

Women's Perception of Attractiveness of Men's Faces Inversely Correlates with Men's Serum Testosterone During the Fertility Phase of the Menstrual Cycle

Jaroslaw Krejza (✉ jkrejza@me.com)

Institute of Innovative Medicine

Rafał Śledziwski

Medical University of Białystok

Marek Tabedzki

University of Technology

Rong Chen

University of Maryland Medical Center

Ewa Krzystanek

Medical University of Silesia

Michał Arkuszewski

Institute of Innovative Medicine

Magdalena Odachowska

Medical University of Białystok

Kinga Kwiatkowska

Medical University of Białystok

Marcelina Marynowska

Medical University of Białystok

Maciej Rogowski

Medical University of Białystok

Andrzej Ustymowicz

Medical University of Białystok

Research Article

Keywords: face, attractiveness, men, testosterone, menstrual cycle

Posted Date: November 17th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1054555/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The attractiveness of the human face may signal the genetic suitability of a mate. The 'ovulatory shift hypothesis' postulates that women in the fertile phase of the menstrual cycle prefer faces of masculine men that signal 'good genes', whereas in the non-fertile phase they prefer good parental providers. We studied relationships between serum total testosterone and face attractiveness of 77 healthy men (20-29 years, mean±SD 22.44±1.79) as rated by 19 healthy women (20-27 years, mean±SD 22.84±1.96) on day 13 of their menstrual cycle. Using advanced Bayesian multilevel modeling we showed that the attractiveness of faces is negatively associated with the concentration of serum testosterone in the men, even taking into account the concentration of serum estrogen in the raters. The average face composited from images of 39 faces rated above pool median attractiveness rate, was slightly narrower than the average face composited from 38 less attractive faces. Our results challenge the 'ovulatory shift hypothesis' as faces of males with higher circulating testosterone were rated as less attractive than faces of males with lower testosterone by women on the fertile phase of the cycle.

Introduction

The attractiveness of the human face may signal the genetic suitability of a potential mate. Facial symmetry, skin's texture, expressions of happiness, or hints of trustworthiness reflect developmental stability, health status, and ability to shape positive personal and social interactions, respectively [1–4]. Darwin suggested that individuals with genetic cues essential in the struggle for existence mate more often and produce better-adapted offspring. Increasing occurrences of such favorable characteristics in the population of a geographic region eventually develop them into dominant or alike average phenotypic features over time. It shall not be a surprise therefore that an average face, created digitally by morphing many random faces together, is universally rated as attractive by the opposite sex [5].

The pace and direction of evolution of certain traits can offer a hint which features have advanced survival and reproductive successes. The remarkably rapid development of intelligence in humans signals that it has been a preferential trait. Species with larger brains, which is a proxy of intelligence, do appear to have survival and reproductive advantages [6–8]. Archeological studies demonstrated that the human skull that contains the brain expanded while the size of the face shrank [6]. Preferences for larger brains and smaller, women-like faces across cultures can explain why intelligence correlates with facial attractiveness, as both are highly heritable [9, 10, 11–16].

Women are more selective in choosing mates because they are more biologically limited in the number of offspring they can produce. They invest more into pregnancy, nursing, and nurturing [17]. Hence, women may prefer long-term effective caretakers and are less receptive to short-term sexual offers [17–19]. Modern women prefer qualities such as education, wealth, social status, and health, which signal the ability of a man to provide material support for their children [20, 21].

Several mating selection theories are trying to explain women's preferences [22]. A recent hypothesis called the "ovulatory shift hypothesis" [23–28] postulates that women's preferences change across the menstrual cycle. Specifically, women desire men with traits signaling 'good genes' in the fertile phase, whereas in the non-fertile phase they prefer 'good fathers' with traits signaling willingness to provide parental care. As men with 'good genes' may not be willing to invest in a long-term relationship [29, 30], women tend to obtain these 'good genes' through short-term sexual affairs during the fertile phase, regardless of long-term relationships. Such a cyclic shift of preferences supposes to justify women's desire to obtain the best genes for their children and to ensure long-term access to material resources from 'good fathers'.

The ovulatory shift hypothesis is being disputed [31–35]. Very low rates of cuckoldry at the range from 0.73 to 3% in western societies and historical populations challenge the main concept of the hypothesis [36, 37]. The theory assumes that women can identify 'good genes' based on men's phenotypic testosterone-related features such as facial masculinity and dominance [38]. Testosterone is supposed to suppress the immunological system and only men with 'good genes', which encode resistance to common pathogens, would afford to maintain a high testosterone concentration and develop masculinity features without being infected [39, 40]. However, dominance [41], facial masculinity [42, 43], and higher testosterone are not related to superior actual health or stronger disease resistance [44, 45] and overall mating success [46]. There is little support for claims that health increases mating success in relatively healthy humans [47]. This negates the role of testosterone as an indicator of 'good genes'.

Given the importance of the males' testosterone in the mating theory, we examined relationships between men's serum concentration of testosterone and women's preferences for men's faces on the fertile day of their menstrual cycle. We also tried to identify preferential face features by graphic comparison of composite images obtained from groups of attractive and unattractive men as rated by women.

Subjects And Methods

Study design

This was a prospective analytical observational study to explore associations of the attractiveness of male faces rated by female participants in the 13th day of their menstrual cycle, with male's serum total testosterone (TT) taking into account the female's serum sex hormones concentration. The protocol of the study was approved by the Bioethics Committee of the Medical University of Bialystok, Poland. Each participant gave written informed consent to take part in the study as well as for the publication of identifying images in an on-line open-access publication and was not paid. All research was performed in accordance with the Declaration of Helsinki.

Population and study participants

We recruited healthy women based on the most recent general medical examination and standard laboratory tests, with no major health problems in the past, not pregnant, not used hormonal

contraceptives, who had regular, non-complicated menstrual cycles (from 25 to 35 days) in the last 6 months, were heterosexual, never smoked and never took drugs, had not abused alcohol and were not taking any concomitant medication. Only participants 19 to 31 years old were included to capture those with the highest reproductive potential. We included men based on self-reported health status and no smoking habit. Women were not familiar with men.

We recruited 77 men (aged 20-29 years, median 22, mean \pm SD 22.44 \pm 1.79) and 19 women (aged 20-27 years, median 23, mean \pm SD 22.84 \pm 1.96) from local universities.

Hormonal status

A blood sample for measurements of 17 β -estradiol (E2) and progesterone (PG) concentrations was drawn on 13th day of the menstrual cycle, after 12 hours of fasting, 72 hours abstain from vigorous physical exercises and alcohol intake. In males, a blood sample for measurements of TT concentrations was taken at the university hospital laboratory between 7 a.m. and 9 a.m. after 12 hours of fasting, 72 hours abstain from vigorous physical exercises and alcohol intake. The concentration of TT was measured according to a standard clinical practice using Chemiluminescent Microparticles Immunoassay and device by Architect ci100 (Abbott, USA), and Alinity 2nd Generation Testosterone Reagent Kit (Abbott, Germany). Intra-assay coefficients of variation (CV) given by the manufacturer were 3.5%, 2.5%, and 2.3% for mean TT concentration of 0.32, 2.53, and 8.43 nmol/L of in control samples, respectively. Abbott reports an inter-assay CV of 8.1%, 3.8% and 2.6%, respectively, for analogous low, medium, and high TT-level of control samples. Intra- and inter-assay CV for human serum panel was calculated at 2.8% and 3.1%, respectively, for mean TT concentration of 2.30 nmol/L. The assay is linear across the measuring interval of 0.15 to 64.57 nmol/L. We employed ARCHITECT Estradiol and Progesterone (Abbott, Ireland) kits to measure E2 and PG, respectively.

Assessment of face attractiveness

A professional photographer took frontal view pictures of male faces in standardized studio conditions on a white background and the same lighting conditions. Clean shaved male participants were asked to maintain a neutral facial expression with their mouths closed while seated upright on a chair. They were photographed using a professional digital camera (ILCE-6500, Sony, Japan), with a horizontal and vertical picture resolution of 350x350 DPI and matrix size 6000x3376 pixels, with an exposure time of 0.01s. Flashlights were assembled to yield relatively natural lighting conditions. The distance of each face from the camera was constant at 1.8m.

We used a program, developed locally for this purpose (Inspeerity, Bialystok, Poland), to present pictures of faces on a computer screen and capture assessment of attractiveness by each female rater in an ordinal scale from 1 to 10 (10 indicates the most attractive). The raters were instructed how to operate the program and they assessed 10 phantom faces as a warm-up just before starting the trial. Thereafter, pictures of all faces were displayed in one session to each rater independently, in random order and in standardized conditions between 8:00 and 12:00am. Seven seconds were allotted for each face assessment while time flow was graphically displayed on the bottom of each display.

Creation of composite images of faces

We used programs written in Python programming language (OpenCV, <https://opencv.org>; and Dlib, <http://dlib.net>) to combine images of individual faces and generate composite faces for comparisons. First, we detected and extracted facial landmarks (salient regions of the face, such as: eyes, eyebrows, nose, mouth, jawline) using Multi-PIE land-marking scheme from iBUG 300-W dataset from every image [48, 49]. We applied the pre-trained face shape predictor to the face on an image, to estimate the location of 68 (x, y)-coordinates that map to facial structures on the face. The predictor was trained on the iBUG 300-W face landmark dataset. We used the facial landmark detector, which is an implementation of the One Millisecond Face Alignment with an Ensemble of Regression Trees approach [50].

In the second step, we normalized the faces and placed them in the same reference frame. Then we aligned the facial features of all input images. Subsequently, we calculated the average of all landmarks in the output image coordinates by simply averaging the x and y values. Then we used 68 feature points of each image, as well as points on the boundary of the image, to calculate a Delaunay Triangulation, which allows us to break the image into triangles. Then we aligned these regions to the average face. The vertices of the corresponding triangles were used to calculate an affine transform (the composition of transformations such as translations, scaling, and rotations). This affine transform then was used to transform all pixels inside the triangle. This procedure was repeated for every triangle in the input image [51]. To calculate the average composite image, we averaged pixel intensities of all warped images.

Statistical analysis

We employed a statistical software (SYSTAT 12, Systat Software, USA) to perform basic data analyses. To rank facial attractiveness we used the median value of all scores assigned to the individual male face by raters, and then we classified all faces into attractive and non-attractive ones based on the median value from those assigned median values to each face. We calculated mean, median, and SD of serum hormone concentrations. We used the Shapiro-Wilk test to determine the normality of the continuous data.

We used advanced Bayesian multilevel modeling (BMM) to model the association between rate, which is the primary outcome variable (response variable), and 1) TT, and 2) TT adjusted for effects of E2, as exploratory variables. The ordinal response is assumed to originate from the categorization of a latent continuous variable. Cases (data points) from the same rater may be correlated. BMM can model the data measured on different levels at the same time, thus taking a complex dependence structure into account. In multilevel modeling, the response variable y follows a probability distribution from a family, $y_i \sim D(\eta_i, \theta)$, where y_i is the value of y of case i , θ is additional model parameters. The linear predictor can be written as $\eta = X\beta + Z\mu$. In this equation, β and μ are coefficients at the population-level and cluster-level, respectively. X and Z are the corresponding design matrices, μ follows a multivariate Gaussian distribution with mean zero and an unknown covariance matrix. Model parameters are (β, μ, θ) .

In the first multilevel model, the rate was the response variable with family cumulative and TT was a predictor. We also incorporated an intercept varying by raters. That is, our model was $\text{Rate} \sim \text{TT} + (1|\text{rater})$. In the second model, we included TT and E2 and incorporated an intercept varying by raters.

In BMM, the inference is based on the posterior distributions which are generated by Markov Chain Monte Carlo (MCMC) [52]. Relative to non-Bayesian multilevel models, BMM can draw inference for any pattern based on posterior distributions. The classic MCMC methods converge slowly. To address this limitation, we used Hamilton Monte Carlo, which converges quickly because of its ability to avoid the random walk behavior [53]. The MCMC parameters are: 4 chains, each with 2000 iterations, and the first 1000 are warm-up.

Results

Table 1 provides pooled attractiveness' scores of male's faces and serum TT concentrations. Serum concentration of E2 [pg/mL] and PG [ng/mL] in women were within normal range at day 13th of the menstrual cycle: mean±SD and range limits: 158.3±99.1, 29.0-392.0; and 0.2±0.1, 0.1-0.4, respectively.

Table 1

Median pooled attractiveness' scores of natural faces of 77 males aged 20-29 years, as assessed by 19 female raters aged 20-27 years, along with males' serum total testosterone (TT) concentrations (nmol/L).

Parameter		All males	(A) More attractive >median	(B) Less attractive <median	Difference A – B
Score of face attractiveness*	mean	3.99	4.85	2.92	1.93
	median	4	4	3	1
	N	77	39	38	
Serum TT concentration	median	22.77	22.55	23.93	-1.38
	mean±SD	24.26±9.07	22.74±6.74	25.83±10.84	-3.09
*Ordinal scale from 1 to 10 (10 indicates the most attractive).					

Histograms (figure 1) shows skewed distributions of rates and TT, thus we applied a \log_{10} transformation to fulfill the condition of Gaussian distribution in modeling.

We found a significant association of males' serum TT concentrations with attractiveness ratings. The coefficient of TT was -1.05 (estimate error was 0.3, 95% confidence interval (CI) was [-1.64, -0.48]). The higher the TT concentration, the lower attractiveness score men received from women (figure 2).

After adjusting for E2, the association of TT with scoring was still significant. For TT, coefficient was -1.05 (estimate error 0.29, 95% CI: -1.63, -0.50). The effect of E2 was marginal, coefficient was 0.56

(estimate error 0.52, 95% CI: -0.50, 1.54). Lower scores of face attractiveness were assigned to images of faces of men with higher TT concentration, regardless of the E2 concentration in female raters.

Figure 3 shows three faces composited from images of all 77 male faces, 39 faces rated above and 38 faces rated below pool median value.

To facilitate geometric comparison of composite faces we measured the width and length of composite faces in relation to the composite face from 77 faces (figure 4). The more attractive faces were slightly smaller than average while the less attractive were slightly bigger.

Discussion

We have shown that faces of men with a higher concentration of circulating testosterone were rated as significantly less attractive by young women, on the 13th fertile day of their menstrual cycle. This result challenges the hypothesis that women in the fertile phase would prefer masculine, high testosterone men [54]. If such men would have been preferred by hundreds of generations of ancestral women, as the hypothesis suggests, androgen associated traits such as the massive jaw, cheekbones, and prominent brow ridges, increase musculature, and dominant aggressive behavior would have been amplified by forces of natural intra-sexual selection [54, 55]. In contrast, several archeological and anthropometric studies provided compelling evidence that the Homo genus, which includes Humans, Neanderthals, and other ancestors, has experienced an evolution toward smaller faces over time, with Homo sapiens showing the greatest reduction in size [56, 57]. Fast shrinking face, the rapid evolution of mandibular shape and size, and rampant brain growth are suggested from analysis of numerous human fossils [57–60]. Hence the forces of evolution worked against face masculinization and rather favored the development of cognitive abilities.

We showed that the composite attractive male face is slightly smaller than the less attractive one. Humans can detect facial attractiveness very quickly and reliably based on subtle geometric differences in facial features along specific axes in face space [61–62]. Smaller faces and lower TT concentration in the attractive group are mutually consistent findings, which could be related to diminished androgenic effects in earlier periods of life [63].

A broader framework of the evolution of advanced intelligence in humans connects drastic reduction of testosterone-related aggression with diminished sexual dimorphism, which may have begun as early as 3 million years ago [6]. Aggressive behavior can be even predicted based on testosterone-related structural brain phenotype [64]. Lower testosterone most likely favored increased cooperation between individuals for better survival and facilitated departure from a chimpanzee-like mating system with high levels of male-male physical competitive violence, requiring a large muscle body mass, and sexual coercion of females [65–67]. Accelerated natural selection against aggression in the last 200,000 years coincided with the further feminization of human faces as evidenced by a reduction in brow ridge projections and a shortening of the upper facial region [63, 67, 68]. Given that testosterone is linked to dominance, egoistic choices, aggressive and unfaithful behavior in men, the feminization of male faces may indicate

preferential selection for increased social tolerance that allowed humans to work more productively [66, 69, 70].

To the best of our knowledge, we provide for the first time empirical support for the inverse relationship between serum TT in men and their facial attractiveness as scored by young women during their fertile phase. In contrast, Roney et al. reported that women prefer natural faces of men with high saliva testosterone [71]. Their raters scored men's faces in random days of their menstrual cycle, however, and their fertility window was merely guessed based on the calendar method and estimates of hormonal concentration from data published in other reports. Despite each woman rated faces only once Roney et al. did not provide crucial information on the distribution of ratings across the cycle. Especially, they did not provide how many women were in the fertile days. Also, we studied Caucasian men and women while Roney et al.'s studied samples of different ethnic groups.

We measured serum concentration of TT using a standardized well-validated clinical methodology between 7 a.m. and 9 a.m. because the concentration can fall up to 30-35% in the late afternoon [72]. Roney et al. measured saliva 'free' testosterone at various times of the day and regressed transformed testosterone concentrations onto the individual rates of the attractiveness of faces. A simple linear regression may not be the right tool to explore the association of ordinal correlated scores of unknown distribution with testosterone values. Furthermore, they did not explain how regression coefficients and variance from all 75 individual linear regression analyses were pooled.

A more sophisticated statistical approach is needed to model positive correlation among raters' attractiveness scores of the presented facial images, which could be the result of sequentially dependent attractiveness perception or sequentially dependent response bias [73, 74]. For example, if a rater's scoring criterion gradually changes over time (e.g., the rater tends to give higher ratings at the beginning of the experiment and lower ratings at the end of the experiment), then the autocorrelation of the rater's scoring criteria will lead to a positive correlation between current and previous ratings.

Peters et al. also investigated an association between saliva testosterone and attractiveness and masculinity of face and body on photographs of 119 young men, rated by two independent groups of 12 females [75]. Most of these women were taking hormonal contraceptives that could have altered any potential menstrual cycle effects on ratings. Testosterone was not correlated with either attractiveness or masculinity, however, they implied that the correlation between testosterone and attractiveness was more likely to be negative. Also, Neave et al., who used natural photographs of 48 men and 36 female raters, did not find an association between salivary testosterone and attractiveness or masculinity [76]. They did not report, however, at which phase of the menstrual cycle their female raters were during scoring.

We measured the concentration of serum TT, while most investigators rely on measurements of 'free' testosterone in saliva [71, 75, 76]. Despite the speed, easy saliva collection, and avoidance of stress of vein puncture, multidisciplinary clinical experts recommend serum TT as the first-line, reliable indicator of the physiological function of gonads, with low analytical variation (precision about 4–10%) and close correlation with calculated bio-testosterone and free testosterone [77, 78]. It is uncertain if saliva

testosterone represents the so-called 'free' fraction, due to its binding to albumin, proline-rich proteins, and steroid hormones-binding globulin [79, 80]. Hence, there is a poor agreement of saliva testosterone with serum values, even in laboratory-controlled conditions [81]. Substantial analytical errors and biological variability of saliva testosterone should be accounted for when attempting to determine biologically relevant a 'least significant difference' threshold [82, 83]. Even in highly controlled conditions the threshold for salivary testosterone is large (78–90%) [84–86].

We allowed for each rater 7 seconds to assess each photograph of men's face to capture the first impression of attractiveness, which is most likely rapid, automatic, and mandatory [87]. Rates could have reliably judged the attractiveness of faces presented for just 13 ms [61], but we selected longer time and random display of images to minimize the systematically biased perception of face attractiveness toward faces seen up to several seconds before. Xia et al. demonstrated that perceived face attractiveness was pulled by the attractiveness level of facial images encountered up to 6 s after the previous image [88]. By giving only one task to our raters we avoided multi-task cognitive overload, which could have affected raters' ability to intentionally form attractiveness impressions, and automaticity of impression formation. An additional bias could have been introduced if raters simultaneously assessed masculinity and man's attractiveness for a short or long-term partnership as in [89].

The theory of shifting women's preferences for facial masculinity across the menstrual cycle is being promoted [23, 25, 33]. We did not assess the masculinity of participants' faces, but the smaller size of the composite face of our attractive group suggests that women prefer rather more feminized male faces. The theory, however, is based on a premise that there is a significant association between testosterone and masculinity [31–33, 35, 44]. Androgenic influence during puberty likely shapes "high-testosterone" face, which is purported to be an honest indicator of health and male fitness because testosterone enhances sexual signals, but suppresses immune function [39]. A recent meta-analysis found little evidence that testosterone suppressed immune function [42], whereas another meta-analysis identified an opposite effect - a strong suppressive effect of experimental immune activation on testosterone [90].

Surprisingly, there are contradictory reports that substantiate associations of testosterone with masculinity [54, 63, 89, 91]. The association remains an issue because of unresolved discrepancies between structural and masculinity ratings and methodological shortcomings of studies that used abstracted computer-manipulated images of 'high and low testosterone faces' or saliva testosterone measurements [25, 75, 89, 91, 92]. Recent findings do not support preferences for male masculinity traits either at low- or high-conception probability groups of women [34, 35, 44, 93, 94]. Pound et al. [54] suggest that raters may attribute 'masculine' ratings to faces they find attractive irrespective of the objective sexual dimorphism, due to stereotypical associations between the term 'masculinity' and attractiveness. Furthermore, perceived masculinity may not correlate with attractiveness [46, 95], especially in view of an association of perception of sexual unfaithfulness with face masculinity [4].

Females' preferences for types of male faces could differ among populations due to ethnic, socio-cultural, and human development factors [9, 96]. Recent reports showed women's preferences for more

feminized faces of Caucasian men [9, 10, 34, 97]. Kocnar et al. reported that feminized male faces were preferred over masculinized faces by women in most European populations, especially in countries with high human development index [9, 10]. Perrett et al. [96] reported preferences for more face feminization of Japanese and Scottish participants compared to Caucasian North Americans faces among the raters, whereas Harris et al. have found the opposite pattern [35]. Studies comparing preferences among populations provide contradictory results, hence, differences in women's preferences for masculinity or femininity between populations remain an open topic.

The selection of participants of different ages to conduct a research study on physical attraction to the opposite sex could have an impact on the ratings and comparability of results [98]. Two studies of large cross-cultural samples found that males prefer females considerably younger than themselves and females prefer males considerably older than themselves [99, 100]. Most studies are based on samples of male students 18-20 years old, who may still be developing their ultimate secondary sexual characteristics and face masculinity. In a study when the mean age of men was 18, female raters favored faces of men of higher testosterone levels [71, 101]. Male teenagers are less likely to be rated as attractive masculine men by women in their late twenties or older [96]. This could be a factor in studies, where the mean age of female raters was substantially higher than the mean age of men whose computer-modified or even natural faces were used as stimuli [25, 26]. Likewise, female teenagers may give very different rates when assessing masculinity in teenagers' male faces compared to women over 30 years old [89, 101]. Individual variation in evaluations of trustworthiness, dominance, and attractiveness, is largely shaped by people's personal experiences and the rapidity of their sexual development [98, 102]. Thus, including teenage participants may lead to doubts whether all participants of a study are biologically and psychologically suitable to provide generalizable assessments, especially when taking into account the age disparity in sexual relationships. In our study, we opted for the age-matched sample.

We used natural non-edited images of male faces to explore associations between male's facial attractiveness and TT. Such an approach links realistic TT measures in men with individual rates of women. Many other studies on preferences of women across the menstrual cycle in context men's perceived attractiveness or masculinity opted to use computer simulations or heavily edited pictures or constructs of male's face [23, 25, 34, 91, 92, 101, 103]. Certainly, computer alterations of face images to obtain a certain degree of masculinity or femininity facilitate face ratings and reduce the cost of the study, however, at the same time the differences between composite images of faces can be unnaturally high leading to believe that certain face images are more appealing to women than raw non-edited images of male faces. Ratings of images of real faces may present a more appropriate assessment of women's preferences than ratings of unrealistic computer-manipulated visuals.

Several studies that use surveys of females to determine the fertile time window reported contradictory results [25, 26, 34]. The timing of women's fertile window is often unpredictable. Wilcox et al. showed that only in about 30% of women is the fertile window entirely within the days of the menstrual cycle identified by clinical guidelines—that is, between days 10 and 17 [104]. A study, in which daily values of sex

hormones in female participants across the cycle were determined, showed that ovulation occurred as late as 7 days before menses, and as early as on day 8th of the cycle [105]. Thus, we were fortunate enough to target accurately the fertile window as female participants were still before ovulation on day 13 of the cycle as indicated by the low concentration of serum progesterone and high concentration of estrogen. Despite we did not determine if the women were fertile, we assumed that healthy, young regularly menstruating women are fertile.

Our sample size of female participants was rather small due to the budget constraints of the project. The sample size is comparable to previous reports, however. In contrast to previous studies, the group of our raters was more homogenous and was studied in the same day of the menstrual cycle to reduce sex hormone related variability. Also, we used a state of the art statistical modeling to further reduce the probability of erroneous inferences. Our preliminary report can help to plan more powerful studies.

In summary, we have demonstrated that young healthy women prefer images of natural faces of young men with lower concentrations of total testosterone in serum on day 13 of the menstrual cycle. The size of the preferred male face tends to be smaller.

Declarations

Author contributions

RS, JK, AU and RC formulated the hypotheses for this study. RS, AU, MO, KK, MM and MR were involved in data collection (blood sample collection and assessment of faces attractiveness) and literature search. MT and RS generated and analyzed composite images of male faces. RC, RS, JK, EK and MA performed sophisticated statistical analysis of the results. RS and AU - secured funding of the project. All authors were closely involved in manuscript preparation, manuscript writing, final review of the literature and the provision of critical feedback.

Competing Interests Statement

There is no conflict of interests among all authors.

References

1. Scheib, J. E., Gangestad, S. W. & Thornhill, R. Facial attractiveness, symmetry and cues of good genes. *Proc. R. Soc. Lond. B.* **266**, 1913-1917 (1999)
2. Henderson, A. J., Holzleitner, I. J., Talamas, S. N. & Perrett, D. I. Perception of health from facial cues. *Phil. Trans. R. Soc. Lond. B*, **371**, 20150380 <https://doi.org/10.1098/rstb.2015.0380> (2016).
3. Oosterhof, N. N. & Todorov, A. Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, **9**, 128–133 (2009).
4. Sutherland, C. A. M. *et al.* Impressions of sexual unfaithfulness and their accuracy show a degree of universality. *PLoS ONE*, **13**, e0205716 <https://doi.org/10.1371/journal.pone.0205716> (2018).

5. Langlois, J. H. & Roggman, L. A. Attractive faces are only average. *Psychol. Sci*, **1**, 115–121 (1990).
6. MacLean, E. L. Unraveling the evolution of uniquely human cognition. *Proc. Natl Acad. Sci. USA*. **113**, 6348-6354 (2016)
7. Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. & Malone, T. W. Evidence for a collective intelligence factor in the performance of human groups. *Science*, **330**, 686–688 (2010).
8. Kotrschal, A., Buechel, S. D., Zala, S. M., Corral-Lopez, A. & Penn, D. J. Kolm, N. Brain size affects female but not male survival under predation threat. *Ecol. Lett*, **18**, 646–652 (2015).
9. Moore, F. R. *et al.* Cross-cultural variation in women's preferences for cues to sex- and stress-hormones in the male face. *Biol. Lett*, **9**, 20130050 <https://doi.org/10.1098/rsbl.2013.0050> (2013).
10. Kocnar, T., Saribay, S. A. & Kleisner, K. Perceived attractiveness of Czech faces across 10 cultures: Associations with sexual shape dimorphism, averageness, fluctuating asymmetry, and eye color. *PLoS ONE*, **14**, e0225549 <https://doi.org/10.1371/journal.pone.0225549> (2019).
11. Davies, G. *et al.* Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol. Psychiatr*, **16**, 996–1005 (2011).
12. Mitchem, D. G. *et al.* Estimating the sex-specific effects of genes on facial attractiveness and sexual dimorphism. *Behav. Genet*, **44**, 270–281 (2014).
13. Nedelec, J. L. & Beaver, K. M. Physical attractiveness as a phenotypic marker of health: An assessment using a nationally representative sample of american adults. *Evol. Hum. Behav*, **35**, 456–463 (2014).
14. Langlois, J. H. *et al.* Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull*, **126**, 390–423 (2000).
15. Zebrowitz, L. A., Hall, J. A., Murphy, N. A. & Rhodes, G. Looking smart and looking good: Facial cues to intelligence and their origins. *Pers. Soc. Psychol. B*, **28**, 238–249 (2002).
16. Kanazawa, S. Intelligence and physical attractiveness. *Intelligence*, **39**, 7–14 (2011).
17. Trivers, R.. Syntax of referencing in *Sexual selection and the descent of man 1871–1971* (ed. Campbell, B.),136–208(Aldine Transaction, 1972)
18. Greitemeyer, T. Receptivity to sexual offers as a function of sex, socioeconomic status, physical attractiveness, and intimacy of the offer. *Pers. Relationsh*, **12**, 373–386 (2005).
19. Buss, D. M. & Schmitt, D. P. Sexual strategies theory - an evolutionary perspective on human mating. *Psychol. Rev*, **100**, 204–232 (1993).
20. Miller, G. F. & Todd, P. M. Mate choice turns cognitive. *Trends Cogn. Sci*, **2**, 190–198 (1998).
21. Gueguen, N. & Lamy, L. Men's social status and attractiveness women's receptivity to men's date requests. *Swiss J. Psychol*, **71**, 157–160 (2012).
22. Schmitt, D. Syntax of referencing in *Evolutionary perspectives on human sexual psychology and behavior. Evolutionary psychology*. 3-39 (Springer, 2014)
23. Penton-Voak, I. S. *et al.* Menstrual cycle alters face preference. *Nature*, **399**, 741–742 (1999).

24. Gangestad, S. W., Thornhill, R. & Garver, C. E. Changes in women's sexual interests and their partners' mate-retention tactics across the menstrual cycle: Evidence for shifting conflicts of interest. *Proc. R. Soc. Lond. B.* **269**, 975-982 (2002)
25. Penton-Voak, I. S. & Perrett, D. I. Female preference for male faces changes cyclically: Further evidence. *Evol. Hum. Behav.*, **21**, 39–48 (2000).
26. Little, A. C., Jones, B. C. & DeBruine, L. M. Preferences for variation in masculinity in real male faces change across the menstrual cycle: Women prefer more masculine faces when they are more fertile. *Pers. Individ. Differ.*, **45**, 478–482 (2008).
27. Jones, B. C. *et al.* Effects of menstrual cycle phase on face preferences. *Arch. Sex. Behav.*, **37**, 78–84 (2008).
28. Puts, D. A. Cyclic variation in women's preferences for masculine traits - potential hormonal causes. *Hum. Nature-Int. Bios*, **17**, 114–127 (2006).
29. Gangestad, S. W. & Simpson, J. A. The evolution of human mating: Trade-offs and strategic pluralism. *Behav. Brain Sci.* **23**, 573-587; discussion 587-644 (2000)
30. Puts, D. A. Beauty and the beast: Mechanisms of sexual selection in humans. *Evol. Hum. Behav.*, **31**, 157–175 (2010).
31. Wood, W., Kressel, L., Joshi, P. D. & Louie, B. Meta-analysis of menstrual cycle effects on women's mate preferences. *Emot. Rev.*, **6**, 229–249 (2014).
32. Gildersleeve, K., Haselton, M. G. & Fales, M. R. Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychol. Bull.*, **140**, 1205–1259 (2014).
33. Gildersleeve, K., Haselton, M. G. & Fales, M. R. Meta-analyses and p-curves support robust cycle shifts in women's mate preferences: Reply to Wood and Carden (2014) and Harris, Pashler, and Mickes (2014). *Psychol. Bull.* **140**, 1272-1280 (2014)
34. Marcinkowska, U. M., Jasienska, G. & Prokop, P. A comparison of masculinity facial preference among naturally cycling, pregnant, lactating, and post-menopausal women. *Arch. Sex. Behav.*, **47**, 1367–1374 (2018).
35. Harris, C. R. Menstrual cycle and facial preferences reconsidered. *Sex Roles*, **64**, 669–681 (2011).
36. Voracek, M., Haubner, T. & Fisher, M. L. Recent decline in nonpaternity rates: A cross-temporal meta-analysis. *Psychol. Rep.*, **103**, 799–811 (2008).
37. Greeff, J. M. & Erasmus, J. C. Three hundred years of low non-paternity in a human population. *Heredity (Edinb)*, **115**, 396–404 (2015).
38. Thornhill, R. & Gangestad, S. W. Facial attractiveness. *Trends Cogn. Sci.*, **3**, 452–460 (1999).
39. Folstad, I. & Karter, A. J. Parasites, bright males, and the immunocompetence handicap. *Am. Nat.*, **139**, 603–622 (1992).
40. Tybur, J. M. & Gangestad, S. W. Mate preferences and infectious disease: Theoretical considerations and evidence in humans. *Phil. Trans. R. Soc. Lond. B.*, **366**, 3375–3388 (2011).

41. Habig, B. & Archie, E. A. Social status, immune response and parasitism in males: A meta-analysis. *Phil. Trans. R. Soc. Lond. B*, **370**, 20140109 <https://doi.org/doi:10.1098/rstb.2014.0109> (2015).
42. Roberts, M. L., Buchanan, K. L. & Evans, M. R. Testing the immunocompetence handicap hypothesis: A review of the evidence. *Anim. Behav*, **68**, 227–239 (2004).
43. Boothroyd, L. G. *et al.* Facial masculinity is related to perceived age but not perceived health. *Evol. Hum. Behav*, **26**, 417–431 (2005).
44. Harris, C. R. Shifts in masculinity preferences across the menstrual cycle: Still not there. *Sex Roles*, **69**, 507–515 (2013).
45. Rantala, M. J. *et al.* Adiposity, compared with masculinity, serves as a more valid cue to immunocompetence in human mate choice. *Proc. R. Soc. Lond. B*. **280**, 20122495; [doi:10.1098/rspb.2012.2495](https://doi.org/doi:10.1098/rspb.2012.2495) (2013)
46. Rhodes, G., Chan, J., Zebrowitz, L. A. & Simmons, L. W. Does sexual dimorphism in human faces signal health? *Proc. R. Soc. Lond. B*. **270**, S93-S95 (2003)
47. Foo, Y. Z., Simmons, L. W. & Rhodes, G. The relationship between health and mating success in humans. *R. Soc. Open. Sci*, **4**, 160603 <https://doi.org/doi:10.1098/rsos.160603> (2017).
48. Rosebrock, A. Facial landmarks with dlib, opencv, and python. <https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/> (2017)
49. Sagonas, C., Antonakos, E., Tzimiropoulos, G., Zafeiriou, S. & Pantic M. 300 faces in-the-wild challenge: Database and results. *Image Vis. Comp*, **47**, 3–18 (2016).
50. Kazemi, V. & Sullivan, J. Syntax of referencing in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1867-1874 (IEEE, 2014)
51. Mallick, S. Average face: OpenCV (C++/Python) tutorial. <https://www.learnopencv.com/average-face-opencv-c-python-tutorial/> (2016)
52. Duane, S., Kennedy, A. D., Pendleton, B. J. & Roweth, D. Hybrid Monte-Carlo. *Phys. Lett. B*, **195**, 216–222 (1987).
53. Hoffman, M. D. & Gelman, A. The no-u-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res*, **15**, 1593–1623 (2014).
54. Pound, N., Penton-Voak, I. S. & Surridge, A. K. Testosterone responses to competition in men are related to facial masculinity. *Proc. R. Soc. Lond. B*. **276**, 153-159 (2009)
55. Grey, P. B., Straftis, A. A., Bird, B. M., McHale, T. S. & Zilioli, S. Human reproductive behavior, life history, and the Challenge Hypothesis: A 30-year review, retrospective and future directions. *Horm. Behav.* **123**, 104530; [doi:/10.1016/j.yhbeh.2019.04.017](https://doi.org/doi:10.1016/j.yhbeh.2019.04.017) (2019)
56. Lieberman, D. E., McBratney, B. M. & Krovitz, G. The evolution and development of cranial form in Homo sapiens. *Proc. Natl Acad. Sci. USA*. **99**, 1134-1139 (2002)
57. Lacruz, R. S. *et al.* The evolutionary history of the human face. *Nat. Ecol. Evol*, **3**, 726–736 (2019).
58. Raia, P. *et al.* Unexpectedly rapid evolution of mandibular shape in hominins. *Sci. Rep*, **8**, 7340 <https://doi.org/doi:10.1038/s41598-018-25309-8> (2018).

59. Schoenemann, P. T. Evolution of the size and functional areas of the human brain. *Annu. Rev. Anthropol*, **35**, 379–406 (2008).
60. Bastir, M. *et al.* Effects of brain and facial size on basicranial form in human and primate evolution. *J. Hum. Evol*, **58**, 424–431 (2010).
61. Olson, I. R. & Marshuetz, C. Facial attractiveness is appraised in a glance. *Emotion*, **5**, 498–502 (2005).
62. Chang, L. & Tsao, D. Y. The code for facial identity in the primate brain. *Cell*. **169**, 1013-1028 e 1014; doi:10.1016/j.cell.2017.05.011 (2017)
63. Whitehouse, A. J. *et al.* Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood. *Proc. R. Soc. Lond. B*. **282**, 20151351; doi:10.1098/rspb.2015.1351 (2015)
64. Nguyen, T. V. *et al.* A testosterone-related structural brain phenotype predicts aggressive behavior from childhood to adulthood. *Psychoneuroendocrinology*, **63**, 109–118 (2016).
65. Plavcan, J. M. & van Schaik, C. P. Interpreting hominid behavior on the basis of sexual dimorphism. *J. Hum. Evol*, **32**, 345–374 (1997).
66. Muller, M. N. Testosterone and reproductive effort in male primates. *Horm. Behav*, **91**, 36–51 (2017).
67. Cieri, R. L. *et al.* Craniofacial feminization, social tolerance, and the origins of behavioral modernity. *Curr. Anthropol*, **55**, 419–443 (2014).
68. Klein, R. G. Anatomy, behavior, and modern human origins. *J. World Prehist*, **9**, 167–198 (1995).
69. Klimas, C., Ehlert, U., Lacker, T. J., Waldvogel, P. & Walther, A. Higher testosterone levels are associated with unfaithful behavior in men. *Biol. Psychol*, **146**, 107730 <https://doi.org/doi:10.1016/j.biopsycho.2019.107730> (2019).
70. Wright, N. D. *et al.* Testosterone disrupts human collaboration by increasing egocentric choices. *Proc. R. Soc. Lond. B*. **279**, 2275-2280(2012)
71. Roney, J. R. & Simmons, Z. L. Women's estradiol predicts preference for facial cues of men's testosterone. *Horm. Behav*, **53**, 14–19 (2008).
72. Kempnaers, B., Peters, A. & Foerster, K. Sources of individual variation in plasma testosterone levels. *Phil. Trans. R. Soc. Lond. B*, **363**, 1711–1723 (2008).
73. Kondo, A., Takahashi, K. & Watanabe, K. Sequential effects in face-attractiveness judgment. *Perception*, **41**, 43–49 (2012).
74. Pegors, T. K., Mattar, M. G., Bryan, P. B. & Epstein, R. A. Simultaneous perceptual and response biases on sequential face attractiveness judgments. *J. Exp. Psychol. Gen*, **144**, 664–673 (2015).
75. Peters, M., Simmons, L. W. & Rhodes, G. Testosterone is associated with mating success but not attractiveness or masculinity in human males. *Anim. Behav*, **76**, 297–303 (2008).
76. Neave, N., Laing, S., Fink, B. & Manning, J. T. Second to fourth digit ratio, testosterone and perceived male dominance. *Proc. R. Soc. Lond. B*. **270**, 2167-2172 (2003)
77. Collier, C. P. *et al.* Functional testosterone: Biochemical assessment of hypogonadism in men-report from a multidisciplinary workshop hosted by the Ontario Society of Clinical Chemists. *Aging Male*,

- 10, 211–216 (2007).
78. Trost, L. W. & Mulhall, J. P. Challenges in testosterone measurement, data interpretation, and methodological appraisal of interventional trials. *J. Sex. Med*, **13**, 1029–1046 (2016).
79. Fiers, T. *et al.* A critical evaluation of salivary testosterone as a method for the assessment of serum testosterone. *Steroids*, **86**, 5–9 (2014).
80. Selby, C., Lobb, P. A. & Jeffcoate, W. J. Sex hormone binding globulin in saliva. *Clin. Endocrinol. (Oxf)*, **28**, 19–24 (1988).
81. Hayes, L. D. *et al.* Salivary testosterone measurement does not identify biochemical hypogonadism in aging men: A ROC analysis. *Endocrine*, **50**, 256–259 (2015).
82. Lewis, J. G. Steroid analysis in saliva: An overview. *Clin. Biochem. Rev*, **27**, 139–146 (2006).
83. Prasad, S., Lassetter, B., Welker, K. M. & Mehta, P. H. Unstable correspondence between salivary testosterone measured with enzyme immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology*, **109**, 104373 <https://doi.org/doi:10.1016/j.psyneuen.2019.104373> (2019).
84. Fraser, C. G. & Fogarty, Y. Interpreting laboratory results. *BMJ*, **298**, 1659–1660 (1989).
85. Valero-Politi, J. & Fuentes-Arderiu, X. Within- and between-subject biological variations of follitropin, lutropin, testosterone, and sex-hormone-binding globulin in men. *Clin. Chem*, **39**, 1723–1725 (1993).
86. Hayes, L. D., Sculthorpe, N., Young, J. D., Baker, J. S. & Grace, F. M. Critical difference applied to exercise-induced salivary testosterone and cortisol using enzyme-linked immunosorbent assay (ELISA): Distinguishing biological from statistical change. *J. Physiol. Biochem*, **70**, 991–996 (2014).
87. Ritchie, K. L., Palermo, R. & Rhodes, G. Forming impressions of facial attractiveness is mandatory. *Sci. Rep*, **7**, 469 (2017).
88. Xia, Y., Leib, A. Y. & Whitney, D. Serial dependence in the perception of attractiveness. *J. Vis*, **16**, 28 (2016).
89. Roney, J. R., Hanson, K. N., Durante, K. M. & Maestripieri, D. Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proc. R. Soc. Lond. B*. **273**, 2169-2175 (2006)
90. Boonekamp, J. J., Ros, A. H. & Verhulst, S. Immune activation suppresses plasma testosterone level: A meta-analysis. *Biol. Lett*, **4**, 741–744 (2008).
91. Penton-Voak, I. & Chen, J. High salivary testosterone is linked to masculine male facial appearance in humans. *Evol. Hum. Behav*, **25**, 229–241 (2004).
92. DeBruine, L. M. *et al.* Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proc. R. Soc. Lond. B*. **273**, 1355-1360(2006)
93. Marcinkowska, U. M., Galbarczyk, A. & Jasienska, G. La donna e mobile? Lack of cyclical shifts in facial symmetry, and facial and body masculinity preferences—a hormone based study. *Psychoneuroendocrinology*, **88**, 47–53 (2018).

94. Peters, M., Simmons, L. W. & Rhodes, G. Preferences across the menstrual cycle for masculinity and symmetry in photographs of male faces and bodies. *PLoS ONE*, **4**, e4138 <https://doi.org/doi:10.1371/journal.pone.0004138> (2009).
95. Stephen, I. D. *et al.* Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evol. Hum. Behav*, **33**, 260–267 (2012).
96. Perrett, D. I. *et al.* Effects of sexual dimorphism on facial attractiveness., **394**, 884–887 (1998).
97. Flegr, J., Blum, A., Nekola, O. & Kroupa, S. What people prefer and what they think they prefer in short- and long-term partners. The effects of the phase of the menstrual cycle, hormonal contraception, pregnancy, and the marital and the parenthood status on partner preferences. *Evol. Hum. Behav*, **40**, 112–125 (2019).
98. Cornwell, R. E. *et al.* Reproductive strategy, sexual development and attraction to facial characteristics. *Phil. Trans. R. Soc. Lond. B*, **361**, 2143–2154 (2006).
99. Buss, D. M. Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behav. Brain Sci*, **12**, 1–14 (2010).
100. Schwarz, S. & Hassebrauck, M. Sex and age differences in mate-selection preferences. *Hum. Nat*, **23**, 447–466 (2012).
101. Roney, J. R., Simmons, Z. L. & Gray, P. B. Changes in estradiol predict within-women shifts in attraction to facial cues of men's testosterone. *Psychoneuroendocrinology*, **36**, 742–749 (2011).
102. Sutherland, C. A. M. *et al.* Individual differences in trust evaluations are shaped mostly by environments, not genes. *Proc. Natl Acad. Sci. USA* **117**, 10218-10224(2020)
103. Swaddle, J. P. & Reiersen, G. W. Testosterone increases perceived dominance but not attractiveness in human males. *Proc. R. Soc. Lond. B*. **269**, 2285-2289 (2002)
104. Wilcox, A. J., Dunson, D. & Baird, D. D. The timing of the "fertile window" in the menstrual cycle: Day specific estimates from a prospective study. *BMJ*, **321**, 1259–1262 (2000).
105. Krejza, J. *et al.* Oscillations of cerebrovascular resistance throughout the menstrual cycle in healthy women. *Ultrasound. Obstet. Gynecol*, **22**, 627–632 (2003).

Figures

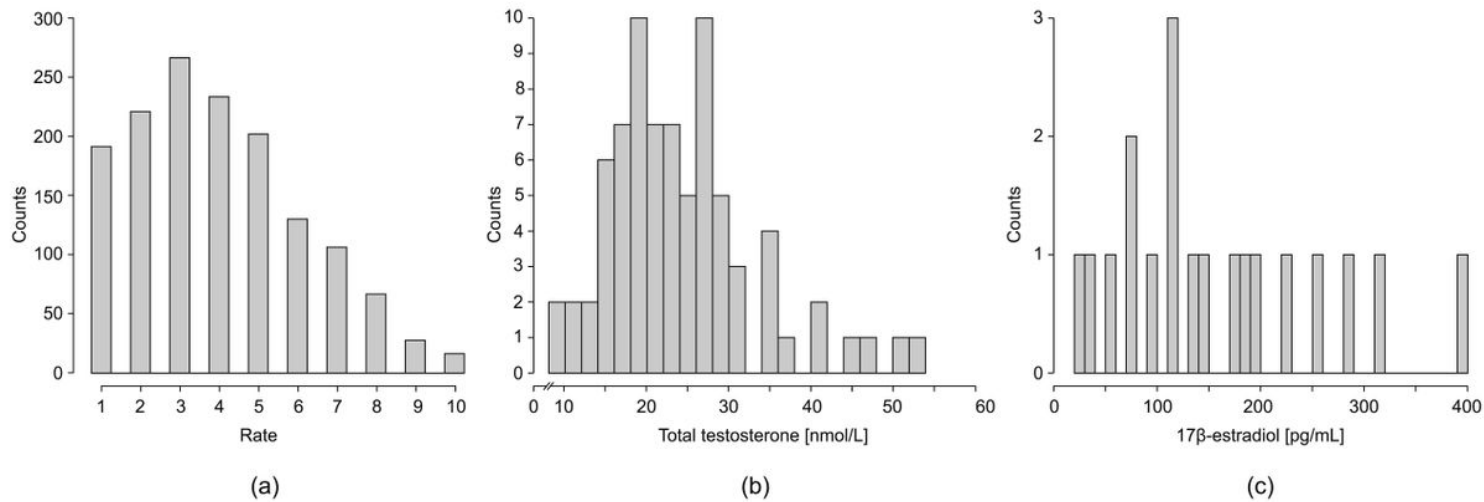


Figure 1

Histograms show distributions of a) rate of attractiveness of 77 men as assessed by 19 women, b) serum total testosterone concentration in 77 men, and (b) serum 17-β estradiol concentrations in 19 women.

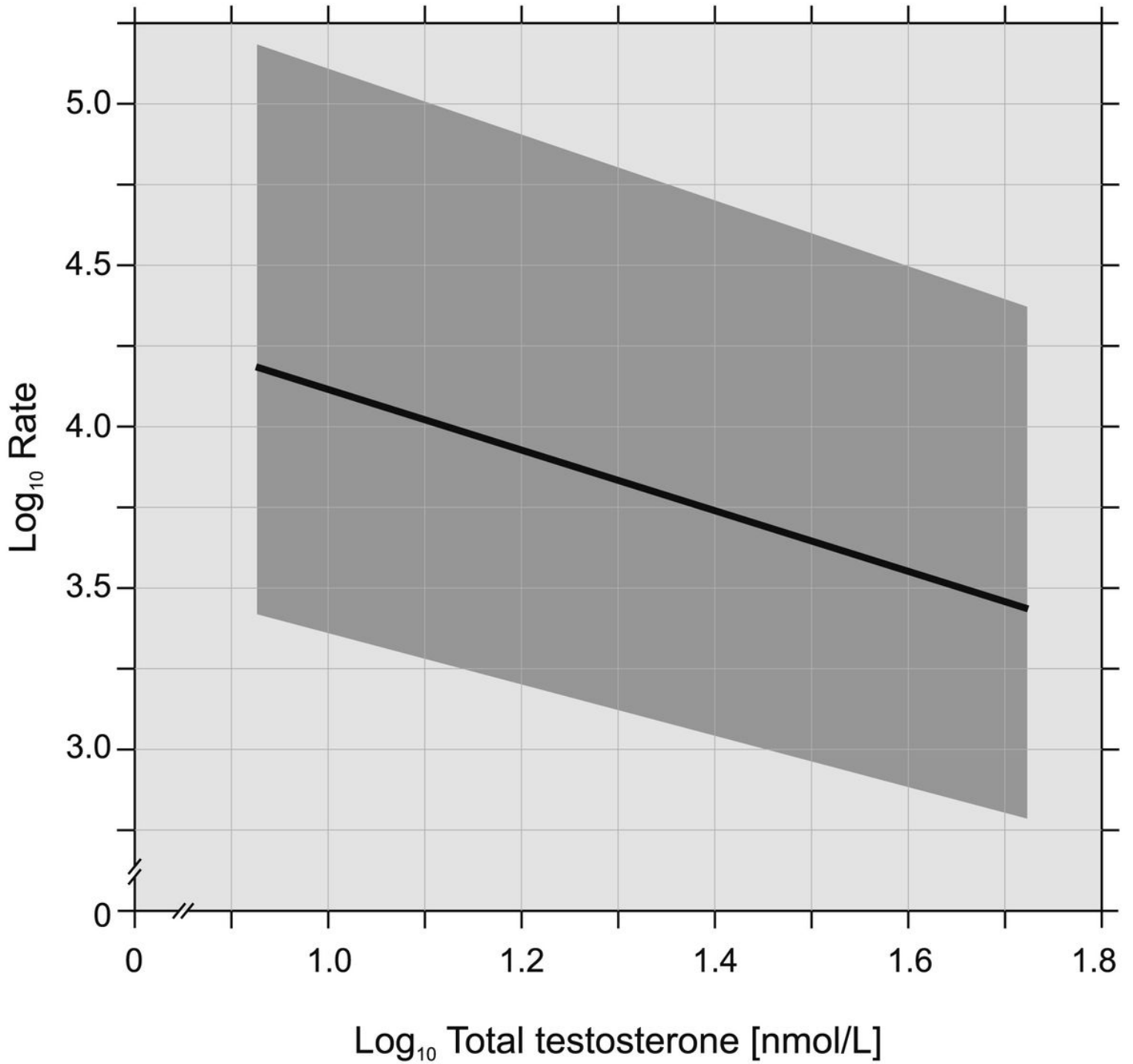


Figure 2

The graph shows the inverse relationship between logarithmic values of serum total testosterone concentration in 77 men and the rate of their attractiveness as assessed by 19 women.

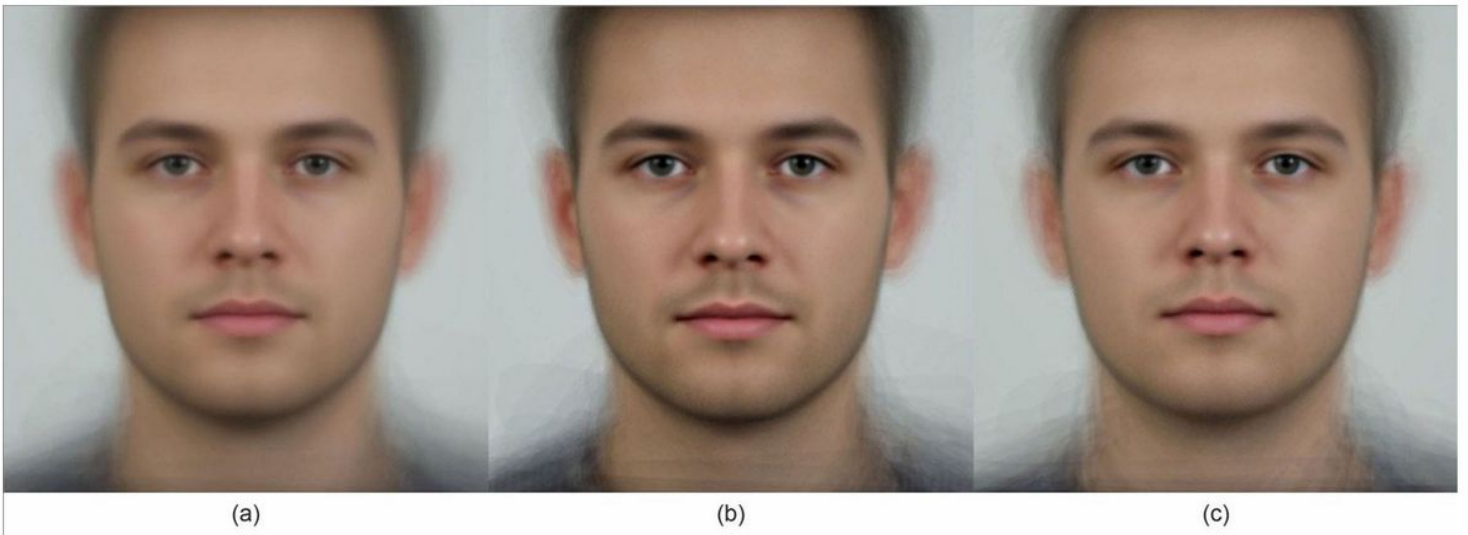


Figure 3

The images of a composite face of all 77 men (a), of 39 men scored as more attractive than median (b), and of 38 men scored as less attractive than median (c) by 19 female raters. Note that compared to the attractive one the less attractive composite face is slightly bigger and was scored on average by 1.93 points lower. The men from the attractive group had lower testosterone concentration than from the less attractive group.

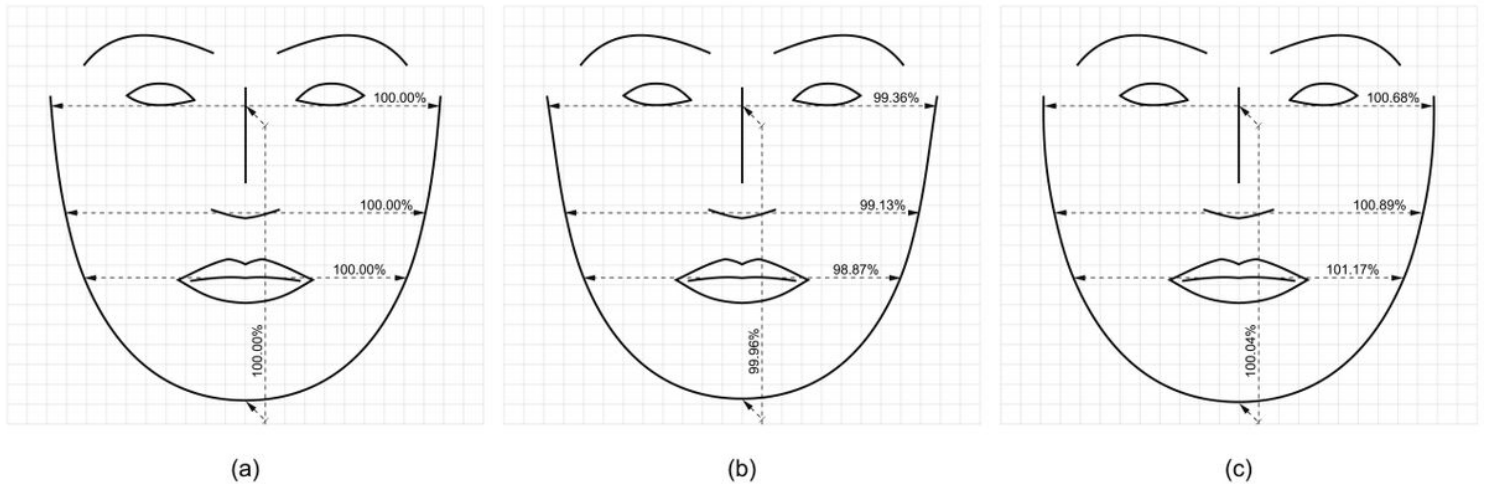


Figure 4

Drawings of extracted facial landmarks such as eyes, eyebrows, nose, mouth, and jawline from a composite image of faces of all 77 men (a), of 39 men scored as more attractive than median (b), and of 38 men scored as less attractive than median (c) by 19 female raters. Note that compared to the attractive one the less attractive composite face is wider by 1.32% at the level of eyes, 1.76% at the level of the lower ridge of the nose, and 2.30% at the level of mouth, and shorter by about 0.08% in midline. The average composite image of 77 men has all the measures set at 100%.