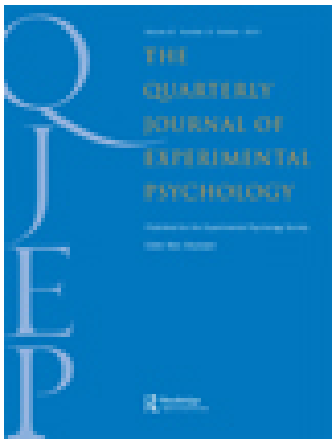

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Testosterone-to-Estradiol Ratio is Associated with Female Facial Attractiveness

Fabian Probst[ab], Cora Bobst[ab] & Janek S. Lobmaier[ab]
[a] Institute of Psychology, University of Bern, Switzerland
[b] Center for Cognition, Learning, and Memory, University of Bern, Switzerland
Accepted author version posted online: 02 Mar 2015.

Publisher: Taylor & Francis & The Experimental Psychology Society

Journal: *The Quarterly Journal of Experimental Psychology*

DOI: 10.1080/17470218.2015.1024696

Running head: Hormones and Female Facial Attractiveness

Testosterone-to-Estradiol Ratio is Associated with Female Facial Attractiveness

Fabian Probst, Cora Bobst, Janek S. Lobmaier

Institute of Psychology, University of Bern, Switzerland.

Center for Cognition, Learning, and Memory, University of Bern, Switzerland

Correspondence concerning this article should be addressed to: Fabian Probst, University of Bern, Institute of Psychology, Fabrikstrasse 8, 3012 Bern, Switzerland. Email: fabian.probst@psy.unibe.ch.

Acknowledgments

FP thanks Simonetta Barbero for comments on an earlier version of the manuscript. This study was supported by a grant from the Swiss National Science Foundation awarded to JL (grant number PP00P1_139072/1).

Abstract

The relationship between facial shape and attractiveness has been extensively studied, yet few studies have investigated the underlying biological factors of an attractive face. Many researchers have proposed a link between female attractiveness and sex hormones, but there is little empirical evidence in support this assumption. In the present study we investigated the relationship between circulating sex hormones and attractiveness. We created prototypes by separately averaging photographs of 15 women with high and low levels of testosterone, estradiol, and testosterone-to-estradiol ratio levels, respectively. An independent set of facial images was then shape transformed toward these prototypes. We paired the resulting images in such a way that one face depicted a female with high hormone level and the other a low hormone level. Fifty participants were asked to choose the more attractive face of each pair. We found that low testosterone-to-estradiol ratio and low testosterone were positively associated with female facial attractiveness. There was no preference for faces with high estradiol levels. In an additional experiment with 36 participants we confirmed that a low testosterone-to-estradiol ratio plays a larger role than low testosterone alone. These results provide empirical evidence that an attractive female face is shaped by interacting effects of testosterone and estradiol.

Keywords: facial attractiveness, face structure, sexual dimorphism, estrogen, testosterone

Testosterone-to-Estrogen Ratio is Associated with Female Facial Attractiveness

Introduction

Physical appearance plays an important role in mate choice for human as well as non-human animals, since attractiveness allegedly reflects biological quality. Much of the research on human mate preferences has focused on facial attractiveness (for reviews see Little, Jones, & DeBruine, 2011). For human faces, at least four parameters influence facial attractiveness: symmetry, averageness, healthy appearance and sexual dimorphism (e.g., Little et al., 2011). In female faces sexually dimorphic characteristics include feminine features such as less prominent brow ridge and jaws, full lips and a small chin. Indeed, female faces with feminized traits are perceived as being more attractive than faces with average traits (Perrett et al., 1998; Rhodes, Hickford, & Jeffery, 2000).

Feminine features become evident after the onset of puberty. During puberty the availability of endogenous sex hormones (such as testosterone, estrogens and progesterone) increases, affecting the development of secondary sexual features. Testosterone promotes the growth of lower face and jaw, cheekbones and brow ridges to more masculine features in men (Verdonck, Gaethofs, Carels, & de Zegher, 1999). In contrast, oestrogen is thought to be related to small bony features of the lower face, a flat middle face and large lips (Thornhill & Grammer, 1999), all of which are considered typically female traits. Faces of women with high estradiol levels should hence be particularly attractive (cf. Law-Smith et al., 2006).

Johnston and Franklin (1993) hypothesized that testosterone in relation to oestrogen is responsible for the sexual dimorphism and that this ratio might be a reliable marker of female attractiveness: Females with low testosterone-to-oestrogen ratio may have the most attractive faces since masculinizing of facial features is prevented and instead lip size is increased,

resulting in a more feminine looking face (Enlow, 1990; Thornhill & Grammer, 1999). However, empirical evidence is scarce regarding the relation between female facial attractiveness and sex hormones.

Virtually all studies that have investigated the relationship between female appearance and hormones measured only estradiol and progesterone (e.g., Jasienska, Ziomkiewicz, Ellison, Lipson, & Thune, 2004; Law-Smith et al., 2006; Puts et al., 2013). For example, Jasienska et al. (2004) investigated body characteristics such as breast size and waist-to-hip ratio (WHR) in relation to hormone levels and found that women with higher levels of trait estradiol have a more feminine body shape (large breasts, low WHR; but see Grillot, Simmons, Lukaszewski, & Roney, 2014). Law-Smith et al. (2006) compared concentrations of urinary oestrogen and progesterone metabolites with face ratings. Results showed that women with higher late follicular oestrogen levels have more feminine and more attractive looking faces than women with lower oestrogen levels. The authors conclude that oestrogen may be a marker for health and fertility. However they focused only on the female sex hormones and did not consider testosterone and its interaction effect with oestrogen.

To our knowledge no previous study has looked at the influence of testosterone on female facial attractiveness. Testosterone is typically considered as the prototypical male sex hormone, while oestrogens and progesterone are considered as female sex hormones. Indeed, testosterone and oestrogens play a critical role in the respective sex-differentiation during neuro-development (Arnold & Gorski, 1984). As a result, virtually all studies that investigated the relationship between male facial appearance and hormones measured testosterone, while studies with females looked at oestrogens and progesterone. Yet, testosterone as well as oestrogens and progesterone are synthesized in both men and women. Hence, not only the availability of the respective sex

hormone, but in particular the right balance of hormones might be fundamental for the development and maintenance of physiological processes, for example in the regulation of body bone mass growth and maintenance (Callewaert, Boonen, & Vanderschueren, 2010; Oury & Nyas, 2012).

The present work aims at assessing the relative role of testosterone, estradiol (a particular oestrogen that is found in humans) and the testosterone-to-estradiol (T/E) ratio on the shape of the attractive female face. Based on the reasoning mentioned above, we hypothesize that female facial attractiveness will be related to high levels of estradiol and low levels of testosterone. Specifically, we expect female faces with a lower T/E ratio to be perceived as more attractive than those with a higher T/E ratio.

We experimentally mimicked the effect that these hormones have on female face shape using computer graphics software and asked participants to choose the faces they like best in a two alternative forced choice (2AFC) task. In Experiment 1 we separately compare attractiveness ratings of (i) faces of women with high levels of estradiol to faces of women with low estradiol levels, (ii) faces of women with high testosterone to faces of women with low testosterone levels, and (iii) faces of women with a high T/E ratio to faces of women with a low T/E ratio in order to establish whether the respective hormone levels and ratios are related to facial attractiveness. In Experiment 2 we established which of these hormonal markers is most strongly involved in female facial attractiveness.

EXPERIMENT 1

Methods

Stimuli

Photography procedure

Fifty-one naturally cycling Caucasian women aged between 19-34 years (M = 24.704, SD = 3.965) were recruited for the following inclusion criteria: (a) no use of hormonal contraceptives for at least 3 previous months, (b) no suspected or confirmed pregnancy, (c) no breastfeeding, (d) regular menstrual cycle between 25 and 35 days of length, and (e) no medication.

Each woman agreed to have her photograph and saliva sample taken twice, once during ovulation and once during the luteal phase. This was done to control for cyclic variations in hormone levels and facial appearance(Bobst & Lobmaier, 2012). The time of ovulation was determined using WH Ovultell^{TM} ovulation test strips, which measure the concentration of metabolite from luteinizing hormone (LH) in urine. The surge in LH provokes ovulation in the next 24-36 hours. This test is a valid assessment to determine the onset of ovulation (Miller & Soules, 1996). We used ovulation test strips rather than the backward-counting-method because the backward counting method has low validation (Wilcox, Dunson, & Baird, 2000) resulting from wide inter- and intra-individual variations concerning the onset of ovulation (Park, Goldsmith, Skurnick, Wojtczuk, & Weiss, 2007; Wilcox et al., 2000). Additionally, the backward-counting-method does not detect anovulatory cycles. For photographs taken during ovulation, women reported to the lab immediately after a positive ovulation test strip result and were photographed within 24 hours of the positive test result. The luteal phase photography session took place 7 days after the LH surge. Order of photography session was counter-balanced across women. Specifically, half of the women were scheduled to be photographed within 48

hours of LH surge and then again 7 days later. The other half of the women were scheduled to have their first photograph taken 7 days after the measured peak of the LH surge. These latter women then again assessed the LH surge in the following cycle and were then photographed for the second time within 48 hours of the next LH peak. In both phases pictures and saliva were taken at the same time of the day to control for circadian changes in hormone levels. All participants were photographed without wearing glasses and they were instructed to look directly into the camera with a neutral expression. Photographs were taken under standard conditions with lighting from six Philips TL-D 90 Graphica 18W 950 – 59 cm fluorescent lamps. The distance between the camera and participant was 2 m. Images were captured on a Nikon D90 digital camera with a 50 mm lens and a resolution of 2163x3216 pixels. All women also completed the general health questionnaire (GHQ; Banks et al., 1980).

Saliva samples were collected using commercially available sampling devices (Salivette; Sarstedt, Rommelsdorf, Germany and SaliCaps; IBL, Hamburg, Germany). Estradiol and testosterone concentrations were analyzed from the saliva samples by an independent laboratory (Dresden Lab Service GmbH, Dresden, Germany). For every individual woman, we averaged the levels of estradiol and testosterone concentrations of the two photography sessions from which we calculated the testosterone-to-estradiol ratios. Each of these hormone measures was rank-ordered from the highest levels to the lowest levels. The 15 highest and lowest ranks of each respective hormone measure were used to create hormone level specific prototypes. Because saliva samples were collected using two different sampling devices ($N_{Salivette} = 23$; $N_{Salicaps} = 28$) and because they were analyzed in two separate batches, the absolute values of the hormone assays were not comparable. We hence z-transformed the values separately for each sampling device before rank-ordering the values. Importantly, the methodological procedure was identical

for both samples, the only difference being that we used different sampling devices (Salicaps and Salivette).

Stimulus creation

In a first step we created prototypes of high and low testosterone (T), estradiol (E) and T/E by averaging faces of 15 females with the highest and the 15 females with the lowest levels of T, E or T/E, respectively. The rationale here is that averaging a group of faces into one image reveals consistent characteristics of this group while characteristics that are not shared are averaged out. Using PsychoMorph computer graphics software (Tiddeman, Burt, & Perrett, 2001), the shape of each face was manually defined with 178 facial landmark points, marking the shape and position of eyes and brows, nose, mouth, ears, cheekbones, chin, as well as the outer face shape (for a more detailed description of prototype creation see Tiddeman et al., 2001). We then averaged the face shapes with the highest and lowest levels of T, E or T/E. Note that the selection of faces that were included in the prototype was based on the average hormone levels of the two cycle phases; hence two photographs of each woman (one taken during ovulation and one taken during the luteal phase) were included in the prototype. T and E were not correlated (r = .085, p = .563) but T and T/E levels were generally highly correlated (r = .499, p < .001). Nevertheless, a relevant proportion of individual faces (40%) in the high T group differed from those in the high T/E group. The same applies to the low T versus the low T/E groups. Prototypes are shown in Figure 1a.

Controlling for potential confounds

Testosterone and estradiol levels underlie fluctuation during the day; they start out at high levels in the morning and then decline across the course of the day (Dabbs & Delarue, 1991; Goji, 1993). In order to ensure that creation of prototypes was not confounded with circadian

fluctuation or age, we run a regression analyses with hormonal levels as dependent variable and age, general health and time of testing as predictor variables. Separate analyses were done for T and E levels. We found no significant association between T levels and age, general health questionnaire and circadian fluctuation (all *p*'s > .105). The same pattern was found in E levels (all *p*'s > .372). Additional analysis showed that high and low hormone groups did not significantly differ in age (all *p*'s > .669).

Stimulus Transformation

In the second step, these prototypes were taken as the endpoints on a high-low hormone level continuum. We then mimicked the effects that these hormones have on facial shape by shape-transforming unrelated faces towards the respective high and low prototypes. Sixty new frontal photographs of female Caucasian faces ranging in age between 18 and 35 years (M = 24.016, SD = 4.474) and showing a neutral expression were randomly selected from the CAL/PAL Face Database (Minear & Park, 2004). Twenty of these were shape-transformed towards the high and low estradiol prototypes, 20 were shape-transformed towards the high and low testosterone prototypes, and 20 towards the high and low T/E prototypes. Specifically, 75% of the linear differences in 2D shape between the high and low hormone prototype were added to face images of the new faces. Note that the transformation was shape transformation only, meaning that the stimuli differed only in shape while keeping other dimensions such as colour or luminance constant. This was done to avoid confounding effects due to different skin appearance (Stephen, Smith, Stirrat, & Perrett, 2009). The high and low versions of the same female face were paired and each pair was shown twice, once with the "high" face to the left, and once to the right. Faces were masked around the face line (for an example see Figure 1b). In total we used 3

(testosterone, estradiol, T/E) x 2 (lateralization on screen) x 20 (faces) = 120 face pairs in this experiment.

Attractiveness Task and Participants

Face pairs were randomly presented in two-alternative forced choice format on a computer screen. Participants were seated at a distance of 50 cm from a 17" monitor with a resolution of 1280 x 1024 pixels on which face pairs appeared at a visual angle of approximately 11.42° x 8° and were asked to choose the more attractive female face.

Fifty participants (29 female), ranging in age between 18 and 35 years (M = 24.54, SD = 3.81) took part in this study. For female participants we collected information about hormonal contraception use (21 contraceptive pill users, 8 naturally cycling) and menstrual cycle (length, onset of last menstruation, regularity). All participants provided written informed consent to take part in this study and were treated in accordance with the ethical protocol approved by local ethics committee and conformed with *The Code of Ethics of the World Medical Association* (Declaration of Helsinki).

Statistical analysis

Statistical analyses were performed using SPSS 22.0 with the significance level set at p = .05. According to our predictions, we coded answers as "correct" if participants chose the face that was transformed towards the prototype with the high E-level, low T-level and low T/E, respectively. A Shapiro-Wilk test showed that proportions of correctly chosen faces were normal distributed (p = .553). One-sample t-tests were used to test whether the proportion of correctly chosen faces significantly differ from chance (.50). Cohen's d was used as a measure of effect size. Additionally, the proportion of correctly chose face was analyzed using repeated-measures ANOVA with hormone (T, E vs. T/E) as within-subject factor and participant sex (female vs.

male) as between subject factor. The Huynh-Feldt epsilon correction for heterogeneity of covariances (Huynh & Feldt, 1976) was used when sphericity could not be assumed. In the ANOVA partial eta squared (η_p^2) was used a as measure of effect size. All post hoc pairwise comparisons were Bonferroni corrected.

Results

One-sample *t*-tests showed that for the T and T/E conditions the proportions of correctly chosen faces significantly differed from what would have been expected from chance (both *t's* > 6.116, *p's* < .001, *d's* > 1.75). In the E condition, the percentage of trials in which the higher estradiol face was chosen as more attractive did not significantly differ from chance level, $t(49) = -1.530$, $p = .132$, $d = -0.437$.

Repeated measures analysis of variance (ANOVA) showed a significant effect of hormone ($F_{1.390,\ 66.746} = 22.728$, $p < .001$, $\eta_p^2 = .321$). Post-hoc pairwise comparisons revealed that the proportion of correctly chosen faces significantly differed between T and E ($p < .001$) and between T/E and E ($p < .001$), but no differed between T and T/E ($p = .537$). There was no effect of participants sex ($F_{1,\ 48} = .618$, $p = .436$, $\eta_p^2 = .013$) and no participant sex x hormone interaction ($F_{1.39,\ 66.75} = 2.322$, $p = .122$, $\eta_p^2 = .046$). Additional analyses revealed that for female participants hormonal contraception had no impact on the face preferences (*p's* > .25). The results are shown in Figure 2. To investigate whether low T/E has a stronger impact on female attractiveness than low T, we conducted Experiment 2.

EXPERIMENT 2

The results of Experiment 1 indicate that faces of women with low testosterone levels are more attractive than faces of women with high testosterone levels. Likewise, faces of women with low testosterone-to-estradiol ratios were preferred over faces of women with high

testosterone-to estradiol ratios. This experimental design can not differentiate between preferences for facial cues associated with low testosterone levels and low testosterone-to-estradiol ratios. Experiment 2 was designed to directly test the hypothesis that faces of women with a low testosterone-to-estradiol ratio are perceived as being more attractive than faces of women with low testosterone levels, as suggested by Johnston and Franklin (1993).

Methods

Stimuli and Attractiveness task

The task and stimuli were very similar to Experiment 1, except that we shape-transformed faces between the low testosterone and low T/E prototypes. Specifically, a further twenty frontal photographs of Caucasian female faces ranging in age between 18 and 35 years (M = 24.45, SD = 5.041) and showing a neutral expression were randomly selected from the CAL/PAL Face Database (Minear & Park, 2004). In contrast to Experiment 1, faces were shape-transformed towards the low testosterone prototype and low testosterone-to-estradiol ratio prototype. In particular, 100% of the linear differences in 2D shape between the low T and low T/E hormone prototype were added to face images of the new faces. As in Experiment 1, the transformation was shape transformation only. We used 100% of difference in shape in Experiment 2 to enhance the differences between low testosterone and low testosterone-to-estradiol ratios. This was necessary because of the considerably large overlap between low T and low T/E faces. The low T and low T/E versions of the same female CAL/PAL face were paired and each pair was shown twice, counterbalancing the side on which each face appeared on the screen. Faces were masked around the face line in order to reduce peripheral cues such as clothes and hairdo. In total we used 2 (lateralization on screen) x 20 (faces) = 40 face pairs in this experiment. Participants were asked to choose the more attractive female face.

Participants

Thirty-six students (22 females) aged between 19 and 42 years (M = 24.75, SD = 6.34) took part in this study for course credit. Similar to Experiment 1, for female participants, we collected information about hormonal contraception use (11 contraceptive pill users, 10 naturally cycling, 1 preferred not to report) and menstrual cycle (length, regularity and onset of last menstruation). All gave written informed consent.

Results

Answers were coded as correct if participants chose the face that was transformed towards the T/E prototype (according to our prediction). One-sample t-tests showed that low T/E faces were preferred over low T faces (M = 61.6 %, SD = 11.79 %, $t(35) = 5.90$, $p < .001$, $d = 2.0$). There was no effect of participant sex ($t(34) = -.434$, $p = .667$). Additionally, there was a tendency that naturally cycling women more often chose the low T/E face (M = 0.66, SD = 0.11) than women using hormonal contraception (M = 0.56, SD = 0.11; $t(19) = -2.056$, $p = .054$, $d = -0.944$).

Discussion

The aim of this study was to determine a possible hormonal marker for female facial attractiveness. We found that a low testosterone-to-estradiol ratio (T/E) seems to have a strong impact on female facial attractiveness, followed by low T levels. We found no influence of estradiol levels on facial attractiveness.

The present study is the first to provide empirical evidence for an association between female facial attractiveness and the relation between circulating testosterone and estradiol levels (low T/high E) which was already suggested two decades ago (Johnston & Franklin, 1993). Low T/E ratio may inhibit masculinization of facial feature and instead might contribute to the increase of lip size, resulting in a more feminine appearance (Thornhill & Grammer, 1999).

Interestingly, faces with high E were not seen as more attractive than faces with low E when not controlling for T concentration. This is in contrast to previous findings reporting positive correlations between oestrogen levels and both attractiveness and femininity (Law-Smith et al., 2006). By means of urinary oestrogen assays and facial photographs Law-Smith et al. (2006) investigated the relationship between circulating oestrogen, attractiveness and femininity. In contrast to our data, they found positive correlations between oestrogen levels and both attractiveness and femininity. There are two main differences that may account for this divergent result. First, Law-Smith et al. (2006) determined the oestrogen level in the late-follicular stage of the menstrual cycle. In our study we took the average of two cycle phases to cover the hormonal level of two distinct phases of the female cycle. More importantly, our stimuli were created by transforming only shape information, resulting in stimuli that differ only in shape but not in any other dimension, such as colour and luminance. Please note that the correlation between female attractiveness and oestrogen level was only present when the women did not wear make-up (Law-Smith et al., 2006). It is therefore plausible that the positive relation between oestrogen level and female attractiveness found by(Law-Smith et al., 2006) was mainly due to facial skin coloration and not to facial shape. Various researchers found that the rating of facial attractiveness is related to skin colour (Fink, Grammer, & Thornhill, 2001; Jones, Little, Burt, & Perrett, 2004; Stephen et al., 2009). At the same time endocrinological studies have shown that oestrogens have a protective influence on the skin thickness, wrinkling and moisture (Kanda & Watanabe, 2005; Sator, Schmidt, Sator, Huber, & Honigsmann, 2001; Verdier-Sevrain, Bonte, & Gilchrest, 2006). We suggest that estradiol levels may have a positive influence on attractiveness by making the skin appear healthier but are less responsible for the face shape.

An alternative explanation for why we found no effect of estradiol on facial shape may be that the body mass indexes (BMI) of the women might have obscured the relationship between estradiol and facial attractiveness. Indeed, a recent study by Grillot et al. (2014) found significant positive interaction between sex hormone levels and women's body attractiveness only after controlling for BMI. Both of these speculative explanations do not detract from our main finding that the interaction between testosterone and estradiol (T/E ratio) plays a key role in shaping an attractive female face.

Participant sex had no effect on the choice of the more attractive woman, indicating that men and women strongly agree on which females they find attractive. While this finding is consistent with previous literature (Penton-Voak & Chen, 2004), it must be taken with caution because of the relative low number of male raters (especially in Experiment 2).

The present study shows a positive relationship between T/E and female attractiveness. The question remains why current sex hormone levels should be associated with female attractiveness, since sexually dimorphic traits mainly develop during puberty (e.g., Brook, 1981) and do not depend on momentary hormone levels. Given the assumption that levels of sex hormones remain relatively stable between puberty and menopause, current sex hormone levels can be seen as a proxy for pubertal hormone levels, mirroring the hormonal state of the critical pubertal period. Similar assumptions have been made in research looking at associations between hormone levels and facial appearance in men (e.g., Lefevre, Ewbank, Calder, von dem Hagen, & Perrett, 2013; Penton-Voak & Chen, 2004; Pound, Penton-Voak, & Surridge, 2009). Future directions should involve measuring the relationship between current female and male sex hormones during puberty.

We note that in the present study we focused only on shape information. Several studies suggest that other information, such as for example facial surface information (skin colour or texture) also influences perceptions of attractiveness (Fink et al., 2001; Lefevre et al., 2013). It seems that, particularly in women, both shape and surface information are relevant for an attractive appearance (Torrance, Wincenciak, Hahn, DeBruine, & Jones, 2014). It will be important to replicate these findings using other experimental approaches. Because some inconsistent findings in face preference have been found due to differences in the methods used to manipulate face images (see DeBruine, 2013; Scott, Clark, Boothroyd, & Penton-Voak, 2012 for disscussion) it will be a necessary next step to test whether the effects found in this experimentally controlled paradigm survive in more ecologically valid, real-life settings where shape cues are confounded with colour and texture information (such as correlational studies using unmanipulated faces). A further limitation of the present study is that we did not control for BMI. Future studies are needed to clarify what effect BMI has on the association between sex hormones and facial appearance.

In conclusion, we found that T/E is a valid hormonal maker for female attractiveness. We assume that adult hormone levels reflect hormone levels that were available during puberty, when important changes in facial appearance occur. The link between hormones and human facial features is a much-discussed topic in human evolutionary psychology, yet the relation between testosterone-to-estradiol ratio and facial appearance has not been quantified previously. We provide empirical evidence that an attractive female face has been shaped under the influence of optimally balanced androgenic and oestrogenic hormones. Our findings are in line with assumptions from evolutionary psychology stating that attractiveness is a reliable cue for women's fertility and health. High estradiol levels accompanied by low testosterone levels (low

T/E ratio) are putatively valid signs of female fertility and faces that were shaped by particularly high estradiol/low testosterone levels hence appear to be especially attractive. Last but not least, our findings underline the importance to consider both androgenic and oestrogenic influences when studying female facial attractiveness.

Reference

Arnold, A. P., & Gorski, R. A. (1984). Gonadal-steroid induction of structural sex-differences in the central nervous-system. *Annual Review of Neuroscience, 7*, 413-442. doi: 10.1146/annurev.neuro.7.1.413

Banks, M. H., Clegg, C. W., Jackson, P. R., Kemp, N. J., Stafford, E. M., & Wall, T. D. (1980). The use of the general health questionnaire as an indicator of mental-health in occupational studies. *Journal of Occupational Psychology, 53*(3), 187-194.

Bobst, C., & Lobmaier, J. S. (2012). Men's preference for the ovulating female is triggered by subtle face shape differences. *Hormones and Behavior, 62*(4), 413-417. doi: 10.1016/j.yhbeh.2012.07.008

Brook, C. G. D. (1981). Endocrinological control of growth at puberty. *British Medical Bulletin, 37*(3), 281-285.

Callewaert, F., Boonen, S., & Vanderschueren, D. (2010). Sex steroids and the male skeleton: a tale of two hormones. *Trends in Endocrinology and Metabolism, 21*(2), 89-95. doi: 10.1016/j.tem.2009.09.002

Dabbs, J. M., & Delarue, D. (1991). Salivary testosterone measurements among women - relative magnitude of circadian and menstrual cycles. *Hormone Research, 35*(5), 182-184.

DeBruine, L. M. (2013). Evidence versus speculation on the validity of methods for measuring masculinity preferences: comment on Scott et al. *Behavioral Ecology, 24*(3), 591-593. doi: 10.1093/beheco/ars098

Enlow, D. H. (1990). *Facial growth* (3rd ed.). Philadelphia: Harcourt Brace.

Fink, B., Grammer, K., & Thornhill, R. (2001). Human (Homo sapiens) facial attractiveness in relation to skin texture and color. *Journal of Comparative Psychology, 115*(1), 92-99. doi: 10.1037//0735-7036.115.1.92

Goji, K. (1993). 24-hour concentration profiles of gonadotropin and estradiol (e(2)) in prepubertal and early pubertal girls - the diurnal rise of e(2) is opposite the nocturnal rise of gonadotropin *Journal of Clinical Endocrinology & Metabolism, 77*(6), 1629-1635. doi: 10.1210/jc.77.6.1629

Grillot, R. L., Simmons, Z. L., Lukaszewski, A. W., & Roney, J. R. (2014). Hormonal and morphological predictors of women's body attractiveness. *Evolution and Human Behavior, 35*(3), 176-183.

Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics, 1*, 69-82.

Jasienska, G., Ziomkiewicz, A., Ellison, P. T., Lipson, S. F., & Thune, I. (2004). Large breasts and narrow waists indicate high reproductive potential in women. *Proceedings of the Royal Society B-Biological Sciences, 271*(1545), 1213-1217. doi: 10.1098/rspb.2004.2712

Johnston, V. S., & Franklin, M. (1993). Is beauty in the eye of the beholder? *Ethology and Sociobiology, 14*(3), 183-199. doi: 10.1016/0162-3095(93)90005-3

Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception, 33*(5), 569-576. doi: 10.1068/p3463

Kanda, N., & Watanabe, S. (2005). Regulatory roles of sex hormones in cutaneous biology and immunology. *Journal of Dermatological Science, 38*(1), 1-7. doi: 10.1016/j.jdermsci.2004.10.011

Law-Smith, M. J., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., . . . Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings of the Royal Society B-Biological Sciences, 273*(1583), 135-140. doi: 10.1098/rspb.2005.3296

Lefevre, C. E., Ewbank, M. P., Calder, A. J., von dem Hagen, E., & Perrett, D. I. (2013). It is all in the face: carotenoid skin coloration loses attractiveness outside the face. *Biology Letters, 9*(6), 4. doi: 10.1098/rsbl.2013.0633

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philosophical Transactions of the Royal Society B-Biological Sciences, 366*(1571), 1638-1659. doi: 10.1098/rstb.2010.0404

Miller, P. B., & Soules, M. R. (1996). The usefulness of a urinary LH kit for ovulation prediction during menstrual cycles of normal women. *Obstetrics and Gynecology, 87*(1), 13-17. doi: 10.1016/0029-7844(95)00352-5

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods Instruments & Computers, 36*(4), 630-633. doi: 10.3758/bf03206543

Oury, F., & Nyas. (2012). A crosstalk between bone and gonads *Blavatnik Awards for Young Scientists 2011* (Vol. 1260, pp. 1-7). Oxford: Blackwell Science Publ.

Park, S., Goldsmith, L. T., Skurnick, J. H., Wojtczuk, A., & Weiss, G. (2007). Characteristics of the urinary luteinizing hormone surge in young ovulatory women. *Fertility and Sterility, 88*(3), 684-690. doi: 10.1016/j.fertnstert.2007.01.045

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior, 25*(4), 229-241. doi: 10.1016/j.evolhumbehav.2004.04.003

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., . . . Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*(6696), 884-887. doi: 10.1038/29772

Pound, N., Penton-Voak, I. S., & Surridge, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society B-Biological Sciences, 276*(1654), 153-159. doi: 10.1098/rspb.2008.0990

Puts, D. A., Bailey, D. H., Cardenas, R. A., Burriss, R. P., Welling, L. L. M., Wheatley, J. R., & Dawood, K. (2013). Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle. *Hormones and Behavior, 63*(1), 13-19. doi: 10.1016/j.yhbeh.2012.11.007

Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex-typicality and attractiveness: Are supermale and superfemale faces super-attractive. *British Journal of Psychology, 91*, 125-140. doi: 10.1348/000712600161718

Sator, P. G., Schmidt, J. B., Sator, M. O., Huber, J. C., & Honigsmann, H. (2001). The influence of hormone replacement therapy on skin ageing - A pilot study. *Maturitas, 39*(1), 43-55. doi: 10.1016/s0378-5122(00)00225-5

Scott, I. M. L., Clark, A. P., Boothroyd, L. G., & Penton-Voak, I. S. (2012). Do men's faces really signal heritable immunocompetence? *Behavioral Ecology*. doi: 10.1093/beheco/ars092

Stephen, I. D., Smith, M. J. L., Stirrat, M. R., & Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *International Journal of Primatology, 30*(6), 845-857. doi: 10.1007/s10764-009-9380-z

Thornhill, R., & Grammer, K. (1999). The body and face of woman: One ornament that signals quality? *Evolution and Human Behavior, 20*(2), 105-120. doi: 10.1016/s1090-5138(98)00044-0

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *Ieee Computer Graphics and Applications, 21*(5), 42-50. doi: 10.1109/38.946630

Torrance, J. S., Wincenciak, J., Hahn, A. C., DeBruine, L. M., & Jones, B. C. (2014). The Relative Contributions of Facial Shape and Surface Information to Perceptions of Attractiveness and Dominance. *Plos One, 9*(10), e104415. doi: 10.1371/journal.pone.0104415

Verdier-Sevrain, S., Bonte, F., & Gilchrest, B. (2006). Biology of estrogens in skin: implications for skin aging. *Experimental Dermatology, 15*(2), 83-94. doi: 10.1111/j.1600-0625.2005.00377.x

Verdonck, A., Gaethofs, M., Carels, C., & de Zegher, F. (1999). Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *European Journal of Orthodontics, 21*(2), 137-143. doi: 10.1093/ejo/21.2.137

Wilcox, A. J., Dunson, D., & Baird, D. D. (2000). The timing of the "fertile window" in the menstrual cycle: day specific estimates from a prospective study. *British Medical Journal, 321*(7271), 1259-1262. doi: 10.1136/bmj.321.7271.1259

Figure captions

Figure 1. (a) Hormonal prototypes. Composite of 15 woman with lowest and highest levels of estradiol (E), composite of 15 woman with lowest and highest levels of testosterone (T), and composite of 15 woman with lowest and highest T/E ratios. (b) Example of a female face in the T/E condition. The face on the left is shape transformed to reduce T/E; the image on the right is a shape transformed to enhance T/E.

Figure 2. Experiment 1: Proportion correct for estradiol transformation (E, white bar), testosterone (T, grey bar) and testosterone-to-estradiol ratio (T/E ratio, dark bar). Answers were coded as correct if participants chose the face that was transformed towards the prototype with the high E-level, low T-level and low T/E, respectively. Error bars depict standard errors *(*** $p < .001$).*