Hormonal and morphological predictors of women’s body attractiveness

Rachel L. Grillot, Zachary L. Simmons, Aaron W. Lukaszewski, James R. Roney

1. Introduction

Functional approaches to understanding women’s body attractiveness posit the evolution of specialized mechanisms in perceivers that hone in on bodily features that would have predicted reproductively valuable qualities in human ancestral environments, such as health or fecundity (e.g., Gangestad & Scheyd, 2005; Symons, 1995). A low waist-to-hip ratio (WHR), for instance, has been proposed to signal qualities such as health, fecundity, and greater specialized fat stores for healthy fetal brain development (Lassek & Gaulin, 2008; Singh, 1993b; Singh & Singh, 2006; Singh & Singh, 2011), and, as such, men’s preference for this trait in mating partners (e.g., Furnham, Tan, & McManus, 1997; Singh, 1993a; Streeter & McBurney, 2003) may provide an example of specialized preference mechanisms honing in on reproductively valuable traits in others. Complicating this issue, however, are findings from some non-Western cultures that suggest preferences for larger body size and associated higher WHRs in women (Marlowe & Wetsman, 2001; Wetsman & Marlowe, 1999; Yu & Shephard, 1998; c.f. Marlowe, Apicella, & Reed, 2005; Sugiyama, 2004; Swami & Tovee, 2007); some have argued from such findings that preferences for traits such as low WHR are not products of specialized preference mechanisms but are instead attributable to Western media influences (Yu & Shephard, 1998).

One strategy for testing whether attractiveness judgments are generated by specialized preference mechanisms is to assess whether such judgments correlate with biological markers of health or fecundity, since positive correlations would be difficult to explain if attractiveness standards were culturally arbitrary. Women’s concentrations of estradiol and progesterone appear to act as biological markers of fecundity given evidence that these concentrations are positively correlated with conception probabilities (see Baird et al., 1999; Lipson & Ellison, 1996; Stewart et al., 1993; Venners et al., 2006). Jasienska, Ziomkiewicz, Ellison, Lipson, and Thune (2004) demonstrated that low WHR and large breast size predicted higher concentrations of these ovarian hormones across broad regions of the menstrual cycle, suggesting that these body shape characteristics may be valid cues of fecundity, at least within their sample of well-nourished Polish women. These authors did not report associations between hormone concentrations and ratings of the women’s body attractiveness, but such associations would more directly test whether preference mechanisms are attuned to physical cues of fecundity.
A few previous studies have provided evidence regarding the relationship between women's hormone concentrations and perceptions of their attractiveness. Law Smith et al. (2006) reported that ratings of women's facial attractiveness were positively correlated with the women's late follicular estradiol concentrations (see also Puts et al., 2013). Durante and Li (2009) observed a positive association between the mean of two estradiol measurements – one from the late follicular phase and the other from the luteal phase – and ratings of women's combined facial and body attractiveness. Rilling, Kaufman, Smith, Patel, and Worthman (2009) collected ratings of women's body attractiveness without face information, but failed to find a significant correlation between such ratings and a single measure of the women's estradiol that did not control for cycle day. In summary, with respect to hormonal correlates of women's body attractiveness, one study has reported significant correlations between ovarian hormone concentrations and both WHR and breast size, but no previous study has tested for hormonal correlates of isolated body attractiveness ratings when hormones were sampled frequently across broad regions of the menstrual cycle.

In the present research, we obtained salivary measurements of estradiol, progesterone, and testosterone across 1–2 menstrual cycles from a sample of young women; collected ratings of the women's body attractiveness from photos of them in standardized clothing (with faces obscured); and also obtained measurements of body mass index (BMI), breast size, and WHR. We hypothesized replication of higher estradiol and progesterone among women with lower WHR and larger breast size (Jasienska et al., 2004), and also predicted that concentrations of these hormones would positively predict body attractiveness ratings. Although not a primary purpose of the study, we also expected replication of negative correlations (in Western cultures) between body attractiveness ratings and both WHR and breast size, but no previous study has tested for hormonal correlates of isolated body attractiveness ratings when hormones were sampled frequently across broad regions of the menstrual cycle.

In the present research, we obtained salivary measurements of estradiol, progesterone, and testosterone across 1–2 menstrual cycles from a sample of young women; collected ratings of the women's body attractiveness from photos of them in standardized clothing (with faces obscured); and also obtained measurements of body mass index (BMI), breast size, and WHR. We hypothesized replication of higher estradiol and progesterone among women with lower WHR and larger breast size (Jasienska et al., 2004), and also predicted that concentrations of these hormones would positively predict body attractiveness ratings. Although not a primary purpose of the study, we also expected replication of negative correlations (in Western cultures) between body attractiveness ratings and both BMI and WHR (e.g., Rilling et al., 2009; Singh & Singh, 2011; Streeter & McBurney, 2003; Tovee & Cornelissen, 2001).

2. Methods

2.1. Body stimuli

2.1.1. Stimulus participants

Body photographs were obtained from a sample of women who participated in a larger study on the relationship between ovarian hormones and sexual psychology and behavior within natural menstrual cycles (see Roney & Simmons, 2013). Women participants provided daily saliva samples each morning (to control for diurnal variation in hormones; see Bao et al., 2003) across 1–2 menstrual cycles. Although 52 total women participated in the study, saliva samples were not sent for assay for women with many missing samples, and hormone data were ultimately obtained for 43 women; 41 of these women provided consent for use of their photographs in research. Of those women, 33 were judged to have experienced at least one ovulatory menstrual cycle (see below). These 33 women comprise the final stimulus sample (Mean age ± SD = 18.85 ± 1.28 years). Nineteen of the women self-identified as White, seven as Asian, five as Hispanic, and two as mixed ethnicity; none of the hormone variables, body dimensions, or attractiveness ratings differed significantly across ethnic categories.

2.1.2. Anthropometry

Participants attended four laboratory sessions per menstrual cycle; anthropometric measurements were obtained in one of the sessions from the first cycle. Weight, muscle mass, body fat, visceral fat, and water percentage were measured using a Tanita electrical impedance scale (Tanita BC-573), and height was self-reported via questionnaire. The values for height and weight were used to calculate body mass index (BMI). Women research assistants used measuring tapes to measure breast size (the widest circumference at the level of the chest) and underbreast circumference; following Jasienska et al. (2004), the ratio of these two values was employed as a measure of relative breast size. Bras were not removed before measurement, which may have introduced measurement error, although the average relative breast size in our sample (Mean breast size ± SD = 1.15 ± 0.04) was very similar to that reported by Jasienska et al. (2004; Mean breast size ± SD = 1.16 ± 0.04). WHR was initially measured using measuring tapes but a number of values appeared implausible when compared to the photographs and it appears that our research assistants identified waists in inconsistent ways. We therefore attempted to obtain reliable measurements of WHR from the women's photographs using a technique for photo measurements that was validated against more standardized direct body measurements (Steve Gaulin, personal communication, September 2012): the waist was defined as the narrowest point on the torso below the breasts, and the hips were defined as the widest point below the waist. Two research assistants independently measured these using Adobe Photoshop Elements 3.0, and computed the ratio of the two; these measurements were highly reliable (r = 0.97) and the means of the two ratios were used for data analyses.

2.1.3. Hormone measures

Morning saliva samples were first stored in women's home freezers and then delivered weekly to our research lab, after which they were stored at −80 °C until shipping for assay (for full details of the collection procedure, see Roney & Simmons, 2013). We initially estimated the day of ovulation as 15 days prior to the end of each cycle, and sent for assay all samples in a nine day window centered on the estimated day of ovulation, as well as samples from alternating days outside of this window. Samples were shipped on dry ice to the Endocrine Core Laboratory at the California Regional Primate Research Center, Davis, CA, where they were assayed for concentrations of estradiol, testosterone, and progesterone. Full details of the assay procedures can be found in Roney and Simmons (2013); intra- and inter-assay CVs were below 10% for each of the hormones.

Hormone data were used to re-estimate the day of ovulation based on the mid-cycle estradiol drop following the procedures described in Jasienska et al. (2004) and Lipson and Ellison (1996) (see Electronic Supplementary Material for the specific algorithm, available on the journal's website at www.ehbonline.org). Following Jasienska et al. (2004), we computed cycle mean estradiol as the mean estradiol concentration for the 18 cycle days centered on the estimated day of ovulation (days −8 to +9 relative to ovulation as day zero), whereas cycle mean progesterone was computed as the average concentration of progesterone in the final 14 days of the cycle; although Jasienska et al. did not measure testosterone, we computed cycle mean testosterone the same way as cycle mean estradiol (i.e. an average of the 18 cycle days centered on ovulation), given similarities in the secretion patterns of these hormones. Because identification of the day of ovulation was not possible in anovulatory cycles, we restricted data analyses to ovulatory cycles in order to ensure that similar cycle regions were being compared across women. Following Ellison, Lager, and Calef (1987), we defined as anovulatory any cycle that did not achieve a maximum progesterone value of at least 300 pmol/L.

Among the 41 women with both photo consent and hormone data, eight did not experience an ovulatory cycle based on the above criterion. Among the remaining women, 18 had hormone data for two ovulatory cycles, 10 women participated in both cycles but only one of the two was judged ovulatory, and five women participated in a single cycle that was also judged ovulatory; as such, the final sample included hormone data from two cycles for 18 women and from one cycle for 15 women. Subject mean hormone concentrations were computed from a single cycle mean (as defined above) for the 15 women with one ovulatory cycle and as the average of the two cycle
means for the 18 women with two ovulatory cycles (this procedure entailed that some women had more reliable mean hormone values than others due to the larger number of sample days; however, a set of mixed regression models that treated daily hormone concentrations as dependent variables and body dimensions and attractiveness ratings as higher level predictor variables – and thereby weighted women with more hormone data more heavily due to the more reliable estimates of their hormone concentrations – produced identical statistical conclusions to those presented below using subject mean hormone values). Data analyses tested associations between these subject mean hormone values and both body shape dimensions and mean body attractiveness ratings (see Section 2.2.3).

2.1.4. Stimulus photos

During the third laboratory session of the first cycle (typically within the luteal phase), each woman was photographed in standardized dress comprised of grey gym shorts and a blue tank top shirt. Photos were taken with a digital camera at a standard distance in a windowless room with artificial lighting. For each woman, photos were taken from front-facing, back-facing, and side-facing perspectives; these three photos were placed together onto a single stimulus array for each woman, with an opaque mask blocking the head area in each photo. An example stimulus array appears in Fig. 1.

2.2. Stimulus ratings

2.2.1. Rating participants

Raters were UCSB students who participated in exchange for partial course credit. The primary 39 raters were 23 men (Mean age ± SD = 19.17 ± 1.50 years) and 16 women (Mean age ± SD = 18.81 ± 1.22 years), but an additional batch of 19 raters comprised of 11 women (Mean age ± SD = 19.64 ± 0.67 years) and 8 men (Mean age ± SD = 19.38 ± 1.30 years) was recruited in order to obtain ratings for five stimulus photos that were previously omitted due to a clerical error. Participants provided written, informed consent for their participation, and all procedures were approved by the UCSB Institutional Review Board.

2.2.2. Rating procedures

Raters viewed the stimulus photos one at a time on a computer and were asked: “How physically attractive do you find this woman, relative to other women of the same age?” (1–7 scale). After rating all of the stimuli for general attractiveness, participants read the following: “We will now be focusing on the woman’s attractiveness as a LONG-TERM [SHORT-TERM] mate,” and ratings of either long- or short-term attractiveness followed on the same scale, with the order of these two rating dimensions counterbalanced across raters. The order of photo presentation was randomized within each rating dimension.

There was high between-rater agreement for each of the three rating dimensions (all ICCs > 0.90); thus, ratings were aggregated across raters to give each woman a mean rating for each rating dimension. The three rating dimensions also had high reliability (α = 0.99 for the mean ratings) and were therefore averaged to create a composite attractiveness variable that was used in subsequent data analyses. Male and female raters were in high agreement regarding their perceptions of the women’s attractiveness (ICC = 0.92 for the composite mean attractiveness ratings); in addition, for all of the correlations between composite attractiveness ratings and other variables presented in the Results, there were no significant differences between correlations computed using only male raters and those computed using only female raters (Fisher’s z-test; all p ≥ 0.40). The average attractiveness rating was just below the midpoint of scale (composite attractiveness mean = 3.92, S.D. = 1.05).

2.2.3. Data analyses

Pearson correlation, partial correlation, and multiple regression were employed to test relationships between women’s mean hormone concentrations (as defined in 2.1.3), body dimensions, and rated attractiveness. Following Jasienska et al. (2004), we also

Fig. 1. Sample stimulus photo.
constructed categorical body dimension groups (top vs. bottom quartile of WHR and breast size, as well as combinations of above and below average WHR with above and below average breast sizes) and t-tests and one-way ANOVAs were used to test whether such groups differed in mean hormone concentrations. Bias-corrected, nonparametric bootstrapping procedures (see Preacher & Hayes, 2008) were employed as tests of whether specific body dimensions statistically mediated relationships between hormone concentrations and attractiveness ratings. This analysis essentially tests whether a third variable is related to both the hormones and attractiveness ratings such that its addition to the model significantly diminishes the direct effect of hormones on attractiveness ratings; mediation is established if the 95% confidence interval for the unstandardized indirect effect does not include zero.

Measured variables more than three standard deviations from their respective means were excluded to avoid undue influence of outliers; one subject mean testosterone concentration and one BMI value were thus excluded (effect sizes for significant effects were generally larger with the outliers included). After outlier removal, all mean hormone and body dimension variables were approximately normally distributed by visual inspection and the Shapiro–Wilk test.

3. Results

3.1. Hormones

Excluding the one woman whose mean testosterone concentration was an outlier, the 32 women in the sample provided 798 saliva samples from the middle 18 days of their respective cycles out of 900 eligible cycle days (89% compliance rate). After selection of saliva samples from alternating days outside of the nine day window surrounding the initial estimate of mid-cycle, measured hormone concentrations were available for 565 and 577 of these days for estradiol and testosterone, respectively (insufficient remaining quantity of saliva for assay accounted for the difference given that testosterone was assayed first). With respect to the final 14 days of the cycle, 631 saliva samples were collected out of 700 eligible cycle days (90% compliance rate); progesterone assay values were obtained for 388 of these days.

3.2. Hormones and body dimensions

Table 1 presents correlations between mean hormone concentrations, body dimensions, and body attractiveness ratings. Contrary to previous findings (Jasienska et al., 2004), there were null zero-order correlations between body dimensions and hormones; neither WHR nor breast size was significantly associated with mean estradiol, progesterone, or testosterone. Null results persisted in one-way ANOVAs that tested for differences in mean hormone concentrations across the four body shape categories (large and small WHR crossed with large and small breast size) defined by Jasienska et al. (2004) (all ps > 0.46). Likewise, a series of t-tests found no differences in mean hormones when comparing women in the top and bottom quartiles of breast size and WHR, respectively (all ps > 0.27). Jasienska et al. (2004) also tested associations between body dimensions and hormone concentrations within narrower ranges of cycle days (e.g., estradiol concentrations on the day of ovulation); we again found only null results when we tested the same correlations presented in Table 1 within these narrower cycle day windows (see Electronic Supplementary Material, available on the journal’s website at www.ehbonline.org).

3.3. Predictors of body attractiveness ratings

3.3.1. Morphological predictors

Consistent with previous research, body attractiveness was significantly negatively associated with both WHR and BMI (see Table 1). A multiple regression with WHR and BMI entered together as predictors of body attractiveness ratings revealed a strong independent effect of BMI (β = −0.79, p < 0.001) and a null effect of WHR (β = −0.02, p = 0.87). BMI accounted for approximately 64% of the variance in women’s body attractiveness.

3.3.2. Hormonal predictors

As can be seen from Table 1, there were no significant zero-order correlations between subject mean hormone concentrations and body attractiveness ratings, although power limitations may have prevented detection of a small association between estradiol and attractiveness (r = 0.23). The large association between BMI and attractiveness may have obscured the influence of smaller predictor variables, however, and we therefore tested whether hormone concentrations were correlated with attractiveness ratings after controlling for the influence of BMI. Table 2 demonstrates that subject mean estradiol and testosterone both exhibited significant partial correlations with body attractiveness ratings after controlling for BMI. Progesterone was not a significant independent predictor of the body attractiveness residuals from BMI, and neither WHR nor breast size had residual variance from BMI that was significantly associated with any hormone. A multiple regression analysis testing the partial effects of BMI, testosterone, and estradiol revealed independent effects of BMI (β = −0.83, p < 0.001), mean estradiol (β = 0.20, p = 0.055), and mean testosterone (β =

### Table 1

<table>
<thead>
<tr>
<th>Attractiveness</th>
<th>Estradiol</th>
<th>Testosterone</th>
<th>Progesterone</th>
<th>WHR</th>
<th>Breast size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estradiol</td>
<td>0.23</td>
<td>0.14</td>
<td>−0.21</td>
<td>0</td>
<td>0.13</td>
</tr>
<tr>
<td>Testosterone</td>
<td>0.14</td>
<td>0.15</td>
<td>0.01</td>
<td>0</td>
<td>0.13</td>
</tr>
<tr>
<td>Progesterone</td>
<td>−0.45</td>
<td>0</td>
<td>−0.07</td>
<td>0.23</td>
<td>0.55</td>
</tr>
<tr>
<td>Breast size</td>
<td>−0.14</td>
<td>−0.01</td>
<td>0.24</td>
<td>−0.16</td>
<td>−0.14</td>
</tr>
<tr>
<td>BMI</td>
<td>−0.80</td>
<td>0</td>
<td>0.24</td>
<td>−0.16</td>
<td>−0.14</td>
</tr>
</tbody>
</table>

* Estradiol and testosterone concentrations represent subject means for 18 days surrounding ovulation; progesterone concentrations represent subject means for the last 14 days of the cycle.

* p < 0.05.

** p < 0.01.

*** p < 0.001.

### Table 2

<table>
<thead>
<tr>
<th>Hormone</th>
<th>Attractiveness</th>
<th>WHR</th>
<th>Breast size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estradiol</td>
<td>0.39*</td>
<td>0</td>
<td>−0.01</td>
</tr>
<tr>
<td>Testosterone</td>
<td>0.42*</td>
<td>−0.17</td>
<td>0.22</td>
</tr>
<tr>
<td>Progesterone</td>
<td>−0.06</td>
<td>0.14</td>
<td>−0.21</td>
</tr>
</tbody>
</table>

* p < 0.05.
or testosterone and women

None of these variables signi

water percentage were signi

hormone production across broad regions of the menstrual cycle.


discussion between women

concentrations and residual attractiveness ratings (all CIs included 0).

angled inward more sharply from their upper torsos, we also

women with higher residual attractiveness ratings had waists that

the mediators were tested separately or jointly (all CIs for the indirect

was a signi

attractiveness. Neither shoulder width nor waist width on its own

mediator of the relationship between estradiol and testosterone with ratings of their body attractiveness. As can be seen from

furthermore, these relationships held across broad regions of the menstrual cycle. These patterns thus provide some evidence that

perceivers’ attractiveness judgments may in fact hone in on cues of fecundity in young women’s bodies, although interpretive questions

are raised by the necessity of holding BMI constant in order to demonstrate robust relationships between hormones and attractiveness

(see discussion of this issue in Section 4.2 below).

Given previous research demonstrating higher estradiol and progesterone among women with lower WHR and larger breast size

(Jasienska et al., 2004), WHR and breast size were expected to mediate any relationship between body attractiveness and hormone

concentrations. However, there was no evidence for this in our study. Neither breast size nor WHR was associated with subject mean

concentrations of estradiol, progesterone, or testosterone; nor did they predict any hormone after controlling for variability in these

body shapes due to BMI.

Differences in study samples or measurement techniques may help account for inconsistencies between results of the current study and that of Jasienska et al. (2004). Whereas Jasienska et al. (2004) investigated over a hundred Polish women (mean age = 29 years), our sample was younger (mean age = 18 years), more ethnically heterogeneous, and much smaller. Menstrual cycles are notably less investigated over a hundred Polish women (mean age = 29 years), although it is important to note that the correlations between WHR and attractiveness in our sample were virtually identical to those reported elsewhere (compare Table 1 to findings in Cornelissen, Hancock, Kiviniemi, George, & Tovee, 2009; Faries & Bartholomew, 2012), which suggests that our measurements were consistent with others in this literature. We measured WHR from photographs in our sample vs. direct body measurements in Jasienska et al. (2004), although it is important to note that the correlations between WHR and attractiveness in our sample were virtually identical to those reported elsewhere (compare Table 1 to findings in Cornelissen, Hancock, Kiviniemi, George, & Tovee, 2009; Faries & Bartholomew, 2012), which suggests that our measurements were consistent with others in this literature. Although our sample size was less than ideal, low power is unlikely to explain the null relationships between hormones, WHR, and breast size given the absence of even trend-level effects in the relevant analyses (see Table 1). Furthermore, our sample size was sufficient to detect relationships between estradiol, testosterone, and residual variance in body attractiveness not accounted for by BMI.

The lack of relationships between hormone concentrations and either WHR or breast size suggested that at least one other physical cue was mediating the relationship between both estradiol and testosterone and the body attractiveness residuals from BMI.

Contrary to our predictions, there were no significant zero-order correlations between hormone concentrations and attractiveness ratings. However, after controlling for BMI, which was strongly negatively associated with attractiveness, women’s concentrations of estradiol and testosterone were significantly positively correlated with ratings of their body attractiveness. As can be seen from

Given the estradiol and testosterone measurements represented

subject means for 18 days surrounding ovulation, it is possible that their associations with body attractiveness could have been driven by effects in a narrow region of the cycle. To assess this, Fig. 2 plots hormone concentrations against day of the cycle (aligned on the estimated day of ovulation as day zero) with separate curves for women who were above and below the mean residual attractiveness rating after controlling for BMI. It can be seen that estradiol was consistently higher across the entire cycle among women who were rated more attractive than predicted by their BMI alone (Fig. 2A); this pattern was less consistent for testosterone, but still visible across broad regions of the cycle (Fig. 2B); whereas the curves were very similar across the entire cycle for progesterone (Fig. 2C).

The patterns depicted in Fig. 2 suggest that, after controlling for BMI, other observable cues in women’s bodies both contribute to attractiveness judgments and predict concentrations of estradiol and testosterone. In an exploratory attempt to identify such cues, we employed nonparametric bootstrapping methods to first test whether scale measures of women’s muscle mass, visceral fat, body fat, or water percentage were significant mediators between either estradiol or testosterone and women’s body attractiveness, controlling for BMI. None of these variables significantly mediated the relationship between either of the hormones and attractiveness ratings, whether the mediators were tested separately or jointly (all CIs for the indirect effects included zero). Based on the subjective impression that women with higher residual attractiveness ratings had waists that angled inward more sharply from their upper torsos, we also computed a ratio of shoulder width (measured from front-facing photos) to waist width and tested it as a mediator of the hormone effects. This shoulder-to-waist (SWR) ratio was in fact a significant mediator between residual variance in women’s body attractiveness from BMI and both their estradiol (Indirect Effect = 0.118, SE = 0.080, 95% CI = 0.016–0.417) and testosterone (Indirect Effect = 0.018, SE = 0.01, 95% CI = 0.004–0.051) concentrations, with larger SWR associated with both higher hormone concentrations and greater attractiveness. Neither shoulder width nor waist width on its own was a significant mediator of the relationship between hormone concentrations and residual attractiveness ratings (all CIs included 0).

4. Discussion

4.1. Hormones, body dimensions, and body attractiveness

The present research provided an initial, direct test of the possible relationship between women’s body attractiveness and their ovarian hormone production across broad regions of the menstrual cycle.

0.22, \( p = 0.04 \)); the two hormones jointly explained an additional 10% of the variance in body attractiveness beyond that explained by BMI alone (change in \( R^2 F (2, 27) = 5.24, p = 0.01 \)).
Exploratory analyses revealed the shoulder-to-waist ratio (SWR) as a statistical mediator of the effects of both estradiol and testosterone on attractiveness ratings. These results should be interpreted with caution, however, given both the number of potential mediators tested (see Section 3.3.2) and the fact that we had no way of testing whether observers actually used this ratio as a perceptual cue that contributed to their attractiveness judgments. SWR might correlate inversely with android fat depositions (i.e. fat in the abdomen and upper torso) since such fat will cause the waist to spread out toward the width of the shoulders and thus reduce this ratio (WHR may not capture quite the same variable given cases of wide waists but even wider hips); android fat deposits, in turn, have been shown to be strong negative predictors of body attractiveness ratings (e.g., Faries & Bartholomew, 2012; Rilling et al., 2009). Ideally, android fat would be measured more directly via tools such as dual-energy X-ray absorptiometry scans (see Faries & Bartholomew, 2012; Sowers, Beebe, McConnell, Randolph, & Jannausch, 2001), and future research that combined such measurements with hormone assays would allow for more precise tests of which body dimensions may account for relationships between endocrine variables and body attractiveness ratings.

4.2. BMI, hormone concentrations, and specialized preference mechanisms

Why was it necessary to control for BMI in order to see clear relationships between ovarian hormone concentrations and body attractiveness ratings? If specialized preference mechanisms track cues of fecundity as indexed by hormone concentrations, then one might expect positive zero-order associations between hormones and attractiveness without the need to control for other variables. We offer two conjectures regarding this issue.

First, BMI may predict other fitness-relevant traits aside from fecundity that are also relevant to attractiveness judgments. Higher BMI is strongly predictive of a wide array of health problems in industrialized countries (e.g., Calle, Rodriguez, Walker-Thurmond, & Thun, 2003; Gilmore, 1999; Manson et al., 1995; Willett et al., 1995). Although many of those health problems may not have been relevant to reproductive fitness in ancestral environments, higher BMI has also been associated with greater fluctuating asymmetry (Hume & Montomgery, 2001; Losken, Fishman, Denson, Moyer, & Carlson, 2005; Manning, 1995; Milne et al., 2003) and higher rates of inflammation (e.g., Festa et al., 2001; Panagiotakos, Pittavos, Yannakoulia, Chrysohoou, & Stefanadis, 2005; Trayhurn & Wood, 2005), suggesting that greater BMI may predict greater developmental instability and reduced immunocompetence, both of which likely entailed fitness costs to mates even independent of any effects on fecundity. These inverse associations of BMI with health and developmental stability – at least in industrialized nations – may lead cues of high BMI to become associated with poor health, thus partly explaining the negative effect of BMI on attractiveness. In addition, BMI is on average positively correlated with age in the United States (Brown, Kaye, & Folsom, 1992; Fryar, Gu, & Ogden, 2012; Laske & Gaulin, 2006), such that high BMI may become associated with declining reproductive value and thereby reduce attractiveness via that association even among young women (see Wells, 2010). These associations of BMI with health and age appear to be reversed under conditions of food shortage (e.g., Evans, Hoffman, Kalkhoff, & Kissebah, 1983; Sowers et al., 2001) and be associated with reduced fecundity (e.g., Okon, Laird, Tuckerman, & Li, 1998; Steinberger, Smith, Tcholakian, & Rodriguez-Rigau, 1979). Many of the negative effects of testosterone on reproductive functioning are associated with obesity (Clark et al., 1995; Kiddy et al., 1992; Pasquali, 1982; Turcato et al., 1997; Tworoger et al., 2006). These patterns suggest that higher BMI is likely to be associated with artificially inflated measures of salivary, free hormones relative to the total ovarian hormone production consistent with this, in a large study of premenopausal women, BMI was significantly inversely correlated with total estradiol but was uncorrelated with free estradiol (Tworoger et al., 2006). This in turn implies that when two women have the same free hormone concentrations but differ in BMI, the woman with lower BMI is likely to have greater ovarian hormone production since a greater fraction of her hormones will be bound to SHBG. Likewise, when two women have the same BMI but differ in free hormone concentrations, the woman with greater free hormone concentrations should have higher ovarian production since the effect of BMI on SHBG will be held constant. As such, if perceivers’ attractiveness judgments specifically track cues of ovarian hormone production, then BMI should negatively predict attractiveness when free hormones are held constant and free hormones should positively predict attractiveness when BMI is held constant, which is exactly the pattern produced by the regression models in Section 3.3.2. In short, controlling for BMI may increase the size of correlations between free hormone concentrations and attractiveness ratings by removing the variability in hormone concentrations that is associated with binding proteins and is thus potentially unrelated to fecundity. This idea could be tested more directly in future research that used blood samples in order to test associations between body attractiveness and both total and free hormone concentrations.

4.3. Independent effects of testosterone on attractiveness

The positive effect of testosterone on attractiveness after controlling for BMI was surprising given evidence that elevated testosterone in women may promote visceral fat deposition (e.g., Evans, Hoffman, Kalkhoff, & Kissebah, 1983; Sowers et al., 2001) and be associated with reduced fecundity (e.g., Okon, Laird, Tuckerman, & Li, 1998; Steinberger, Smith, Tcholakian, & Rodriguez-Rigau, 1979). Many of the negative effects of testosterone on reproductive functioning are associated with obesity (Clark et al., 1995; Kiddy et al., 1992; Pasquali, Casimirri, & Vicennati, 1997) and associated reductions in SHBG (see above), however, such that controlling for BMI may more uniquely capture follicle-derived sources of testosterone that could in principle be associated with higher fecundity. Testosterone acts a precursor to estradiol produced by the dominant follicle, for instance, and periovulatory peaks in estradiol are typically accompanied by concomitant peaks in testosterone (e.g., Abraham, 1974; Campbell & Ellison, 2003).
1992; Roney & Simmons, 2013) such that larger dominant follicles that produce higher estradiol in more fertile cycles may likewise produce higher testosterone. As such, the combination of estradiol and testosterone concentrations may better predict dominant follicle production within ovulatory cycles than does the concentration of either hormone alone, thus potentially explaining the independent effects of the two hormones on attractiveness ratings. This is speculation, of course, and the unexpected association of attractiveness with testosterone concentrations warrants replication before assigning much confidence to the robustness of this finding.

4.4. Conclusion

The present study is to our knowledge the first to demonstrate a link between women's body attractiveness and concentrations of ovarian hormones measured across broad regions of the menstrual cycle. Both estradiol and testosterone independently predicted body attractiveness ratings after controlling for the effects of BMI, which suggests that preference mechanisms may indeed track cues of attractiveness ratings after controlling for the effects of BMI, which suggests that preference mechanisms may indeed track cues of attractiveness signals different aspects of “quality” in women and men. Evolution and Human Behavior, 22, 93–112.


