlated with the Friend factor ($r=-.08$). These relationships indicate that the personality attributes associated with desirable STMs are quite different between the high Bem-M and low Bem-M participants.

4. Discussion

The morph movie used in the current experiment offers a sensitive tool for examining facial preferences by providing participants with a choice among male and female facial images that have been masculinized and feminized to varying degrees. Unlike caricaturing techniques that masculinize or feminize faces on the basis of questionable mathematical assumptions, the morph movie systematically changes the major facial features and proportions that are perceived to characterize human maleness and femaleness following puberty. Furthermore, because such secondary sexual characteristics are mainly a consequence of different levels of pubertal hormones (Grumbach, 2000; Tanner, 1978), this methodology provides a basis for interpreting how facial preferences are related to the degree to which such hormonal markers are displayed on the faces of men and women.

In agreement with prior studies, the current results support the conclusion that women prefer male faces that are more masculinized than an average male face (Mealey et al., 1999; Penton-Voak & Perrett, 2000; Thornhill & Gangestad, 1993). That is, the attractive male face possesses more extreme testosterone markers, such as a longer, broader, lower jaw, and more pronounced brow ridges and cheekbones than the average male face. These same hormone markers are also associated with good health. That is, when participants were required to select a healthy male face, their choice was not significantly different from the attractive male face. This suggests that women consider such testosterone markers to be an index of good health and that important health considerations may underlie their aesthetic preference. The relationship between attractiveness and perceived health is also evident in the analysis of the participants' attribute ratings of male faces. On the basis of their close correlation, the factor analysis grouped together physical attractiveness, sexually exciting, masculinity, healthy, and protectiveness as a single factor, F3, that changed systematically with increasing masculinization (Fig. 3). An examination of the F3 factor scores, however, reveals that the relationships among health, attractiveness, and hormone markers are not linear. Although the "Lover" scores initially increase with facial masculinization, they reach a maximum value and then decline with further masculinization. This observation suggests that either negative attributes are associated with high levels of testosterone, or positive attributes are perceived to decline under these conditions; both hypotheses appear to be true. An examination of Fig. 3 reveals that the Enemy factor (F1) increases and the Friend factor (F2) decreases at high levels of masculinity. Clearly, the women in this study viewed pronounced testosterone facial markers to be associated with dominance, unfriendliness, and a host of negative traits (threatening, volatile, controlling, manipulative, coercive, and selfish). The causal connection between testosterone levels and these behavioral attributes is still controversial (see review by Mazur & Booth, 1998), but if such relationships are valid, then the aesthetic preference of human females could be viewed as an adaptive compromise between the positive attributes associated with higher-than-average testosterone (health cues) and the negative attributes associated with more extreme masculinization.

The changes in females' preferences over their menstrual phase provide additional support for an adaptive model of aesthetic preference. First, of all the preferences that were examined, only the attractive male face varied as a function of menstrual phase. Second, this shift in preference occurred under very specific circumstances: during the 9 days prior to ovulation, when conception risk is highest. This is also the menstrual phase when there is a unique hormonal mix — high estrogen levels accompanied by low progesterone levels. Third, the overall shift in preference was very specific; a more masculine male face was preferred when conception risk was high. Taken together, these three observations are strong evidence for adaptive design. It appears that a female's attraction to the testosterone markers on a male's face may be influenced by her estrogen/progesterone ratio. This suggests that the neural mechanism responsible for generating such positive feelings is sensitive to these circulating hormone levels. The observation that the preference change is restricted to the time of high conception risk indicates that the "healthy male" preference, discussed above, may involve genetic factors (e.g., immunocompetence) rather than, or in addition to, phenotypic healthiness. Indeed, Scheib et al. (1999) have shown that a generally accepted measure of genotypic quality across species, FA, is correlated with both attractiveness and a composite index of masculine features that included the length of the lower jaw and prominent cheekbones. From this perspective, the positive feelings evoked by testosterone markers are a fitness-enhancing adaptation.

The experimental participants showing the largest menstrual shift in their attractive male preference were those scoring lowest on androgyny: more specifically, lowest on the Bem masculinity scale (Bem-M). The observed correlation between Bem-M scores and the size of the shift suggests that this relationship should be viewed as a continuum rather than a dichotomy. For the purpose of discussion, however, it is useful to view the behaviors of these two groups as representing two different reproductive strategies. In contrast to the high Bem-M group, whose STM and LTM preferences are consistent and relatively stable across their menstrual cycle, the low Bem-M participants shift their preference toward a more masculinized male, both for STMs, and during the high conception risk phase of their cycle. That is, the low Bem-M group may be characterized as employing an opportunistic strategy rather than a stable mate choice strategy. An opportunistic strategy offers a solution to what Cashdan (1996) describes as a woman's conflict between finding a mate who will invest (LTM) and securing good genes for her offspring (STM); different males are preferred for different functions. It has been suggested that such mating strategies arise as a function of the security of attachment to primary caregivers during childhood (Bowlby, 1969, 1973, 1980; Draper & Harpending, 1982). Girls who grow up without fathers, for example, are more likely to mature earlier, exhibit "precocious" sexuality, have low self-esteem, and have difficulty forming long-term relationships (Chisholm, 1993; Draper & Harpending, 1982; Jones, Leeton, McLeod, & Wood, 1972; Moffitt, Caspi, & Belsky, 1992; Surbey, 1990). Although the current research findings do not reflect directly on these hypotheses, it is of interest that the "opportunistic" females were lower in self-esteem and more sensitive to male dominance cues than those with a stable mate choice strategy. Perhaps, father–daughter bonding enhances a female's self-esteem and reduces her sensitivity to male dominance cues, while a lack of attachment has the reverse effect. Future research should attempt to further characterize these "opportunistic" females by directly measuring their parental bonding and their performance on cognitive tasks, such as mental rotations, that could provide a second measure of "masculinization" for comparison with their low Bem-M scores. In this regard, it is noteworthy that males' performance on such mental rotation tasks has already been shown to decline with father absence during childhood (Draper & Harpending, 1982).

In the current study, participants (a) expressed a preference for male faces that were more masculine than the average, (b) shifted their preference toward a more masculine face during periods of high conception risk, and, in some cases, (c) rated very masculine faces as most desirable for STMs. All such preferences and shifts in preference, however, were toward male faces that were less symmetrical than the average male face. This is convincing evidence that participants' choices were strongly influenced by hormone markers, and enhanced measured symmetry, as an index of immunocompetence, cannot account for these preference shifts. It is still possible, however, that the facial hormone markers serve as an honest index of immunocompetence and explain why females are attracted to such features, but there may be another reason for this preference. This alternative hypothesis arises from an examination of the specific facial and bodily features that are influenced by testosterone at puberty.

Boys and girls enter puberty with almost identical proportions of muscle, fat, and bone, but they exit this critical phase of development as reproductive adults with very different body shapes and compositions. During puberty, a male's body undergoes a pronounced adolescent growth spurt regulated by testosterone and aromatized testosterone in the presence of growth hormone (Grumbach, 2000; Tanner, 1978). By the end of puberty, men have about 1.5 times the skeletal and muscle mass of women, whereas women have stored twice as much body fat as men (Forbes, 1975). Facial changes parallel these modifications to a male's body. On average, men have more pronounced brow ridges, sunken eyes, and bushier eyebrows set closer to the eyes (Farkas, 1981). Both the nose and the mouth are wider in the male face, while the lower jaw is both wider and longer than that of an average female face (Farkas, 1981). All of these male characteristics appear to have little relevance in today's world, but they would have been useful during his long ancestral history as a hunter. The enlarged openings of the mouth and nostrils provide effective passageways for the rapid transport of air to and from the lungs. This enhanced airflow, together with the larger vital capacity of his lungs, is necessary for an adequate supply of oxygen to support the higher metabolic rate and hemoglobin level required for the efficient use of his larger muscle mass. These attributes are clearly advantageous for a hunting lifestyle. Less obvious are the adaptations around the eyes. Hunting requires a high degree of energy expenditure that inevitably entails profuse sweating from the brow and other regions of the body in order to regulate body temperature. Large bushy eyebrows set close to the eye on protruding brow ridges provide an effective method for excluding sweat from the eye sockets as well as offering protection from an overhead sun. Given these considerations, a female's preference for a more masculinized male during periods of high conception risk may reflect an ancestral adaptive strategy for acquiring the genes of a healthy hunter — a direct benefit to her male offspring.

The healthy hunter hypothesis gains additional support from examining the origins of other human aesthetic preferences. Why does sugar have a pleasant taste, or rotten eggs have an unpleasant odor, or tissue damage feel unpleasant? These evaluative sensory feelings are present at birth, but they do not reflect the inherent properties of events in the external world. That is, hydrogen sulfide gas does not possess an unpleasant odor; rather, it is the human brain that has evolved to generate an unpleasant sensation for an event that is biologically dangerous, in this case, contaminated food. Indeed, Johnston (1999) has argued that all pleasant or unpleasant sensory feelings (and emotions) have evolved as proximate evaluations of events in the physical (or social) world that consistently posed threats or benefits to gene survival in ancestral environments. In some instances, such as the pleasant taste of sugar, these preferences may have little current value, and may even be maladaptive in the modern world. Prior to sugar refineries, however, individuals with a sugar preference ate ripe fruit, the major available source of sugars, and benefited from both a healthy diet and a rich energy supply for manufacturing ATP. When viewed from this perspective, male facial beauty may be an evaluative feeling evoked by attributes that were more important in a hunter–gatherer era than in today's environment. Indeed, it is difficult to comprehend the modern female's preference for males with prominent gluteus maximus muscles (buttock muscles) without considering the ancestral environment when such preferences had clear fitness value.

Facial beauty is a positive feeling generated within the brains of men and women, in response to a complex configuration of visual cues displayed on the faces of males and females. It now appears that such features and proportions are predominantly the result of pubertal hormones, and these hormone markers may indicate fitness-enhancing attributes that are, or were, of importance to the opposite sex. It is also apparent that the hormonal state of viewers, as determined by their sex and/or reproductive status, may modulate the intensity of the feelings evoked by these hormone markers (Krug, Plihal, Fehm & Born, 2000). When viewed within this framework, facial beauty is not a trivial matter, but rather, it is the product of the co-evolution of fitness cues (hormone markers) and feelings (fitness monitors) that have, or had, important reproductive consequences for both males and females.

References

Alley, T. R., & Cunningham, M. R. (1991). Averaged faces are attractive, but very attractive faces are not average. *Psychological Science*, *2*, 123–125.

Barrett, J. C., & Marshall, J. (1969). The risk of conception on different days of the menstrual cycle. *Population Studies*, *23*, 455–461.

Beier, T., & Neely, S. (1992). Feature-based image metamorphosis. *Computer Graphics*, *26*, 35–42.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*, 155–162.

Bowlby, J. (1969). Attachment and loss. *Attachment*, *vol. 1*. New York: Basic Books.

Bowlby, J. (1973). Attachment and loss. *Separation: anxiety and anger vol. 2*. New York: Basic Books.

Bowlby, J. (1980). Attachment and loss. *Loss*, *vol. 3*. New York: Basic Books.

Cashdan, E. (1996). Women's mating strategies. *Evolutionary Anthropology*, *5*, 134–143.

Chisholm, J. S. (1993). Death, hope, and sex. *Current Anthropology*, *34*, 1–24.

Cunningham, M. R., Barbee, A. P., & Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of Personality and Social Psychology*, *59*, 61–72.

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. (1995). "Their ideas of beauty are, on the whole, the same as ours": consistency and variability in the cross-cultural perception of female physical attraction. *Journal of Personality and Social Psychology*, *68*, 261–279.

Dalton, K. (1982). Premenstrual tension: an overview. In: R. C. Friedman (Ed.), *Behavior and the menstrual cycle* (pp. 276–290). New York: Marcel Dekker.

Draper, P., & Harpending, H. (1982). Father absence and reproductive strategy: an evolutionary perspective. *Journal of Anthropological Research*, *38*, 255–273.

Farkas, L. G. (1981). *Anthropometry of the head and face in medicine*. New York: Elsevier.

Fluhmann, C. F. (1957). Irregularities of the stages of the menstrual cycles. *Western Journal of Surgery, Obstetrics, and Gynecology*, *65*, 265–277.

Forbes, G. B. (1975). Puberty: body composition. In: S. R. Berenson (Ed.), *Puberty* (pp. 132–145). Leiden: Stenfert-Kroese.

Gangestad, S. W., & Thornhill, R. (1997). The evolutionary psychology of extrapair sex: the role of fluctuating asymmetry. *Evolution and Human Behavior*, *18*, 69–88.

Gangestad, S. W., Thornhill, R., & Yeo, R. A. (1994). Facial attractiveness, developmental stability, and fluctuating asymmetry. *Ethology and Sociobiology*, *15*, 73–85.

Gomes, J., Darsa, L., Costa, B., & Velho, L. (1999). *Warping and morphing of graphical objects*. San Francisco: Morgan Kaufmann Publishers.
Gomi, A. (1998). *Americans 1.0, Los Angeles 1994*. Bunkasha, Japan: Digitalogue (CD-ROM).
Grammer, K., Fieder, M., & Fink, B. (1998). *Facial explorer, version 1.0*. Vienna: Ludwig-Boltzmann Institute for Urban Ethology (Computer software).
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.
Grumbach, M. M. (2000). Pubertal maturation in aromatase deficiency and resistance to estrogen. In: J. P. Bourguignon & T. M. Plant (Ed.), *The onset of puberty in perspective* (pp. 247–267). Amsterdam: Elsevier.
Johnston, V. S. (1994). Methods and apparatus for generating composites of human faces. US Patent no. 5,375,195.
Johnston, V. S. (1999). *Why we feel: the science of human emotions*. Reading, MA: Perseus Press.
Johnston, V. S. (2000). Female facial beauty: the fertility hypothesis. *Pragmatics and Cognition*, *8*, 107–122.
Johnston, V. S., & Franklin, M. (1993). Is beauty in the eye of the beholder? *Ethology and Sociobiology*, *14*, 183–199.
Johnston, V. S., & Oliver-Rodriguez, J. C. (1997). Facial beauty and the late positive component of event-related potentials. *The Journal of Sex Research*, *34*, 188–198.
Jones, B., Leeton, J., McLeod, I., & Wood, C. (1972). Factors influencing the age of menarche in a lower socio-economic group in Melbourne. *Medical Journal of Australia*, *21*, 533–535.
Katchadourian, H. A. (1980). *Fundamentals of human sexuality* (3rd ed.). New York: Holt, Rinehart, and Winston.
Keating, C. F. (1985). Gender and physiognomy of dominance and attractiveness. *Social Psychology Quarterly*, *48*, 61–70.
Krug, R., Plihal, W., Fehm, H. L., & Born, J. (2000). Selective influence of the menstrual cycle on perception of stimuli with reproductive significance: an event-related potential study. *Psychophysiology*, *37*, 111–122.
Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Review*, *1*, 115–121.
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.
Langlois, J. H., Roggman, L. A., Musselman, L., & Acton, S. (1991). A picture is worth a thousand words: reply to "On the difficulty of averaging faces". *Psychological Science*, *2*, 354–357.
Lein, A. (1979). *The cycling female: her menstrual rhythm*. San Francisco: Freeman.
Matsumoto, S. Y., Nogami, Y., & Ohkuri, S. (1962). Statistical studies on menstruation: a criticism on the definition of normal menstruation. *Gunma Journal of Medical Science*, *11*, 297–309.
Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences*, *21*, 353–363.
Mealey, L., Bridgstock, R., & Townsend, G. C. (1999). Symmetry and perceived facial attractiveness. *Journal of Personality and Social Psychology*, *76*, 151–158.
Miller, G. (2000). *The mating mind: how sexual choice shaped the evolution of human nature*. New York: Doubleday.
Moffitt, T. E., Caspi, A., & Belsky, J. (1992). Childhood experience and the onset of menarche: a test of a socio-biological model. *Childhood Development*, *63*, 47–58.
Møller, A. P., & Swaddle, J. P. (1997). *Asymmetry, developmental stability, and evolution*. Oxford: Oxford Univ. Press.
Møller, A. P., & Thornhill, R. (1997). Bilateral symmetry and sexual selection: a meta-analysis. *American Naturalist*, *151*, 174–192.
Møller, A. P., & Thornhill, R. (1998). Male parental care, differential parental investment by females and sexual selection. *Animal Behavior*, *55*, 1507–1515.
Oliver-Rodriguez, J. C., Guan, Z., & Johnston, V. S. (1999). Gender differences in late positive components evoked by human faces. *Psychophysiology*, *36*, 176–185.
Penton-Voak, I. S., & Perrett, D. I. (2000). Female preference for male faces change cyclically: further evidence. *Evolution and Human Behavior*, *21*, 39–48.
Penton-Voak, I. S., Perrett, D. J., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, *399*, 741–742.

Perrett, D. I., Burt, M. D., 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. L., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, *394*, 884–887.

Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, *368*, 239–242.

Rikowski, A., & Grammer, K. (1999). Human body odour, symmetry and attractiveness. *Proceedings of the Royal Society of London, Series B*, *266*, 869–874.

Rosengerg, M. (1965). *Society and the adolescent self-image*. New Jersey: Princeton University Press.

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London, Series B*, *266*, 1913–1917.

Surbey, M. (1990). Family composition, stress, and human menarche. In: F. Bercovitch, & T. Zeigler (Eds.), *The socio-endocrinology of primate reproduction*. New York: Liss.

Tanner, J. M. (1978). *Foetus into man: physical growth from conception to maturity*. London: Open Books.

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: averageness, symmetry, and parasite resistance. *Human Nature*, *4*, 237–269.

Thornhill, R., & Gangestad, S. W. (1994). Human fluctuating asymmetry and sexual behavior. *Psychological Science*, *5*, 297–302.

Waynforth, D. (1998). Fluctuating asymmetry and human male life history traits in rural Belize. *Proceedings of the Royal Society of London, Series B*, *265*, 1497–1501.

Williams, C. G. (1966). *Adaptation and natural selection: a critique of some evolutionary thought*. Princeton: Princeton University Press.

Wolberg, G. (1990). *Digital image warping*. Los Alamitos, CA: IEEE Computer Society Press.