

Original Article

Physical Cues of Ovulatory Status: A Failure to Replicate Enhanced Facial Attractiveness and Reduced Waist-to-Hip Ratio at High Fertility

April Bleske-Rechek, Psychology Department, University of Wisconsin-Eau Claire, Eau Claire, WI, USA.
Email: bleskeal@uwec.edu (Corresponding author).

Heather D. Harris, Psychology Department, University of Wisconsin-Eau Claire, Eau Claire, WI, USA.

Kelly Denkinger, Psychology Department, University of Wisconsin-Eau Claire, Eau Claire, WI, USA.

Rose Mary Webb, Psychology Department, Appalachian State University, Boone, NC, USA.

Leah Erickson, Psychology Department, University of Wisconsin-Eau Claire, Eau Claire, WI, USA.

Lyndsay A. Nelson, Psychology Department, Appalachian State University, Boone, NC, USA.

Abstract: We investigated women's facial attractiveness and body shape as a function of menstrual cycle phase, with the expectation from previous research that both would be enhanced during the high fertile phase. To control for the effects of women's daily behaviors on their appearance and waistline, we visited 37 normally cycling women twice in their dorm, where we photographed and measured them at low and high fertile days of their cycle immediately upon their waking. Seventy-four judges from a separate institution chose, for each woman, the picture they thought was more attractive. We analyzed a subset of 20 women who, by forward counting, had a High Fertility visit between Days 10-13 and a Low Fertility visit between Days 20-23; and we also analyzed a subsample of 17 women who, by reverse counting, had a High Fertility visit on the days leading to ovulation and a Low Fertility visit one week after ovulation. In neither set of analyses were women's waist-to-hip ratios lower nearer ovulation, and in neither set were women's high fertile pictures chosen at an above-chance rate by either male or female judges. We did not find evidence that facial attractiveness and waist-to-hip ratio are reliable physical cues of ovulatory status.

Keywords: Attractiveness, face, body shape, waist-to-hip ratio, menstrual cycle, ovulation

Introduction

Human females are typically portrayed as unique among primates in their lack of conspicuous visual signals of their ovulatory status. Sexual intercourse occurs throughout the human female's menstrual cycle (Brewis and Meyer, 2005), which may promote both social monogamy through constant sexual receptivity (Lovejoy, 1981) and paternal care because of paternity uncertainty (Hrdy, 1981). A variety of recent studies, however, indicate that although women's ovulatory status is not conspicuous, it is not entirely hidden (Thornhill and Gangestad, 2008). For example, multiple studies have shown that women's sexual desires, mate preferences, and overt behaviors vary across the cycle. When fertile (as compared to not fertile), women report feeling more sexually desirable (Haselton and Gangestad, 2006), respond more to sexual stimuli (Laeng and Falkenberg, 2007) and the sight of attractive male features (Gangestad, Thornhill, and Garver-Apgar, 2010), report a heightened desire to flirt with men (Haselton and Gangestad, 2006), are more responsive to romantic overtures (Gueguen, 2009), prefer more revealing clothing (Durante, Li, and Haselton, 2008), take more care in their appearance (Haselton, Mortezaie, Pillsworth, Bleske-Rechek, and Frederick, 2007), show systematic preferences for men who display good-genes indicators such as a masculine face, dominance, and symmetry (Gangestad, Simpson, Cousins, Garver-Apgar, and Christensen, 2004; Gangestad and Thornhill, 1998; Havlicek, Roberts, and Flegr, 2005; Johnston, Hagel, Franklin, Fink, and Grammer, 2001; Jones et al., 2008; Penton-Voak and Perrett, 2000), and engage in systematic avoidance of their fathers, with whom sexual activity can have negative genetic consequences (Lieberman, Pillsworth, and Haselton, 2011). Women also report more love and attention from their romantic partners when they are fertile (Gangestad, Thornhill, and Garver, 2002; Haselton and Gangestad, 2006), and erotic dancers earn more tips in the fertile phase relative to the luteal phase (Miller, Tybur, and Jordan, 2007). These findings on variation in men's behavior toward women bring to the forefront the question of whether women *behave* differently when fertile and thus evoke responses from men around them, or whether women actually are *physically* different when fertile, apart from any subtle behavioral or dress displays they may engage in.

Some research suggests that women *are* physically different when fertile, and that others are sensitive to those differences. For example, men rate the odor of fertile women as more attractive than that of non-fertile women, a finding that has been replicated on multiple occasions (Gildersleeve, Larson, Pillsworth, and Haselton, 2010; Havlicek, Dvorakova, Bartos, and Flegr, 2006; Kuukasjarvi et al., 2004; Singh and Bronstad, 2001). Further, men's testosterone levels decrease after exposure to the scent of non-fertile women but stay constant after exposure to the scent of fertile women (Miller and Maner, 2010), and men's mating-related behaviors and cognitions are positively enhanced by exposure to the scent of fertile women (Miller and Maner, 2011). Because odor studies utilize ratings of plain gauze pads or T-shirts worn by women overnight under strict guidelines of no co-sleeping or use of perfume, their findings provide strong evidence for the proposal that women emit physical cues of their ovulatory status.

Women appear to experience other physical changes over the cycle that might render them more attractive. For example, in a laboratory study designed to minimize

women's opportunities for behavioral displays by having them complete a mundane task of counting from 1 to 10, women's vocal samples taken at high fertility were rated as more attractive than their vocal samples taken at low fertility (Pipitone and Gallup, 2008), perhaps because women's voices increase in pitch as they near ovulation (Bryant and Haselton, 2009). A study of women's breast asymmetry over the menstrual cycle found lower levels of asymmetry near ovulation (Manning, Scutt, Whitehouse, Leinster, and Walton, 1996). In fact, two studies have shown that other soft tissue traits, including ears, wrists, and digit length, become less asymmetrical approaching the day of ovulation (Manning et al., 1996; Scutt and Manning, 1996).

In the current research, we focus on two physical cues of fertility status that, to our knowledge, have each been reported just once in the literature: lower waist-to-hip ratio (WHR) near ovulation (Kirchengast and Gartner, 2002), and enhanced facial attractiveness at high fertility (Roberts et al., 2004).

Waist-to-Hip Ratio

In a study designed to test the hypothesis that WHR decreases near ovulation as an indicator of reproductive status, Kirchengast and Gartner (2002) assessed change in basal body temperature and WHRs for 32 regularly cycling women (24 not on contraceptives and eight contraceptive-using controls) over the course of one menstrual cycle. In the study, women took their own basal body temperature and measured their own waist and hips each morning over one complete cycle. In a subgroup of 15 women with cycle lengths between 27 and 30 days, WHRs were significantly lower, by a difference of approximately .015, around the day at which their basal body temperature increased (signaling ovulation). A group of four women with cycle lengths between 23 and 26 days showed no discernible pattern of change in WHRs over the cycle, and a group of five women with cycle lengths between 31 and 34 days showed a significant increase in WHR during the latter three weeks of their cycle compared to the first 10 days.

Facial Attractiveness

Roberts and colleagues (2004) documented that naïve judges who compared two pictures of the same woman, one taken during the luteal phase and another near ovulation, chose the high fertile picture more often than expected by chance (50-50). The research involved two separate samples of women and judges, one from Prague and one from Newcastle. Women came into the lab for a high fertility photo between Days 8 and 14 and a low fertility photo between Days 17 and 25. (Data on cycle day appear to have been obtained via women's self-report.) At the lab, women removed their makeup and were photographed from a set distance under constant lighting. In one set of photos, the "masked" set, women's hair and ears were masked from the photo so that only an oval face was revealed; in the other, "unmasked" set, women's hair and ears were retained in the photos. Male and female judges from Prague selected, for each woman from Prague, the masked picture they found more attractive and the unmasked picture they found more attractive. Judges from Newcastle completed the same tasks for the women from Newcastle. Overall, the researchers found a statistically significant effect in six of the eight comparisons involved. In Prague, both male and female judges preferred the photos, both

masked and unmasked, that had been taken in the fertile window; across the four conditions, judges chose the high-fertile picture approximately 56% of the time. The results for Newcastle were less robust, although in the same direction. In Newcastle, male but not female judges chose the masked high fertility images at above-chance levels (53%), and female but not male judges chose the unmasked high fertility images at above-chance levels (55%). Across conditions, judges in Newcastle chose the high-fertility picture approximately 53% of the time.

The Current Study

The objective of the current research was to provide an operational replication of the finding that women's faces are more attractive at high fertility (Roberts et al., 2004) and of the finding that women's WHRs are lower at high fertility (Kirchengast and Gartner, 2002). We pursued facial attractiveness because women's faces contribute substantially to their overall attractiveness (Peters, Rhodes, and Simmons, 2007); and we pursued WHR because previous research has implicated it as a prominent locus of men's perceptions of female bodily attractiveness (Singh, 1993; Streeter and McBurney, 2003). We reasoned that replicating a lower WHR at high fertility would encourage subsequent research comparing men's attractiveness ratings of the same woman's body shape at low and high fertility.

As in previous research (e.g., Roberts et al., 2004), we assessed women at low and high fertile phases of their cycle. We extended prior protocols to control for the possibility that women's behaviors over the course of the day, such as exercising, eating (Fessler, 2003), or flirting, may exert subtle influences on women's facial appearance or waistline. That is, we photographed and measured women upon waking. Women slept in a plain black tank top we provided and immediately upon waking, prior to washing or even brushing their teeth, went down to the basement floor of their residence hall to be photographed and measured.

Materials and Methods

Participants

Participants came from an initial set of 52 women who completed all phases of a study on women's face and body across the menstrual cycle. Women were residents of an all-female residence hall at a regional public university in the United States; the study was advertised to all women in the hall, and hence we retained those who were currently on hormonal contraceptives despite our presumption that we might not obtain enough of them to include in analyses. The women were representative of the university population; 90% were Caucasian and 10% were Asian, Latino, or Native American. Women ranged from 18 to 22 years of age (Mean = 19.04), all but one were heterosexual (one was lesbian), 50% were involved in a serious romantic relationship, and 25% were currently sexually active. Non-contraceptive users did not differ from contraceptive users in age, sexual orientation, relationship status, or sexual activity status, all $ps > .11$.

Of the 52 women who participated, 37 were normally cycling (not on a hormonal contraceptive) and 15 were on a hormonal contraceptive. The study included an initial information-and-questionnaire session and two photo-measurement visits. As described

below, however, only some of the original 52 completers were included in data analyses. Following previous researchers' selection criteria, women who attended the initial information session were not retained for subsequent visits if they used tobacco, were pregnant, were currently taking a prescription medication or sleep aid, or reported a highly irregular cycle.

Procedure and Measures

Intake session. Women were first recruited, with paper advertisements and food incentives, to an initial information and data collection session. At this session women were told that the researchers were interested in how women's body and face change, if at all, over the menstrual cycle. The researchers would be housed in a small room in the dorm basement, where they would photograph and measure each woman first thing in the morning on two occasions over the next six to eight weeks. After giving their consent to participate, women reported on their demographics (above) and on their menstrual cycle length, regularity, and bleeding duration. Participants then used a calendar to mark the first day of their most recent period and the expected first day of their next period, and they reported their daily wake-up time. Finally, they provided their room number, cell phone number, and email address to aid researchers in scheduling visits and sending reminders. At this session, each participant received a pre-washed tight black tank top, which they were instructed to wear to bed the night before each visit with the researchers, and to the visit upon wake-up. Women were also given a copy of rules to follow the night before each visit: attempt to get eight hours of sleep, no alcohol after 8:00 pm, no sexual intercourse, and no co-sleeping. Women were asked to go to the research room immediately upon waking; they were instructed explicitly to avoid washing their face, showering, brushing their teeth, or grooming in any way. Women received an email reminder the day before each visit in which they were reminded of these rules. At the initial information session and throughout the study, researchers made no reference to low or high fertility, and it became clear through various participant emails and questions throughout the study that many defaulted to an assumption that the focus was on menstruation. They were debriefed fully in writing upon study completion.

After the initial information session, the researchers established where each woman was in her cycle and scheduled the relevant visits using a forward day cycle count. Depending on their current cycle position, women completed either the high fertility visit followed by the low fertility visit or the low fertility visit followed by the high fertility visit. Women who had just finished menstruating were scheduled immediately for their first visit (high fertility); women who were currently menstruating were scheduled for their high visit in the next week or so; and so on. Each woman was asked to send an email to the researchers upon beginning a period (menstruation), so visits could be revised as needed. Researchers stayed in email/text contact with each woman for up to 60 days to verify the onset of their next menstruation.

Photo and measurement visits. Each time women reported to the research room, they completed three tasks. First, they pulled their hair back in a head band or hair binder and sat in a chair against a white wall to be photographed from the neck up; these photos were subsequently masked digitally by ovals that allowed display of the face only (see Fig.

1). Women were instructed to look straight at the camera without smiling. Pictures were taken at a fixed distance with a Canon EOS digital camera that was stationed for the duration of the study on a tripod.

Second, participants stood for measurements of their waist and hips. We improved upon Kirchengast and Gartner's (2002) protocol on waist and hip measurements by having two researchers perform independent measurements, thus providing objective measurements and inter-rater reliability. Inter-rater reliability coefficients for measurements of hips and waist for low and high fertile visits were all $\geq .988$. To evaluate the validity of our measurement protocol, we also measured breast circumference above the breast, at mid-breast, and below the breast because previous research has consistently implicated breast volume (via water content) and breast density as lower during the follicular (high fertility) phase of the cycle relative to luteal and secretory (low fertility) phases (Buist, Aiello, Miglioretti, and White, 2006; Fowler, Casey, Cameron, Foster, and Knight, 1990; Ramakrishnan, Khan, and Badve, 2002).

Third, participants completed a brief questionnaire. They reported their behavior on the previous day and night, including the hours of sleep obtained, co-sleeping, sexual intercourse, and alcohol use. All women reported abiding by the rules set forth in the information session, although on average women did not obtain eight hours of sleep per night: Mean hours of sleep was 7.0 hours for low fertile and 7.1 hours for high fertile visits. Women also reported whether or not they had been tanning or used self-tanning lotions (and if so, when). Finally, women reported their current mood. In response to the question, "How do you feel RIGHT NOW?", women rated themselves on a five point scale (*not at all* to *extremely*) on the following moods: happy, tired, confident, upset, scared, sexy, proud, attractive, irritable, alert, sad, flirtatious, ugly.

Upon completing the questionnaire and being photographed and measured, participants were thanked with a breakfast muffin and dismissed.

Photo judgments. Participants' low and high fertile photos were compiled into one PowerPoint slide show, with the high fertile photo for each woman placed randomly on the left or right side of the slide (see Fig. 1 for an example). Women were entered into the slide show in random order. The slide show was sent to researchers at another public university across the country. There, 34 male and 40 female college students viewed the slide show as part of a lower level psychology course research participation requirement. Student judges attended one of 11 different sessions offered, with number of judges in each session varying from 1 to 20. Judges were instructed to not converse with one another during the session. For each woman in the slide show, judges recorded which picture they viewed as most attractive.

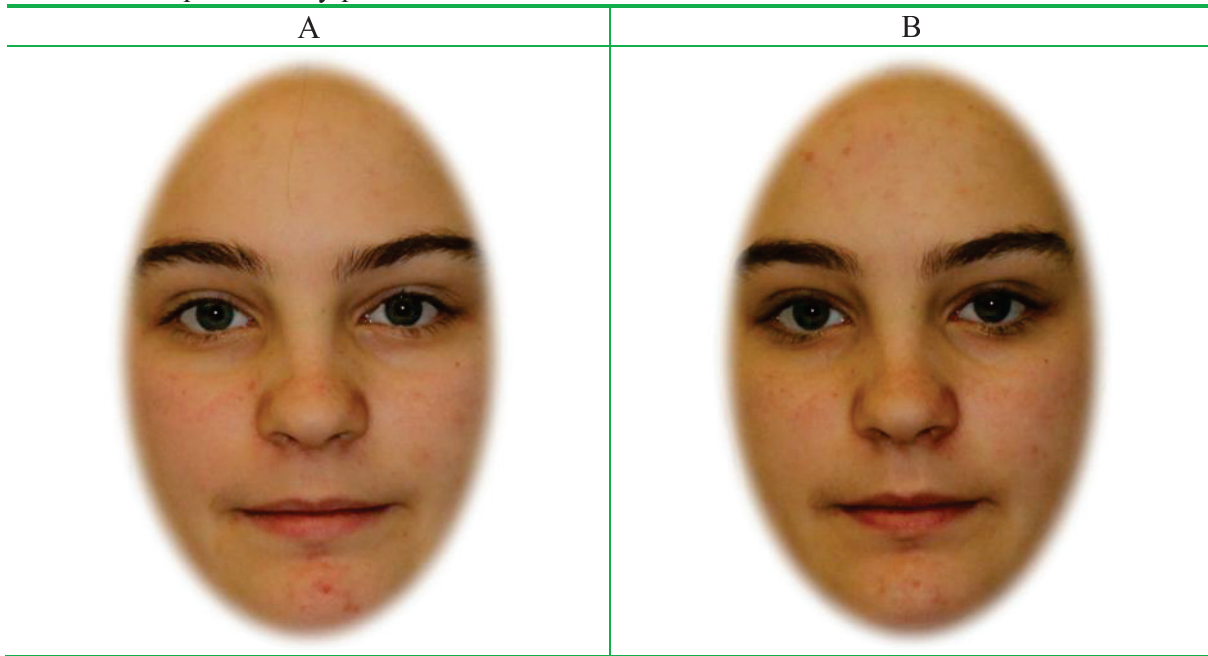
Results

Sample Selection

For the 52 women who completed the study, high fertile phase visits occurred between Days 8 and 13 ($M = 10.75$, $SD = 1.79$), and low fertile phase visits occurred between Days 17 and 25 ($M = 21.06$, $SD = 1.74$), with one woman on Day 17 and one on Day 25. For some women, the selection of visit day was not ideal because of two types of

scheduling challenges we encountered. First, to maximize compliance with our behavioral rules, we did not schedule photo-measurement visits on weekends or Friday mornings (students commonly go out on Thursday nights and thus potentially would have consumed alcohol or engaged in co-sleeping). Second, we had to reschedule a number of visits when women failed to show up for their planned visit.

Figure 1. Sample slide of the faces shown to judges. For each woman, judges chose A or B to select the picture they perceived as more attractive



We retained two partially overlapping subsets of women for data analysis. First, using a forward cycle day count with Day 1 as day of menstrual onset, we retained women whom we visited on Day 10, 11, 12, or 13 for their high fertility session and on Day 20, 21, 22, or 23 for their low fertility visit. This forward cycle day subset included 20 normally cycling women not on contraceptives (and another six women on hormonal contraceptives whom we did not analyze due to their limited number). Of the 20 women, 15 completed the high fertile visit first and five completed the low fertile visit first. Second, using a reverse cycle day count by which we estimated ovulation as 14 days prior to subsequent menstrual onset, we retained women whom we visited within five days prior to their predicted day of ovulation up through their predicted day of ovulation for their high fertility session, and at least one week after the estimated day of ovulation for their low fertility session. This reverse cycle day subset included 17 normally cycling women not on contraceptives (and another eight women on hormonal contraceptives whom we did not analyze due to their limited number). Of the 17 women, 11 completed the high fertile visit first and six women completed the low fertile visit first. A total of 23 normally cycling women were included in one of these two subsets, with 14 women meeting the criteria for both forward count and reverse count subsets. Cycle length for the 23 women included in analyses, from onset of one menstruation to the next by our records, was between 20 and 38 days in length (Median Evolutionary Psychology – ISSN 1474-7049 – Volume 9(3). 2011. -342-

= 28); 16 women had a cycle length of 27 to 30 days. Three of the women reported tanning: one woman had tanned within three days of each photo session and the other two women tanned more than a week prior to each visit. Because tanning reports were consistent within each woman, we retained these women for analysis. No participant completed photo-measurement sessions during menses.

Participant Mood

Table 1 displays descriptive statistics for women’s mood reports at low and high fertility, for the forward count and reverse count subsets. In both subsets, women reported significantly lower levels of felt ugliness at high fertility than at low (Forward count: two-tailed paired samples $t_{19} = 2.67, p = .015, d = .60$; Reverse count: $t_{16} = 2.40, p = .029, d = .60$). No other cycle phase differences were statistically significant.

Table 1. Means (*M*) and standard deviations (*SD*) for self-reported mood at low and high fertility from normally cycling women

	Forward Count (<i>n</i> =20)		Reverse Count (<i>n</i> =17)	
	Low <i>M</i> (<i>SD</i>)	High <i>M</i> (<i>SD</i>)	Low <i>M</i> (<i>SD</i>)	High <i>M</i> (<i>SD</i>)
Happy	2.95 (0.83)	2.95 (0.61)	2.82 (0.73)	2.88 (0.60)
Tired	3.20 (1.06)	2.85 (0.88)	3.12 (1.05)	2.88 (0.99)
Confident	2.90 (0.72)	2.65 (0.75)	2.76 (0.75)	2.65 (0.79)
Upset	1.05 (0.22)	1.05 (0.22)	1.00 (0.00)	1.06 (.024)
Scared	1.15 (0.37)	1.10 (0.31)	1.12 (0.33)	1.12 (0.33)
Sexy	1.95 (0.76)	2.00 (0.86)	1.94 (0.75)	2.06 (0.83)
Proud	2.20 (0.83)	2.55 (0.83)	2.06 (0.83)	2.47 (0.94)
Attractive	2.30 (0.98)	2.45 (0.95)	2.12 (1.05)	2.47 (0.94)
Irritable	1.75 (1.02)	1.55 (0.83)	1.71 (0.99)	1.53 (0.87)
Alert	2.30 (0.80)	2.80 (0.95)	2.29 (0.85)	2.76 (0.97)
Sad	1.50 (1.00)	1.15 (0.49)	1.53 (1.07)	1.18 (0.53)
Flirtatious	1.70 (0.73)	1.75 (1.02)	1.71 (0.69)	1.71 (1.05)
Ugly*	1.70 (0.80)	1.35 (0.49)	1.65 (0.86)	1.29 (0.47)

*Low-high comparisons significant for both forward count and reverse count subsets, $ps < .05$. All mean comparisons were two-tailed. Moods were reported on five-point scales (1 to 5). Values for the forward cycle day and reverse cycle day subsets are not independent because 14 women met criteria for both subsets.

Participant Body Measurements

Breast measurements. Table 2 displays descriptive statistics for women’s body measurements at low and high fertility, for the forward count and reverse count subsets. Data from the forward count met our expectation that women’s breast circumference across the middle would be smaller at high fertility, $t_{19} = 2.22, p = .039, d = .50$. A similar pattern was revealed in the reverse count subset, but it was not statistically significant, $t_{16} = 1.43, p = .17, d = .36$. Measurements taken above the breasts and just below the breasts did not differ by cycle phase for either subset, $ps > .33$.

Waist and hip measurements. Women’s waist measurements did not differ by cycle

phase. For both the forward count subset ($t_{19} = .17, p = .864, d = .04$) and the reverse count subset ($t_{16} = .55, p = .59, d = .14$), women’s waist circumference at high fertility was not systematically different from waist circumference at low fertility. Within the forward count subset, women’s hip measurements differed by cycle phase, with hip measurements actually *lower* at high fertility than at low fertility, $t_{19} = 2.72, p = .014, d = .61$). This pattern was not replicated in the reverse count subset, however, $t_{16} = 1.17, p = .259, d = .29$; and in neither subset did women’s WHR differ significantly by cycle phase: Forward count subset $t_{19} = -1.69, p = .11, d = -.38$ (the trend was in the direction of being reduced at *low* fertility); Reverse count subset $t_{16} = -.53, p = .60, d = -.13$. When we isolated analyses to the 16 women with cycle lengths between 27 and 30 days, WHRs still did not differ by cycle phase, $t_{15} = -1.17, p = .26, d = -.29$. Thus, we did not observe the same pattern reported in Kirchengast and Gartner (2002).

Of the 20 women included in the forward count subset, only five had a lower WHR at high fertility. Actuarial conception risk (Wilcox, Dunson, Weinberg, Trussell, and Baird, 2001) on the cycle day of their high fertile visit was not associated with change in WHR from low to high fertility, $r = .15, N = 20, p = .53$. Of the 17 women included in the reverse count subset, only eight had lower waist-to-hip ratios at high fertility; days to ovulation at women’s high fertility visit did not significantly predict the difference in waist-to-hip ratio between visits, $r = .004, N = 17, p = .99$. We confirmed these non-significant associations by scatter plot (see Fig. 2).

Table 2. Means (*M*) and standard deviations (*SD*) for body shape measurements at low and high fertility from normally cycling women

	Forward Count (<i>n</i> =20)		Reverse Count (<i>n</i> =17)	
	Low <i>M</i> (<i>SD</i>)	High <i>M</i> (<i>SD</i>)	Low <i>M</i> (<i>SD</i>)	High <i>M</i> (<i>SD</i>)
Breast (above)	85.06 (5.51)	84.80 (5.64)	85.65 (5.36)	85.62 (5.72)
Breast (mid)*	87.66 (7.15)	86.42 (6.79)	88.80 (6.81)	87.81 (6.65)
Breast (below)	75.53 (5.03)	75.15 (5.02)	76.38 (5.05)	76.23 (5.16)
Waist	78.07 (7.29)	77.98 (6.50)	78.99 (6.55)	78.64 (5.36)
Hips*	87.06 (5.99)	85.85 (5.10)	87.07 (5.92)	86.28 (4.85)
Waist-to-hip ratio	.90 (.04)	.91 (.04)	.91 (.04)	.91 (.04)

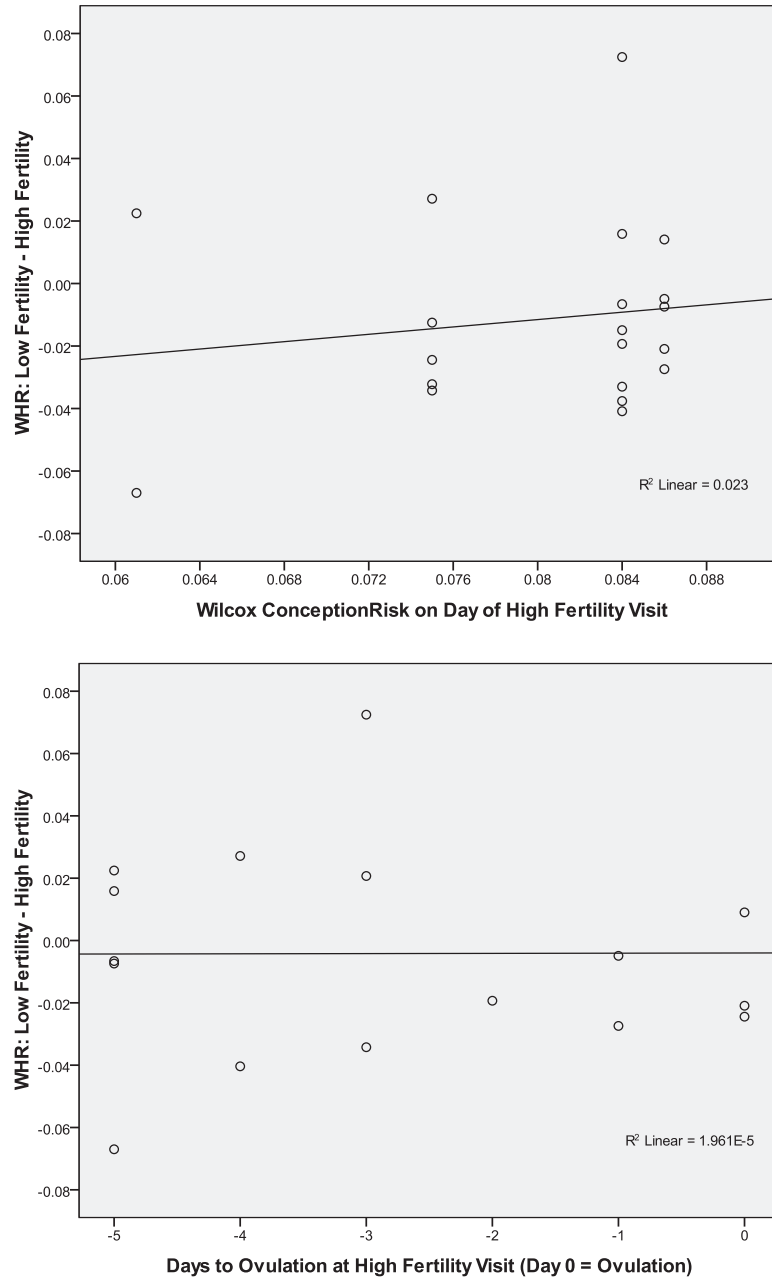
*For the Forward Count subset only, breast circumference across the middle was larger at low fertility than at high, $p = .039$, and hip circumference was larger at low fertility than at high, $p = .014$. All mean comparisons were two-tailed. Values (with exception of waist-to-hip ratio) are circumferences in centimeters. Values for the forward cycle day and reverse cycle day subsets are not independent because 14 women met criteria for both subsets.

Judges’ Photo Choices

Each judge viewed two facial shots of each woman and selected the one they thought was more attractive. Judges did not choose high fertility photos significantly more than half the time. For the forward count subset of women, judges selected the high fertile phase photo as more attractive between 20% and 75% of the time ($M = 45.61\%, SD = 11.04$; two-sided 95% CI = 35% to 57%). Male judges chose the high fertile phase photo 47.65% of the time, and female judges chose the high fertile phase photo 43.88% of the time; the sexes did not differ in their rate of high fertile picture choice, $t_{72} = 1.48, p = .144$,

$d = .35$. For the reverse count subset of women, judges selected the high fertile phase photo as more attractive between 24% and 65% of the time, ($M = 43.64\%$, $SD = 18.07$; two-sided 95% CI = 33% to 55%). Again, male judges (45.16%) and female judges (42.35%) did not differ in the frequency with which they chose the high fertile phase photo, $t_{72} = 1.10$, $p = .275$, $d = .26$.

Figure 2. Wilcox conception risk was not associated with a decrease in WHR from low fertility to high fertility for the forward count subset (*top*). Likewise, days-to-ovulation was not related to change in WHR for the reverse count subset (*bottom*)



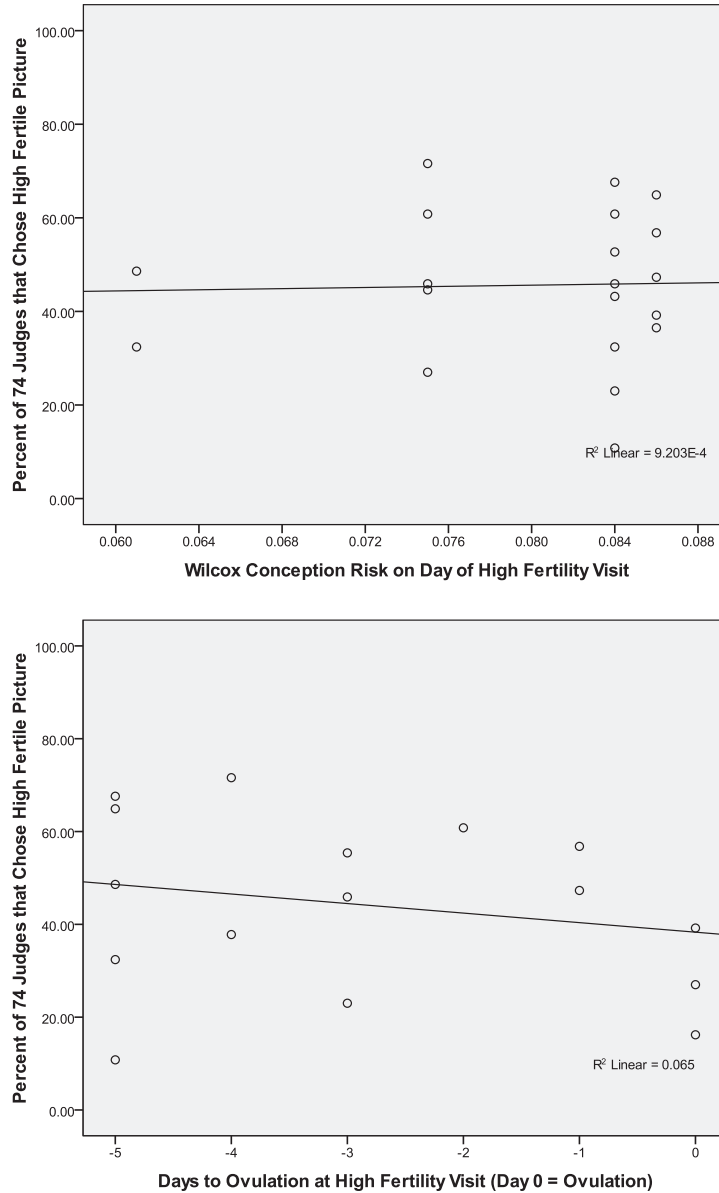
Women varied widely in the percent of judges who selected their high fertile phase photo. For both the forward and reverse count subsets, the percent ranged from 11% to 72% of judges: Forward count $M = 45.61\%$, $SD = 15.78$; Reverse count $M = 43.64\%$, $SD = 18.07$. Contrary to expectation, women's conception risk (Wilcox et al., 2001) on the day of their high fertile visit was not significantly related to the frequency with which their high fertile photo was chosen, $r = .03$, $N = 20$, $p = .899$. Moreover, women's proximity to estimated day of ovulation was not significantly related to the frequency with which judges chose the high fertile photo, $r = -.26$, $N = 17$, $p = .32$. These non-significant associations are displayed in Fig. 3. Day of low fertility visit, as measured by number of days prior to menstruation for the reverse count subset, also was not associated with judges' photo choice, $r = .35$, $N = 17$, $p = .18$.

Finally, we pursued associations between the degree to which cycling women felt more positive about themselves at their high (compared to low) fertile visit and the percent of judges who preferred their high fertile picture. These analyses were limited because the majority of women did not differ in their self-reported attractiveness, sexiness, or ugliness at the two visits. Regardless, differences in mood were not significantly associated with judges' photo choice in either the forward count subset ($N = 20$) or reverse count subset ($N = 17$), all r s $< .30$, p s $> .19$, with one exception: To the extent that participants reported more irritation at their low visit than at their high fertility visit, more judges chose their high fertility picture, $r = .44$, $N = 20$, $p = .05$.

One of the reviewers questioned whether the ratio of high-to-low fertility photos on the left versus right side of the slides was a clean 50:50. Research suggests that people may prefer items on their right (Nisbett and Wilson, 1977), and that many people associate positive concepts with the right rather than the left (Casasanto, 2009). Therefore, if the majority of high fertility photos happened to be on the left side, raters may have had position effects operating against any effect of high fertility. In fact, when we investigated our original slide show, we found that for the forward counting subset, fully 80% of the high fertility photos were on the left; and for the reverse counting subset, 71% of the high fertility photos were on the left. To address the possibility of position effects, we created three new slide shows of the 23 women in the reverse and forward counting subsets, with a near 50-50 split for each show (48:52 and 52:48). Then, we showed one of the three slide shows to each of 54 new raters (15 men, 37 women, 1 gender neutral, 1 unstated). The pattern of findings was essentially identical to that for our first set of raters who had viewed the original slide show. Specifically, the raters who judged the new slide shows did not choose high fertility photos significantly more than half the time. For the forward count subset of women, judges selected the high fertile phase photo as more attractive between 20% and 70% of the time ($M = 47.13\%$, $SD = 10.22$). Male judges chose the high fertile phase photo 49.00% of the time, and female judges chose the high fertile phase photo 46.49% of the time; the sexes did not differ in their rate of high fertile picture choice, $t_{50} = .79$, $p = .435$, $d = .22$. For the reverse count subset of women, judges selected the high fertile phase photo as more attractive between 18% and 71% of the time ($M = 44.89\%$, $SD = 12.22$). Male judges (47.06%) and female judges (43.40%) did not differ in the frequency with which they chose the high fertile phase photo, $t_{50} = .98$, $p = .334$, $d = .28$. Finally, although women varied widely in the percent of judges who selected their high fertile phase

photo (Range = 13% to 76%; forward count $M = 47.13\%$, $SD = 19.48$; reverse count $M = 44.89\%$, $SD = 20.87$), women's conception risk on the day of their high fertile visit was not related to the frequency with which their high fertile photo was chosen, $r = .07$, $N = 20$, $p = .773$, and women's proximity to estimated day of ovulation was not related to the frequency with which judges chose the high fertile photo, $r = -.25$, $N = 17$, $p = .325$.

Figure 3. Wilcox conception risk was not associated with judge preference for high fertile photo (*top*). Likewise, days-to-ovulation was not related to judge preference for the reverse count subset (*bottom*)



Discussion

Ovulation in humans is not perfectly concealed. When fertile, women are more receptive to romantic overtures (Gueguen, 2009), engage in more self-ornamentation (Haselton et al., 2007), and wear more revealing clothing (Durante et al., 2008). In addition, women experience increased feelings of flirtatiousness and sexual attraction to men other than their primary partner (Gangestad et al., 2002; Haselton and Gangestad, 2006), as well as changes in odor (Gildersleeve et al., 2010; Havlicek et al., 2006; Kuukasjarvi et al., 2004; Singh and Bronstad, 2001), voice (Bryant and Haselton, 2009; Pipitone and Gallup, 2008), and symmetry (Manning et al., 1996; Scutt and Manning, 1996). Together, this suite of effects might explain why female lap dancers earn larger tips during the fertile days of their cycle (Miller et al., 2007) and why at high fertility women report higher levels of mate guarding behaviors from their partners (Gangestad et al., 2002; Haselton and Gangestad, 2006). In short, there seems to be substantial evidence now that some female fertility cues are not completely suppressed and that males may detect those leaked cues.

We designed the current study to replicate two other potential cues of women's ovulatory status: lower waist-to-hip ratios at high fertility and enhanced facial attractiveness at high fertility. We failed to replicate either effect. Normally cycling women's waist-to-hip ratios did not differ by cycle phase, and women's high fertile pictures were not chosen by judges more often than expected by chance alone. When we pursued differences among women – analyses not pursued in previous research – judges' likelihood of choosing women's high fertile photo was not associated with women's conception risk or likelihood of being in the fertile window when the high fertile photo was taken.

Waist-to-Hip-Ratio

To our knowledge, only one other team of researchers has investigated waist-to-hip ratios over the menstrual cycle (Kirchengast and Gartner, 2002). We interpret their findings cautiously for three reasons. First, women took their own basal body temperatures, and thus may have been aware of their ovulatory status. Second, the women measured their own waist-to-hip ratios, which may have introduced bias into the process. Although participants were instructed to measure themselves without clothes each morning, there would have been no guarantees with self-measurements that that was so. Third, the effect documented in their study was not robust - differences in waist-to-hip ratio over the cycle occurred only in a group of 15 women with regular-length cycles. When we isolated our analyses to women with regular-length cycles, with a sample size slightly larger than that of Kirchengast and Gartner, we were unable to replicate the waist-to-hip ratio effect.

In the current study, two researchers independently measured women's waist and hips (and bust), and inter-rater reliability coefficients were near perfect. In partial validation of our technique, we documented in the forward count subset a finding expected from the medical literature (Buist, Aiello, Miglioretti, and White, 2006; Fowler, Casey, Cameron, Foster, and Knight, 1990; Ramakrishnan, Khan, and Badve, 2002) -- a lower breast circumference during the follicular phase. However, our sample itself differed from

Kirchengast and Gartner's in an important respect: The mean WHR in their sample was .74, whereas ours was .90 – far less desirable overall in terms of health and fat distribution, although not atypical. Kirchengast and Gartner's sample was comprised of university women, as was ours, but in Austria rather than the United States. Perhaps women's eating and exercise habits in the United States disrupt changes in fat distribution over the menstrual cycle. Our findings implicate a need for more research designed to determine if women's body shape changes systematically over the cycle, and how individual health and initial body shape might moderate those changes.

Facial Attractiveness

Roberts et al. (2004) are frequently cited for documenting evidence that women's facial attractiveness increases at ovulation. Using a 2 (Prague sample $N=25$ /Newcastle sample $N=23$) x 2 (masked vs. unmasked images) x 2 (male/female raters) mixed design, they found a significant effect of fertility for six of eight comparisons. The effect was small but intriguing because an increase in facial attractiveness at high fertility might be detected by male partners as well as attract attention from other desirable men when women are most receptive to it.

In our study, we attempted to acquire a larger sample of women to boost our power to detect the effect of enhanced facial attractiveness near ovulation; and we originally obtained data from 37 normally cycling women. However, we retained for analyses only 23 of those women across two partially overlapping subsets of 20 and 17 women each. Although we retained for analyses a smaller sample than anticipated and a smaller sample than those obtained by Roberts et al., the women we did retain were those who were most likely to have been in appropriate points in their cycle for detecting the effects we expected. We also attempted to isolate facial cues associated with ovulation that might translate to a mateship: Just as men in sexual relationships see their female partners first thing in the morning, we photographed women first thing in the morning. By going to women's dorms immediately upon their waking instead of having them come to us in a lab, we minimized their opportunities to engage in potentially confounding behaviors, such as exercising or primping or socializing, prior to taking photos and measurements. An important advantage of our analysis was that we looked not only at judges' photo choices, but also differences among women as predictors of judge choice. Our failure to replicate the original effect of heightened facial attractiveness at high fertility implicates a need for more research designed to determine if women's faces do change systematically over the cycle and if so, by what mechanisms.

Limitations

In the current study, we replicated previous research suggesting that women feel more attractive (or less ugly) at high fertility and have larger breasts at low fertility. However, we failed to replicate previous research suggesting enhanced facial attractiveness and reduced waist-to-hip ratio at high fertility. These null findings are difficult to interpret. It is possible that there are systematic changes but they are too subtle to detect reliably in a small sample or in a sample of certain characteristics (e.g., baseline body shape or level of attractiveness). Despite the methodological controls we implemented in the current study,

there are several weaknesses we could improve on to provide a more accurate test of change in face and body shape. First, we would utilize urine tests to confirm day of ovulation. Although we used both forward and reverse counting methods to hone in on low and high fertility days, we did not obtain hormonal verifications of ovulatory status. Urine tests would have defrayed a second limitation, which is that two high-conception-risk days, Days 14 and 15, were not included in the forward-count high fertility scheduling window. (That said, Roberts et al. did not schedule on Day 15 either.) Third, the majority of women in our sample completed their high fertility session first, followed by their low. Ideally, half of women would attend high then low, and half low then high. It is likely that effects of visit order were minimal in the current study because participants met us for an extensive intake session and their levels of fear, upset, and irritation were near floor for both visits. Moreover, even though most women completed the High Fertility session first, their levels of felt ugliness were significantly lower at the High Fertility session. Further, it seems that any differences in nervousness or fear would not have been clearly expressed in waist and hip measurements, nor in the facial shot we took because women were not allowed to smile. Another limitation is that we did not explicitly request women to remove their makeup the night before each photo-measurement visit. The typical woman washes before bed and, upon reviewing our photos, we found no woman who appeared to have any makeup on besides a couple with eyeliner remnants (presumably from daily wear). However, it would have been advantageous to explicitly ask women to wash thoroughly the night prior to each visit. A final potential limitation is that the format in which judges viewed the pictures, through a large screen several feet away rather than on an LCD display directly in front of them, may have limited their ability to detect subtle changes in skin tone or color. We kept the rooms dark during the rating sessions to enhance the displays, but future researchers could utilize Roberts et al.'s choice of a personal LCD display to minimize concerns about resolution.

Conclusion

There is unequivocal evidence that females leak a variety of cues to their ovulatory status. Some cues, however, have only been documented once in the peer-reviewed literature, and some scientists have charged that the first article that gets published in a given domain tends to be one that has documented the most extreme effect – an effect that is bigger than the “true” effect if it actually exists (Ioannidis, 2005; Young, Ioannidis, and Al-Ubaydli, 2008). Lessons well-learned in genetics are that one-time research findings are more likely to actually be false when samples are small, effects are small, or when the field of research involves many scientists competing to be the first to document an effect (Young et al., 2008). For physical and behavioral changes across the menstrual cycle, which is a “hot” area of research, the best way to determine the real state of affairs is to engage in both operational and constructive replications (Lykken, 1991; Rosenthal, 1979). As behavioral geneticist David Lykken (1991, p. 33) quoted Ronald Fisher saying back in 1929, “significance testing may make a finding more intriguing, but it takes replication to make it believable.” Given the theoretical importance of evidence to suggest women experience physical changes over the menstrual cycle that leak cues to their ovulatory status, we encourage future replications to determine the magnitude and relational impact

of those changes.

Acknowledgements: This research was supported by a Faculty/Student Collaborative Research grant from the Office of Research and Sponsored Programs at the University of Wisconsin-Eau Claire.

Received 25 March 2010; Revision submitted 28 June 2011; Accepted 02 August 2011

References

- Brewis, A., and Meyer, M. (2005). Demographic evidence that human ovulation is undetectable (at least in pair bonds). *Current Anthropology*, *46*, 465-471.
- Bryant, G. A., and Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biology Letters*, *5*, 12-15.
- Buist, D. S., Aiello, E. J., Miglioretti, D. L., and White, E. (2006). Mammographic breast density, dense area, and breast area differences by phase in the menstrual cycle. *Cancer Epidemiology, Biomarkers, and Prevention*, *15*, 2303-2306.
- Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General*, *138*, 351-367.
- Durante, K. M., Li, N. P., and Haselton, M. G. (2008). Changes in women's choice of dress across the ovulatory cycle: Naturalistic and laboratory task-based evidence. *Personality and Social Psychology Bulletin*, *34*, 1451-1460. doi: 10.1177/0146167208323103
- Fessler, D. M. T. (2003). No time to eat: An adaptationist account of periovulatory behavioural changes. *Quarterly Review of Biology*, *78*, 3-21.
- Fowler, P. A., Casey, C. E., Cameron, G. G., Foster, M. A., and Knight, C. H. (1990). Cyclic changes in composition and volume of the breast during the menstrual cycle, measured by magnetic resonance imaging. *British Journal of Obstetrics and Gynaecology*, *97*, 595-602.
- Gangestad, S. W., Simpson, J. A., Cousins, A. J., Garver-Apgar, C. E., and Christensen, P. N. (2004). Women's preferences for male behavioral displays change across the menstrual cycle. *Psychological Science*, *15*, 203-207.
- Gangestad, S. W., and Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society of London B*, *265*, 927-933.
- Gangestad, S. W., Thornhill, R., and Garver, C. E. (2002). Changes in women's sexual interests and their partners' mate retention tactics across the menstrual cycle: Evidence for shifting conflicts of interest. *Proceedings of the Royal Society of London B*, *269*, 975-982.
- Gangestad, S. W., Thornhill, R., and Garver-Apgar, C. E. (2010). Fertility in the cycle predicts women's interest in sexual opportunism. *Evolution and Human Behavior*, *31*, 400-411.
- Gildersleeve, K. A., Larson, C., Pillsworth, E., and Haselton, M. (2010). *Detection of ovulation through scent cues: A study of women with hormonally-confirmed*

- ovulation. Paper presented at the 22nd annual meeting of the Human Behavior and Evolution Society, University of Oregon, Eugene, Oregon, USA.
- Gueguen, N. (2009). Menstrual cycle phases and female receptivity to a courtship solicitation: An evaluation in a nightclub. *Evolution and Human Behavior*, *30*, 351-355.
- Haselton, M. G., and Gangestad, S. W. (2006). Conditional expression of women's desires and men's mate guarding across the menstrual cycle. *Hormones and Behavior*, *49*, 509-518.
- Haselton, M. G., Mortezaie, M., Pillsworth, E. G., Bleske-Rechek, A., and Frederick, D. A. (2007). Ovulatory shifts in human female ornamentation: Near ovulation, women dress to impress. *Hormones and Behavior*, *51*, 4-45.
- Havlicek, J., Dvorakova, R., Bartos, L., and Flegr, J. (2006). Non-advertised does not mean concealed: Body odour changes across the menstrual cycle. *Ethology*, *112*, 81-90.
- Havlicek, J., Roberts, S. C., and Flegr, J. (2005). Women's preference for dominant male odour: Effects of menstrual cycle and relationship status. *Biology Letters*, *1*, 256-259.
- Hrdy, S. B. (1981). *The woman that never evolved*. Cambridge, MA: Harvard University Press.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*, 0696-0701.
- Johnston, V. S., Hagel, R., Franklin, M., Fink, B., and Grammer, K. (2001). Male facial attractiveness: Evidence for hormone-mediated adaptive design. *Evolution and Human Behavior*, *22*, 251-267.
- Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., and Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, *37*, 78-84.
- Kirchengast, S., and Gartner, M. (2002). Changes in fat distribution (WHR) and body weight across the menstrual cycle. *Collegium Anthropologicum*, *26 (Suppl)*, 47-57.
- Kuukasjarvi, S., Eriksson, C. J. P., Koskela, E., Mappes, T., Nissinen, K., and Rantala, M. J. (2004). Attractiveness of women's body odors over the menstrual cycle: The role of oral contraceptives and receiver sex. *Behavioral Ecology*, *15*, 579-584.
- Laeng, B., and Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormonal cycle. *Hormones and Behavior*, *52*, 520-530.
- Lieberman, D., Pillsworth, E. G., and Haselton, M. G. (2011). Kin affiliation across the ovulatory cycle: Females avoid fathers when fertile. *Psychological Science*, *22*, 13-18.
- Lovejoy, C. O. (1981). The origin of man. *Science*, *211*, 341-350.
- Lykken, D. T. (1991). What's wrong with psychology, anyway? In D. Cicchetti and W. M. Grove (Eds.), *Thinking clearly about Psychology: Essays in honor of Paul E. Meehl, Vol. 1: Matters of public interest* (pp. 1-39). Minneapolis, MN: University of Minnesota Press.
- Manning, J. T., Scutt, D., Whitehouse, G. H., Leinster, S. J., and Walton, J. M. (1996). Asymmetry and the menstrual cycle in women. *Ethology and Sociobiology*, *17*, 129-143.

- Miller, G., Tybur, J. M., and Jordan, B. D. (2007). Ovulatory cycle effects on tip earnings by lap dancers: Economic evidence for human estrus? *Evolution and Human Behavior*, *28*, 375-381.
- Miller, S. L., and Maner, J. K. (2010). Scent of a woman: Men's testosterone responses to olfactory ovulation cues. *Psychological Science*, *21*, 276-283.
- Miller, S. L., and Maner, J. K. (2011). Ovulation as a male mating prime: Subtle signs of women's fertility influence men's mating cognition and behavior. *Journal of Personality and Social Psychology*, *100*, 295-308.
- Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- Penton-Voak, I. S., and Perrett, D. I. (2000). Female preference for male faces changes cyclically – further evidence. *Evolution and Human Behavior*, *21*, 39-48.
- Peters, M., Rhodes, G., and Simmons, L. W. (2007). Contributions of the face and body to overall attractiveness. *Animal Behaviour*, *73*, 937-942.
- Pipitone, R. N., and Gallup Jr, G. G. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior*, *29*, 268-274.
- Ramakrishnan, R., Khan, S. A., and Badve, S. (2002). Morphological changes in breast tissue with menstrual cycle. *Modern Pathology*, *15*, 1348-1356.
- Roberts, S. C., Havlicek, J., Flegr, J., Hruskova, M., Little, A. C., Jones, B. C., Perrett, D. I., and Petrie, M. (2004). Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proceedings of the Royal Society B (Suppl.)*, *271*, S270-S272.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*, 638-641.
- Scutt, D., and Manning, J. T. (1996). Symmetry and ovulation in women. *Human Reproduction*, *11*, 2477-2480.
- Singh, D. (1993). Adaptive significance of waist-to-hip ratio and female physical attractiveness. *Journal of Personality and Social Psychology*, *65*, 293-307.
- Singh, D., and Bronstad, P. M. (2001). Female body odour is a potential cue to ovulation. *Proceedings of the Royal Society London B*, *268*, 797-801.
- Streeter, S. A., and McBurney, D. H. (2003). Waist-hip ratio and attractiveness: New evidence and a critique of a critical test. *Evolution and Human Behavior*, *24*, 88-98.
- Thornhill, R., and Gangestad, S. W. (2008). *The evolutionary biology of human female sexuality*. New York, NY: Oxford University Press.
- Wilcox, A. J., Dunson, D. B., Weinberg, C. R., Trussell, J., and Baird, D. D. (2001). Likelihood of conception with a single act of intercourse: Providing benchmark rates for assessment of post-coital contraceptives. *Contraception*, *63*, 211-215.
- Young, N. S., Ioannidis, J. P. A., and Al-Ubaydli, O. (2008). Why current publication practices may distort science. *PLoS Medicine*, *5*, 1418-1422.