Psychoneuroendocrinology (2011) 36, 742—749

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/psyneuen

Changes in estradiol predict within-women shifts in attraction to facial cues of men's testosterone

James R. Roney [a,*], Zachary L. Simmons [a], Peter B. Gray [b]

[a] Department of Psychology, University of California, Santa Barbara, CA 93106-9660, USA
[b] Department of Anthropology and Ethnic Studies, University of Nevada, Las Vegas, 4505 Maryland Parkway, Box 455003, Las Vegas, NV 89154-5003, USA

Received 10 July 2010; received in revised form 9 September 2010; accepted 12 October 2010

KEYWORDS
Menstrual cycle;
Mate preferences;
Estradiol;
Testosterone;
Attractiveness;
Sex hormones

Summary Many studies have demonstrated that women express stronger attraction to androgen-related traits when tested near ovulation than when tested at other times in the cycle. Much less research, however, has directly addressed which hormonal or other physiological signals may regulate these temporal shifts in women's attractiveness judgments. In the present study, we measured women's preferences for facial cues of men's testosterone concentrations on two occasions spaced two weeks apart, while also measuring women's salivary estradiol and testosterone concentrations at each testing session. Changes in women's estradiol concentrations across sessions positively predicted changes in their preferences for facial cues of high testosterone; there was no such effect for changes in women's testosterone concentrations. For the subset of women who had a testing session fall within the estimated fertile window, preferences for high testosterone faces were stronger in the fertile window session, and change in estradiol from outside to inside the fertile window positively predicted the magnitude of the ovulatory preference shift. These patterns were not replicated when testing preferences for faces that were rated as high in masculinity, suggesting that facial cues of high testosterone can be distinguished from the cues used to subjectively judge facial masculinity. Our findings suggest that women's estradiol promotes attraction to androgen-dependent cues in men (similar to its effects in females of various nonhuman species), and support a role for this hormone as a physiological regulator of cycle phase shifts in mating psychology.

1. Introduction

An expanding research literature has demonstrated shifts in women's attractiveness judgments across the menstrual cycle, with the time near ovulation associated with stronger attraction to more masculine morphological and behavioral traits in men (for reviews, see Gangestad and Thornhill, 2008; Jones et al., 2008; Thornhill and Gangestad, 2008). Stronger preferences for facial masculinity, for instance, have been reported when women are tested in higher vs. lower fertility regions of the cycle (Johnston et al., 2001; Little et al., 2008; Penton-Voak and Perrett, 2000; Penton-Voak et al., 1999), and facial masculinity has in turn been proposed to index

* Corresponding author. Tel.: +1 805 893 4871; fax: +1 805 893 4303.
E-mail address: roney@psych.ucsb.edu (J.R. Roney).

0306-4530/$ — see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.psyneuen.2010.10.010

heritable components of men's health via its association with testosterone exposure (e.g., Penton-Voak et al., 1999). Consistent with the possibility that cycle phase shifts are in fact tracking cues of androgen exposure, a recent study also reported mid-cycle peaks in women's preferences for the faces of men with higher circulating testosterone concentrations (Roney and Simmons, 2008).

A first goal of the present research was to test replication of cycle phase shifts in preferences for facial cues of men's testosterone. The Roney and Simmons (2008) paper tested women's responses to a sample of natural faces, which has advantages in terms of ecological validity, but also leaves ambiguous the nature of the cues that may be driving the results. Here, we tested the effects of variations in 2D face shape associated with testosterone concentrations by warping individual face identities into the shapes characteristic of high and low testosterone men. Women participants were then asked to judge whether the high vs. low testosterone version of the same face was more attractive. This method holds roughly constant a variety of cues associated with individual face identities (e.g., skin texture, skin color, symmetry), and thus allows tests of whether cycle phase shifts in preferences for testosterone are still found when cues are restricted to variations in face shape.

A second goal of this research was to provide further evidence regarding the hormonal signals that may regulate temporal shifts in women's attractiveness judgments. In a between-subjects design, Roney and Simmons (2008) found that women's salivary estradiol concentrations (but not their progesterone or testosterone concentrations) on the day of testing positively predicted their degree of preference for the faces of higher testosterone men; when plotted against day of cycle, furthermore, estradiol concentrations and testosterone preferences fluctuated in concert across most regions of the cycle. Welling et al. (2007), by contrast, measured women's preferences for facial masculinity on 2–4 occasions per woman and found stronger preferences on the test day when women's salivary testosterone was highest relative to the test day when salivary testosterone was lowest; no such effects were found for salivary progesterone or estradiol. Finally, Rosen and Lopez (2009) recently demonstrated that changes in attention to courtship language across two test sessions (within-women) were strongly predicted by changes in estradiol but were unrelated to changes in either progesterone or testosterone.

In addition to the above studies that directly measured women's hormone concentrations, a number of other studies have estimated hormone concentrations from day of cycle in order to test which hormonal signals may regulate cycle phase shifts. Two such studies reported stronger preferences for masculinized traits on cycle days when progesterone is typically lower (Jones et al., 2005; Puts, 2006), a third study reported positive effects of estimated estradiol and negative effects of estimated progesterone on preferences for the scents of more symmetrical men (Garver-Apgar et al., 2008), and a fourth study reported positive relationships between estimated estradiol and women's preferences for dominant personality traits (Lukaszewski and Roney, 2009). In summary, studies that have actually measured women's hormones have found evidence that estradiol (Roney and Simmons, 2008; Rosen and Lopez, 2009) or testosterone (Welling et al., 2007) may regulate cycle phase shifts in mating psychology, whereas studies that have estimated hormones from cycle day have reported positive effects of estradiol (Garver-Apgar et al., 2008; Lukaszewski and Roney, 2009) and negative effects of progesterone (Garver-Apgar et al., 2008; Jones et al., 2005; Puts, 2006) on preferences for masculine traits.

In light of these mixed findings, further evidence regarding the potential hormonal regulators of cycle phase effects appears necessary. In addition, no study with actual hormone measurements has ever directly tested whether within-women changes in specific hormones predict ovulatory shifts in attractiveness judgments.[1] The present study was designed to provide additional evidence on these questions by assessing whether changes in specific hormones across two test sessions predict changes in face preferences. In addition to testing such correlations for all participants, we also planned tests of the hormonal predictors of ovulatory preference shifts via within-women comparisons of test sessions falling inside vs. outside the estimated fertile window.

We measured two hormones from women raters in the present investigation: salivary estradiol and salivary testosterone. Progesterone was not measured primarily because the cost of assaying a third analyte was prohibitive given our funding for this project. Although studies with estimated hormones have suggested that preferences for masculine traits may be lower when progesterone is high (Garver-Apgar et al., 2008; Jones et al., 2005; Puts, 2006), this hormone could not regulate the increased preferences for specific traits that are found when moving from the early follicular to the periovulatory phase (e.g., Gangestad et al., 2004; Roney and Simmons, 2008) since progesterone tends to be uniformly low across this time period. At most, then, progesterone could supplement or interact with other signals that regulate cycle phase shifts. Furthermore, no study with actual hormone measurements has ever demonstrated correlations between progesterone and cycle phase shifts in mating psychology, whereas the studies reviewed above support roles for both testosterone and estradiol.

Based on the between-subject findings in Roney and Simmons (2008), we hypothesized that within-women changes in estradiol across test sessions would positively predict within-women changes in preferences for high testosterone faces. For the subset of cases with a testing session near ovulation, we also predicted that women would exhibit stronger preferences for facial cues of men's testosterone when tested inside the estimated fertile window than when tested outside this window, and that changes in estradiol across sessions would positively predict the magnitude of these ovulatory preference shifts. In addition to constructing face stimuli that varied in cues of testosterone, we also created stimuli that varied in shapes associated with subjective ratings of masculinity (see methods). If what distinguishes the faces of

[1] Welling et al. (2007) showed within-women that the highest testosterone test day (out of 2–4 days) was associated with stronger preferences for facial masculinity than the lowest testosterone test day, but they did not specifically test whether fertile window preference shifts were predicted by changes in testosterone. In addition, the highest testosterone test day was on average over 7 days from the estimated day of ovulation (in either direction), suggesting that this day was often outside the fertile window.

high testosterone men is that they appear subjectively masculine, then women's choice patterns (and their hormonal predictors) may be similar for the testosterone-based and masculinity-based stimuli; different patterns across the two types of stimuli, on the other hand, would suggest that preferences for cues of elevated testosterone can be discriminated from preferences for cues of masculinity.

2. Methods

2.1. Face stimuli

The faces used for stimulus construction were drawn from a database of men's face photographs taken in previous studies (n = 153; mean age = 18.88 $\pm$ 0.10 years). Photos were taken at a distance of 1.25 m under standard lighting conditions with subjects instructed to adopt a neutral facial expression. Baseline salivary testosterone concentrations were available from 76 of the men who were photographed (for sample collection and assay details, see Roney et al., 2007).

For construction of the testosterone-based stimuli, we identified the photos of the men with the 12 highest and 12 lowest testosterone concentrations (these were actually residual values after controlling for time of day and assay batch). From each set of 12, subsets of four faces were randomly selected and then morphed together to produce three high testosterone composite faces (mean testosterone values 1.36 SD above the full sample mean) and three low testosterone composite faces (mean testosterone values 0.95 SD below the full sample mean). Composites were created using the Fantamorph 3.1 morphing software (Abrosoft Co.), which averages the shape, color, and texture features of facial photographs.

In order to create face pairs in which the same identity was stretched into both a low and high testosterone face shape, we next applied a technique known as warping to other faces from our face database. In warping, landmark points are first placed on both an identity face and a target composite face, and the morphing program then stretches the shape of the identity face to the corresponding point locations on the composite face. This process preserves the color and texture of the identity faces and thus varies only the shape cues associated with the high and low testosterone composites. Eighteen faces not used in construction of the composites were chosen from the middle of the testosterone range for use as identity faces. Pairs of these faces were first morphed together (this produces faces that are less idiosyncratic and thus appear more realistic when warped) and then each of the nine two-face morphs was warped into the shape of one low and one high testosterone composite. Each of the three high and low testosterone composites served as target faces for the warping process three times (each low composite was paired once as a target face with each high composite), producing an initial pool of nine stimulus pairs, but errors in stimulus construction led to the use of only seven of these pairs in the rating study. Fig. 1 presents an example of a two-face morph that has been warped into the shapes of high and low testosterone composites, respectively. Because the warped faces preserved the same identity, only the warped faces and not the composites were used in the forced choice trials of the main study (see below).

The same procedures were used to create the masculinity-based stimuli. Each of the 153 faces in our face bank was rated for masculinity on a 1—7 likert-type scale. Undergraduate women (mean age = 18.64 $\pm$ .13 years) performed the ratings across three different studies (ns from 40 to 81) in which each woman rated an average of 51 faces. Ratings were standardized within raters and the average z-score for each face was computed. The 12 highest and 12 lowest rated faces were identified and subsets of four faces were randomly chosen from each set of 12 and morphed together to form three high masculinity composites (mean rating 1.07 SD above the full sample mean) and three low masculinity composites (mean rating 0.81 SD below the full sample mean). As with the testosterone-based stimuli, other faces from our face database were warped into the shapes of both a low and a high masculinity composite; each high composite was paired as a target in the warping process three times with each low

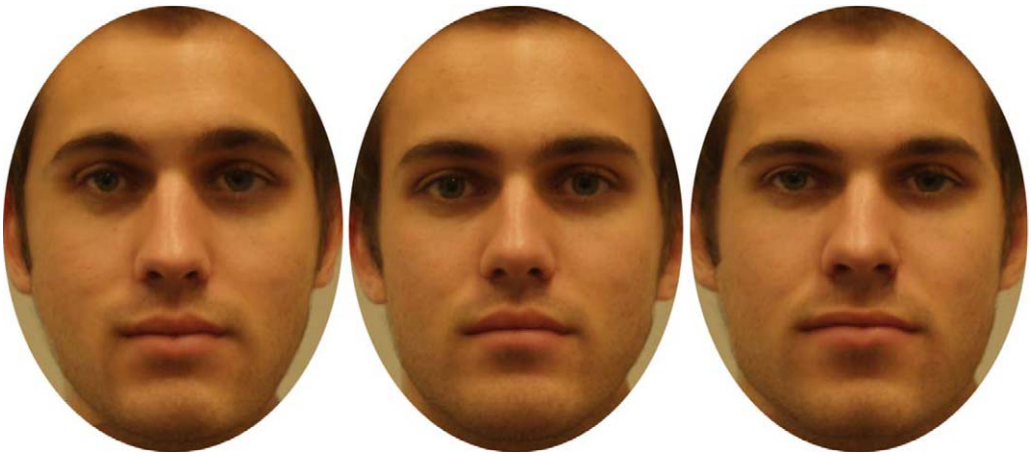

Figure 1 Example stimuli for the testosterone-based forced choice trials. The center image is a composite of two natural faces that have been morphed together. The image on the left is the center image warped into the shape of a low testosterone composite face, and the image on the right is the center image warped into the shape of a high testosterone composite face. Participants viewed the warped images in pairs (e.g., the leftmost and rightmost images in the figure) and indicated which one they found more physically attractive.

composite, resulting in a total of 27 face pairs representing high and low masculinity versions of the same faces.

2.2. Participants

Women raters were UCSB students recruited conditional on not using hormonal contraceptives. In order to reach a target sample size of 40 women tested across two sessions, an initial sample of 61 women was recruited for a first rating session, of whom 40 agreed to return for a second session. Thirty-six of these 40 women actually returned two weeks later and completed the second testing, though missing data from one woman reduced the sample size to $n = 35$ (mean age of final sample = 19.46 ± 0.29 years). Participation was in exchange for partial fulfillment of a course requirement.

2.3. Procedures

Participants viewed pairs of faces in which the same identity had been warped into the shape of one low and one high testosterone/masculinity composite (see Fig. 1). For each face pair, participants made forced-choice judgments of which face was more physically attractive. The testosterone- and masculinity-based stimuli were intermingled across trials and presented in a random order via a computer program.

The participants also provided saliva samples via passive drool at the beginning and end of each testing session (approximately 30 min. apart), and completed surveys at the end of each session. In one of these surveys, women indicated the first day of their last menses. In addition, the participants were asked to send an electronic mail confirming the onset of their next menses after the second session. These procedures were approved by the UCSB Institutional Review Board.

2.4. Fertile window estimation

The fertile window is defined as the days of the cycle in which intercourse can produce conception, and has been shown to extend from five days before ovulation through the day of ovulation itself (Wilcox et al., 1995). The day of ovulation, in turn, has been shown to occur most frequently at about 14 days before the onset of next menses (Baird et al., 1995), although there is certainly variability in this timing (see Fehring et al., 2006; Wilcox et al., 2000). Based on these findings, and following previous cycle phase studies in the mating literature (e.g., Garver-Apgar et al., 2008; Rosen and Lopez, 2009), we employed a backward counting technique in which we estimated the day of ovulation as 14 days prior to the end of the cycle and then estimated the fertile window as that day and the previous five days. For cases in which both testing sessions occurred in the same cycle, the end of cycle was determined from the e-mail indicating date of next menses onset after the second session; for cases in which the sessions spanned two cycles, the end of the first cycle was determined as the date of previous menses onset reported in the second session.

2.5. Hormone assays

Women's saliva samples were stored at $-80\ ^{\circ}C$ before being shipped on dry ice to the Human Behavioral Endocrinology Lab directed by Peter Gray at the University of Nevada, Las Vegas. Samples were assayed in duplicate using Salivary Estradiol Enzyme Immunoassay Kit 1-3702 and Salivary Testosterone Enzyme Immunoassay Kit 1-1402, both from Salimetrics, LLC. Each participant's saliva samples were run in the same assay. For estradiol, interassay coefficients of variation (CVs) were 15.1% and 15.4% for high and low controls, respectively, and the intraassay CV was 7.0%. For testosterone, interassay CVs were 4.7% and 21.4% for high and low controls, respectively, and the intraassay CV was 8.1%. The two saliva samples collected in each testing session produced hormone concentrations that were highly correlated with one another ($rs > 0.80$ for both testosterone and estradiol), and we therefore used the mean of these two values in our data analyses.

2.6. Data analyses

Within each testing session, the percentages of high testosterone and high masculinity stimuli chosen as more attractive were converted to *z*-scores. This transformation efficiently controls for any order of testing effects (without losing degrees of freedom) since preferences are then measured relative to other scores in the same testing session.

Pearson correlation coefficients and linear regression analyses were used to test associations between changes in hormones and changes in preferences across testing sessions for the full sample of women (without reference to the fertile window).[2] Fertile window effects were analyzed using repeated measures ANOVA. For cases in which one of the two testing sessions fell within the estimated fertile window ($n = 18$), the within-subjects factor compared preferences in the fertile window session to those in the non-fertile window session. The effects of changes in hormones on fertile window preferences were then tested by adding hormone change scores (fertile window minus non-fertile window) as continuous covariates to the repeated measures ANOVA: the *F* test for the interaction between the covariate and the within-subjects factor tests whether the size of the fertile window effect depends on the magnitude of hormone change across sessions. Significant interaction effects were followed up with correlation analyses. Finally, we used paired *t*-tests to analyze pre-ratings of our stimuli as manipulation checks (see Section 3.1). Reported significance levels are all two-tailed.

All preference and hormone change scores were normally distributed by visual inspection and the Shapiro–Wilk test (all $ps > 0.30$) after the exclusion of one 3 SD outlier for change in testosterone preference. Exclusion of this case reduced the sample size to $n = 34$ for analyses involving testosterone preference in the full sample, but left unchanged the sample size for fertile window cases since this woman did not have a session that fell within the estimated fertile window.

[2] Because change scores are often associated with baseline values, we re-computed our change score analyses while controlling for baseline (i.e. session 1) scores: in every case, statistical conclusions were unchanged.

3. Results

3.1. Manipulation checks

Fifty-eight undergraduate women (mean age = 18.93 ± 0.17 years) who were not participants in the main study rated each of the testosterone- and masculinity-based composite stimuli for both masculinity and physical attractiveness on 1–7 likert-type scales. For each rater, we computed mean ratings for composites belonging to the same class (e.g., high testosterone, low masculinity, etc.), and then compared ratings of high and low composites using paired *t*-tests. Providing evidence for the effectiveness of our masculinity manipulation, the high masculinity composites were rated more masculine (mean = 5.05 ± 0.13) than the low masculinity composites (mean = 3.03 ± 0.11), paired t (57) = 14.76, $p < 0.001$. The high masculinity composites were also rated more physically attractive (mean = 3.94 ± 0.14) than the low masculinity composites (mean = 3.02 ± 0.14), paired t (57) = 6.13, $p < 0.001$. The high testosterone composites were likewise rated more masculine (mean = 4.52 ± 0.10) than the low testosterone composites (mean = 3.99 ± 0.10), paired t (57) = 5.17, $p < 0.001$, which replicates previous reports showing that women perceive the faces of higher testosterone men as more subjectively masculine (Penton-Voak and Chen, 2004; Roney et al., 2006; cf. Pound et al., 2009). However, the high testosterone composites were not rated more physically attractive (mean = 3.39 ± 0.14) than the low testosterone composites (mean = 3.36 ± 0.11), paired t (57) = 0.33, $p = 0.74$, which suggests that the testosterone-based stimuli vary cues of testosterone concentrations while holding physical attractiveness fairly constant.

3.2. Preferences for testosterone-based stimuli

For the full sample of women, as predicted, change in preference for the high testosterone faces from session 1 to session 2 was positively correlated with change in estradiol across the same sessions (see Fig. 2 and Table 1). Change in testosterone, on the other hand, was uncorrelated with change in testosterone preference (see Table 1). A multiple regression analysis with changes in estradiol and testosterone entered together as predictors of change in testosterone preference confirmed a significant influence of estradiol, $\beta = 0.42$, $p = 0.02$, and a null effect for testosterone, $\beta = -0.13$, $p = 0.43$.

Preference data were re-analyzed for the subset of women who had one session fall within the estimated fertile window. Repeated measures ANOVA revealed that, as predicted, women exhibited stronger preferences for the high testosterone faces when tested inside the estimated fertile window (mean $z = 0.19 \pm 0.27$; mean raw percentage of high testosterone faces chosen = 52%) than when tested in their other session (mean $z = -0.28 \pm 0.18$; mean raw percentage of high testosterone faces chosen = 43%), F (1, 17) = 4.84, $p = 0.04$.[3] When changes in estradiol and testosterone from outside to inside the fertile window were simultaneously added to the model as covariates, only change in estradiol interacted with the fertile window factor, F (1, 15) = 6.12, $p = 0.026$ (for change in testosterone, F (1, 15) = 2.97, $p = 0.11$). More positive changes in estradiol were associated with more positive changes in preferences for facial cues of high testosterone, $r = 0.52$, $p = 0.03$. Although not significant, changes in testosterone from outside to inside the fertile window were actually negatively correlated with changes in preferences for the high testosterone face shapes, $r = -0.37$, $p = 0.13$.

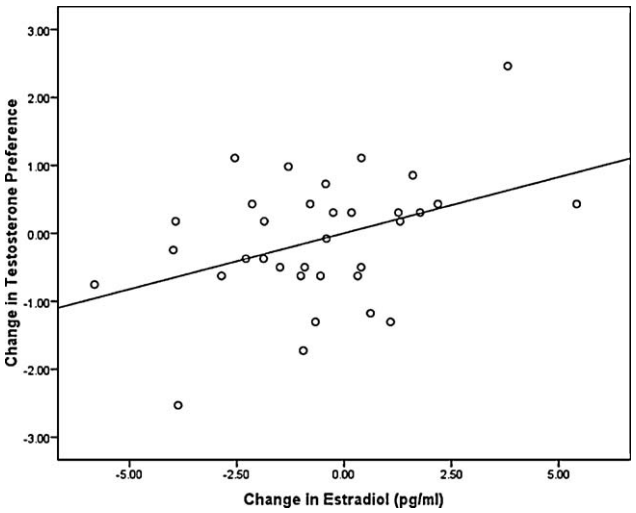

Figure 2 Change in women's testosterone preferences (session 2 preference *z*-score minus session 1 preference *z*-score) plotted against change in their estradiol concentrations (session 2 minus session 1).

3.3. Preferences for masculinity-based stimuli

In contrast to the findings for the testosterone-based stimuli, preferences for the high masculinity versions of the faces were not higher when women were tested inside the estimated fertile window (mean $z = -0.32 \pm 0.20$; mean raw percentage of high masculinity faces chosen = 53%) as opposed to when they were tested in their other session (mean $z = -0.07 \pm 0.23$; mean raw percentage of high masculinity faces chosen = 56%), F (1, 17) = 2.22, $p = 0.16$. For the full sample of women, neither change in estradiol nor change in testosterone significantly predicted change in

Table 1 Zero-order correlation matrix for face preference and hormone change scores (session 2 minus session 1) for the full sample of women.

	Δ Testosterone preference	Δ Masculinity preference	Δ Estradiol
Δ Masculinity preference	−0.35*		
Δ Estradiol	0.40*	−0.27	
Δ Testosterone	−0.05	0.28	0.23

* $p < .05$.

[3] As expected given the *z*-score transformation (see Section 2.6), order of testing (i.e. whether the fertile window was in the first or second session) did not interact with the within-subjects factor, F (1, 16) = 0.03, $p = 0.88$.

preference for high masculinity (see Table 1). Table 1 also demonstrates that change in testosterone preference was negatively correlated with change in masculinity preference, despite the fact that, between-subjects, preferences for the two types of stimuli showed signs of positive association within testing sessions (for session 1: $r = 0.10$, $p = 0.57$; for session 2: $r = 0.43$, $p = 0.01$).

4. Discussion

The present results demonstrate that changes in women's estradiol concentrations across a two-week period positively predict changes in their preferences for facial cues of high testosterone, which represents a within-subjects replication of a similar effect demonstrated between-subjects in a separate sample (Roney and Simmons, 2008). In addition, women in the current study exhibited stronger preferences for facial cues of high testosterone when tested inside the estimated fertile window than when tested in their other session, and change in salivary estradiol (but not testosterone) positively predicted the size of this fertile window preference shift. Surprisingly, the fertile window findings appear to be the first within-women demonstration of changes in measured hormone concentrations predicting an ovulatory shift in attractiveness judgments. As such, these results make unique contributions to an expanding cycle phase literature, and join other studies (Garver-Apgar et al., 2008; Lukaszewski and Roney, 2009; Roney and Simmons, 2008; Rosen and Lopez, 2009) in supporting a role for estradiol as an important physiological regulator of menstrual phase shifts in women's mating psychology.

A role for estradiol in modulating attraction to androgen-dependent cues is also consistent with findings in nonhuman species. In rodents, ovariectomy eliminates and estradiol treatment restores female preferences for the odors of testosterone-treated vs. castrated males (e.g., Xiao et al., 2004). Likewise, in songbirds, estradiol treatment modulates female neural responses to the auditory perception of androgen-dependent male song (Maney et al., 2008). The present findings can thus be seen within the context of a broader phylogenetic pattern in which sex hormones in perceivers promote attraction to cues of elevated sex hormone concentrations in members of the opposite sex.

The functional logic of using estradiol to calibrate attraction to androgen-dependent cues may in turn derive from the ability of estradiol to index fertility across a range of timescales. If in fact elevated testosterone signals heritable fitness — as has been theorized (e.g., Folstad and Karter, 1992) — then scrutiny of androgen-dependent cues should be most important when fertility is higher and the acquisition of genetic benefits is thus physically possible. Estradiol peaks near ovulation and can therefore index within-cycle fertility, of course, but it is also the case that estradiol tends to be higher across most days of higher vs. lower fertility cycles (Lipson and Ellison, 1996; Venners et al., 2006). Because women in ancestral environments likely experienced fertile cycles quite rarely (see Lancaster and Kaplan, 2009; Strassmann, 1997), it may have been functional to down-regulate attention to heritable fitness indicators during long stretches of infertility but then up-regulate such attention when fertility returned since any mates chosen during the rare fertile cycles would have a greater probability of fathering offspring in the near future (see also Roney, 2009; Roney and Simmons, 2008). By indexing both between- and within-cycle fluctuations in fertility, then, estradiol can efficiently signal the presence of those circumstances in which genetic benefits are most relevant, thus potentially explaining the evolution of brain mechanisms that use estradiol as a modulator of attraction to androgen-dependent cues.

Results from the present study also demonstrated dissociations between the hormonal predictors of preferences for the testosterone- and masculinity-based face stimuli. Neither change in estradiol nor change in testosterone significantly predicted change in preference for the masculinized versions of men's faces, nor was preference for masculinity higher inside vs. outside the estimated fertile window. In addition, despite some evidence that preferences for high testosterone and high masculinity were positively correlated between subjects, there was actually a negative correlation between changes in preferences for the respective stimuli, such that women who exhibited an increase in preference for the high testosterone faces across sessions were likely to exhibit a decrease in preference for the high masculinity faces (see Table 1). Thus, although the high testosterone faces were perceived as more masculine than the low testosterone faces, the above dissociations suggest that facial cues of high circulating testosterone must be different in some respects from the cues used to subjectively judge facial masculinity. These cue differences could be traceable to individual feature sizes, or to more subtle differences in relational configurations between features, but future research will be necessary to test these possibilities.

We can only speculate as to why our findings did not replicate ovulatory shifts in preferences for masculine faces. Two aspects of our methods may be relevant here. First, we manipulated masculinity based on women's subjective ratings, and implicit definitions of masculinity may vary across distinct samples of raters. Some raters may employ subjective definitions of masculinity that closely track cues of sexual dimorphism (e.g., DeBruine et al., 2006 in one sample demonstrated similar preferences for faces masculinized based on measured sexual dimorphism vs. based on subjective ratings), but this may not be true in all cases since raters may vary in the extent to which they fold concepts like attractiveness into their implicit definitions. Little et al. (2008), for instance, found for one sample of women that natural faces rated the most masculine did not differ significantly in attractiveness ratings from faces rated the least masculine, whereas Peters et al. (2009) in a different sample found a strong positive relationship between rated attractiveness and rated masculinity. Differences between raters in their implicit conceptions of masculinity, then, might help explain why some studies that have employed subjective ratings of masculinity have provided positive evidence for ovulatory preference shifts (Johnston et al., 2001; Little et al., 2008), whereas other studies have not (the present study; Peters et al., 2009). Studies that have masculinized faces based on measured sexual dimorphism have been more consistent in demonstrating ovulatory shifts (e.g., Jones et al., 2005; Penton-Voak and Perrett, 2000; Penton-Voak et al., 1999), and it is possible that these measures more reliably capture androgen-dependent cues than do more variable measures based on subjective ratings.

The second aspect of our method that may be relevant to the null results for the masculinity-based stimuli is our construction of faces stretched to shapes associated with the extreme ends of the rated masculinity distribution. Previous studies that have presented continuous frames running from less to more masculine face transforms have found that fertile window preferences are further toward the masculine end of continua than non-fertile window preferences, but are still far from the most masculine faces (e.g., Johnston et al., 2001; Penton-Voak et al., 1999). This raises the possibility that our masculinity transformations were too extreme, such that more moderate variations in masculinization may have replicated previous findings. Such methodological differences from previous studies complicate the interpretation of the null results for our masculinity-based stimuli, and further research is clearly necessary to test these conjectures.

Another interpretive issue regarding our findings concerns the meaning of facial cues of circulating testosterone concentrations. Men's testosterone concentrations can fluctuate in a state-like manner in response to circumstances such as exposure to potential mates (e.g., Roney et al., 2007, 2010) or current relationship status (e.g., Gray et al., 2002, 2004; Mazur and Michalek, 1998), but facial structure intuitively seems far more stable. Importantly, however, men's rank-ordering in terms of testosterone production stays fairly stable over time despite state-like fluctuations, with correlations in the 0.50—0.80 range for testosterone concentrations measured even years apart (Dabbs, 1990; Mazur and Michalek, 1998; Vermeulen and Verdonck, 1992). In addition, our use of multiple faces in the construction of the testosterone composites may have dampened the effects of transient hormone fluctuations in any given individual and thus abstracted out those features more reliably associated with high or low rates of testosterone production. As such, there is a reasonable likelihood that our positive results for the testosterone-based stimuli reflect reactions to facial cues of long-term, trait-like rates of testosterone production. If in turn high long-term testosterone production is sustainable only by healthier men (see Folstad and Karter, 1992), then facial cues of high testosterone may in fact indicate heritable fitness. However, we cannot rule out the possibility that facial cues of testosterone are more strongly associated with social dominance than with heritable fitness (see Puts, 2010), which might call into question whether cycle phase shifts in preferences for these cues reflect greater attraction to fitness indicators when the acquisition of genetic benefits is physically possible. Thus, although there is now fairly strong evidence that fluctuations in women's estradiol concentrations predict their attraction to facial cues of men's testosterone, the precise information signaled by such cues is a topic that requires further research (see also Ellison, 2008).

In conclusion, the present findings make contributions to an expanding cycle phase literature by demonstrating that within-women changes in salivary estradiol predict ovulatory shifts in attraction to face shapes characteristic of men with higher testosterone. Although we cannot exclude the importance of other hormonal predictors of women's attractiveness judgments — especially progesterone, which was not measured in the current study — the results are consistent with findings from the nonhuman literature in which estradiol has been demonstrated to regulate many components of female mating behavior (for a review, see Thornhill and Gangestad, 2008). Likewise, recent reports in humans have implicated estradiol in women's mating motivations (Durante and Li, 2009), emotional processing (Guapo et al., 2009), and spatial cognition (Resnick et al., 2009), and the present study adds further evidence for the importance of this hormone in explaining dynamic changes in women's psychology and behavior.

Role of funding source

Funding for this study was provided by a UCSB Academic Senate grant to James Roney; the UCSB Academic Senate had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

We thank Daniel Schantz for assistance with construction of the face stimuli, and Freidun Hadi for assistance with the hormone assays.

References

Baird, D.D., McConnaughey, R., Weinberg, C.R., Musey, P.I., Collins, D.C., Kesner, J.S., Knecht, E.A., Wilcox, A.J., 1995. Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. Epidemiology 6, 547—550.

Dabbs, J.M., 1990. Salivary testosterone measurements: reliability across hours, days, and weeks. Physiol. Behav. 48, 83—86.

DeBruine, L.M., Jones, B.C., Little, A.C., Boothroyd, L.G., Perrett, D.I., Penton-Voak, I.S., Cooper, P.A., Penke, L., Feinberg, D.R., Tiddeman, B.P., 2006. Correlated preferences for facial masculinity and ideal or actual partner's masculinity. Proc. R. Soc. Lond. B: Biol. Sci. 273, 1355—1360.

Durante, K.M., Li, N.P., 2009. Oestradiol level and opportunistic mating in women. Biol. Lett. 5, 179—182.

Ellison, P.T., 2008. Beauty: in the gonads of the beholder—and the beheld. Horm. Behav. 53, 11—13.

Fehring, R.J., Schneider, M., Raviele, K., 2006. Variability in the phases of the menstrual cycle. J. Obstet. Gynecol. Neonat. Nurs. 35, 376—384.

Folstad, I., Karter, A.J., 1992. Parasites, bright males, and the immunocompetence handicap. Am. Nat. 139, 603—622.

Gangestad, S.W., Simpson, J.A., Cousins, A.J., Garver-Apgar, C.E., Christensen, P.N., 2004. Women's preferences for male behavioral displays change across the menstrual cycle. Psychol. Sci. 15, 203—207.

Gangestad, S.W., Thornhill, R., 2008. Human oestrus. Proc. R. Soc. Lond. B: Biol. Sci. 275, 991—1000.

Garver-Apgar, C.E., Gangestad, S.W., Thornhill, R., 2008. Hormonal correlates of women's mid-cycle preferences for the scent of symmetry. Evol. Hum. Behav. 29, 223—232.

Gray, P.B., Chapman, J.F., Burnham, T.C., McIntyre, M.H., Lipson, S.F., Ellison, P.T., 2004. Hum. Nat. 15, 119—131.

Gray, P.B., Kahlenberg, S.M., Barrett, E.S., Lipson, S.F., Ellison, P.T., 2002. Marriage and fatherhood are associated with lower testosterone in human males. Evol. Hum. Behav. 23, 193—201.

Guapo, V.G., Graeff, F.G., Zani, A.C.T., Labate, C.M., dos Reis, R.M., Del-Ben, C.M., 2009. Effects of sex hormonal levels and phases of

the menstrual cycle in the processing of emotional faces. Psychoneuroendocrinology 34, 1087—1094.

Johnston, V.S., Hagel, R., Franklin, M., Fink, B., Grammer, K., 2001. Male facial attractiveness: evidence for hormone-mediated adaptive design. Evol. Hum. Behav. 22, 251—267.

Jones, B.C., DeBruine, L.M., Perrett, D.I., Little, A.C., Feinberg, D.R., Law Smith, M.J., 2008. Effects of menstrual phase on face preferences. Arch. Sex. Behav. 37, 78—84.

Jones, B.C., Perrett, D.I., Little, A.C., Boothroyd, L., Cornwell, R.E., Feinberg, D.R., Tiddeman, B.P., Whiten, S., Pitman, R.M., Hillier, S.G., Burt, D.M., Stirrat, M.R., Law Smith, M.J., Moore, F.R., 2005. Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. Horm. Behav. 48, 283—290.

Lancaster, J.B., Kaplan, H.S., 2009. The endocrinology of the human adaptive complex. In: Ellison, P.T., Gray, P.B. (Eds.), Endocrinology of Social Relationships. Harvard University Press, Cambridge, MA, pp. 95—118.

Lipson, S.F., Ellison, P.T., 1996. Comparison of salivary steroid profiles in naturally occurring conception and non-conception cycles. Hum. Reprod. 11, 2090—2096.

Little, A.C., Jones, B.C., DeBruine, L.M., 2008. 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. Indiv. Differ. 45, 478—482.

Lukaszewski, A.W., Roney, J.R., 2009. Estimated hormones predict women's mate preferences for dominant personality traits. Pers. Indiv. Differ. 47, 191—196.

Maney, D.L., Goode, C.T., Lange, H.S., Sanford, S.E., Solomon, B.L., 2008. Estradiol modulates neural responses to song in a seasonal songbird. J. Comp. Neurol. 511, 173—186.

Mazur, A., Michalek, J., 1998. Marriage, divorce, and male testosterone. Soc. Forces 77, 315—330.

Penton-Voak, I.S., Chen, J.Y., 2004. High salivary testosterone is linked to masculine male facial appearance in humans. Evol. Hum. Behav. 25, 229—241.

Penton-Voak, I.S., Perrett, D.I., 2000. Female preference for male faces changes cyclically: further evidence. Evol. Hum. Behav. 21, 39—48.

Penton-Voak, I.S., Perrett, D.I., Castles, D.L., Kobayashi, T., Burt, D.M., Murray, L.K., Minamisawa, R., 1999. Menstrual cycle alters face preference. Nature 399, 741—742.

Peters, M., Simmons, L.W., Rhodes, G., 2009. Preferences across the menstrual cycle for masculinity and symmetry in photographs of male faces and bodies. Plos One 4, e4138.

Pound, N., Penton-Voak, I.S., Surridge, A.K., 2009. Testosterone responses to competition in men are related to facial masculinity. Proc. R. Soc. Lond. B: Biol. Sci. 276, 153—159.

Puts, D., 2006. Cyclic variation in women's preferences for masculine traits. Hum. Nat. 17, 114—127.

Puts, D.A., 2010. Beauty and the beast: mechanisms of sexual selection in humans. Evol. Hum. Behav. 31, 157—175.

Resnick, S.M., Espeland, M.A., Yang, A.N., Maki, P.M., Coker, L.H., Jackson, R., Stefanick, M.L., Wallace, R., Rapp, S.R., 2009. Effects of conjugated equine estrogens on cognition and affect in postmenopausal women with prior hysterectomy. J. Clin. Endocrinol. Metab. 94, 4152—4161.

Roney, J.R., 2009. The role of sex hormones in the initiation of human mating relationships. In: Ellison, P.T., Gray, P.B. (Eds.), Endocrinology of Social Relationships. Harvard University Press, Cambridge, MA, pp. 246—269.

Roney, J.R., Hanson, K.N., Durante, K.M., Maestripieri, D., 2006. Reading men's faces: women's mate attractiveness judgments track men's testosterone and interest in infants. Proc. R. Soc. Lond. B: Biol. Sci. 273, 2169—2175.

Roney, J.R., Lukaszewski, A.W., Simmons, Z.L., 2007. Rapid endocrine responses of young men to social interactions with young women. Horm. Behav. 52, 326—333.

Roney, J.R., Simmons, Z.L., 2008. Women's estradiol predicts preference for facial cues of men's testosterone. Horm. Behav. 53, 14—19.

Roney, J.R., Simmons, Z.L., Lukaszewski, A.W., 2010. Androgen receptor gene sequence and basal cortisol concentrations predict men's hormonal responses to potential mates. Proc. R. Soc. Lond. B: Biol. Sci. 277, 57—63.

Rosen, M.L., Lopez, H.H., 2009. Menstrual cycle shifts in attentional bias for courtship language. Evol. Hum. Behav. 30, 131—140.

Strassmann, B.I., 1997. The biology of menstruation in *Homo sapiens*: total lifetime menses, fecundity, and nonsynchrony in a natural-fertility population. Curr. Anthropol. 38, 123—129.

Thornhill, R., Gangestad, S.W., 2008. The Evolutionary Biology of Human Female Sexuality. Oxford University Press, New York.

Venners, S.A., Liu, X., Perry, M.J., Korrick, S.A., Li, Z., Yang, F., Yang, J., Lasley, B.L., Xu, X., Wang, X., 2006. Urinary estrogen and progesterone metabolite concentrations in menstrual cycles of fertile women with non-conception, early pregnancy loss or clinical pregnancy. Hum. Reprod. 21, 2272—2280.

Vermeulen, A., Verdonck, G., 1992. Representativeness of a single point plasma testosterone level for the long-term hormonal milieu. J. Clin. Endocrinol. Metab. 74, 939—942.

Welling, L.L.M., Jones, B.C., DeBruine, L.M., Conway, C.A., Law Smith, M.J., Little, A.C., Feinberg, D.R., Sharp, M.A., Al-Dujaili, E.A.S., 2007. Raised salivary testosterone in women is associated with increased attraction to masculine faces. Horm. Behav. 52, 156—161.

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. Br. Med. J. 321, 1259—1262.

Wilcox, A.J., Weinberg, C.R., Baird, D.D., 1995. Timing of sexual intercourse in relation to ovulation. N. Engl. J. Med. 333, 1517—1521.

Xiao, K., Kondo, Y., Sakuma, Y., 2004. Sex-specific effects of gonadal steroids on conspecific odor preference in the rat. Horm. Behav. 46, 356—361.