ELSEVIER

Evolution and Human Behavior 25 (2004) 229–241

Evolution and Human Behavior

High salivary testosterone is linked to masculine male facial appearance in humans

Ian S. Penton-Voak[a,*], Jennie Y. Chen[b]

[a] Department of Psychology, University of Stirling, Stirling FK9 4LA, UK
[b] Department of Psychology, University of Texas, Austin, TX, USA

Received 29 September 2003; accepted 28 April 2004

Abstract

Although many studies of male facial attractiveness assume that facial masculinity is related to circulating testosterone levels in adult males, there is little empirical evidence in support of this assumption. Here, we used salivary testosterone assays to investigate the relationship between circulating testosterone and both masculinity and attractiveness of facial appearance by (1) constructing digital composites from the faces of men with high and low testosterone, which were presented using a forced-choice task to subjects and (2) using a forced-choice task in which participants judged the masculinity of pairs of original photographs. Composites from high-testosterone men were judged to be more masculine than those from low-testosterone men. Evidence that high-testosterone composites are considered more attractive than low-testosterone composites was equivocal. The forced-choice task using the original face images indicated that participants identified faces associated with relatively high circulating testosterone as being more masculine than faces of men with lower circulating testosterone. This effect was more pronounced when the faces in the pair were from men who differed greatly in testosterone levels. These preliminary findings provide support for the underlying assumptions of much attractiveness research, particularly studies that have identified systematic variation in female preferences for masculine faces.

Keywords: Facial attractiveness; Masculinity; Testosterone

* Corresponding author.
E-mail address: isp1@stir.ac.uk (I.S. Penton-Voak).

1090-5138/04/$ – see front matter
doi:10.1016/j.evolhumbehav.2004.04.003

1. Introduction

In many species, males advertise their condition through sexually dimorphic traits that are testosterone dependent (e.g., plumage patches in birds, colouration in some fish species; see Andersson, 1994, for review). As testosterone has immunosuppressive effects, these dimorphic traits have been proposed to be honest signals of heritable quality to potential mates (Folstad & Karter, 1992). Although this remains a controversial proposal (Getty 2002), offspring viability has been demonstrated to covary with the condition of male sexually dimorphic traits in some species (Jennions & Petrie, 2000, for review). Recently, researchers interested in biologically based explanations of human behaviour have attempted to examine the influence of sexually dimorphic facial characteristics on judgements of male and female attractiveness. Facial attractiveness has considerable influence on mating and other social behaviours in our own species (see Fink & Penton-Voak, 2002, for review).

Various techniques have been used to manipulate sexual dimorphism (masculinity/femininity) in facial attractiveness research. For example, Perrett et al. (1998) developed caricaturing techniques to exaggerate the differences between average male and female faces, creating artificially masculinised and feminised images for use in preference tests. Johnston, Hagel, Franklin, Fink, and Grammer (2001) used rater perceptions of masculinity and femininity (rather than male–female shape differences) to drive an algorithm that generated masculine and feminine faces. More recently, Swaddle and Reierson (2002) used facial metric growth data to estimate the effects of pubertal testosterone on male facial growth, and then developed a morphing technique that attempted to manipulate only those facial features that appeared to be testosterone dependent. Whereas all of these techniques generate faces that differ in perceived masculinity, each method has drawbacks (e.g., for criticisms of caricaturing methods, see Meyer & Quong, 1999; for criticisms of facial metrics as a measure of masculinity, see Penton-Voak et al., 2001).

The results of these studies are somewhat inconsistent, with some finding an overall preference for masculine faces (e.g., Johnston et al., 2001) and others finding preferences for average (Swaddle & Reierson, 2002) or feminised male faces (Perrett et al., 1998). Studies using real rather than manipulated faces also fail to find consensus in attractiveness judgements of masculine or "dominant" faces. Some studies find preferences for dimorphic characteristics in faces (Keating, 1985; Grammer & Thornhill, 1994), whereas others find preferences for less masculine/dominant faces (Berry & McArthur, 1985) or preferences for faces that present a combination of masculine and feminine features (Cunningham, Barbee, & Pike, 1990).

The discrepancies between these results are further complicated by recent findings demonstrating that preferences for sexual dimorphism can vary systematically across individuals, time, and with the experimental task. For example, preferences for facial masculinity vary with conception risk across the menstrual cycle (Johnston et al., 2001; Penton-Voak et al., 1999) as a function of rater attractiveness (Little, Burt, Penton-Voak, & Perrett, 2001; Penton-Voak et al., 2003) and differ when a face is being judged in the context of a short- or long-term relationship (Little, Jones, Penton-Voak, Burt, & Perrett, 2002).

One coherent theoretical model to account for these findings suggests that shifting preferences represent adaptive trade-offs in mate choice (Gangestad & Simpson, 2000; Penton-Voak & Perrett, 2001) Masculinised faces elicit stereotypical attributions of coldness, dominance, and dishonesty (for a review, see Penton-Voak & Perrett, 2001). The association between testosterone and antisocial behaviour (Dabbs, 1993) suggests that even if masculine faces predict heritable health benefits for offspring, they may indicate poor paternal social behaviour. Indeed, a recent related finding is that whereas masculine male faces appear (and may actually be) more healthy, they are not necessarily more attractive than relatively feminine male faces (Rhodes, Chan, Zebrowitz, & Simmons, 2003). Lower testosterone men may have more feminine faces that attract positive personality attributions, but are predicted by theory to be less healthy than high-testosterone men (see Fink & Penton-Voak, 2002; Penton-Voak & Perrett, 2001). The relative attractiveness of cues to "prosociality" and "good genes" may vary according to the situation in which a choice is being made. For example, cues to health and, hence, heritable benefits for offspring may be of primary importance in a short-term or extra-pair partner, whereas cues to future paternal investment and prosociality may be favoured in a potential long-term partner (Fink & Penton-Voak, 2002; Penton-Voak & Perrett, 2001).

A weakness of all of the techniques used to generate stimuli in the above studies, however, is that relationships between any measure of circulating testosterone, facial structure and social perception have not, to our knowledge, been empirically demonstrated in adult faces (although a relationship between facial expression and testosterone has been reported previously; Dabbs, 1997). This relationship also needs to be demonstrated to support the link between facial appearance, testosterone, and behaviour that underpins the adaptive trade-off model summarised above. We used salivary testosterone assays to examine the relationship between circulating testosterone, facial structure, and the social attributions made to the faces of men with high and low testosterone levels.

One method of assessing the facial correlates of circulating testosterone is to use digital composites (averages) of multiple individual faces. Computer graphic composites can be used to extract the defining characteristics of a group, whilst losing the characteristics that make each face look individual (Rowland & Perrett, 1995). If high- and low-testosterone men have faces that differ in systematic ways (e.g., if high-testosterone men have larger jaws, on average, than low-testosterone men), this technique should extract visibly different composite faces. Although the differences between subgroup composites are subtle, they reliably generate differing social perceptions (Penton-Voak et al., 2001). This technique also avoids making numerous measurements on the face, which may miss differences that drive social perceptions. Accumulating evidence suggests that face perception does not operate on feature-by-feature analysis, but that faces are perceived more holistically, with the global configuration of features influencing face perception (Hancock, Bruce, & Burton, 2000; Tanaka & Farah, 1993). Such configurations are difficult to detect with measurement alone.

A further advantage of using composite images is that they tend towards symmetry regardless of the fluctuating asymmetries in each individual image. Thus the use of composite faces controls for possible differences in fluctuating asymmetry (although any directional asymmetries will remain) between high- and low-testosterone individuals, allowing simpler comparisons between this study and previous studies of masculinity in face shape (e.g., Johnston et al. 2001; Penton-Voak et al., 1999, Perrett et al., 1998; Swaddle & Reierson, 2002). Study 1 below employs salivary testosterone assays and a composite image method to examine the relationship between circulating testosterone and facial appearance. Salivary assays allow for the noninvasive assessment of circulating testosterone levels because testosterone concentrations in saliva correlate highly with free serum concentrations (Granger, Schwartz, Booth, & Arentz, 1999).

For all their potential advantages, composite methods can be criticised on the grounds that the faces are not "real" and so lack ecological validity. Hence, we performed Study 2, in which the masculinity and attractiveness of pairs of unmanipulated, original face images is assessed with respect to the measured salivary testosterone of the face's owner.

2. Study 1 methods

2.1. Participants and photography

Fifty Caucasian males (mean age 20 years, S.D. 2.3) were photographed and had their circulating testosterone levels measured using a salivary assay. The participants were recruited from introductory psychology classes and upper-division classes at the University of Texas at Austin for either course requirement or monetary payment. These participants were screened for hormone or steroid consumption and a history of facial surgeries by survey. Men were excluded from participation if they had facial hair or facial hair and/or ear piercing that could not be removed. Participants were provided with a grey tee shirt and baseball cap worn backwards when posing for the pictures against a dark-coloured sheet. The caps and tee shirts were used to hide hair and clothing cues that might affect the ratings of the photos. Researchers used a 3-megapixel Fuji digital camera to take the photographs.

2.2. Testosterone assays

Participants were given a piece of chewing gum to facilitate saliva flow before providing a saliva sample for testosterone assay. All samples were collected at 10:00 a.m. to attempt to control for diurnal variations in testosterone production. Immunoassays were performed on the saliva samples collected using assay kits from Assay Designs. Sensitivity of the kit is reported at 3.82 pg/ml, and the assay range is reported as 7.81–2000 pg/ml. Performance specification of the assay kit is 7.8–10.8% for intra-assay variance and 9.3–14.6% for interassay variance. The samples were first vortexed to homogenize the liquid so testosterone would not bind with the protein particles in the samples, and then centrifuged to separate the saliva from solid particles. Mean testosterone levels were 164.4 pg/ml (S.D. = 88.7) with a maximum of 533.9 pg/ml and a minimum of 50.1 pg/ml. Full details of the procedure and specification of the kit can be found at http://www.assaydesigns.com.

2.3. Stimuli construction

To construct composites, 172 feature points are marked on facial landmarks on each face (for details on the choice of these landmark points, see Rowland & Perrett, 1995). The mean XY position of each delineated feature point on each face is then calculated to generate average shape information across the face set. "Average" colour is generated by rendering colour information from each individual into this average shape, and calculating mean RGB colour values across the face set for each pixel location. The results of the salivary testosterone assays were used to construct a "high-testosterone" composite face from the shape and colour of the 13 faces in the top quartile of the measured testosterone distribution, and a "low-testosterone" face from the 13 faces in the bottom quartile. The high- and low-testosterone groups did not differ significantly in any of the other physical traits measured (height, weight, and body mass index).

Comparison of high- and low-testosterone composites (Fig. 1) showed clear visible differences between the faces, with the high-testosterone composite appearing more masculine and more attractive. To enable a more robust test of the hypothesis that facial masculinity is associated with circulating testosterone a further 20 composites were generated. To generate these 20 faces, a median split was performed on the salivary testosterone measures, giving high- and low-testosterone groups with 25 faces in each. Ten random samples of 13 faces were drawn from each of these two 25-face groups, and composite faces were made from each of these samples. In combination with the two composites from the high- and low-testosterone quartiles, this made a total of 22 composites for use in more formal testing. There was no significant difference in the age of the men in the low- and high-testosterone groups (low, 20.2 years, S.D. = 2.3; high, 20.4, S.D. = 2.4; $t(48) = -0.24$, NS). Nor did the low- and high-testosterone composites differ in perceived age. Twenty-two raters (6 male, 16 female) were asked to estimate the age of one of the two composites. The mean estimated ages were 20.6 years, S.D. = 1.8, for the low-testosterone composite and 21.18 years, S.D. = 2.7, for the high-testosterone composite [$t(20) = -5.6$, NS].

2.4. Procedure

To assess whether composites made from the high-testosterone half of the 50 faces were perceived as being more masculine than the composites from the low-testosterone group, a forced-choice testing procedure was employed. One face was randomly selected from each group and the pair was presented together side by side (left–right counterbalancing). Each image was 200 by 267 pixels, and the screen resolution was set to 640×480. Participants were asked to choose which of the two faces presented was "more masculine." Responses were made by pressing keys that corresponded to the left or right faces. The task was self-paced. The faces were chosen without replacement, meaning that after 11 trials, each face had been seen once. This procedure was then repeated in a second block (giving new random high- and low-testosterone pairings, with each face being randomly assigned to the left or right position), making 22 trials in total. The percentage of trials that each participant chose a high-testosterone face was recorded as the dependent variable.

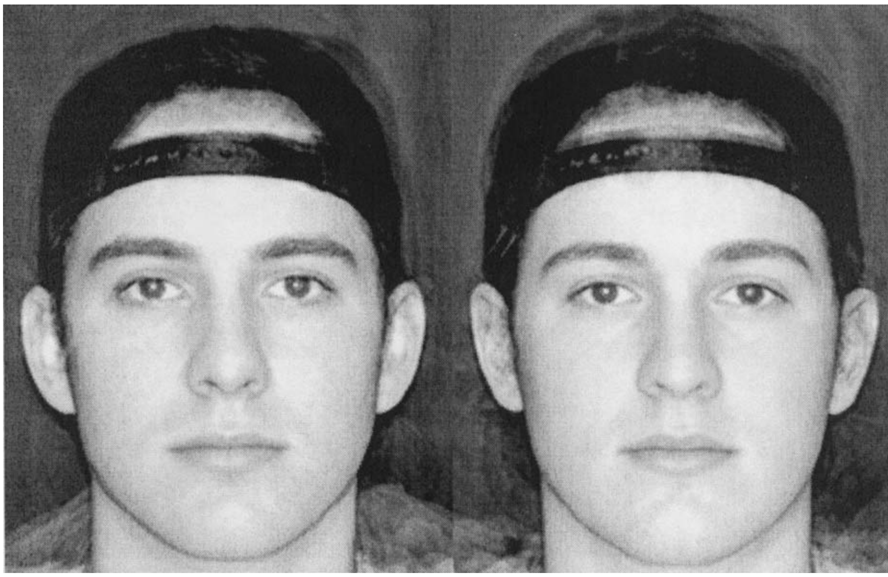

Fig. 1. Composites constructed from the faces of men in the high (left) and low (right) quartiles ($N = 13$ in each composite) of salivary testosterone.

To assess whether high-testosterone faces were perceived as more attractive than low-testosterone faces, an identical experiment was performed, with the participants asked to make an attractiveness judgement rather than a masculinity judgement. In this experiment, the percentage of trials in which a high-testosterone face was judged most attractive was recorded for each participant.

2.5. Raters

Thirty-three participants (mean age 23 years, 14 male) performed the masculinity judgement task. A further 40 participants (mean age 21 years, 20 male) performed the attractiveness task. These raters were all recruited from the student and staff population at Stirling University, and were all of White British ethnicity.

3. Results

3.1. Masculinity judgements

Across subjects, the percentage of trials in which the high-testosterone face was picked as the most masculine was 57% (S.D. 16%). A one-sample *t* test indicated that this value was significantly different from the chance value of 50% [$t(32) = 2.42$, $p = .021$]. Participants considered composites derived from faces in the top half of the salivary testosterone distribution to be more masculine than composites derived from faces in the bottom half.

To determine whether this effect was more robust when the most extreme high- and low-testosterone faces were present in a trial, a new mean was calculated for each subject using only trials in which the high-testosterone quartile face, the low-testosterone quartile face, or both, were presented. So, across participants, in trials involving the top quartile composite versus any low-testosterone composite, the bottom quartile composite versus any high-testosterone composite, or the top quartile composite versus the bottom quartile composite, the higher testosterone composite was picked as the most masculine 65% (S.D. 29%) of the time. This value was also significantly different from the chance rate of 50% [one-sample t test, $t(32) = 12.75$, $p < .001$]. Male and female raters produced statistically indistinguishable results [male mean = 57%, S.D. = 16%; female mean = 57%, S.D. 17%, $t(31) = -0.02$, NS].

3.2. Attractiveness judgements

An identical analysis was performed for the attractiveness data. Participants judged composites derived from faces in the top half of the salivary testosterone distribution as the most attractive of the pair significantly more often than composites derived from faces in the bottom half [59% of trials (S.D. 18%); one-sample t test against chance value of 50%, $t(39) = 3.14$, $p = .003$]. As before, to see if this result was exaggerated in trials when the highest and lowest testosterone faces were presented (as seemed to be the case for masculinity judgements), a new mean was calculated for each subject using only trials in which either the high-testosterone quartile face, the low-testosterone quartile face, or both, were presented. Across participants, the higher testosterone composite was picked as most attractive 53% (S.D. 27%) of the time. This suggests that in the trials where either the high- or low-testosterone quartile faces were presented, participants showed no reliable preference for high- or low-testosterone faces [one-sample t test against chance value of 50%, $t(39) = 0.74$, NS]. Again, there were no significant differences in the preferences of male (mean = 57%, S.D. = 18%) and female raters [mean = 61%, S.D. = 17%; $t(38) = 0.62$, NS].

4. Study 2 methods

4.1. Participants and testosterone assays

The same set of photographs and testosterone assays as employed in Study 1 were used in this study.

4.2. Procedure

To assess whether the individual faces of men with high and low salivary testosterone differ in masculinity, a similar forced-choice experiment to that used in Study 1 was devised. Pairs of images from the original face set were selected randomly and presented side by side. In each block (of 50 trials), each face appeared twice, once on each side of the screen (giving left–right counterbalancing). Participants were presented with two blocks, meaning each participant saw each face four times over the course of 100 trials.

Each image was 200×267 pixels, and the screen resolution was set to 640×480. Participants had to choose which of the two faces presented was "more masculine." Responses were made by pressing keys that corresponded to the left or right faces. The task was self-paced. The percentage of trials that each participant chose the face with the higher salivary testosterone measure of the pair was recorded as the dependent variable.

To assess whether the individual faces of men differed reliably in attractiveness, an identical experiment was conducted, with participants asked to decide which of the two faces was most attractive rather than most masculine.

4.3. Raters

Thirty-six participants (18 female) with a mean age of 20.6 years assessed the masculinity of the individual faces. Thirty-one different participants (13 female), mean age 20.8 years, assessed the attractiveness of the same individual faces.

5. Results

5.1. Masculinity judgements

Across subjects, the percentage of trials in which the higher testosterone face was picked as the most masculine was 53% (S.D. 6.4%). A one-sample *t* test indicated that this value was significantly different from the chance value of 50% [$t(35)=2.82$, $p=.008$]. There was no difference between responses made by male (mean = 53%, S.D. = 6%) and female (mean = 53%, S.D. = 6%) participants [$t(34)=-0.18$, NS]. Overall, participants reliably considered the face with relatively higher salivary testosterone of the pair to be more masculine than the face with relatively lower salivary testosterone. This result, although statistically significant, is a small effect. If circulating testosterone has facial correlates associated with attributions of masculinity, it is likely that masculinity discriminations between faces differing greatly in testosterone will be easier than those between faces close in testosterone.

To further investigate this possibility, a correlational analysis was performed. The individual faces were ranked 1–50 according to the subjects' salivary testosterone levels (the lowest testosterone face receiving a rank of 1, the highest a rank of 50). For each of the 3600 masculinity judgement trials (36 subjects, 100 trials each) the rank difference between the two faces presented was calculated. Data from pairs with equivalent rank differences were pooled and the mean probability of the higher ranked face being identified as the most masculine was calculated from participant responses (where less than 10 observations were available for a rank difference, that data point was excluded, as the S.E.s for such data points were high or unavailable. This applied to three rank differences: 47—nine observations, 48—six observations, and 49—one observation). A plot of rank difference against the probability of masculine identification of the high-testosterone face is shown in Fig. 2. There is a

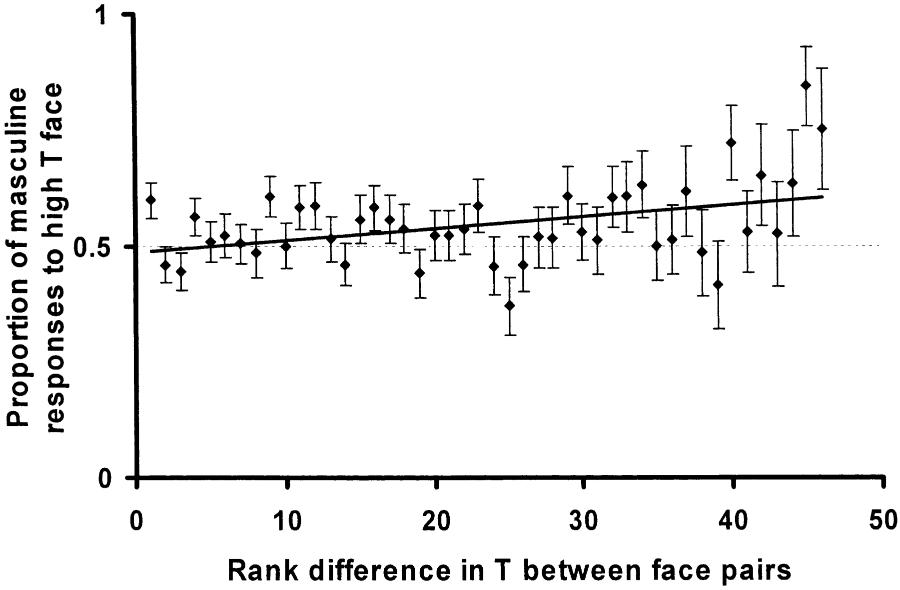

Fig. 2. Rank difference in salivary testosterone (T) measures between face pairs, and proportion of responses that identified the higher testosterone face of a presented pair to be more masculine. Error bars show S.E.M of responses to a given rank difference.

significant positive relationship between the rank difference of the two faces presented and the probability that the high-testosterone face is identified as more masculine (Spearman's $\rho = .33$, $N = 46$, $p = .024$).[1] There is no relationship between testosterone rank and age of the participant ($\rho = .16$, $N = 50$, NS).

5.2. Attractiveness judgments

Across subjects, the percentage of trials in which the higher testosterone face was picked as the most attractive was 50% (S.D. 4.6%). Clearly, participants are responding at chance levels [confirmed by a one-sample t test: $t(30) = 0.00$, NS]. There was no difference between responses made by male (mean = 51%, S.D. = 5%) and female (mean = 49%, S.D. = 4%) participants [$t(29) = 1.12$, NS]. Participants do not reliably consider the face with relatively higher salivary testosterone of the pair to be more attractive than the face associated with relatively lower salivary testosterone.

[1] Including the excluded rank differences (47,48, and 49) made no difference to this result (Spearman's $\rho = .33$, $N = 49$, $p = .019$).

6. Discussion

These data suggest that circulating testosterone is associated with facial masculinity because when composites are made from high- and low-testosterone faces, the former are perceived to be more masculine. Moreover, when the faces of individual men with different salivary testosterone measures are presented together, the face of the higher testosterone man is more likely to be judged as more masculine, an effect that is particularly pronounced when the testosterone differences between the men are large. These findings support the hypothetical link between facial appearance and testosterone proposed in earlier studies of human facial sexual dimorphism (Johnston et al., 2001, Perrett et al., 1998; Swaddle & Reierson, 2002).

Similarly, the data from Study 1 provide some evidence that circulating testosterone and facial attractiveness are positively correlated. This finding conflicts with some previous literature (Perrett et al., 1998; Swaddle & Reierson, 2002), but is consistent with other studies (e.g., Grammer & Thornhill, 1994; Johnston et al., 2001). No preference for high-testosterone faces was found in trials in which the high- and low-testosterone composites were presented, suggesting that a simple linear relationship between circulating testosterone and attractiveness is unlikely. Study 2 indicated that when presented with individual faces from men with differing salivary testosterone, participants did not associate either high- or low-testosterone measures with attractiveness. As mentioned in the introduction, the relationship between masculinity and attractiveness appears to vary with a number of contextual, hormonal, and life history factors. The link between circulating testosterone and masculinity appears to be more reliable than the link between circulating testosterone and attractiveness in this sample.

Although the data presented here are consistent with most assumptions concerning testosterone and facial appearance, some caution is advised given the nature of salivary testosterone assays, the effects of testosterone on facial growth, and the type of response we have collected. In addition to seasonal and diurnal variations, circulating testosterone levels can change relatively quickly in response to social interactions and internally generated cues. The reliability of repeated samples from an individual varied from $r = .52$ to $r = .64$ in one study, depending on the amount of time separating the two points of data collection (Dabbs, 1990). Although work on behavioural correlates of testosterone often relies on a single assay of each participant, a replication and extension of the current findings, preferably with repeated testosterone assays, would be desirable.

The assumption that testosterone is directly related to facial growth has been challenged by recent discoveries in endocrinology indicating that oestrogens (from direct secretion or the aromatization of testosterone) have a critical role in skeletal growth in males and females (Frank, 2003; Grumbach, 2000). Thus, although testosterone still has a direct role in stimulating the greater growth of the male skeleton (Riggs, Khosla, & Melton, 2002), the “influence” of testosterone on facial growth is likely to be less direct than previously supposed.

One recent study (Neave, Laing, Fink, & Manning, 2003) also employed salivary testosterone assays and facial photographs to examine the relationship between circulating testosterone and facial masculinity, but failed to find any relationships between perceived masculinity, attractiveness, and testosterone measures in a similar-sized sample. Why this may be is unknown, but there are clear differences between the experimental procedure in the Neave et al. (2003) and current studies that may be informative. Firstly, the Neave et al. study used a wider age range of participants than the current study, and reported a negative relationship between testosterone and age. This relationship was not present in our more homogenous sample, removing a potential confound. Secondly, the current study used a forced-choice methodology, which is probably a more sensitive measure of differences in masculinity between faces than the rating scale test employed by Neave et al. It is worth noting that our data from the individual face experiment, though demonstrating a statistically significant effect, shows that raters picked only the higher testosterone individual from a face pair 53% of the time. It seems likely that whereas individuals who vary greatly in salivary testosterone may well appear to differ in facial masculinity (as indicated by the correlational analysis), those in the middle of the distribution are more difficult to discriminate. Given the variability in testosterone assays, and the indirect link between facial growth and testosterone, this is perhaps unsurprising. In practical terms, only men at the extremes of the salivary testosterone continuum may be distinguishable from facial appearance.

This study gives little information as to the cues used by participants to make masculinity judgements. It is possible that expression cues may vary with circulating testosterone—high-testosterone men have been reported to smile less than low-testosterone men (Dabbs, 1997). Examination of the composites, however, suggests that no large differences in smiling are apparent between the high- and low-testosterone quartile faces. Another possibility is that high-testosterone faces appear older than low-testosterone faces, even though the samples are matched for age. As female faces grow less than male faces at puberty, "masculine" faces may appear older for their chronological age than feminine faces (Perrett et al., 1998; Perrett & Penton-Voak, 1999). In this sample, the high- and low-testosterone composites do not differ in perceived age, so masculinity and age appear at least to some extent independent here.

If the major finding of these studies is reliable, and circulating testosterone is associated with facial masculinity in men, then these results may carry some importance, as facial masculinity has been repeatedly demonstrated to influence social perception. Although facial masculinity is associated with dominance and poor social qualities, and these attributions appear to be cross-culturally stable (Keating et al., 1981), a relationship between testosterone and facial appearance in adults has not, to our knowledge, been demonstrated until now. Testosterone in males is associated with marital dissatisfaction and poor father–child relationships (Julian & McKenry, 1989); low scores on a "spousal investment" measure (Gray, Kahlenberg, Barrett, Lipson, & Ellison, 2002) is related to antisocial behaviour (Dabbs, 1993), and correlates negatively with self-reported prosocial personality characteristics (Harris, Rushton, Hampson, & Jackson, 1996). A cue to such behaviours/personality characteristics would be useful to participants in many social situations, including a female considering any given male as a partner. The link between facial masculinity and circulating testosterone reported here suggests that facial features may facilitate both accurate social attributions and the identification of appropriate mates.

Acknowledgments

A University of Texas Undergraduate Research Fellowship to J.Y.C funded the testosterone assays in this research. We would like to thank Devendra Singh, Yvon Delville, Christina Chou, Julie Grinsfelder, Jason King, and Ja-Young Song for assistance during the running of this experiment. Thanks also to Nicholas Pound and the referees for comments and improvements on the manuscript.

References

Andersson, M. (1994). *Sexual selection*. Princeton, NJ: Princeton University Press.

Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, *48* (2), 312–323.

Cunningham, M. R., Barbee, A. P., & Pike, C. L. (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* (1), 61–72.

Dabbs, J. M. (1990). Salivary testosterone measurements: reliability across hours, days, and weeks. *Physiology and Behavior*, *48*, 83–86.

Dabbs, J. M. (1993). Salivary testosterone measurements in behavioral research. In D. Malamud, & L. Tabak (Eds.), Saliva as a diagnostic fluid, *Annals of the New York Academy of Sciences 694*, 177–183.

Dabbs, J. M. (1997). Testosterone, smiling, and facial appearance. *Journal of Nonverbal Behavior*, *21*, 45–55.

Fink, B., & Penton-Voak, I. S. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, *11*, 154–158.

Folstad, I., & Karter, A. J. (1992). Parasites, bright males and the immunocompetence handicap. *American Naturalist*, *139*, 603–622.

Frank, G. R. (2003). Role of estrogen and androgen in pubertal skeletal physiology. *Medical and Pediatric Oncology*, *41* (3), 217–221.

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: trade-offs and strategic pluralism. *Behavioral and Brain Sciences*, *23*, 573–644.

Getty, T. (2002). Signaling health versus parasites. *The American Naturalist*, *159*, 363–371.

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.

Granger, D. A., Schwartz, E. B., Booth, A., & Arentz, M. (1999). Salivary testosterone determination in studies of child health and development. *Hormones and Behavior*, *35*, 18–27.

Gray, P. B., Kahlenberg, S. M., Barrett, E. S., Lipson, S. F., & Ellison, P. T. (2002). Marriage and fatherhood are associated with lower testosterone in males. *Evolution and Human Behavior*, *23*, 193–201.

Grumbach, M. M. (2000). Estrogen, bone, growth and sex: a sea change in conventional wisdom. *Journal of Pediatric Endocrinology and Metabolism*, *13*, 1439–1455.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*, 330–337.

Harris, J. A., Rushton, J. P., Hampson, E., & Jackson, D. N. (1996). Salivary testosterone and self-report aggressive and pro-social personality characteristics in men and women. *Aggressive Behavior*, *22*, 321–331.

Jennions, M. D., & Petrie, M. (2000). Why do females mate multiply? A review of the genetic benefits. *Biological Reviews*, *75*, 21–64.

Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness: evidence for hormone mediated adaptive design. *Evolution and Human Behavior*, *22*, 251–267.

Julian, T., & McKenry, P. C. (1989). Relationship of testosterone to men's family functioning at mid-life: a research note. *Aggressive Behavior*, *15*, 281–289.

Keating, C. F. (1985). Gender and the physiognomy of dominance and attractiveness. *Social Psychology Quarterly*, *48*, 61–70.

Keating, C. F., Mazur, A., & Segall, M. H. (1981). A cross-cultural exploration of physiognomic traits of dominance and happiness. *Ethology and Sociobiology*, *2*, 41–48.

Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society of London B*, *268*, 1–6.

Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., & Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proceedings of the Royal Society of London B*, *269*, 1095–1100.

Meyer, D. A., & Quong, M. W. (1999). The bio-logic of facial geometry. *Nature*, *397*, 661–662.

Neave, N., Laing, S., Fink, B., & Manning, J. T. (2003). Second to fourth digit ratio, testosterone and perceived male dominance. *Proceedings of the Royal Society of London B*, *270*, 2167–2172.

Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions, and male facial attractiveness. *Proceedings of the Royal Society of London B*, *268*, 1617–1623.

Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P., & Perrett, D. I. (2003). Female condition influences preferences for sexual dimorphism in faces of male humans (*Homo sapiens*). *Journal of Comparative Psychology*, *117*, 264–271.

Penton-Voak, I. S., Perrett, D., Castles, D., Burt, M., Koyabashi, T., & Murray, L. K. (1999). Menstrual cycle alters face preference. *Nature*, *399*, 741–742.

Penton-Voak, I. S., & Perrett, D. I. (2001). Male facial attractiveness: perceived personality and shifting female preferences for male traits across the menstrual cycle. *Advances in the Study of Behavior*, *30*, 219–260.

Perrett, D., & Penton-Voak, I. (1999). The bio-logic of facial geometry. *Nature*, *397*, 662.

Perrett, D. I., Lee, K. J., Penton-Voak, I. S., Rowland, D. R., 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.

Rhodes, G., Chan, J., Zebrowitz, L., & Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proceedings of the Royal Society of London, B*, *270* (Supplement), S93–S95.

Riggs, B. L., Khosla, S., & Melton, L. J. (2002). Sex steroids and the construction and conservation of the adult skeleton. *Endocrine Reviews*, *23* (3), 279–302.

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, *15*, 70–76.

Swaddle, J. P., & Reierson, G. W. (2002). Testosterone increases perceived dominance but not attractiveness in human males. *Proceedings of the Royal Society of London B*, *269*, 2285–2289.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A, HEP*, *46*, 225–245.