Research Article

Psychological Science
1–13

DOI: 10.1177/0956797619882022
www.psychologicalscience.org/PS
SAGE

Probing Ovulatory-Cycle Shifts in Women's Preferences for Men's Behaviors

TC

Julia Stern, Tanja M. Gerlach, and Lars Penke
Department of Psychology and Leibniz ScienceCampus Primate Cognition, University of Goettingen

Abstract

The existence of ovulatory-cycle shifts in women's mate preferences has been a point of controversy. There is evidence that naturally cycling women in their fertile phase, compared with their luteal phase, evaluate specific behavioral cues in men as more attractive for sexual relationships. However, recent research has cast doubt on these findings. We addressed this debate in a large, preregistered, within-participants study using salivary-hormone measures and luteinizing-hormone tests. One hundred fifty-seven female participants rated the sexual and long-term attractiveness of 70 men in dyadic intersexual interactions in natural videos. Multilevel comparisons across two ovulatory cycles indicated that women's mate preferences for men's behaviors did not shift across the cycle for either competitive or courtship behavior. Within-women hormone levels and relationship status did not affect these results. Hormonal mechanisms and implications for estrus theories are discussed.

Keywords
ovulatory cycle, mate preferences, steroid hormones, fertility, attractiveness, open data, open materials, preregistered

Received 3/1/18; Revision accepted 12/8/19

Scientific interest in whether women experience systematic psychological changes across their ovulatory cycle has increased in recent years. A substantial amount of research indicates that women's sexual interests change across the ovulatory cycle. Although cycle shifts in sexual desire appear robust, with higher levels of desire during women's fertile phase (e.g., Arslan, Schilling, Gerlach, & Penke, 2018; Grebe, Thompson, & Gangestad, 2016; Jones, Hahn, Fisher, Wang, Kandrik, & DeBruine, 2018; Roney & Simmons, 2013, 2016), there is ongoing discussion whether there are changes in mate preferences as well. According to the good-genes-ovulatory-shift hypothesis (GGOSH; Gangestad, Garver-Apgar, Simpson, & Cousins, 2007; Gangestad, Thornhill, & Garver-Apgar, 2005), women's mate preferences should differ according to the mating context: When fertile, women should prefer men with characteristics indicative of good genes for sexual relationships. These preferences should be absent in the luteal phase (i.e., between ovulation and menstrual onset) and when evaluating men for long-term relationships (given that long-term bonding with these men can be costly, because they may be less willing to provide parental effort; Gangestad & Simpson, 2000).

Evidence for this hypothesis is mixed. Previous research has documented cycle shifts in women's mate preferences for several physical and behavioral traits (for an overview, see Gildersleeve, Haselton, & Fales, 2014). However, changes in preferences for masculine faces, bodies, and voices did not replicate in more recent studies (e.g., Jones, Hahn, Fisher, Wang, Kandrik, Han, et al., 2018; Jünger, Kordsmeyer, Gerlach, & Penke, 2018; Jünger, Motta-Mena, et al., 2018; Marcinkowska, Galbarczyk, & Jasienska, 2018; Muñoz-Reyes et al., 2014). Moreover, two meta-analyses came to strikingly diverging conclusions on whether cycle effects exist (Gildersleeve et al., 2014; Wood, Kressel, Joshi, & Louie, 2014). Additionally, previously conducted studies have been criticized for potentially serious methodological problems, such as inappropriate sample sizes, use of between-participants designs, lack of direct assessments of steroid hormones, and not using luteinizing-hormone (LH) tests for validating women's fertile phase (Blake,

Corresponding Author:
Julia Stern, University of Goettingen, Gosslerstrasse 14, 37073 Goettingen, Germany
E-mail: julia.stern@psych.uni-goettingen.de

Dixson, O'Dean, & Denson, 2016; Gangestad et al., 2016). In sum, to clarify the scientific discourse about the existence of ovulatory-cycle shifts, there is strong need for adequately designed and powered replications conducted in different interpersonal contexts.

Overview of the Current Study and Hypotheses

In the current study, we set out to directly probe the GGOSH for men's behaviors while overcoming previously reported methodological problems. In particular, we aimed to clarify (a) whether there are preference shifts for men's behaviors across the ovulatory cycle, (b) which hormonal mechanisms might potentially mediate these effects, and (c) which moderators affect them.

Investigating ovulatory shifts in preferences for men's behaviors

Several studies have found that women's preferences for men displaying behavioral dominance, confidence, and social presence change across the ovulatory cycle (Gangestad et al., 2007; Gangestad, Simpson, Cousins, Garver-Apgar, & Christensen, 2004; Lukaszewski & Roney, 2009). Moreover, previous research has also reported changes in women's flirting behavior and behavioral engagement with men displaying purported markers of genetic fitness (Cantú et al., 2014; Flowe, Swords, & Rockey, 2012). Women also seem to show preferences for flirtatious facial movement in the fertile phase of the ovulatory cycle (Morrison, Clark, Gralewski, Campbell, & Penton-Voak, 2010). Yet there are a lack of well-powered replications investigating preference shifts for men's behavior (Jones, Hahn, & DeBruine, 2019). Therefore, we decided to investigate cycle shifts in preferences for men's flirting behavior and dominance-related cues found in such behavior. It has been suggested that flirting behavior exaggerates one's qualities as a mate (Back et al., 2011). Behavioral attractiveness can be seen as an effort on the part of men to appear attractive to women in a more subtle and indirect manner than via flirting. Self-display behaviors have been seen as an attempt to impress the conversation partner, are correlated with higher testosterone levels in men, and appear to index courtshiplike behavior (Roney, Lukaszewski, & Simmons, 2007; Roney, Mahler, & Maestripieri, 2003). The amount of eye contact (i.e., *gazes*) has been reported to be an indicator of *social presence*, a behavioral display for which women's mate preferences may change across the ovulatory cycle (Gangestad et al., 2004). Men's dominance, arrogance, assertiveness, confrontativeness, social respectability, and likelihood of winning a physical fight are behaviors more directly related to intrasexual competitiveness and social presence, behaviors for which women's mate preferences have been reported to change across the cycle (Gangestad et al., 2004; Gangestad et al., 2007).

In light of previous findings on ovulatory-cycle shifts, we hypothesized that fertile women, compared with women in their luteal phase, evaluate men's behaviors as more attractive for sexual relationships[1] (Hypothesis 1). Moreover, women's mate preferences should shift across the cycle: When fertile, women should be more sexually attracted to men who show more overt flirting behavior, more self-displays, more direct gazes toward the women they are talking to, and more behavior that is consensually perceived as attractive (*behavioral attractiveness*; Hypothesis 2a). Furthermore, although this hypothesis was not preregistered,[2] we expected comparable findings for behavioral cues of dominance, arrogance, assertiveness, confrontativeness, respectability, and likelihood of winning a physical fight. When evaluating long-term attractiveness, we expected preference shifts to be absent or only weakly present (Hypothesis 3). We predicted that our findings would be robust when we controlled for men's age, physical attractiveness, and voice attractiveness. We also formed the alternative hypothesis that women' mate preferences for sexual relationships would not shift (Hypothesis 2b).

Hormones as mediators, relationship status as a moderator

The female ovulatory cycle is regulated by shifts in hormone concentrations. Although estradiol rises in the fertile phase, it decreases during the luteal phase, but with a second smaller peak in the midluteal phase. Progesterone levels are usually lower in the fertile phase and higher in the luteal phase. Therefore, cycle shifts in mate preferences should be mediated by natural within-women changes in hormone levels: higher estradiol and lower progesterone (Hypothesis 4). Because recent research suggests that progesterone effects on mate preferences are between women rather than within women (DeBruine, Hahn, & Jones, 2019; Marcinkowska, Kaminski, Little, & Jasienska, 2018), we also tested between-women hormone effects in an exploratory manner. An important variable that might affect the strengths of ovulatory-cycle shifts is women's relationship status. According to the dual-mating hypothesis, women may receive fitness benefits when forming a relationship with a reliably investing man while seeking good genes from other men through extra-pair sexual encounters (Pillsworth & Haselton, 2006). Because it remains unclear whether singles also pursue different mating strategies across the cycle, we formed two alternative hypotheses: Cycle shifts in preferences for short-term mates will be larger for partnered women than for single women (Hypotheses 5a), or, alternatively, relationship status will not affect the strengths of cycle shifts in preferences for short-term mates (Hypotheses 5b).

Method

Our hypotheses, the study design, and the sampling and the analysis plan were preregistered online at the Open Science Framework (OSF; https://osf.io/m6pnz/) before any data on the women were collected or analyzed. This preregistration also contains further hypotheses that are not part of the present article, because they were written for a larger project. Data, analysis scripts, and instruction materials are also provided on our OSF project (https://osf.io/8ntuc/). All participants signed a written consent form, and the local ethics committee approved the study protocol (No. 144).

Participants and recruitment

We used the same sample as did Jünger and colleagues (Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018). Participants were recruited following the same inclusion criteria of other ovulatory-cycle studies and had to fit the following preregistered criteria: female, between 18 and 30 years old, and naturally cycling (i.e., not having taken hormonal contraception for at least 3 months, not unexpectedly switching to hormonal contraception during the study, not currently pregnant or breastfeeding, not having given birth or breastfed during the previous 3 months, not taking hormone-based medication or antidepressants). Additionally, they had to report that their ovulatory cycles had a regular length between 25 and 35 days during the last 3 months.

In total, we recruited 180 participants, of whom 23 could not be included in the final sample: 17 women who attended only the introductory session of the study dropped out before participation (6 failed one of the inclusion criteria above, 4 quit the study, 4 did not respond to e-mails, and 3 had scheduling problems). Another 6 dropped out during the study because they completed only the first testing session (4 had scheduling problems, 2 did not respond to e-mails after the first session). One of the participants later reported that she was 35 years old. We included her data for robustness checks because she met all other including criteria and had positive LH tests. Excluding her data did not change the results. One hundred fifty-seven heterosexual female participants (age: range = 18–35 years, M = 23.3, SD = 3.4) finished all sessions and could therefore be included in further analyses. At the beginning of the study, 75 of these participants reported that they were in a relationship, 82 reported that they were single. Our sample vastly exceeded the size required to achieve 80% power given a within-participants design and anticipated effects of moderate magnitude (Cohen's d = 0.5 with N = 48 for LH-test-validated cycle phases and two testing sessions per participant, suggesting that there was sufficient power to detect much smaller effect sizes in our study), as suggested by recent guidelines for sample sizes in ovulatory-shift research (Gangestad et al., 2016). After completing all sessions, participants received a payment of 80€ or course credit.

Procedure

All participants took part in five individually scheduled sessions. In the first introductory session, participants received detailed information about the general procedure, the duration of the study, and compensation. Furthermore, the experimenter explained the ovulation tests and checked the inclusion criteria. To count the days to the next ovulation and to plan the dates of the experimental sessions, we assessed cycle length as well as the dates of the last and the next menstrual onset. Finally, demographic data were collected.

Sessions two to five were computer-based testing sessions and took place once during the fertile phase and once during the luteal phase for two consecutive cycles per participant. To control for possible effects of diurnal changes in hormone levels, we scheduled all sessions in the second half of the day (mainly between 11:30 a.m. and 6:00 p.m.). After arriving at the lab, participants first completed a screening questionnaire that assessed their eligibility and some control variables for saliva sampling (e.g., the last time participants had eaten something; Schultheiss & Stanton, 2009). Saliva samples were collected via passive drool before the participants started the first rating task. Participants also completed two other rating tasks in which they had to rate the attractiveness of men's bodies or voices (see Jünger, Kordsmeyer, et al., 2018, and Jünger, Motta-Mena, et al., 2018, for detailed descriptions of these tasks). The order in which participants completed all rating tasks (the videos described in the current study, as well as bodies or voices as described by Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018) was randomized between participants and sessions. Additionally, anthropometric data were collected between these tasks (a) to make sure that participants got breaks between the rating tasks and (b) as part of a larger study (see the preregistration).

In the first testing session, participants saw a short preview video that presented facial pictures of all men they were about to rate for 1 s each. Participants were then instructed to evaluate the men in the following videos, which were the actual stimulus material, according to their attractiveness as they perceived it "in that moment," independently of their own current relationship status or general interest in other men, and to rate the attractiveness of the men by focusing only on the behavior exhibited in the videos.

Video clips were presented in a randomized order using the experimental software Alfred (Treffenstaedt & Wiemann, 2018), which is based on the programming language Python (Version 2.7; http://www.python.org). After watching each sequence, participants were to separately rate each individual man's sexual attractiveness (to assess short-term attraction) and attractiveness for long-term relationships. Ratings were made on 11-point Likert scales from −5 (*extremely unattractive*) to 5 (*extremely attractive*), including 0 as a neutral point. Definitions of sexual attractiveness and attractiveness for a long-term relationship were provided prior to the rating task. *Sexually attractive* was defined as follows: "Men that score high would be very attractive for a sexual relationship that can be short-lived and must not contain any other commitment. Men scoring low would be very unattractive for a sexual relationship." *Attractiveness for a long-term relationship* was defined as follows: "Men that score high would be very attractive for a committed relationship with a long-term perspective. Men that score low would be very unattractive as a long-term partner." After each session, the appointment for the next session was arranged individually on the basis of the participant's ovulatory cycle.

Measures

Ovulatory-cycle phase. Women's cycle phase was determined by the reverse-cycle-day method, which is based on the estimated day of the next menstrual onset (Gildersleeve, Haselton, Larson, & Pillsworth, 2012) and confirmed by highly sensitive (10 mIU/ml) urine ovulation test strips (Purbay Ovulation Tests, MedNet, Muenster, Germany), which measure LH. These LH tests had to be done at home at the estimated day of ovulation and the 4 days prior to that. We investigated two ovulatory cycles, in which each participant reported to the lab twice: once while being fertile (at the days immediately preceding ovulation, usually Reverse Cycle Day 16 to 18, with Reverse Cycle Day 16 as the most ideal date) and once when not fertile (during the luteal phase, after ovulation and prior to the next menstrual onset, usually Reverse Cycle Day 4 to 11, with Reverse Cycle Days 6 to 8 as the most ideal dates). Out of all participants who finished every session, 66 participants started the first session in their luteal phase, and 91 started in the fertile phase.

For analyses of the main cycle phase, we excluded 45 participants because of negative LH tests in both cycles, irregular ovulatory cycles, or inappropriate scheduling of testing sessions (see Preliminary Analyses for more details), leaving a final sample of 112 women. Of these participants, 46 started with the first session in their luteal phase, and 66 started in their fertile phase. However, all 157 women were included in the robustness checks.

Hormone assessments. For hormone assays, we collected four saliva samples from each participant, one per testing session. Contamination of saliva samples was minimized by asking participants to abstain from eating, drinking (except plain water), smoking, chewing gum, or brushing their teeth for at least 1 hr before each session. Samples were visually inspected for blood contamination and stored at −80 °C directly after collection until shipment on dry ice to the Kirschbaum Lab at the Technical University of Dresden, where estradiol, progesterone, testosterone and cortisol were assessed via liquid chromatography–mass spectrometry (LCMS). Estradiol levels could be detected by LCMS analysis in only 22% of the hormone samples. Therefore, all samples were reanalyzed using a highly sensitive 17β-estradiol enzyme immunoassay kit (IBL International, Hamburg, Germany). These latter estradiol values were used in subsequent analyses. We centered all hormone values on their participant-specific means and scaled them afterwards (i.e., divided them by a constant), so that the majority of the distribution for each hormone varied from −0.5 to 0.5, to facilitate calculations in linear mixed models (e.g., Jones, Hahn, Fisher, Wang, Kandrik, & DeBruine, 2018; Jones, Hahn, Fisher, Wang, Kandrik, Han, et al., 2018). This is a common procedure to isolate effects of within-participants changes in hormones, avoiding the influence of outliers on results and dealing with the nonnormal distribution of hormone levels. It is also in line with the procedure followed by Jünger and colleagues (Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018). Hormone levels were nearly normally distributed afterward; Figure S1 in the Supplemental Material available online shows the distribution of hormone levels after this procedure. Importantly, this procedure did not change any findings compared with analyses with untransformed hormone values. The R code for this procedure can be found on our OSF project (https://osf.io/8ntuc/). One woman had extremely high levels of progesterone and could be considered an outlier. However, results remained virtually identical when we excluded her from all hormone analyses. All analyses excluding this woman can be found on our OSF project.

Stimuli and behavioral ratings. Thirty-second-long sequences of videos of men in dyadic interactions, recorded in a study on sociosexuality (Penke & Asendorpf, 2008), were presented. We selected the videos of 70 men who were single at the time of the initial study out of a larger pool of 283 videos in total. For every video, a male participant was seated in a room with an attractive female confederate. They were instructed to get to know each other, while the experimenter left the room (see Penke & Asendorpf, 2008, for details). From each conversation, we extracted the sequence from 2 min to 2.5 min to avoid the potential awkwardness of the first moments and ensure that the interaction was in full flow. The participants saw the conversation from a camera recording over the shoulder of the female confederate, so they saw a frontal view of only the man in each interaction.

To assess the behaviors of all men, four independent, trained raters (two women, two men) who were unacquainted with the participants first rated the videos. Ratings were done using 7-point Likert scales for the 30-s sequences on the following behavioral dimensions: flirting behavior, self-displays, and behavioral attractiveness. Ratings were collected in two rounds, the first based on recordings from a side perspective, and the second based on the frontal recordings that were used as stimuli in the present study. In both rounds, videos were presented with audio. Interrater agreement was high (side perspective: αs = .84–.88; frontal perspective: αs = .85–.90); thus, ratings of all raters and both perspectives were aggregated.

Further ratings of the male behavior were collected separately later on. The following dimensions were rated: dominance, arrogance, assertiveness, confrontativeness, social respectability, and likelihood of winning a physical fight. Each dimension was separately rated by 10 independent raters (5 women, 5 men) on the basis of the 30-s videos using 7-point Likert scales. Interrater agreement was high (dominance: α = .88, arrogance: α = .71, assertiveness: α = .89, confrontativeness: α = .83, social respectability: α = .89, likelihood of winning a physical fight: α = .86); thus, ratings of all raters were aggregated for each dimension.

In addition, two trained research assistants coded objective male gazes (percentage of total amount of time the man looked the confederate directly in the face) using Observer software (Noldus, Leesburg, VA). Intraclass correlations were high (.99); thus, codings from both assistants were averaged. Additionally, men's facial and vocal attractiveness were rated on 7-point Likert scales as control variables. For facial attractiveness, frontal face pictures with neutral facial expressions were rated by 15 independent undergraduate students. Interrater reliabilities were high (α = .91), so ratings were aggregated after z scoring. For vocal attractiveness, voice recordings (counting from 1 to 10) were rated by six trained research assistants, and ratings were aggregated afterward (α = .80). Behaviors varied substantially among the videos; descriptive values for all can be found in the Supplemental Material. More details about the rating and coding procedures can be found in Penke and Asendorpf (2008).

Statistical analyses

In our preregistration, we explicitly stated our hypotheses, methods, recruitment strategy, and stopping rule. However, our preregistration was not fully explicit about our statistical analyses. Hence, we decided to run a number of robustness checks, which consist of analyses combining various reasonable analytical decisions. To substantiate that these choices were reasonable, we based them on procedures followed in previously published studies investigating cycle shifts in mate preferences or on suggestions we received during the review process. As described in our preregistration, all data analyses were done using multilevel modeling. Details can be found on our OSF project.

Results

Preliminary analyses

First, we checked how many of the participants' ovulatory cycles had positive LH tests (showing an LH surge) in the calculated fertile phase to detect nonovulatory cycles. Twelve participants reported negative LH test results for both investigated cycles, and 9 reported negative LH tests results for one cycle. In total, LH tests in 33 of all 314 cycles (10.5%) were negative. Next, we counted how many cycles were reported as being irregular, that is, where days of the testing sessions deviated from the prior defined phase of appropriate testing days by more than 3 days (see Ovulatory-Cycle Phase). Eight women reported irregular cycles in both investigated cycles, and 32 reported one cycle being irregular, resulting in 48 out of 314 (15.3%) cycles being irregular (despite all participants reporting having regular ovulatory cycles in the introductory session prior to the testing sessions). Additionally, we checked the temporal relationship between the reported day of LH surge and the date of scheduled testing session. Because ovulation usually occurs within 24 to 36 hr after the observed LH surge, testing sessions that were scheduled more than 2 days after the surge might have already been in the early luteal phase. Out of the 281 cycles for which an LH surge was observed, 13 (4.63%) purportedly fertile-phase sessions were scheduled 3 or 4 days after the LH surge. Therefore, 268 (95.37%) were scheduled within an appropriate range of 3 days before to 2 days after the LH surge (in total: M = −0.12, SD = 1.39 days in relation to the day of the observed LH surge). For a histogram showing the distribution of days of fertile-phase testing sessions relative to the observed LH surge, see Figure S2 in the Supplemental Material. Participants with irregular cycles, negative LH tests, or the risk of early luteal phase instead of fertile-phase testing session were excluded from the main analyses (as cycle-phase estimates based on LH have a much higher validity than estimates based on backward counting alone; e.g., Blake et al., 2016; Gangestad et al., 2016) but included in robustness checks described in the Supplemental Material.

Main analyses: cycle shifts in women's attraction and mate preferences

We first tested for possible ovulatory-cycle shifts in women's attractiveness ratings for men's behavior in general (Hypothesis 1). For multilevel analyses with attractiveness rating as the dependent variable (Model 1 with sexual attractiveness, Model 2 with long-term attractiveness), female raters and male stimuli were treated as random effects. Women's cycle phase (0 = luteal phase, 1 = fertile phase) was treated as a fixed effect. We also let participant's slopes vary systematically across cycle phase by modeling cycle phase as a random slope. These models did not converge. We thus followed recommendations from Bates, Kliegl, Vasishth, and Baayen (2015) and tried to reduce random-slope variance in a number of different ways (see https://osf.io/kyqn8/ for details). We decided for the solution with the best model fit (determined using Akaike information criterion, Bayesian information criterion, and log likelihood) by defining simpler scalar random effects for different cycle-phase intercepts per participant. This procedure applies to all of the following models. Importantly, results remained virtually identical, results for the (nonconverging) models with maximum random-slope specification can be found on our OSF project.

Both models showed a significant cycle shift in women's attraction: When women were fertile, their ratings for sexual attractiveness were higher than in the luteal phase of the ovulatory cycle, $\gamma = 0.10$, $SE = 0.05$, 95% confidence interval (CI) = [0.01, 0.20], $t(111) = 2.13$, $p = .035$, providing only modest support for Hypothesis 1, as the effect seems to be rather small and the lower bound of the CI is near 0. Similar results were found for ratings of long-term attractiveness, $\gamma = 0.10$, $SE = 0.05$, 95% CI = [0.00, 0.19], $t(111) = 2.05$, $p = .042$. When women were fertile, their attractiveness ratings of men's flirting behavior increased compared with their ratings in the luteal phase. For robustness checks, we repeated these analyses with estradiol and progesterone, rather than cycle phase, as predictors using the data set of 157 women, not including random slopes. This analysis resulted in seven additional models for sexual and long-term attractiveness each. Significant effects were observed in only six out of these models, four with cycle phase and two with hormone levels as predictors. Details can be found in Tables S1 to S7 in the Supplemental Material.

Next, we analyzed whether women's mate preferences for specific behaviors changed across the cycle (Hypotheses 2a, 2b, and 3). Deviating from our preregistration because of suggestions from reviewers, we ran a factor analysis for the different behavioral tactics, because some of them correlated rather strongly (Table S8 in the Supplemental Material). An exploratory factor analysis with oblimin rotation yielded two factors. Factor 1 (labeled "competitiveness") consisted of positive loadings of dominance, confrontativeness, assertiveness, likeliness of winning a fight, self-display behavior, arrogance, and respectability. Factor 2 (labeled "courtship") consisted of positive loadings of flirting behavior, behavioral attractiveness, respectability, and gazes and negative loadings of arrogance. Factor scores were saved for further computations; detailed factor loadings can be found in Table S9 in the Supplemental Material. Both factors were correlated ($r = .55$, $p < .001$, 95% CI = [.36, .69]) and predicted attractiveness ratings—competitiveness predicting sexual attractiveness: $\gamma = 1.00$, $SE = 0.15$, 95% CI = [0.71, 1.29], $t = 6.83$, $p < .001$; competitiveness predicting long-term attractiveness: $\gamma = 0.86$, $SE = 0.15$, 95% CI = [0.57, 1.16], $t = 5.73$, $p < .001$; courtship predicting sexual attractiveness: $\gamma = 1.44$, $SE = 0.10$, $t = 15.06$, $p < .001$, 95% CI = [1.25, 1.63]; courtship predicting long-term attractiveness: $\gamma = 1.40$, $SE = 0.09$, $t = 15.63$, $p < .001$, 95% CI = [1.22, 1.58]. To decrease the length of our manuscript, we report all other results for ratings of long-term attractiveness as outcome variable in the Tables S24 to S30 in the Supplemental Material, as results were mostly virtually identical to sexual-attractiveness ratings.

We computed multilevel models with women's cycle phase and either the competitiveness factor (Model 3) or the courtship factor (Model 4) as fixed effects. Female participants as well as male stimuli were treated as random effects, and two random slopes for cycle phase varying within participants and a behavioral factor varying within participants were included. These models did not converge; thus, random-slope variance was reduced as described above (details can be found on our OSF project). Results revealed no significant interactions of cycle phase and the behavioral factors, indicating that women's mate preferences for specific cues in men's behavior did not shift across the ovulatory cycle, contradicting Hypothesis 2a but supporting alternative Hypothesis 2b. However, there were significant main effects for cycle phase and both factors on sexual-attractiveness ratings (Table 1). The effects were comparable for ratings of long-term attractiveness (Table S24 in the Supplemental Material) and thus in

Table 1. Multilevel Regression Analyses of Sexual-Attractiveness Ratings as a Function of Cycle Phase and the Behavioral Factors

Model and predictor	γ	*SE*	*t*	*p*	95% CI
Competitiveness					
Cycle phase	0.10	0.05	2.13	.035	[0.01, 0.20]
Competitiveness	1.01	0.15	6.84	< .001	[0.72, 1.30]
Cycle Phase × Competitiveness	0.02	0.02	1.06	.290	[–0.02, 0.05]
Courtship					
Cycle phase	0.10	0.05	2.13	.035	[0.01, 0.20]
Courtship	1.44	0.10	15.01	< .001	[1.25, 1.63]
Cycle Phase × Courtship	0.02	0.02	0.85	.395	[–0.02, 0.05]

Note: Each model had 31,360 observations (112 Participants × 4 Test Sessions × 70 Stimuli). CI = confidence interval.

line with Hypothesis 3. Results were virtually identical when we computed a model with a global factor as predictor variable (average of competitiveness and courtship; as suggested in the review process; Table S38 in the Supplemental Material). All results were virtually identical when we controlled for men's age, physical attractiveness, and voice attractiveness (Tables S10 and S27 in the Supplemental Material). However, men's age had a significant main effect on ratings of sexual and long-term attractiveness in all courtship models, and physical attractiveness had a significant main effect on both ratings in all competitiveness models.

Following our preregistration, we also computed separate models for all behavioral cues. None of these models revealed a significant interaction between cycle phase and cue. Details can be found in Tables S11, S12, S28, and S29 in the Supplemental Material. Again, we found some main effects for individual behavioral cues on sexual and long-term attractiveness: Ratings were higher when flirting behavior, self-display behavior, behavioral attractiveness, assertiveness, confrontativeness, dominance, likeliness of winning a fight, and respectability were higher. There were no significant main effects for gazes or arrogance. All results were virtually identical when we controlled for men's age, physical attractiveness, and voice attractiveness (Tables S13–S15 in the Supplemental Material).

Hormonal mechanism potentially underlying preference shifts

To investigate possible effects of steroid hormones underlying cycle shifts in women's mate preferences for sexual attractiveness as a dependent variable (Hypothesis 4), we entered estradiol as well as progesterone within and between women, either the competitiveness factor (Model 5) or the courtship factor (Model 6) as fixed effects to our multilevel model, female participants and male stimuli as random effects, and random slopes for estradiol, progesterone, and the respective behavioral factor varying within participants. These models did not converge; thus, random-slope variance was reduced as described above (details can be found on our OSF project). Results revealed no significant interaction of within-women or between-women estradiol or progesterone and the behavioral factors, indicating that women's mate preferences for specific cues in men's behavior did not shift because of within-women changes in estradiol or progesterone, contradicting Hypothesis 4. However, there were significant main effects for both factors on sexual-attractiveness ratings (Table 2). The effects were comparable for ratings of long-term attractiveness (Table S25B in the Supplemental Material), besides significant interaction effects between the behavioral factors and between-women progesterone levels, in that both factors were rated as being more attractive for long-term relationships when between-women progesterone levels were lower (competitiveness: $p = .029$; courtship: $p = .021$).

Results were virtually identical when we computed the model with a global factor as predictor variable instead (Table S39 in the Supplemental Material) or when we added estradiol-to-progesterone ratio rather than estradiol and progesterone as predictor variables (Table S17 in the Supplemental Material). All results were virtually identical when we controlled for men's age, physical attractiveness, and voice attractiveness (Tables S16 and S30 in the Supplemental Material).

The role of women's relationship status for ovulatory-cycle shifts

In order to analyze whether women's relationship status might moderate ovulatory-cycle shifts (Hypothesis 5), we categorized all women as being in a relationship

Table 2. Multilevel Regression Analyses of Sexual-Attractiveness Ratings as a Function of Within-Women (WW) or Between-Women (BW) Estradiol or Progesterone and Behavioral Factors

Model and predictor	γ	*SE*	*t*	*p*	95% CI
Competitiveness					
Estradiol (WW)	0.06	0.10	0.57	.572	[–0.14, 0.26]
Progesterone (WW)	–0.07	0.07	–1.02	.307	[–0.20, 0.06]
Estradiol (BW)	0.08	0.10	0.77	.443	[–0.12, 0.28]
Progesterone (BW)	–0.07	0.10	–0.70	.488	[–0.27, 0.13]
Competitiveness	1.01	0.15	6.89	< .001	[0.72, 1.29]
Estradiol (WW) × Competitiveness	–0.05	0.03	–1.46	.143	[–0.13, 0.05]
Progesterone (WW) × Competitiveness	0.04	0.02	1.76	.078	[–0.01, 0.09]
Estradiol (BW) × Competitiveness	0.03	0.02	1.74	.082	[–0.01, 0.05]
Progesterone (BW) × Competitiveness	–0.01	0.02	–0.71	.481	[–0.03, 0.01]
Courtship					
Estradiol (WW)	0.06	0.10	0.57	.572	[–0.14, 0.26]
Progesterone (WW)	–0.07	0.07	-1.02	.307	[–0.20, 0.06]
Estradiol (BW)	0.08	0.10	0.77	.443	[–0.12, 0.28]
Progesterone (BW)	–0.07	0.10	-0.70	.488	[–0.27, 0.13]
Courtship	1.45	0.10	15.08	< .001	[1.25, 1.63]
Estradiol (WW) × Courtship	–0.02	0.03	–0.70	.486	[–0.11, 0.07]
Progesterone (WW) × Courtship	–0.00	0.02	–0.03	.975	[–0.06, 0.06]
Estradiol (BW) × Courtship	0.01	0.02	0.69	.493	[–0.01, 0.04]
Progesterone (BW) × Courtship	0.00	0.02	0.19	.847	[–0.02, 0.03]

Note: Each model had 27,300 observations (each 112 Participants × 4 Test Sessions × 70 Stimuli; missing values). CI = confidence interval.

who reported being in an open relationship, in a committed relationship, engaged, or married. However, results did not change when we categorized women who reported being in an open relationship as singles instead. Relationship status changed for 13 women across the study; these cases were categorized according to their relationship status on the particular testing day. Relationship status was effect coded (–1 for single women and 1 for women in relationships). We computed two multilevel models: Model 7 included competitiveness as a predictor, and Model 8 included courtship as a predictor. Both models included women's cycle phase and their relationship status as fixed effects and female participants and male stimuli as random effects. We additionally modeled random-slope variations for cycle phase; relationship status and the respective behavioral factor varied within participants. These models did not converge; thus, random-slope variance was reduced as described above (details can be found on our OSF project). Results showed significant main effects for competitiveness and courtship, in that ratings were generally higher when competitiveness or courtship were higher. There were no other significant effects: Neither the main effects of cycle phase or relationship status were significant, nor was any interaction effect (Table 3), indicating that women's mate preferences did not shift across the cycle and that relationship status did not moderate such shifts, supporting Hypothesis 5b but not Hypothesis 5a. Results for long-term attractiveness revealed comparable results (Tables S26 in the Supplemental Material).

For robustness checks, we repeated our analyses with estradiol and progesterone, rather than cycle phase, as predictor variables. None of the two-way or three-way interactions involving within-women estradiol or progesterone were significant, whereas one significant positive two-way interaction between competitiveness and between-women estradiol levels occurred ($p = .013$). Details can be found in Table S18 in the Supplemental Material. All results were virtually identical when we controlled for men's age, physical attractiveness, and voice attractiveness (Tables S19 and S20 in the Supplemental Material).

Additional robustness checks

We conducted further analyses to probe the robustness of our effects. To rule out that our results might have been caused by order effects of testing sessions, particularly participating in the first session when fertile (Suschinsky, Bossio, & Chivers, 2014), we controlled for initial cycle phase in our main analyses. Initial cycle

Table 3. Multilevel Regression Analyses of Sexual-Attractiveness Ratings as a Function of Cycle Phase, Relationship Status, and the Competitiveness Factor or the Courtship Factor

Model and predictor	γ	*SE*	*t*	*p*	95% CI
Competitiveness					
Cycle phase	0.07	0.07	1.07	.286	[–0.06, 0.20]
Competitiveness	0.98	0.15	6.58	< .001	[0.68, 1.27]
Relationship status	–0.24	0.14	–1.73	.097	[–0.53, 0.03]
Cycle Phase × Competitiveness	0.03	0.03	1.04	.301	[–0.02, 0.08]
Cycle Phase × Relationship Status	0.07	0.09	0.72	.475	[–0.11, 0.24]
Competitiveness × Relationship Status	0.07	0.04	1.73	.084	[0.01, 0.14]
Cycle Phase × Competitiveness × Relationship Status	–0.02	0.04	–0.47	.640	[–0.08, 0.05]
Courtship					
Cycle phase	0.07	0.07	1.07	.286	[–0.06, 0.20]
Courtship	1.42	0.10	14.50	< .001	[1.23, 1.61]
Relationship status	–0.24	0.14	–1.73	.097	[–0.53, 0.03]
Cycle phase × Courtship	0.02	0.03	0.66	.510	[–0.03, 0.07]
Cycle Phase × Relationship Status	0.07	0.09	0.72	.474	[–0.11, 0.24]
Courtship × Relationship Status	0.05	0.04	1.24	.214	[–0.03, 0.13]
Cycle Phase × Courtship × Relationship Status	–0.00	0.04	–0.13	.900	[–0.08, 0.07]

Note: Each model had 31,360 observations (each 112 Participants × 4 Test Sessions × 70 Stimuli). CI = confidence interval.

phase affected attractiveness ratings, as they were higher when participants were tested first when fertile. However, all other results remained stable (Tables S21–S23 in the Supplemental Material). We then repeated all main analyses with all recruited participants (N = 157). Results remained virtually identical and can be found in Tables S31 to S37 in the Supplemental Material.

Additionally, we did further robustness checks that were all requested in the review process. First, we contrasted results for ratings of sexual and long-term attractiveness, as done in previous studies (e.g., Cantú et al., 2014), by adding mating context as an additional predictor. Results revealed that competitiveness and courtship interacted with mating context (ps < .003), in that men who showed more competitive or courtship behavior were rated as being less attractive for long-term relationships. Compared with our main findings, all other results remained virtually identical (e.g., only main effects of cycle phase, competitiveness, and courtship were significant). Details can be found in Tables S40 and S41 in the Supplemental Material. However, when we included within- and between-women hormone levels as predictors, rather than cycle phase, significant negative between-women-hormones three-way interactions occurred (Table S42 in the Supplemental Material), suggesting that men showing more competitiveness or courtship behavior were rated as being less attractive for long-term relationships. These long-term ratings were lower when between-women estradiol or progesterone levels were higher (see also https://osf.io/kyqn8/, pp. 218–257, for separate analyses of long-term attractiveness or relationship status). However, such complex exploratory higher order interaction effects should be replicated before being interpreted further.

Second, we were asked to repeat our main analyses with a difference score of sexual- and long-term attractiveness as the dependent variable (e.g., as done in Gangestad et al., 2007; Gangestad et al., 2004), even though both were highly correlated (r = .87, 95% CI = [.87, .88]). In these analyses, the main effects of cycle phase and courtship disappeared, whereas the main effect of competitiveness remained. Furthermore, analyses revealed significant positive two-way interactions of competitiveness and courtship with between-women estradiol and progesterone (ps < .005), in that women with higher mean estradiol and women with higher mean progesterone levels rated men showing more competitive or courtship behavior as more attractive for sexual than for long-term relationships. The effect of between-women estradiol was not significant in some further robustness checks (Table S47B in the Supplemental Material). However, additional analyses suggested that the interaction of between-women progesterone and the behavioral factors was moderated by relationship status, such that it was significant only for singles (ps < .001) but not for women in relationships (ps > .347). None of the interactions including cycle phase or within-women hormones were significant. Details can be found in Tables S44 to S47B in the Supplemental Material.

Third, we repeated our analyses using log-transformed hormone levels (but see Roney, 2019). Most results remained virtually identical. However, one negative interaction between courtship and between-women progesterone levels suggested that women with generally higher mean progesterone levels evaluated men who showed less courtship behavior as more attractive for long-term relationships (p = .018). Details can be found in Tables S48 to S50 in the Supplemental Material.

In summary, the null results regarding ovulatory-cycle shifts in mate preferences remained robust across all checks, regardless of whether cycle phase or within-women hormone levels were included in the models. Further, the main effects of competitiveness and courtship remained robust across checks. The main effect of cycle phase, but not main effects of between- or within-women hormones, was significant across the majority of robustness checks. Between-women mean progesterone negatively interacted with competitiveness and courtship in some but not all models, although follow-up analyses found significant effects only for singles. Results for between-women estradiol levels were rather mixed in terms of direction and significance.

Discussion

This study investigated ovulatory-cycle shifts in women's mate preferences for men's behaviors in a noncompetitive context. We included different cycle-phase predictors, direct hormonal measures, relationship status as a moderator variable, and a number of different behavioral cues. We did not observe compelling evidence for ovulatory-cycle shifts in women's mate preferences, either across different cycle phases or as predicted by within-women hormone levels. Effects were not influenced by women's relationship status and remained nonsignificant across multiple robustness checks. Results did not differ considerably between ratings of sexual and long-term attractiveness. Thus, the current study's results do not provide evidence for the GGOSH, contradicting previous findings for ovulatory-cycle shifts for men's behaviors (Gangestad et al., 2007; Gangestad et al., 2004; Lukaszewski & Roney, 2009), but are in line with recent nonreplications of cycle shifts in mate preferences for masculine faces, voices, and bodies (Jones, Hahn, Fisher, Wang, Kandrik, Han, et al., 2018; Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018; Marcinkowska, Galbarczyk, & Jasienska, 2018). Although the effect sizes for preference shifts might be small, our study had enough power to detect even rather small effects, employing the largest sample size so far, a within-participants design with four testing sessions, and high cycle-phase validity because of LH tests. Previous studies that have reported evidence for the GGOSH contain various issues in these regards that might have led to overestimation of effect sizes and false-positive results. However, the fact that we did not find support for the GGOSH does not mean that preference shifts do not exist in general. For example, preference shifts for other cue domains (e.g., odor) might be robust, and we do not know whether preference shifts for behaviors occur only under specific conditions (e.g., male intrasexual competition) or only for specific women (e.g., personality differences in preference shifts, influences of partner attractiveness). Given that the current sample is the same as in Jünger, Kordsmeyer, et al. (2018) and Jünger, Motta-Mena, et al. (2018), we cannot rule out that our reported null findings are sample specific and recommend replication in an independent sample.

Instead of preference shifts, shifts in women's general attraction to men were recently reported in the same data (Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018). Here, we observed partial evidence for this effect in that ratings for sexual as well as for long-term attractiveness were higher in the fertile phase, compared with the luteal phase. This effect was significant for the majority of robustness checks but not for models including relationship status and for some models including single behavioral cues (Table S12 in the Supplemental Material). Moreover, the lack of ties to hormone levels is a fairly significant limitation to this finding. Hence, we recommend independent replications, preferably in large, preregistered studies, to probe the robustness of this effect and further investigate under which conditions it occurs.

Moreover, although between-women mean hormone levels did not influence preferences for men's behaviors in our main analyses, they did in some robustness checks contrasting ratings of sexual and long-term attractiveness. In these analyses, women with higher mean progesterone levels rated men who displayed more competitive and courtship behavior as more attractive for sexual than for long-term attractiveness but as less attractive for long-term relationships. However, follow up analyses found significant effects only for singles. Given that these results occurred only in some robustness checks, not in our main analyses, we suggest that further research is needed to investigate the nature of between-women mean hormone effects on their mate preferences.

Further, future research should also aim to directly compare competing theories from the literature against the GGOSH—for example, cycle shifts as vestigial by-products of hormonal changes (Thornhill & Gangestad, 2015), motivational priority shifts (Roney, 2018), or the spandrels hypothesis of cycle shifts as a by-product of between-women hormonal differences (Havlíček, Cobey, Barrett, Klapilová, & Roberts, 2015).

Limitations

We note an important limitation that should be addressed in future research. Previous studies that found evidence for cycle-phase shifts in preferences for men's behaviors focused more directly on behaviors related to dominance and social presence within an intrasexual competitive context. In contrast, we used videos of an intersexual courtship context. It is possible that the behaviors assessed in competitive contexts (e.g., Gangestad et al., 2007; Gangestad et al., 2004) were better indicators of good genes, because they implied a willingness to risk confrontations with other men, whereas flirting with women in the absence of same-sex rivals may not carry similar implications. However, social behaviors such as dominance and social presence are somewhat stable across situations (Funder & Colvin, 1991; with an average reliability of .78 for dominance across situations). Hence, for example, a man who behaves dominantly in intrasexual competitive situations might also show more dominant behavior in flirting situations. Indeed, stimulus ratings showed a high interrater agreement for competitive behaviors and factor analysis yielded a competitiveness factor alongside a courtship factor, indicating that these behaviors could be consensually perceived from our stimulus material. Both factors had a significant effect on attractiveness ratings, suggesting that women do have preferences for men displaying more competitive and courtship behaviors even in the absence of a rival, although they seem not to shift across the cycle. Nevertheless, it remains unclear whether preference shifts would have been observable if women had watched and evaluated an intrasexual competitive scene between two men rather than an intersexual courtship context.

Conclusion

In conclusion, in the largest study conducted so far investigating possible cycle shifts in women's mate preferences for men's behaviors, we did not observe shifting mate preferences across the cycle. Thus, our findings are inconsistent with the GGOSH. It remains unclear whether women's general attraction to men shifts across the cycle and whether preference shifts for other cues are robust. Future studies combining rigorous methods and large sample sizes with precise preregistration will be crucial to further elucidate these issues.

Transparency

Action Editor: Steven W. Gangestad
Editor: D. Stephen Lindsay
Author Contributions
J. Stern and L. Penke developed the study concept and contributed to the study design. Testing and data collection were performed by J. Stern, who also performed data analysis and interpretation under the supervision of L. Penke. J. Stern drafted the manuscript and T. M. Gerlach and L. Penke provided critical revisions. All authors approved the final version of the manuscript for submission.
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
Funding
This research was funded by Deutsche Forschungsgemeinschaft (DFG) Project No. 254142454 / GRK 2070.
Open Practices
All data, analysis scripts, and instruction materials have been made publicly available at the Open Science Framework and can be accessed at https://osf.io/8ntuc/. The study was preregistered on the Open Science Framework (https://osf.io/m6pnz/). All analyses mentioned in the preregistration that are relevant to the present study are reported here; additional hypotheses were included in the preregistration, because it was written for a larger project. All reported analyses that were not preregistered are explicitly described as such. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619882022. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

TC

ORCID iD

Julia Stern https://orcid.org/0000-0001-8749-6392

Acknowledgments

We thank Ruben C. Arslan for valuable statistical advice, Christian Treffenstädt for assistance in the programming process, Kelly Gildersleeve for kindly providing a blueprint of an Excel sheet to calculate fertile windows, Alex Stern for comments on prior versions of this manuscript, and our research assistants Laura Botzet, Silvia Bradatsch, Paulina Lauer, and Laura Specker-Mattißen for data collection.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619882022

Notes

1. During the review process, it was questioned whether Hypothesis 1, as it was framed in our preregistration, related to an interaction effect rather than a main effect. Although we disagree with this interpretation, we agree that the wording in our preregistration was a bit ambiguous and could lead to different interpretations. Hence, besides running a number of robustness checks, we decided to cautiously interpret our findings regarding this hypothesis.

2. We were asked in the review process to include these variables and agree that they better match the kinds of behaviors for which previous studies reported cycle shifts.

References

Arslan, R. C., Schilling, K. M., Gerlach, T. M., & Penke, L. (2018). Using 26,000 diary entries to show ovulatory changes in sexual desire and behavior. *Journal of Personality and Social Psychology*. Advance online publication. doi:10.1037/pspp0000208

Back, M. D., Penke, L., Schmukle, S. C., Sachse, K., Borkenau, P., & Asendorpf, J. B. (2011). Why mate choices are not as reciprocal as we assume: The role of personality, flirting and physical attractiveness. *European Journal of Personality*, *25*, 120–132. doi:10.1002/per.806

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R.H. (2015). Parsimonious mixed models. *Cornell University ArXiv*. Retrieved from http://arxiv.org/abs/1506.04967

Blake, K. R., Dixson, B. J., O'Dean, S. M., & Denson, T. F. (2016). Standardized protocols for characterizing women's fertility: A data-driven approach. *Hormones and Behavior*, *81*, 74–83. doi:10.1016/j.yhbeh.2016.03.004

Cantú, S. M., Simpson, J. A., Griskevicius, V., Weisberg, Y. J., Durante, K. M., & Beal, D. J. (2014). Fertile and selectively flirty: Women's behavior toward men changes across the ovulatory cycle. *Psychological Science*, *25*, 431–438. doi:10.1177/0956797613508413

DeBruine, L. M., Hahn, A. C., & Jones, B. C. (2019). Does the interaction between partnership status and average progesterone level predict women's preferences for facial masculinity? *Hormones and Behavior*, *107*, 80–82. doi:10.1016/j.yhbeh.2018.12.004

Flowe, H. D., Swords, E., & Rockey, J. C. (2012). Women's behavioural engagement with a masculine male heightens during the fertile window: Evidence for the cycle shift hypothesis. *Evolution and Human Behavior*, *33*, 285–290. doi:10.1016/j.evolhumbehav.2011.10.006

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*, 773–794.

Gangestad, S. W., Garver-Apgar, C. E., Simpson, J. A., & Cousins, A. J. (2007). Changes in women's mate preferences across the ovulatory cycle. *Journal of Personality and Social Psychology*, *92*, 151–163. doi:10.1037/0022-3514.92.1.151

Gangestad, S. W., Haselton, M. G., Welling, L. L., Gildersleeve, K., Pillsworth, E. G., Burriss, R. P., . . . Puts, D. A. (2016). How valid are assessments of conception probability in ovulatory cycle research? Evaluations, recommendations, and theoretical implications. *Evolution and Human Behavior*, *37*, 85–96. doi:10.1016/j.evolhumbehav.2015.09.001

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral & Brain Sciences*, *23*, 573–587. doi:10.1017/S0140525X0000337X

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. *Psychological Science*, *15*, 203–207. doi:10.1111/j.0956-7976.2004.01503010.x

Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (2005). Adaptations to ovulation: Implications for sexual and social behavior. *Current Directions in Psychological Science*, *14*, 312–316. doi:10.1111/j.0963-7214.2005.00388.x

Gildersleeve, K., Haselton, M. G., & Fales, M. R. (2014). Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological Bulletin*, *140*, 1205–1259. doi:10.1037/a0035438

Gildersleeve, K. A., Haselton, M. G., Larson, C. M., & Pillsworth, E. G. (2012). Body odor attractiveness as a cue of impending ovulation in women: Evidence from a study using hormone-confirmed ovulation. *Hormones and Behavior*, *61*, 157–166. doi:10.1016/j.yhbeh.2011.11.005

Grebe, N. M., Thompson, M. E., & Gangestad, S. W. (2016). Hormonal predictors of women's extra-pair vs. in-pair sexual attraction in natural cycles: Implications for extended sexuality. *Hormones and Behavior*, *78*, 211–219. doi:10.1016/j.yhbeh.2015.11.008

Havlíček, J., Cobey, K. D., Barrett, L., Klapilová, K., & Roberts, S. C. (2015). The spandrels of Santa Barbara? A new perspective on the peri-ovulation paradigm. *Behavioral Ecology*, *26*, 1249–1260. doi:10.1093/beheco/arv064

Jones, B. C., Hahn, A. C., & DeBruine, L. M. (2019). Ovulation, sex hormones, and women's mating psychology. *Trends in Cognitive Sciences*, *23*, 51–62. doi:10.1016/j.tics.2018.10.008

Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., & DeBruine, L. M. (2018). General sexual desire, but not desire for uncommitted sexual relationships, tracks changes in women's hormonal status. *Psychoneuroendocrinology*, *88*, 153–157. doi:10.1101/155788

Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., . . . DeBruine, L. M. (2018). No compelling evidence that preferences for facial masculinity track changes in women's hormonal status. *Psychological Science*, *29*, 996–1005. doi:10.1177/0956797618760197

Jünger, J., Kordsmeyer, T., Gerlach, T. M., & Penke, L. (2018). Fertile women evaluate male bodies as more attractive, regardless of masculinity. *Evolution and Human Behavior*, *39*, 412–423. doi:10.1016/j.evolhumbehav.2018.03.007

Jünger, J., Motta-Mena, N. V., Cardenas, R., Bailey, D., Rosenfield, K. A., Schild, C., . . . Puts, D. A. (2018). Do women's preferences for masculine voices shift across the ovulatory cycle? *Hormones and Behavior*, *106*, 122–134. doi:10.1016/j.yhbeh.2018.10.008

Lukaszewski, A. W., & Roney, J. R. (2009). Estimated hormones predict women's mate preferences for dominant personality traits. *Personality and Individual Differences*, *47*, 191–196. doi:10.1016/j.paid.2009.02.019

Marcinkowska, U. M., Galbarczyk, A., & Jasienska, G. (2018). La donna è mobile? Lack of cyclical shifts in facial symmetry, and facial and body masculinity preferences—a hormone based study. *Psychoneuroendocrinology*, *88*, 47–53. doi:10.1016/j.psyneuen.2017.11.007

Marcinkowska, U. M., Kaminski, G., Little, A. C., & Jasienska, G. (2018). Average ovarian hormone levels, rather than daily values and their fluctuations, are related to facial preferences among women. *Hormones and Behavior*, *102*, 114–119. doi:10.1016/j.yhbeh.2018.05.013

Morrison, E. R., Clark, A. P., Gralewski, L., Campbell, N., & Penton-Voak, I. S. (2010). Women's probability of conception is associated with their preference for flirtatious but not masculine facial movement. *Archives of Sexual Behavior*, *39*, 1297–1304. doi:10.1007/s10508-009-9527-1

Muñoz-Reyes, J. A., Iglesias-Julios, M., Martín-Elola, C., Losada-Pérez, M., Monedero, I., Pita, M., & Turiégano, E. (2014). Changes in preference for male faces during the menstrual cycle in a Spanish population. *Annals of Psychology*, *30*, 667–675. doi:10.6018/analesps.30.2.145221

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*, 1113–1135. doi:10.1037/0022-3514.95.5.1113

Pillsworth, E. G., & Haselton, M. G. (2006). Women's sexual strategies: The evolution of long-term bonds and extrapair sex. *Annual Review of Sex Research*, *17*, 59–100. doi:10.1080/10532528.2006.10559837

Roney, J. R. (2018). Functional roles of gonadal hormones in human pair bonding and sexuality. In O. C. Schultheiss & P. H. Mehta (Eds.), *Routledge international handbook of social neuroendocrinology* (pp. 239–255). New York, NY: Routledge.

Roney, J. R. (2019). On the use of log transformations when testing hormonal predictors of cycle phase shifts: Commentary on Gangestad, Dinh, Grebe, Del Giudice and Emery Thompson (2019). *Evolution and Human Behavior*, *40*, 526–530. doi:10.1016/j.evolhumbehav.2019.08.006

Roney, J. R., Lukaszewski, A. W., & Simmons, Z. L. (2007). Rapid endocrine responses of young men to social interactions with young women. *Hormones and Behavior*, *52*, 326–333. doi:10.1016/j.yhbeh.2007.05.008

Roney, J. R., Mahler, S. V., & Maestripieri, D. (2003). Behavioral and hormonal responses of men to brief interactions with women. *Evolution and Human Behavior*, *24*, 365–375. doi:10.1016/S1090-5138(03)00053-9

Roney, J. R., & Simmons, Z. L. (2013). Hormonal predictors of sexual motivation in natural menstrual cycles. *Hormones and Behavior*, *63*, 636–645. doi:10.1016/j.yhbeh.2013.02.013

Roney, J. R., & Simmons, Z. L. (2016). Within-cycle fluctuations in progesterone negatively predict changes in both in-pair and extra-pair desire among partnered women. *Hormones and Behavior*, *81*, 45–52. doi:10.1016/j.yhbeh.2016.03.008

Schultheiss, O. C., & Stanton, S. J. (2009). Assessment of salivary hormones. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in social neuroscience* (pp. 17–44). New York, NY: Guilford Press.

Suschinsky, K. D., Bossio, J. A., & Chivers, M. L. (2014). Women's genital sexual arousal to oral versus penetrative heterosexual sex varies with menstrual cycle phase at first exposure. *Hormones and Behavior*, *65*, 319–327. doi:10.1016/j.yhbeh.2014.01.006

Thornhill, R., & Gangestad, S. W. (2015). The functional design and phylogeny of women's sexuality. In T. K. Shackelford & R. D. Hansen (Eds.), *The evolution of sexuality* (pp. 149–184). Cham, Switzerland: Springer.

Treffenstaedt, C., & Wiemann, P. (2018). Alfred - a library for rapid experiment development (Version 0.2b5) [Computer software]. doi:10.5281/zenodo.1437220

Wood, W., Kressel, L., Joshi, P. D., & Louie, B. (2014). Meta-analysis of menstrual cycle effects on women's mate preferences. *Emotion Review*, *6*, 229–249. doi:10.1177/1754073914523073