Hormones and Behavior 106 (2018) 122–134

Contents lists available at ScienceDirect

Hormones and Behavior

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

Do women's preferences for masculine voices shift across the ovulatory cycle?

Julia Jünger[a,*], Natalie V. Motta-Mena[b], Rodrigo Cardenas[b], Drew Bailey[c], Kevin A. Rosenfield[d], Christoph Schild[e], Lars Penke[a,1], David A. Puts[d,1]

[a] Department of Psychology, Leibniz Science Campus Primate Cognition, University of Goettingen, Gosslerstrasse 14, 37073 Goettingen, Germany
[b] Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA
[c] School of Education, University of California, Irvine, Irvine, CA 92697, USA
[d] Department of Anthropology, Center for Brain, Behavior and Cognition, Pennsylvania State University, University Park, PA 16802, USA
[e] Department of Psychology, University of Copenhagen, Øster Farimagsgade 2A, 1353 Copenhagen, Denmark

ARTICLE INFO

Keywords:
Steroid hormones
Fertility
Attractiveness
Voice masculinity
Mate preferences
Ovulatory cycle

ABSTRACT

Are estrous mate preference shifts robust? This question is the subject of controversy within human evolutionary sciences. For nearly two decades, mate preference shifts across the ovulatory cycle were considered an important feature of human sexual selection, directing women's attention toward mates with indicators of "good genes" in their fertile phase, when conception is possible. However, several recent studies on masculine faces, bodies and behaviors did not find evidence supporting this account, known as the *good genes ovulatory shift hypothesis*. Furthermore, evidence that preferences for masculine characteristics in men's voices are related to women's cycle phase and hormonal status is still equivocal. Here, we report two independent within-subject studies from different labs with large sample sizes (N = 202 tested twice in Study 1; N = 157 tested four times in Study 2) investigating cycle shifts in women's preferences for masculine voices. In both studies, hormonal status was assessed directly using salivary assays of steroid hormones. We did not find evidence for effects of cycle phase, conception risk, or steroid hormone levels on women's preferences for masculine voices. Rather, our studies partially provide evidence for cycle shifts in women's general attraction to men's voices regardless of masculine characteristics. Women's relationship status and self-reported stress did not moderate these findings, and the hormonal pattern that influences these shifts remains somewhat unclear. We consider how future work can clarify the mechanisms underlying psychological changes across the ovulatory cycle.

1. Introduction

Whether women's mate preferences change across the ovulatory cycle has been a central question in the human evolutionary sciences over the last decades. While it seems robust that women experience greater levels of sexual desire and interest when fertile (e.g. Arslan et al., 2018; Jones et al., 2018b; Roney and Simmons, 2013, 2016), it remains unclear if any mate preference shifts exist. Recent studies have cast doubt on the existence of cycle shifts in preferences for masculine faces, bodies and behavioral displays (e.g. Jones et al., 2018a; Jünger et al., 2018a, 2018b; Marcinkowska et al., 2016, 2018a; Muñoz-Reyes et al., 2014), and called attention to methodological criticisms of previous studies. Inconsistencies in the literature are reflected, for instance, in the outcome of two recent meta-analyses, which reached opposite conclusions about whether women's ovulatory cycle phase reliably influences their judgments of men's attractiveness (Gildersleeve et al., 2014a; Wood et al., 2014). In the current manuscript, we tested cycle shifts in women's preferences for masculine voices in two large within-subjects studies from different labs, using natural as well as manipulated voice recordings as stimuli, and also examined hormone concentrations and possible moderator variables.

1.1. Theoretical background

Systematic changes in women's sexual interests across the ovulatory cycle have been intensively investigated. In several studies, women experienced heightened sexual interest during their fertile phase, compared to their non-fertile phases (most notably the luteal phase).

* Corresponding author.
E-mail address: julia.juenger@psych.uni-goettingen.de (J. Jünger).
[1] Lars Penke and David A. Puts share the last authorship.

https://doi.org/10.1016/j.yhbeh.2018.10.008
Received 5 June 2018; Received in revised form 15 October 2018; Accepted 16 October 2018

More precisely, when fertile, women reported higher extra-pair desire (Gangestad et al., 2002, 2005; Grebe et al., 2016; Haselton and Gangestad, 2006; Shimoda et al., 2018), in-pair as well as extra-pair desire (Arslan et al., 2018; Roney and Simmons, 2016) or general sexual desire (Jones et al., 2018b; Roney and Simmons, 2013), which was also found to be linked to their ovarian hormone levels (Jones et al., 2018b; Roney and Simmons, 2013, 2016). To describe differences in sexual psychology and behavior on fertile vs. non-fertile days, Thornhill and Gangestad (2008) proposed the concept of *dual sexuality*. While sexual behavior outside the fertile phase may have evolved for pair-bonding purposes (Grebe et al., 2016), the most direct benefit for sexual behavior within the fertile phase is conception (Roney and Simmons, 2013). Women are thus predicted to change their mate preferences across the ovulatory cycle. When fertile, their sexual interests should hypothetically be directed preferentially toward mates who possess indicators of high genetic quality to achieve fitness benefits for their offspring (Haselton and Gangestad, 2006). In contrast, sexual interests within the non-fertile phases should be directed to long-term mates with a high potential and willingness to provide parental effort (Gildersleeve et al., 2014a; Thornhill and Gangestad, 2015). Since ovulatory shifts are predicted to aid in obtaining good genes, potentially from extra-pair copulations (Pillsworth and Haselton, 2006), we will further call this concept the *good genes ovulatory shift hypothesis* (GGOSH; Arslan et al., 2018).

Previous studies found evidence for the GGOSH: in the fertile (late follicular) phase, women reportedly shift their preferences toward putative indicators of men's genetic quality, including masculine, dominant-appearing faces (Penton-Voak et al., 1999; Penton-Voak and Perrett, 2000), voices (Feinberg et al., 2006; Pisanski et al., 2014; Puts, 2005, 2006), bodies (Gangestad et al., 2007; Little et al., 2007), odor (Gangestad and Thornhill, 1998; Havlíček et al., 2005; Thornhill et al., 2013) and behavioral displays (Gangestad et al., 2004, 2007).

However, some purported indicators of good genes are controversial because reported findings challenge the hypothesis that they actually signal heritable fitness benefits and immunocompetence (Scott et al., 2012, 2014; Simmons et al., 2011; Wood et al., 2014). Additionally, the GGOSH itself has been questioned in recent research (Havlíček et al., 2015; Roney and Simmons, 2017). Moreover, several studies raise skepticism about the robustness of preference shifts because of higher powered null replications of prior findings for masculine or symmetrical faces (Harris, 2011, 2013; Jones et al., 2018a; Marcinkowska et al., 2016, 2018a; Muñoz-Reyes et al., 2014; Peters et al., 2009), bodies (Jünger et al., 2018b; Marcinkowska et al., 2018a; Peters et al., 2009), and behaviors (Jünger et al., 2018a). Furthermore, two large recent studies suggest that women's attraction to men in general, rather than their mate preferences, shifts across the ovulatory cycle (Jünger et al., 2018a, 2018b). Additionally, two meta-analyses analyzing mostly the same datasets (Gildersleeve et al., 2014a; Wood et al., 2014) came to opposite conclusions regarding ovulatory cycle shifts in women's mate preferences, although the methods of Wood et al. (2014) have been criticized (Gildersleeve et al., 2014b; Motta-Mena and Puts, 2017). Given this mixed pattern of findings and the centrality of putative ovulatory shifts in current theorizing about human sexual selection, it is clear that there is an urgent need for further research to determine a) the nature of any shifts in women's preferences for masculine features over the ovulatory cycle, and b) the hormonal correlates of any cycle shifts in women's mate preferences.

1.1.1. Preference shifts for voice masculinity

Human voices are highly sexually dimorphic. Sexual dimorphism in vocal anatomy may have been favored by sexual selection because low frequency male vocalizations intimidate rivals and/or attract females (Puts et al., 2016). Masculine voices are characterized by both a lower fundamental frequency and lower, more closely spaced formant frequencies. Fundamental frequency (F_0), the rate of vocal fold vibration during phonation, is the acoustic measure closest to what we perceive as pitch. In males, F_0 is related to testosterone throughout pubertal development (Butler et al., 1989; Harries et al., 1997, 1998; Hodges-Simeon et al., 2015) and during adulthood (Dabbs and Mallinger, 1999; Evans et al., 2008; Puts et al., 2012, 2016). Lower and more closely spaced formant (resonant) frequencies indicate a longer vocal tract and have also been shown to independently increase perceived masculinity (Collins, 2000) and dominance (Cheng et al., 2016; Puts et al., 2006, 2007; Tusing and Dillard, 2000). In such research, formants are often summarized by the composite metric formant dispersion (D_f, the average distance between consecutive formant frequencies computed across the first N, usually four, formants). Hence, the GGOSH would suggest that fertile women should be especially attracted to men with lower F_0 and lower D_f. If preference shifts across the ovulatory cycle for masculine voices occur, then they should be mediated by ovarian hormonal changes. Previous studies report that estradiol, progesterone and the estradiol-to-progesterone-ratio (henceforth E/P) are likely candidates for mediating changes in women's mate preferences for voice masculinity over the cycle (e.g. Feinberg et al., 2006; Pisanski et al., 2014; Puts et al., 2013). Estradiol peaks in women's late follicular (fertile) phase and exhibits a smaller increase during the mid-luteal phase. Progesterone levels are usually lower throughout the follicular phase and increase in the luteal phase.

Surprisingly, although null effects for masculine voices in previous studies were attributed to an underpowered analysis (Gildersleeve et al., 2014a), there is a lack of published large, high-powered, within-subject studies investigating preference shifts for masculine voices. However, there are three prior studies that investigated possible cycle shifts for masculine voices and interpreted their results as evidence for mate preference shifts across the ovulatory cycle: Puts (2005) conducted a between-subject study with $N = 136$ female participants ($n = 38$ in the fertile group, $n = 98$ in the non-fertile group) who rated the attractiveness of men's voice recordings, manipulated (raised or lowered) in both F_0 and D_f (see also Puts et al., 2006). Women's conception risk was assessed as a continuous measure via backward counting, but then participants were categorized to cycle phases. Results showed significant cycle shifts: Women preferred men's lowered pitch voices only when they rated them in their fertile phase and for potential short-term relationships ($p = .020$).

Feinberg et al. (2006) reported a within-subjects study with $N = 26$ female participants who completed four to six testing sessions resulting in a total of 41 fertile phase sessions ($n = 25$) and 86 non-fertile phase sessions ($n = 25$). However, average scores within each phase were used if a woman was tested more than one time per cycle phase. Cycle phase (fertile vs. non-fertile) was classified via backward counting. Participants rated the general attractiveness of voice recordings that were manipulated in voice pitch and formant frequencies. Notably, cycle shifts for masculine voices were reported only when estrone-3-glucuronide concentrations (E3G, the primary urinary metabolite of estradiol) were included as a covariate in the analyses ($p = .012$), showing that shifts are stronger for women with lower E3G concentrations. Effects were not significant when pregnanediol-3-glucuronide concentrations (P3G, the primary urinary metabolite of pregnanediol) was included as a covariate ($p = .063$), or in an analysis without covariates ($p = .253$).

Using a within-subject design with five weekly test sessions per participant, Pisanski et al. (2014) reported that changes in estradiol, but not progesterone, trended toward predicting stronger preferences for manipulated masculine voices in a sample of 62 women ($p = .055$). Crucially, this effect did not reach significance, and the authors also observed no significant effect of progesterone, testosterone or E/P on preferences for manipulated masculine voices.

Taken together, these studies do not provide strong evidence for cycle shifts in preferences for masculine voices. As Gildersleeve et al. (2014a) noted, sample sizes tended to be small, with limited test trials in the experimental designs (e.g. 12 trials; Pisanski et al., 2014). In addition, averaging participant ratings of voices (Feinberg et al., 2006;

Pisanski et al., 2014) further reduces the statistical power. Moreover, recent research has pointed out additional methodological issues underlying prior cycle shift studies (Blake et al., 2016; Gangestad et al., 2016; Harris, 2013; Shimoda et al., 2018). First, although backward counting was used as a superior means of estimating cycle phase compared to forward counting (Gangestad et al., 2016), authors did not use luteinizing hormone (LH) urine tests to validate the fertile phase estimates, even though a preovulatory surge of LH clearly demarcates distinct cycle phases. Second, the only study that reported a significant preference shift for masculine voices (Puts, 2005) lacks a direct assessment of steroid hormones to analyze mediating effects. Third, effect sizes or 95% confidence intervals of the observed preference shifts were not reported, which makes the reported effects harder to interpret. One would expect cycle shift effect sizes to be rather small (Jünger et al., 2018b), but since previous studies worked with relatively small sample sizes, they may not have had the statistical power to reveal such effects. Consequently, published effects might have been false positives or due to publication bias. Fourth, previous studies used manipulated voices or a combination of manipulated and natural voice recording (Puts, 2006) rather than natural voice recordings alone. It is up for debate to what degree computer-manipulated voices have ecological validity, but in any case natural voices should also be used to ensure that results can be transferred to real-life mate preferences. Considering all of these potential methodological problems and the incongruence in reported results, the associations between women's ovulatory cycle, steroid hormone levels, and mate preferences for masculine voices remains unclear.

1.2. Overview over the present studies

In the present studies, we aim to clarify a) whether women's attraction to and/or preferences for masculine voices shift across the ovulatory cycle, b) which hormonal changes might underlie these shifts, and c) which moderators influence these shifts. In what follows, we report two large, independent studies from different labs at two different institutions. Both studies employed a within-subjects design with large sample sizes, direct hormonal assessments across one (Study 1) or two (Study 2) ovulatory cycles, and backward counting methods to estimate women's fertility. Study 1 included ovarian hormones (estradiol, progesterone and their ratio), and used voice recordings that were manipulated in F_0 and D_f, while Study 2 included estimated cycle phase (validated with LH tests) as a dichotomous measure of fertility, ovarian hormones as possible mediators, and used natural stimuli. Women's relationship status and self-reported stress are tested as possible moderator variables of ovulatory cycle shifts in women's preferences in Study 2. Additionally, Study 2 was pre-registered[2]; open data and material for both studies can be found at the Open Science Framework (https://osf.io/a6byr).

2. Study 1

2.1. Method

2.1.1. Participants

A total of 202 women ages 18 to 27 years (M = 19.56 years; SD = 1.59) participated in this study as part of a larger study at Michigan State University. All participants were exclusively or predominantly heterosexual and normally cycling (e.g. not taking any hormonal contraception[3]). They were recruited via print advertisements and the MSU Psychology Department undergraduate subject pool. Informed consent was obtained from participants using procedures approved by the Institutional Review Board of Michigan State University. Participants were scheduled for two laboratory sessions according to self-reported ovulatory cycle length and date of the beginning of last menstrual onset. One laboratory session was scheduled within one day of expected peak estradiol production during the fertile phase, and the other session was scheduled within two days of expected peak progesterone production (mid-luteal phase), as follows: First, information on women's average cycle length and the beginning day of their last menstrual bleeding was collected online before the participant's first session was scheduled. Second, we used this information to estimate the date of their next midcycle LH peak (assuming that the LH peak occurs 15 days prior to the beginning day of their next menstrual bleeding). Third, we used the methods in Puts (2006) to estimate the days of peak estradiol and progesterone levels (approximately the day before the estimated LH peak and 7 days after, respectively). Finally, we scheduled their follicular phase session within one day of their presumptive estradiol peak (i.e., the day of, the day before, or the day after), and we scheduled their luteal phase session within 2 days of their presumptive peak in progesterone. A t-test between E/P across phases showed that E/P was significantly higher in the fertile phase, compared to the luteal phase (t_{48} = 5.70, p < .001, $95\%CI$ = [0.19; 0.40], d = 0.84), validating the scheduling procedure. Session order was counterbalanced across participants, such that half of the participants started in their presumed fertile phase and the other half in their presumed luteal phase. Sessions occurred between 1:00 pm and 4:00 pm in order to minimize the influence of circadian hormonal fluctuations.

2.1.2. Saliva collection and hormonal analysis

Approximately 9 ml of saliva was collected from each participant in sodium azide-treated polystyrene test tubes. Participants were asked not to eat, drink (with the exception of plain water), smoke, chew gum, or brush their teeth for at least 1 h prior to each session to avoid contamination of saliva samples. To stimulate saliva flow, participants rinsed their mouths with water, and were provided with a piece of sugar-free Trident chewing gum (inert in salivary hormone assays). The tube was capped and left upright at room temperature for 18–24 h to allow mucins to settle. Tubes were then frozen at −20 °C until analysis by the Neuroendocrinology Assay Laboratory at the University of Western Ontario, Canada. Progesterone was assayed using 125I Coat-A-Count assay kits (Diagnostic Products Corporation, Los Angeles, CA) modified for use with saliva (e.g. as in Hampson et al., 2005; Oinonen and Mazmanian, 2007). Similar to previous research (e.g., Finstad et al., 2009), estradiol was assayed using 125I Ultra-Sensitive E2 RIA DSL-4800 kit (Diagnostic Systems Laboratories, Webster, TX) modified for use with saliva. Each sample was assayed twice to verify replicability, and average hormone levels for each sample were used in our analyses. Assay sensitivities were 0.65 pg/ml and 5 pg/ml, and intra-assay coefficients of variation (CV) were 5.1% and 10.7%, for estradiol and progesterone, respectively. Seven participants were excluded from subsequent hormone analysis due to not providing a saliva sample in both sessions, leaving a total of 195 women.[4] Hormone values were positively skewed and thus log10-transformed.

[2] This pre-registration (can be found at https://osf.io/egjwv) also contained further hypotheses that are not part of the present paper.

[3] Because other conditions, such as pregnancy or endocrine disorders, can also greatly affect women's hormone levels, we scanned our participant's hormone levels for arbitrary values. All values were in line with previously published level ranges from studies with daily hormone assessments (Connor et al., 1982; Marcinkowska et al., 2018a, 2018b) and below progesterone levels that might indicate pregnancy (Connor et al., 1982), suggesting that current pregnancy or endocrine disorders were rather unlikely.

[4] Excluding another n = 15 women who reported cycle lengths < 25 days or > 35 days did not change any results.

2.1.3. Voice recordings and manipulation

Six male voices were recorded as described in Wolff and Puts (2010), reading an excerpt from a standard voice passage (Fairbanks, 1960). Each voice recording was analyzed and manipulated using Praat (v. 4.4.06; Boersma and Weenink, 2006). Pitch floor and ceiling were 75 Hz and 300 Hz, in accordance with programmers' recommendations; otherwise default settings were used. Formants were measured using the long-term average spectrum (González, 2004; Xue and Hao, 2003), and D_f was computed by taking the average distance between each of the first four formants (Fitch, 1997). For unmanipulated voices, mean F_0 was 109.9 (range = 97.8–122.1, SD = 10.0), and mean D_f was 1003.5 (range = 941.7–1072.7, SD = 51.6). For the current study, each of the six voices was raised and lowered using just-noticeable-difference (JND) parameters from Puts et al. (2007): F_0 was raised and lowered 1.2 semitones, while D_f was manipulated with a 4% change. Thus, from each of the original voices, four versions were produced: raised F_0, lowered F_0, raised D_f, and lowered D_f, for a total of 24 voice recordings. These recordings were distributed into two stimulus sets of 12 recordings, each set comprising 6 raised F_0 with 6 lowered F_0 and 6 raised D_f with 6 lowered D_f.

2.1.4. Procedure

Each participant was seated at a computer station and provided Sennheiser HD280 Pro headphones. The experiment was computerized and participants were instructed using the following script:

> "*Please put on the headphones. You are about to hear voice recordings from several men. Please rate how attractive you think each man would be for a short-term, purely sexual relationship, such as a one-night stand (even if you are not interested in such a relationship).*"

After listening to each voice recording, participants rated each voice on a 10-point Likert-scale, from "extremely attractive" (coded as 1) to "extremely unattractive" (coded as 10). We reverse-coded the scale for our analyses for an easier understanding of the results. In order to reduce the chance that participants would recognize the voices in each of the voice manipulations, the voice clips were presented in two separate blocks, with an unrelated memory task between each block. Each block consisted of 12 trials with 6 F_0 and 6 D_f manipulations and each speaker represented by one F_0 manipulation and one D_f manipulation. Hence, if for example, in the first block the raised F_0 manipulation was presented for a particular speaker, then the lowered F_0 manipulation was presented in the second block Participants rated all 12 recordings during both laboratory visits in the same order.

2.1.5. Statistical analyses

All analyses in the current manuscript were calculated with the statistic software R 3.4.0 (R Core Team, 2016). The following packages were used: lme4 1.1-13 (Bates et al., 2014), lmerTest 2.0-33 (Kuznetsova et al., 2015), psych 1.7.5 (Revelle, 2016), dplyr (Wickham, 2011).

2.2. Results

2.2.1. Ovulatory cycle shifts in women's mate attraction

First, we tested whether ratings were generally related to ovarian hormone levels or estimated conception risk,[5] independent of voice manipulations, in three separate models. All models included attractiveness ratings as the dependent variable, and a random intercept per female rater as well as for male stimulus. Model 1 included estradiol (E) and progesterone (P), and Model 2 included E/P as predictors.[6] Results show no effect of estradiol or E/P, but importantly, a significant negative effect of progesterone, suggesting higher ratings on average when progesterone levels were lower (Table 1).

2.2.2. Ovulatory cycle shifts in women's mate preferences for masculinized voices

Next, we tested if participants showed preference shifts across the ovulatory cycle for voice pitch or formant dispersion across six separate models (discussed below as Models 3 through 6). Again, female raters and male stimuli were treated as random effects. The first two models included women's hormone levels (estradiol and progesterone), voice manipulation (masculinized vs. feminized F_0 in Model 3, D_f in Model 4), as well as their interaction as fixed effects. Then, we additionally calculated two models including E/P, voice manipulation (masculinized vs. feminized F_0 in Model 5, D_f in Model 6), as well as their interaction as fixed effects. Analyses revealed no significant interactions between hormone levels and F_0 or D_f manipulation (Tables 2 and 3), indicating no hormonal regulated preference shifts. Additionally, there were no significant main effects of D_f manipulation, but significant main effects of F_0 (Models 3 and 5), showing that voices with masculinized voice pitch were rated as more attractive than the same voices with feminized voice pitch. For hormone levels, we found a significant negative main effect for progesterone in Model 3 (with manipulated F_0) but not in Model 4 (with manipulated D_f), showing that ratings were higher when progesterone was lower. We, again, did not find a significant effect for estradiol in Model 3 or Model 4. Additionally, we found a significant main effect of E/P in Model 5 (with manipulated F_0) but not in Model 6 (with manipulated D_f), showing that ratings were higher when E/P was higher.

2.2.3. Robustness checks

We conducted further analyses to test the robustness of our results. To ascertain that our results were not driven by order effects of testing sessions or participants' age, we entered session number and participant age in all of our models. The main effect of progesterone from Model 1 disappeared, but the one from Model 3 and the main effect of E/P from Model 5 remained significant. Moreover, there was a main effect of session number, indicating that ratings were on average higher in the second session (p = .02). However, all other results remained virtually identical (significant main effect for F_0 as well as all non-significant effects) and can be found in the supplement (Tables S2–S7). Next, according to a possibly occurring carryover effect of women's hormonal state in the first session that might influencing the ratings in the second session (Wallen and Rupp, 2010), we repeated all analyses including an interaction between session number and hormone levels. Results revealed no interaction between session number and estradiol levels (p = .91) or session number and E/P (p = .15), but a significant interaction between session number and progesterone levels (p = .02), indicating that ratings were higher in the second session, only when progesterone levels were lower. However, this interaction was not robust in all models. Importantly, all interactions between hormone levels and masculine cues remained non-significant, details can be found in the supplement (Tables S8–S11).

3. Study 2

Study 2 was conducted at the University of Goettingen, Germany, independently from Study 1, and differed from Study 1 in several ways. First, Study 2 used unmanipulated voice recordings as stimuli, which enabled us to explore preferences for other acoustic parameters,

[5] Methods and results for the conception risk analyses can be found in the supplementary material. Ratings did not differ with variation in women's estimated conception risk, no interaction between F_0 or D_f manipulations and estimated conception risk were observed.

[6] We decided to analyze the effect of hormones on ratings in two separate models because of possible problems of multicollinearity (r = 0.61 for estradiol and E/P; r = −0.16 for progesterone and E/P).

Table 1
Multilevel regression analyses of attractiveness ratings as a function of estradiol and progesterone (Model 1) or E/P (Model 2).

	γ	*SE*	*t*	*p*	*95%CI*
Estradiol	−0.17	0.11	−1.53	.127	[−0.39; 0.05]
Progesterone	−0.23	0.10	−2.25	.024	[−0.43; −0.03]
E/P	0.05	0.08	0.63	.529	[−0.11; 0.22]

Note. All variables had 8820 observations, (195 participants × 2 test sessions × 12 stimuli × 2 masculinity manipulations − missing values).

Table 2
Multilevel regression analyses of attractiveness ratings as a function of estradiol and progesterone levels and manipulated voice pitch (Model 3) or formant dispersion (Model 4).

	γ	*SE*	*t*	*p*	*95%CI*
Voice pitch model					
F_0	−1.75	0.33	−5.29	< .001	[−2.40; −1.10]
Estradiol	−0.01	0.18	−0.03	.975	[−0.37; 0.36]
Progesterone	−0.52	0.17	−3.16	.002	[−0.84; −0.20]
F_0 × estradiol	−0.20	0.22	−0.92	.358	[−0.64; 0.23]
F_0 × progesterone	0.18	0.19	0.93	.354	[−0.20; 0.55]
Formant model					
D_f	−0.30	0.32	−0.93	.353	[−0.92; 0.33]
Estradiol	−0.24	0.18	−1.35	.178	[−0.59; 0.11]
Progesterone	−0.02	0.16	−0.94	.347	[−0.47; 0.16]
D_f × estradiol	−0.00	0.21	−0.00	.997	[−0.42; 0.41]
D_f × progesterone	0.20	0.18	1.06	.288	[−0.17; 0.56]

Note. F_0 = fundamental frequency (voice pitch), D_f = formant dispersion. All variables in voice pitch model had 4416 observations, formant model 4404 observations (each 195 participants × 2 test sessions × 12 stimuli − missing values).

Table 3
Multilevel regression analyses of attractiveness ratings as a function of E/P and manipulated voice pitch (Model 5) or formant dispersion (Model 6).

	γ	*SE*	*t*	*p*	*95%CI*
Voice pitch model					
F_0	−1.77	0.26	−6.82	< .001	[−2.28; −1.26]
E/P	0.29	0.14	2.11	.034	[0.02; 0.56]
F_0 × E/P	−0.19	0.16	−1.14	.253	[−0.51; 0.13]
Formant model					
D_f	−0.13	0.25	−0.53	.599	[−0.62; 0.36]
E/P	−0.02	0.13	−0.14	.893	[−0.28; 0.25]
D_f × E/P	−0.12	0.16	−0.74	.460	[−0.43; 0.19]

Note. F_0 = fundamental frequency (voice pitch), D_f = formant dispersion. All variables in voice pitch model had 4416 observations, formant model 4404 observations (each 195 participants × 2 test sessions × 12 stimuli − missing values).

including jitter and shimmer (cycle-to-cycle variation in F_0 and amplitude, respectively), which are associated with pathological voice quality (Dejonckere et al., 1996; Michaelis et al., 1998). Second, baseline testosterone levels of the men who provided the voice stimuli were assessed along with the other vocal cues. This provided a direct test of whether preference shifts occur for men with higher baseline testosterone levels, which are generally found to be negatively associated with F_0 (Butler et al., 1989; Dabbs and Mallinger, 1999; Harries et al., 1997, 1998). Third, in addition to estradiol and progesterone, participants' testosterone and cortisol levels were also assessed. Like estradiol, testosterone can show mid-cycle peaks and has been found to predict women's preferences for masculine faces (Bobst et al., 2014; Welling et al., 2007). Recent research also suggests that cortisol and psychological stress should be measured in studies on hormones and female mate preferences. Stress elevates cortisol levels (Herrera et al., 2016), which may inhibit estradiol production in young women (Roney and Simmons, 2015) and decrease women's preferences for male facial masculinity (Ditzen et al., 2017, but see Jones et al., 2018a). Fourth, we ascertained women's relationship status. Recent studies reported ovulatory cycle shifts in attraction to men (Jünger et al., 2018a, 2018b) and in sexual desire (Roney and Simmons, 2016) that were evident only in partnered women. Furthermore, partnered women were found to be more likely to have sexual fantasies about men other than their primary partner (Gangestad et al., 2002), rate the odor of dominant men as sexy (Havlíček et al., 2005), and report stronger masculinity preferences than singles (Jones et al., 2018a). By contrast, Jones et al. (2018b) reported no evidence for a moderating effect of women's relationship status on general sexual desire. The lack of converging evidence in the literature emphasizes the need for further analyses to evaluate the influence of women's relationship status on cycle shifts in preferences and attraction. Fifth, we used cycle phase (validated with LH tests) as a categorical measure, and all participants were investigated in four testing sessions across two ovulatory cycles each (see below for detailed methods). Sixth, besides assessing sexual attractiveness ratings, we also assessed long-term attractiveness ratings for all stimuli.

3.1. Pre-registered hypotheses and research questions

Following previous findings of ovulatory cycle shifts in mate preferences, we hypothesize that women in the fertile phase, compared to their luteal phase, will evaluate men's voices as more attractive for short-term sexual relationships (Hypothesis 1). This effect should be mediated by changes in the steroid hormones estradiol and progesterone (Hypothesis 2). Hormone levels of testosterone and cortisol will be analyzed as possible mediators in an exploratory manner. Building on previous studies, we derived cues for which cycle shifts in mate preferences, if existent, should occur: Women in their fertile window should be more sexually attracted to men with a lower fundamental frequency and formant dispersion, as well as a higher baseline testosterone level, compared to low-fertility days of their cycle (Hypothesis 3a). We predict these findings to be robust when controlling for men's age. We will furthermore analyze women's preferences for the voice parameters jitter and shimmer in an exploratory manner. We also state the alternative hypothesis that women in their fertile window, compared to their luteal phase, will not show cycle shifts in their mate preferences regarding men's voice attractiveness for sexual relationships (Hypothesis 3b). One possible moderator for cycle shifts might be women's relationship status. Since it remains unclear if both single and partnered alter their mating strategies across the cycle, we state two alternative hypotheses: Cycle phase shifts in preferences for short-term mates are larger for partnered women than for single women, or, alternatively, the participant's relationship status does not affect the strength of cycle phase shifts in preferences for short-term mates (Hypotheses 4a and 4b). Moreover, we hypothesized self-reported stress as a moderator of cycle shifts: Cycle shifts should be attenuated when self-reported stress is high (Hypothesis 5). We also predict, as the GGOSH suggests, that preference shifts should be absent or less pronounced when it comes to long-term mate preferences (Hypothesis 6, see Gildersleeve et al., 2014a).

3.2. Methods

3.2.1. Participants and recruitment

A total of 157 heterosexual female participants (aged 18–35 years, M = 23.3, SD = 3.4), out of 180 recruited, finished all sessions and were therefore included in further analyses (this sample is the same as in Jünger et al., 2018a and 2018b). Seventeen women who attended only the introductory session of the study dropped out before participation (six fulfilled one of the exclusion criteria below, four quit the study without further reasons, four did not respond to emails, three had scheduling problems). Another six dropped out during the study

because of completing only the first testing session (four had scheduling problems, two did not respond to emails after the first session). Based on the inclusion criteria of other ovulatory cycle studies, our participants had to fit the following preregistered criteria: female, between 18 and 30 years[7] old, naturally cycling (no hormonal contraception for at least three months, not expected switch to hormonal contraception during the study, no current pregnancy or breastfeeding, no childbirth or breast-feeding during the previous three months, not taking hormone-based medication or anti-depressants). Additionally, participants had to report that their ovulatory cycles had a regular length between 25 and 35 days during the last 3 months. At the beginning of the study, 75 of the participants reported being in a relationship, 82 reported being single. Upon completion of all sessions, participants received a payment of 80€ or course credit, and a 3D printed figure of themselves.

3.2.2. Procedure

All participants took part in five individually scheduled sessions that were scheduled between May 2016 and March 2017. In the first session participants received detailed information about the general procedure, duration of the study and compensation. All participants signed a written consent document, and the ethics committee of the Institute of Psychology at the University of Goettingen approved the protocol. The experimenter explained the ovulation tests and checked the inclusion criteria. To count the days to the next ovulation and plan the dates of the experimental sessions, cycle length as well as the dates of the last and the next menstrual onset were assessed. Finally, demographic data were collected.

Sessions two to five, the computer-based testing sessions, took place across two ovulatory cycles (approx. two months) per participant, once per cycle during the late follicular (fertile) phase and once during the luteal phase. To control for possible effects of diurnal changes in hormone levels, all sessions took place in the second half of the day (mainly between 11.30 am and 6 pm). When arriving at the lab, participants first completed a screening questionnaire, assessing their eligibility and some control variables for the saliva samples (Schultheiss and Stanton, 2009). Next, the saliva samples were collected via passive drool before the participants started their first rating task.[8] In preparation for listening to the unmanipulated voice recordings, participants were instructed to evaluate the men's attractiveness as they perceived it "in that moment", independent of their own current relationship status or general interest in other men. Participants were then presented with the voice recordings in a randomized order. After listening to a voice, participants rated it for sexual attractiveness (assessing short-term attractiveness) and for long-term attractiveness using an eleven-point Likert scale from −5 (*extremely unattractive*) to +5 (*extremely attractive*). Definitions of sexual attractiveness and attractiveness for a long-term relationship were provided prior to the ratings and read as follows:

a) *Sexually attractive*: Men who 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.
b) *Attractive for a long-term partnership*: Men who score high would be very attractive for a committed relationship with a long-term perspective. Men scoring low would be very unattractive as a long-term partner.

After each session, the appointment for the next session was arranged individually based on participant's ovulatory cycle.

Furthermore, all participants of the current study were asked to participate in a separate daily online diary study (Arslan et al., 2016) that was conducted in parallel to the described lab study. Within this diary study, participants had to fill out a questionnaire about daily feelings and behavior across 70 days. We used the stress ratings from this study for further analyses (see below).

3.2.3. Measures

3.2.3.1. Ovulatory cycle phase. Women's cycle phase was determined by the reverse cycle day method, based on the estimated day of the next menstrual onset (Gildersleeve et al., 2012) and confirmed by highly sensitive (10 mIU/ml) urine ovulation test strips from purbay®, which measure luteinizing hormone (LH). These LH tests were conducted privately at home on the estimated day of ovulation and the four days prior to that, and results were self-reported by the participants. The study investigated two ovulatory cycles in which every participant reported to the lab twice: Once while being fertile (at the days prior to ovulation, usually reverse cycle days 16–18, with reverse cycle day 16 as the ideal date) and once when not fertile (during the luteal phase, after ovulation and prior to the next menstrual onset, usually reverse cycle days 4–11, with reverse cycle days 6 to 8 as the ideal dates). An Excel sheet was used to compute the acceptable days for the testing sessions and track whether a participant started in her fertile or luteal phase. Of all participants who finished all sessions, 66 participants started with the first session in their luteal phase, and 91 started in the fertile phase.

For the main cycle phase analyses, we excluded a total of 45 participants due to negative LH tests in both cycles, irregular ovulatory cycles or inappropriate scheduling of testing sessions (see "Preliminary analyses" for more details), resulting in $n = 112$ women. Of these participants, 46 started with the first session in their luteal phase, and 66 started fertile. However, all 157 women were included in the denoted robustness checks.

3.2.3.2. Hormone measures. We collected four saliva samples from each participant (one per testing session) prior to the rating tasks. Contamination of saliva samples was minimized by asking participants to abstain from eating, drinking (except plain water), smoking, chewing gum or brushing teeth for at least 1 h before each session. The samples were stored at −80 °C directly after collection until shipment on dry ice to the Kirschbaum Lab at Technical University of Dresden, Germany, where estradiol, progesterone, testosterone and cortisol were assessed via liquid chromatography mass spectrometry (LCMS; Gao et al., 2015). Because the LCMS analysis of estradiol detected only 22% of all possible values, the samples were reanalyzed using the highly sensitive 17β-estradiol enzyme immunoassay kit (IBL International, Hamburg, Germany). These latter estradiol values were used in subsequent analyses. Hormone levels were skewed, therefore, we centered all hormone values on their subject-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 the linear mixed models (as in Jones et al., 2018b; and congruent with our approach in Jünger et al., 2018a, 2018b). This is a common procedure to isolate effects of within-subject changes in hormones, avoiding the influence of outliers on results and dealing with the non-normal distribution of hormone levels. Hormone levels were nearly normally distributed afterwards, a figure showing the distribution of hormone levels after this procedure can be found in the supplement (Fig. S1). Importantly, this procedure did not change any findings compared to analyses with untransformed hormone values. The R code for this procedure can be found in the open script.

3.2.3.3. Stimuli and masculinity analyses. Seventy-six voices of different men, counting from three to eight in German, recorded as part of the

[7] One of the participants reported being 35 years old. We excluded her data in the main analyses, but included it in the robustness checks because she met all other inclusion criteria and had positive LH tests. Including her data did not alter the results.

[8] The described study on ovulatory cycle shifts for voice masculinity was one part of a larger study (see pre-registration). Participants also had to complete other rating tasks and anthropometric data was collected between these tasks. The duration of one experimental session was approximately 2–2.5 h.

Berlin Speed Dating Study (see Asendorpf et al., 2011 for more details), were presented via headphones (JVC® HA-RX300). We selected recordings from 76 participants out of a pool of 382 by gender (male) and age (between 18 and 30 years old, matching the age of the eligible female participants in the study). Stimulus males' baseline testosterone levels were measured from saliva samples. The samples were analyzed using radioimmunoassay by the Kirschbaum lab at the Technical University Dresden. Each recording was analyzed using Praat software (version 6.0.17). Pitch, floor, ceiling and other settings were set in line with Study 1. Across each recording, we measured mean F_0 (henceforth, F_0; $M = 110.74$, range = 85.30–157.48; $SD = 12.66$) and median formant frequencies from which we computed D_f ($M = 1043.19$ Hz, range = 961.67–1137.68, $SD = 30.30$ Hz) as in Study 1, and measured four measures of jitter and five measures of shimmer. All jitter ($r > 0.97$) and shimmer ($r > 0.31$) variables were correlated and therefore z-standardized and summed (jitter: $M = 0.00$, $SD = 0.99$; shimmer: $M = -0.02$, $SD = 0.84$). Additionally, we computed formant position (P_f; $M = 0.00$, range = −1.36–2.96, $SD = 0.68$), the standardized formant value for the first four formants which has been found to be more sexually dimorphic than D_f (Puts et al., 2012).

3.2.3.4. Stress ratings. Self-reported stress was measured with one item ("*Today I was stressed out*") on a five-point Likert-scale (from "*less than usual*" to "*more than usual*") on a daily basis within the accompanying online diary study (see above) with planned missings.[9] For the analysis, the stress value from the day of the lab testing session was used. If there was no existing value for that specific day, then we averaged the values of the two days before and after the testing day, if available. In total, 54 of the 157 participants were excluded from analyses, 26 because they did not take part in the diary study, 20 because they did not fill out enough days to provide data for at least one fertile and one luteal session, and eight because they took part in the study at another time window (not parallel to the lab study). Sixty-two participants had stress data for at least one fertile and one luteal session, and 41 for all sessions, resulting in an available dataset of 160 cycles (out of 314 possible cycles; 119 cycles out of 224 for $n = 112$) in total.

3.3. Results

3.3.1. Preliminary analyses

First, we counted how many cycles were irregular, so that the day of the testing session was scheduled more than three days apart (before or after) from the defined windows of appropriate testing days (e.g. fertile sessions were defined as being appropriate within reverse cycle days 15–18, luteal sessions were defined as being appropriate within reverse cycle days 4–11, see section "Ovulatory cycle phase"). Even though all participants reported having regular ovulatory cycles in the introductory session, eight women had irregular cycles in both investigated cycles, and 32 reported one cycle being irregular, resulting in 48 out of 314 (15.3%) cycles being irregular. Next we checked how many of the participants' ovulatory cycles had positive LH tests (indicating a LH surge) in the estimated fertile phase to detect non-ovulatory cycles. Twelve participants reported negative LH test results for both investigated cycles, nine reported negative LH tests results for one cycle. In total, the LH tests in 33 of all 314 cycles (10.5%) were negative. 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 24–36 h after the observed LH surge, testing sessions that were scheduled more than two 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, thirteen (4.63%) purportedly fertile phase sessions were scheduled three or four days after the LH surge. Therefore, 268 (95.37%) were scheduled within an appropriate range of three days before to two days after the LH surge (in total: $M = -0.12$, $SD = 1.39$ days in relation to the day of the observed LH surge). A histogram showing the distribution of days of fertile phase testing sessions relative to the observed LH surge can be found in the supplement (Fig. S2). Participants with irregular cycles, negative LH tests or the risk of early luteal phase instead of fertile phase testing session were excluded in the main analyses, but included in denoted robustness checks.

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

We first tested for possible ovulatory cycle shifts in women's attraction to men's voices in general (Hypothesis 1). For the multilevel analyses with attractiveness rating as the dependent variable (Model 1 with sexual attractiveness, Model 2 with long-term attractiveness), female raters as well as the male stimuli were treated as random effects. Women's cycle phase (0 = luteal phase, 1 = fertile phase) was treated as a fixed effect. We additionally let participant's slopes vary systematically across cycle phase by modeling cycle phase as a random slope. This analysis showed a significant cycle shift in women's attraction: Ratings for sexual attractiveness were higher in the fertile phase than in the luteal phase of the ovulatory cycle ($\gamma = 0.10$, $SE = 0.05$, $t = 2.14$, $p = .035$, $95\%CI = [0.01; 0.19]$), supporting Hypothesis 1. We didn't observe differences between fertile phase and luteal phase ratings for long-term attractiveness ($\gamma = 0.06$, $SE = 0.04$, $t = 1.45$, $p = .150$, $95\%CI = [-0.02; 0.15]$). These results indicate the existence of ovulatory cycle shifts on women's mate attraction to men's voices for sexual, but not long-term attractiveness, such that, overall, fertile women rated men's voices as being more attractive.[10]

To analyze whether women's mate preferences for specific vocal cues change across the ovulatory cycle (Hypothesis 3), we calculated additional three multilevel models. In all models, female participants as well as male vocal stimuli were treated as random effects, women's cycle phase was treated as fixed effect and a random slope for cycle phase varying in participants was included. Moreover, the vocal masculinity cues F_0 (Model 3), D_f (Model 4) and men's baseline testosterone levels (Model 5) were treated as fixed effects separately. Further, because recent research suggests P_f as a superior indicator of vocal masculinity compared to D_f (Puts et al., 2012), we also analyzed possible cycle shifts in mate preferences for men's P_f (Model 6). Results show a significant main effect for cycle phase on sexual attractiveness ratings in each model[11] (Table 4), again supporting Hypothesis 1. Women rated men's voices as more attractive when they were fertile. Moreover, there was a significant effect of fundamental frequency and one of formant dispersion on attractiveness ratings: Voices with lower F_0 and voices with lower D_f were rated as more attractive. The effects of P_f or baseline T did not reach statistical significance. We observed a significant interaction effect between cycle phase and baseline T, indicating that fertile women rate lower T men as more attractive, which is the opposite direction as stated in Hypothesis 3. None of the other vocal cues interacted with cycle phase, indicating that women's mate preferences

[9] The participants had to fill out > 100 items per day. Therefore, we decided to reduce the daily items by planned missings to minimize dropouts while obtaining sufficient data for each item. The relevant stress item was shown on about 40% of all days.

[10] In line with Study 1, we also analyzed possible main effects of hormone values (estradiol and progesterone or E/P) on attractiveness ratings separately in an exploratory manner, as they were not part of the preregistration. No significant effects were observed. Details can be found in the supplementary material (Tables S18–S19).

[11] Regarding the length of our manuscript, we decided to report all other results for the long-term attractiveness ratings in the supplementary material (Tables S12–S15), during the review process. Results for long-term ratings all showed null results for preference shifts across the cycle, hence, all results were supporting Hypothesis 6.

Table 4
Multilevel regression analyses of sexual attractiveness ratings as a function of cycle phase and men's voice pitch, formant dispersion, formant position or baseline testosterone levels.

	γ	*SE*	*t*	*p*	*95%CI*
F0 model					
Cycle phase	0.10	0.05	2.14	.035	[0.01; 0.19]
Men's F_0	−0.68	0.12	−5.71	< .001	[−0.92; −0.45]
Cycle phase × men's F_0	0.01	0.02	0.55	.586	[−0.03; 0.06]
Df model					
Cycle phase	0.10	0.05	2.14	.035	[0.01; 0.19]
Men's D_f	−0.28	0.14	−2.04	.045	[−0.56; −0.01]
Cycle phase × men's D_f	0.02	0.02	0.91	.362	[−0.02; 0.06]
Pf model					
Cycle phase	0.10	0.05	2.14	.035	[0.01; 0.19]
Men's P_f	−0.40	0.21	−1.93	.057	[−0.81; 0.01]
Cycle phase × men's P_f	0.02	0.03	0.52	.600	[−0.05; 0.08]
Baseline t model					
Cycle phase	0.10	0.05	2.14	.035	[0.01; 0.19]
Men's baseline testosterone	0.07	0.14	0.47	.639	[−0.21; −0.35]
Cycle phase × men's baseline testosterone	−0.04	0.02	−2.00	.046	[−0.09; −0.00]

Note. F_0 = fundamental frequency (voice pitch), D_f = formant dispersion, P_f = formant position. All variables had 34,048 observations (112 participants × 4 test sessions × 76 stimuli).

do not shift for specific cues in men's voices across the ovulatory cycle,[12] in contrast to Hypothesis 3. Results remained stable when controlling for men's age. Moreover, results remained virtually identical when adding all four vocal masculinity cues to the same model at the same time, details can be found in the supplement (Tables S16 and S17).

We also analyzed influences of men's jitter and shimmer on attractiveness ratings in an exploratory manner. The main effects of cycle phase stayed significant. We found a significant main effect for shimmer ($\gamma = 0.28$, $SE = 0.14$, $t = 2.04$, $p = .045$, $95\%CI = [0.01; 0.56]$), suggesting higher ratings when shimmer was high; but not for jitter ($\gamma = 0.07$, $SE = 0.14$, $t = 0.51$, $p = .609$, $95\%CI = [-0.21; 0.35]$). The interactions of cycle phase with jitter or shimmer were not significant.

Next, we calculated Spearman rank correlations between attractiveness ratings in the fertile and those in the luteal phase to better understand the reported cycle effect. Results from this analysis indicate that ranks of the rated voices (from the most to the least attractive voice) did not differ between the fertile and the luteal phase for sexual attractiveness ($r = 0.99$, $p < .001$). Rather, most of the voices received a slightly better rating in the fertile phase compared to the luteal phase ($M_{fertile} = -0.33$, $SD = 1.23$, $M_{luteal} = -0.40$, $SD = 1.23$, $d = 0.05$). These results indicate that women rated the same men as more or less attractive, independent of their cycle phase, suggesting that differential effects of masculinity cues are rather unlikely.

3.3.3. Hormonal influences on cycle phase shifts

In order to analyze whether steroid hormones mediate effects of cycle phase (Hypothesis 2), we entered cycle phase, estradiol, progesterone, E/P, testosterone, and cortisol as fixed effects into the multilevel model with sexual attractiveness ratings as the outcome variable (Model 7), female participants and male stimuli as random effects and a random slope for cycle phase varying in participants. Results demonstrate that, in contrast with Hypothesis 2, there were no mediating effects of any hormone levels: results of cycle phase remained significant and effects were even larger than in the model without hormone levels (see Table 5), reinforcing the effect that ratings increased in women's fertile phase compared to ratings in the luteal phase. Moreover, there was a significant positive effect of progesterone on sexual attractiveness ratings. Counterintuitively, ratings were higher when progesterone levels were higher. There were no significant effects of estradiol, E/P, testosterone, or cortisol. Again, because of possible problems of multicollinearity (significant negative correlation between E/P and progesterone, significant positive correlations between E/P and estradiol, as well as E/P and cortisol, see Table S66 for all correlation coefficients between hormones), we also calculated additional models with estradiol, progesterone, testosterone and cortisol as fixed effects, but excluding E/P. Results remained virtually identical and can be found in the supplemental material (Table S20). However, in line with our reported results in Jünger et al. (2018b), and because results did not change when analyzing hormone values separately, we decided to report the models with all hormones included here.

3.3.4. Investigating women's relationship status as a possible moderator

To evaluate whether women's relationship status influences ovulatory cycle shifts,[13] we first categorized all women as in a relationship who reported being in an open relationship, in a committed relationship, engaged or married. Relationship status changed for 13 women (for seven of the $n = 112$ cycle phase sample) across the study. Their data were categorized according to their relationship status on the particular testing day. One multilevel model (Model 8) with women's cycle phase and relationship status as fixed effects, a random slope for cycle phase varying in participants, female participants and male stimuli as random intercepts again showed significant main effects of cycle phase (Table 6). Sexual attractiveness ratings were higher in the fertile phase of the ovulatory cycle. There were no significant effects of relationship status or of the cycle phase × relationship status interaction. Therefore, women's relationship status did not moderate the cycle phase effect on attractiveness ratings.

3.3.5. Self-reported stress

Furthermore, we analyzed whether self-reported stress moderated the relationship between cycle phase and attractiveness ratings. We calculated one further multilevel model (Model 9). Again, female raters as well as the male stimuli were treated as random effects. Women's cycle phase and self-reported stress ratings were treated as fixed effects and a random slope for cycle phase varying in participants was included. Because many women did not fill out the self-reported stress item for every testing day due to the planned missings design (see Methods), data for only about half of the sample were available (119 cycles out of 224 assessed cycles). When evaluating sexual attractiveness ratings as the outcome variable, we found a significant main effect of cycle phase, revealing that attractiveness ratings were

[12] In line with Study 1, we also analyzed possible interaction effects of hormone values (estradiol and progesterone or E/P) with all masculine vocal cues (F_0, D_f, P_f, baseline T) separately in an exploratory manner. None of these models revealed any significant interaction effect, again suggesting no preference shifts for masculine voices. Details can be found in the supplementary material (Tables S28–S31).

[13] Although we originally stated the hypothesis that women's relationship status might moderate preference shifts, we decided to rather report our moderator analyses for attraction shifts, because we did not find any hint for an observable preference shift in our analyses. However, we also investigated possible three-way interactions between cycle phase, relationship status and all masculine vocal cues (F_0, D_f, P_f, baseline T) separately. None of these models revealed any significant interaction effect, indicating no compelling evidence for preference shifts for masculine voices and no moderation effects of women's relationship status, in contrast to Hypothesis 4a, but supporting Hypothesis 4b. Details can be found in the supplement (Tables S32–S33).

Table 5
Multilevel regression analyses of sexual attractiveness ratings as a function of cycle phase with hormone levels as possible mediator variables.

	γ	*SE*	*t*	*p*	*95%CI*
Cycle phase	0.13	0.06	2.29	.023	[0.02; 0.25]
Estradiol	−0.06	0.05	−1.09	.276	[−0.16; 0.05]
Progesterone	0.11	0.05	2.41	.016	[0.02; 0.20]
E/P	0.01	0.03	0.23	.822	[−0.05; 0.06]
Testosterone	0.01	0.02	0.58	.561	[−0.03; 0.05]
Cortisol	−0.01	0.04	−0.18	.855	[−0.10; 0.08]

Note. All variables had 28,956 observations (112 participants × 4 test sessions × 76 stimuli − missing values). We dummy-coded the variable cycle phase with 0 = luteal, 1 = fertile. All hormone values were centered to their subject-specific means and then scaled.

Table 6
Multilevel regression analyses of sexual attractiveness ratings as a function of cycle phase and women's relationship status.

	γ	*SE*	*t*	*p*	*95%CI*
Cycle phase	0.13	0.06	2.04	.044	[0.01; 0.25]
Relationship status	−0.12	0.09	−1.31	.189	[−0.30; 0.06]
Cycle phase × relationship status	−0.05	0.09	−0.62	.537	[−0.22; 0.12]

Note. All variables had 34,048 observations (112 participants × 4 test sessions × 76 stimuli). We dummy-coded the variable cycle phase with 0 = luteal, 1 = fertile, and relationship status with 0 = single, 1 = in a relationship.

Table 7
Multilevel regression analyses of sexual attractiveness ratings as a function of cycle phase and women's self-reported stress.

	γ	*SE*	*t*	*p*	*95%CI*
Cycle phase	0.33	0.11	2.95	.003	[0.11; 0.54]
Self-reported stress	−0.03	0.04	−0.73	.467	[−0.10; 0.05]
Cycle phase × self-reported stress	−0.08	0.05	−1.76	.079	[−0.18; 0.01]

Note. All variables had 18,088 observations (75 participants × 4 test sessions × 76 stimuli − missing values). We dummy-coded the variable cycle phase with 0 = luteal, 1 = fertile.

higher in the fertile phase of the cycle. However, the main effect of self-reported stress, as well as the interaction between cycle phase and self-reported stress was not significant (Table 7), indicating that there was no moderation effect of self-reported stress on cycle effects.

3.3.6. Further robustness checks and exploratory analyses

Besides the exploratory analyses we have already reported in footnotes, we conducted further analyses to test the robustness of our effects. To rule out the possibility that the main effect results were driven by order effects of testing sessions (in particular, participating in the first session when fertile; Suschinsky et al., 2014; Wallen and Rupp, 2010), we controlled for initial cycle phase in our analyses. The effect of cycle phase remained stable (γ = 0.10, *SE* = 0.05, *t* = 2.14, *p* = .035, *95%CI* = [0.01; 0.19]). Moreover, initial cycle phase affected sexual attractiveness ratings (γ = 0.36, *SE* = 0.15, *t* = 2.48, *p* = .014, *95%CI* = [0.07; 0.66]), in that ratings were higher when participants started in the fertile phase. We also controlled our analyses for session number. Again, the effect of cycle phase remained stable (γ = 0.10, *SE* = 0.05, *t* = 2.03, *p* = .045, *95%CI* = [0.00; 0.18]) and session number affected sexual attractiveness ratings (γ = −0.04, *SE* = 0.01, *t* = −3.61, *p* < .001, *95%CI* = [−0.06; −0.02]), in that ratings decreased on average by number of the testing session.

Then, to investigate if being tested while fertile in the first session affects ratings in later sessions, we calculated an additional model including an interaction between session number and initial cycle phase. We found a significant interaction between session number and initial cycle phase (γ = −0.07, *SE* = 0.02, *t* = −3.33, *p* < .001, *95%CI* = [−0.12; −0.03]), showing that ratings decreased by ongoing testing sessions when the initial session was fertile, but not when the initial session was scheduled in the luteal phase. Additionally, there was a main effect of initial session (γ = 0.57, *SE* = 0.16, *t* = 3.68, *p* < .001, *95%CI* = [0.26; 0.89]), indicating higher ratings when the first session was fertile, but no main effect of session number (γ = 0.00, *SE* = 0.02, *t* = 0.14, *p* = .892, *95%CI* = [−0.03; 0.04]). Based on these findings, to rule out that our null results for cycle shifts in mate preferences were caused by a carryover effect of the hormonal state in the initial session, we also controlled our main preference shifts models for an interaction effect between session number and initial cycle phase. Results remained virtually identical and can be found in the supplementary material (Table S65).

Next we conducted our analyses with all recruited women[14] (*N* = 157) including (Tables S21–S25) or excluding random slopes (Tables S60–S64). Nearly all results remained robust across all checks. However, the significant interaction of cycle phase and men's baseline testosterone levels (Table S22 and S61) disappeared in all robustness checks. We conducted additional exploratory robustness checks and falsification tests. First, we repeated all of our analyses using sexual minus long-term attractiveness as the dependent variable, to allow for the possibility that differences in estimated effects on sexual- and long-term attractiveness ratings would be difficult to estimate, because of the high correlations between these outcomes (*r* = 0.90). This is a very specific prediction of the GGOSH (see e.g. Gangestad et al., 2004, 2007). Complementary to that, we also ran all analyses with sexual plus long-term attractiveness as the dependent variable, which provides a more aggregated estimation of overall attraction (Gangestad et al., 2004, 2007). Importantly, none of the models revealed any observable preference shifts as a function of cycle phase or hormone levels (see Tables S42–S59 in the supplementary material for detailed results). In summary, we did not observe any preference shift in our robustness checks. The estimated effect size of cycle phase on attractiveness ratings was robust across robustness checks and statistically significant in the vast majority of models.

4. Discussion

We sought to clarify whether women experience hormone related mate preference shifts for male voice masculinity across the ovulatory cycle. We evaluated hormonal influences underlying women's cycle shifts in attraction and preferences for masculine voices and further investigated potential moderators of these effects. We included multiple measures (hormone levels assayed from saliva and cycle phase confirmed via LH tests), investigated preferences for natural as well as manipulated stimuli, and employed within-subject designs in two samples that exceed the sizes from previous studies. In both studies, we did not find compelling evidence to support the hypothesis that women experience (hormone related) cycle shifts in mate preferences for masculine voices. Further, we report that progesterone and E/P influenced attractiveness ratings in Study 1, Study 2 indicated the presence of cycle phase shifts in women's overall attraction to men's voices, but not shifts in preferences for specific vocal characteristics. Women's relationship status or self-reported stress did not moderate attraction shifts. We did not find a clear pattern of hormonal influences on attractiveness ratings across the cycle. In the following, we interpret these findings and highlight implications for future research.

[14] In the previous versions of this manuscript, we reported these analyses as our main analyses and the analyses with those *n* = 112 women who perfectly met all inclusion criteria as robustness checks. We decided to switch these analyses during the review process.

4.1. Preference vs. attraction shifts

As in previous work evaluating shifts in mate preferences for body and facial masculinity (Jones et al., 2018a; Jünger et al., 2018b; Marcinkowska et al., 2018a; Peters et al., 2009), as well as men's behaviors (Jünger et al., 2018a), we report no observed effects of cycle phase, conception risk or steroid hormone levels on women's mate preferences for masculine voices across two independent studies. Therefore, we did not find compelling evidence for the GGOSH, even with large samples, multiple time points (i.e. greater power to detect an effect across testing sessions), and highly reliable estimates of cycle phases compared to previous studies that purportedly found evidence for mate preference shifts for masculine voices across the ovulatory cycle. Indeed, in one analysis for Study 2, we found an interaction between women's cycle phase and men's baseline testosterone levels on sexual attractiveness ratings, but this effect was in the opposite direction from that predicted by our hypotheses and the GGOSH: Ratings were higher in the fertile phase when men's baseline testosterone was low. However, the effect is counter-intuitive and disappeared in all of our robustness checks. We therefore suggest that it is a false positive. Hence, we interpret our findings as null results for mate preference shifts across the cycle. These results and recent studies reporting null results cycle shifts for body or behavior preferences (Jünger et al., 2018a, 2018b; Marcinkowska et al., 2018a) indicate that it no longer appears to be the case that null results are specific to face preferences (e.g. Jones et al., 2018a; Muñoz-Reyes et al., 2014; Peters et al., 2009).

Instead of a cycle phase shift in preferences, Study 2 suggests a shift in women's overall attraction: Sexual attractiveness ratings were higher in the fertile phase, regardless of men's voice parameters. Similarly, a cycle phase attraction shift was recently reported for body masculinity and men's behaviors within the same dataset (Jünger et al., 2018a, 2018b). These attraction shifts might be connected to fertile phase increases in sexual motivation and desire (Arslan et al., 2018; Jones et al., 2018b; Roney and Simmons, 2013, 2016), though they were not fully supported in further exploratory analyses substituting cycle phase estimates with direct steroid hormones measurements. However, we found only partial evidence for an attraction effect in Study 1. Specifically, in Study 1, ratings were higher when progesterone levels were lower (and when E/P was higher in at least one model), which is usually the case in the fertile phase of the ovulatory cycle; hence, these results support the notion of an attraction shift. Importantly, this effect was not significant in all models. There are several possible reasons why these results differed between and within the two studies. First, different methods were used in both studies. Study 1 did use hormone levels rather than cycle phase, it used manipulated voice recordings of men reading a brief passage, and had two testing sessions per participant. Study 2, contrarily, used cycle phase and hormone levels as predictors for fertility, LH tests to validate the fertile phase, unmanipulated voice recordings of men counting, and investigated two ovulatory cycles (four testing sessions) per participant. Nevertheless, the central conclusions remain: no hormone related or cycle phase shifts in preferences were observed in either study. Second, the reported effect sizes for attraction shifts in Study 2 were small (ratings: $\gamma = 0.10$ in the main analyses). Study 1 had a smaller sample size as well as fewer test trials and therefore fewer observations than Study 2, which makes detecting this small effect more difficult. However, the differences in observations may overstate the differences in power, given that test power in both studies was high compared to previous studies. Moreover, hormone analyses in Study 1 indeed provided partial evidence for attraction shifts, observed by generally higher ratings when progesterone was lower or E/P was higher. Third, given that effect size estimates were very small, and that including random slopes might reduce test power, even the power in Study 2 might have been insufficient to detect the effect in all models. However, according to Gangestad et al. (2016), Study 2 should still have > 80% power to even detect small effect sizes (with $n = 112$ women, within-subject design, four testing sessions each, a measurement validity of ~0.85 with using LH tests and a high correlation for ratings across phases). Fourth, although the other explanations seem to be more likely, the attraction shift effect might simply not be robust. Hence, further research should test the reliability of attraction shifts across the ovulatory cycle, investigate under which circumstances they occur and whether they correlate with a general fertile phase increase in sexual desire.

4.2. Hormonal influences

Previous studies have suggested that changes in women's mate preferences and desire are regulated by steroid hormonal changes across the ovulatory cycle (Feinberg et al., 2006; Jones et al., 2018b; Pisanski et al., 2014; Puts, 2005, 2006; Roney and Simmons, 2013, 2016). However, our results did not reveal a clear pattern of hormonal influences on women's attraction across the ovulatory cycle. In fact, we found different results for hormone levels across the two studies.

Progesterone predicted attractiveness ratings in Study 1 and attraction shifts in Study 2, but in different directions: negatively in Study 1 and positively in Study 2. These contradictory results remained significant in the robustness checks. The positive influence of progesterone in Study 2 is particularly counterintuitive, as progesterone levels are generally higher in the luteal phase, but we found generally higher attractiveness ratings in the fertile phase. Typically, this effect has been found in the opposite direction in previous studies and in Study 1. Critically, the negative association reported in Study 1 aligns more closely with the theoretical assumptions and findings of previous work.

Besides the puzzling effects of progesterone, E/P positively influenced attractiveness ratings in Study 1, but only in one out of three models. Regarding the overall unclear pattern of hormonal influences, we interpret these findings with caution and suggest that they need to be replicated before being interpreted further. We therefore focus here on the lack of robust, reliable hormonal influences on attraction shifts: a) The influence of progesterone and E/P remains unclear, b) estradiol did not reliably affect attractiveness ratings, and c) we found no effects of testosterone or cortisol. Therefore, we could not find evidence for hypotheses that were built on the assumptions of clear hormonal influences on cyclic shifts, e.g. the “spandrels hypothesis” that women with higher estradiol levels show stronger preferences for masculine men (Havlíček et al., 2015; Shimoda et al., 2018).

There are several possible explanations for our findings. First, we used a variety of methods across both studies (e.g. hormone analyses via LCMS vs. immunoassays) and tested participants from two populations (differing in culture and age spans) in two different labs. This might have induced important differences in the results between the two studies, and compared to previous studies. Second, perhaps hormonal influences are different for voice attraction than for other attraction to other stimuli or sexual desire, which would explain why we did not find the same hormonal influences as those predicting sexual desire (Jones et al., 2018b; Roney and Simmons, 2013, 2016). There is thus a strong need for continued research to clarify the hormonal influences on attraction shifts across the ovulatory cycle. Furthermore, it should be investigated whether hormonal influences on mate attraction vary across categories of stimuli (e.g. voices, faces, bodies). However, again, the central conclusion remains as we did not observe any hormonal influences on mate preferences for masculine voices.

4.3. No moderating effects of relationship status and perceived stress

In Study 2, we investigated whether women's relationship status or self-reported stress moderate fertile phase attraction shifts. Whereas previous studies reported that cycle shifts in women's attraction for men's bodies or behaviors were found only for partnered women, not for singles (Jünger et al., 2018a, 2018b), we did not replicate this effect for attraction to masculine voices. In line with this, Jones et al. (2018b) found no evidence that hormonally driven shifts in women's general

sexual desire were moderated by their relationship status. However, other studies have reported that only partnered women, not singles, showed increased fertile phase sexual desire (Roney and Simmons, 2016). Thus, the effects of relationship status on psychological changes across the ovulatory cycle remain unclear. Nevertheless, our results do not support the assumptions of a dual mating strategy (that women may receive fitness benefits when forming a relationship with a reliably investing man, while seeking good genes from another man through extra-pair sexual encounters; Pillsworth and Haselton, 2006). We also did not find evidence of preference shifts for masculine voices, or a moderating effect of women's relationship status on preference shifts.

Moreover, self-reported stress did not moderate fertile phase attraction shifts. Previous studies reported different results: Stress inhibited estradiol levels (Roney and Simmons, 2015) and overrode fertile phase attraction shifts for masculine bodies (Jünger et al., 2018b) and faces (Ditzen et al., 2017), but not for men's behaviors (Jünger et al., 2018a). Hence, stress might affect only the perception of visually available cues in bodies and faces. Self-reported stress values are subjective and might not always reflect physiological stress levels (however, cortisol levels did also not influence attractiveness ratings). To investigate the impact of stress on mate attraction directly, future studies should manipulate stress experimentally. In sum, future research should investigate under which conditions and for which traits or cues cycle shifts in attraction are influenced by relationship status or self-reported stress. Additionally, other possible moderator variables should be taken into account to elucidate psychological changes across the ovulatory cycle.

5. Conclusion

In the current studies, we used substantially larger datasets than those in previous studies, as well as robust methods of fertility estimation and hormone assessments to investigate possible shifts in women's mate preferences and attraction to male voices across the ovulatory cycle. We found at least partial supporting evidence for ovulatory cycle shifts in attraction to men's voices, regardless of vocal masculinity, but the lack of ties to hormones is a fairly significant limitation to this finding. Attraction shifts were not moderated by women's relationship status or self-reported stress and require further research to test their robustness. We found no compelling evidence for shifts in preferences for masculine voice characteristics. Our results contrast with previous work on preference shifts for masculine voices (Feinberg et al., 2006; Pisanski et al., 2014; Puts, 2005; see also Puts, 2006), but align with recent reported null replications of cycle shifts for masculine faces, bodies and behaviors (Jones et al., 2018a; Jünger et al., 2018a, 2018b; Marcinkowska et al., 2016, 2018a; Muñoz-Reyes et al., 2014). Hence, the present research provides no compelling evidence for the good genes ovulatory shifts hypotheses and suggests that cycle shifts in preferences or attraction are more complex than previously assumed. Future research is indispensable for clarifying the conditions under which cycle shifts in women's psychology and behavior can be observed.

Declaration of interest

The authors have no conflict of interest to declare.

Acknowledgements

Study 2 was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 254142454/GRK 2070. We thank Steven Gangestad and Alex Stern for comments on previous versions of this manuscript, and Ruben C. Arslan for valuable statistical advice.

For Study 1, we thank Marc Breedlove and Cynthia Jordan for research support, Bradly Alicea, Jill Armington, Emily Barben, Michael Burla, Kathryn Cheney, Michael DiRaimo, Melina Durhal, Rebecca Frysinger, Sana Khan, Mallory Leinenger, Heather Malinowski, Ernestine Mitchell, Joe Morehouse, Charlene Scheld, Kevin Singh, Sara Sutherland, Lisa Vroman, Tyesha Washington, and Molly Zolianbawi for assistance with study preparation and data collection, and Elizabeth Hampson and Bavani Rajakumar for assistance with hormonal assays.

For Study 2, we thank Christian Treffenstädt for assistance in the programming process, Kelly Gildersleeve for kindly providing a blueprint of an Excel sheet to calculate fertile windows and our research assistants Laura Botzet, Silvia Bradatsch, Paulina Lauer and Laura Specker-Mattißen, without whom the large data collection would not have been possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yhbeh.2018.10.008.

References

Arslan, R.C., Jünger, J., Gerlach, T.M., Ostner, J., Penke, L., 2016. Pre-registration. Online diary study of ovulatory cycle shifts on sexual desire, sexual activity, and mating behavior. Retrieved from. osf.io/d3avf.

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. J. Pers. Soc. Psychol. https://doi.org/10.1037/pspp0000208. (Advance online publication).

Asendorpf, J.B., Penke, L., Back, M.D., 2011. From dating to mating and relating: predictors of initial and long-term outcomes of speed-dating in a community sample. Eur. J. Personal. 25, 16–30. https://doi.org/10.1002/per.768.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear Mixed-effects Models Using Eigen and S4. R Package Version, 1.1-13.

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. Horm. Behav. 81, 74–83. https://doi.org/10.1016/j.yhbeh.2016.03.004.

Bobst, C., Sauter, S., Foppa, A., Lobmaier, J.S., 2014. Early follicular testosterone level predicts preference for masculinity in male faces, but not for women taking hormonal contraception. Psychoneuroendocrinology 41, 142–150. https://doi.org/10.1016/j.psyneuen.2013.12.012.

Boersma, P., Weenink, D., 2006. Praat: Doing Phonetics by Computer (Version 4.4. 06) [Computer Program]. (Retrieved January 30, 2006).

Butler, G.E., Walker, R.F., Walker, R.V., Teague, P., Riad-Fahmy, D., Ratcliffe, S.G., 1989. Salivary testosterone levels and the progress of puberty in the normal boy. Clin. Endocrinol. 30, 587–596. https://doi.org/10.1111/j.1365-2265.1989.tb01431.x.

Cheng, J.T., Tracy, J.L., Ho, S., Henrich, J., 2016. Listen, follow me: dynamic vocal signals of dominance predict emergent social rank in humans. J. Exp. Psychol. Gen. 145, 536–547. https://doi.org/10.1037/xge0000166.

Collins, M., 2000. Understanding uncertainties in the response of ENSO to greenhouse warming. Geophys. Res. Lett. 27, 3509–3512. https://doi.org/10.1029/2000GL011747.

Connor, M.L., Sanford, L.M., Howland, B.E., 1982. Saliva progesterone throughout the menstrual cycle and late pregnancy. Can. J. Physiol. Pharmacol. 60, 410–413. https://doi.org/10.1139/y82-060.

Dabbs, J.M., Mallinger, A., 1999. High testosterone levels predict low voice pitch among men. Personal. Individ. Differ. 27, 801–804. https://doi.org/10.1016/S0191-8869(98)00272-4.

Dejonckere, P.H., Remacle, M., Fresnel-Elbaz, E., Woisard, V., Crevier-Buchman, L., Millet, B., 1996. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Rev. Laryngol. Otol. Rhinol. 117, 219–224. Retrieved from. http://europepmc.org/abstract/med/9102729/.

Ditzen, B., Palm-Fischbacher, S., Gossweiler, L., Stucky, L., Ehlert, U., 2017. Effects of stress on women's preference for male facial masculinity and their endocrine correlates. Psychoneuroendocrinology 82, 67–74. https://doi.org/10.1016/j.psyneuen.2017.05.006.

Evans, S., Neave, N., Wakelin, D., Hamilton, C., 2008. The relationship between testosterone and vocal frequencies in human males. Physiol. Behav. 93, 783–788. https://doi.org/10.1016/j.physbeh.2007.11.033.

Fairbanks, G., 1960. Voice and Articulation Drillbook, 2nd ed. Harper & Row, New York.

Feinberg, D.R., Jones, B.C., Smith, M.L., Moore, F.R., DeBruine, L.M., Cornwell, R.E., Hillier, S.G., Perrett, D.I., 2006. Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. Horm. Behav. 49, 215–222. https://doi.org/10.1016/j.yhbeh.2005.07.004.

Finstad, S.E., Emaus, A., Tretli, S., Jasienska, G., Ellison, P.T., Furberg, A.-S., ... Thune, I., 2009. Adult height, insulin, and 17β-estradiol in young women. Cancer Epidemiol. Biomark. Prev. 18, 1477–1483. https://doi.org/10.1158/1055-9965.EPI-08-0972.

Fitch, W.T., 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. J. Acoust. Soc. Am. 102, 1213–1222. https://doi.org/10.1121/1.421048.

Gangestad, S.W., Thornhill, R., 1998. Menstrual cycle variation in women's preferences for the scent of symmetrical men. Proc. R. Soc. B 265, 927–933. https://doi.org/10.1098/rspb.1998.0380.

Gangestad, S.W., Thornhill, R., Garver, C.E., 2002. Changes in women's sexual interests and their partner's mate-retention tactics across the menstrual cycle: evidence for shifting conflicts of interest. Proc. R. Soc. B 269, 975–982. https://doi.org/10.1098/rspb.2001.1952.

Gangestad, S.W., Simpson, J.A., Cousins, A.J., Garver-Apgar, C.E., Christensen, P.N., 2004. Women's preferences for male behavioral displays change across the menstrual cycle. Psychol. Sci. 15, 203–207. https://doi.org/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. Curr. Dir. Psychol. Sci. 14, 312–316. https://doi.org/10.1111/j.0963-7214.2005.00388.x.

Gangestad, S.W., Garver-Apgar, C.E., Simpson, J.A., Cousins, A.J., 2007. Changes in women's mate preferences across the ovulatory cycle. J. Pers. Soc. Psychol. 92, 151–163. https://doi.org/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. Evol. Hum. Behav. 37, 85–96. https://doi.org/10.1016/j.evolhumbehav.2015.09.001.

Gao, W., Stalder, T., Kirschbaum, C., 2015. Quantitative analysis of estradiol and six other steroid hormones in human saliva using a high throughput liquid chromatography–tandem mass spectrometry assay. Talanta 143, 353–358. https://doi.org/10.1016/j.talanta.2015.05.004.

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. Horm. Behav. 61, 157–166. https://doi.org/10.1016/j.yhbeh.2011.11.005.

Gildersleeve, K., Haselton, M.G., Fales, M.R., 2014a. Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. Psychol. Bull. 140, 1205–1259. https://doi.org/10.1037/a0035438.

Gildersleeve, K., Haselton, M.G., Fales, M.R., 2014b. Meta-analyses and p-curves support robust cycle shifts in women's mate preferences: reply to Wood and Carden (2014) and Harris, Pashler, and Mickes (2014). Psychol. Bull. 140, 1272–1280. https://doi.org/10.1037/a0037714.

González, J., 2004. Formant frequencies and body size of speaker: a weak relationship in adult humans. J. Phon. 32, 277–287. https://doi.org/10.1016/S0095-4470(03)00049-4.

Grebe, N.M., Emery Thompson, M., Gangestad, S.W., 2016. Hormonal predictors of women's extra-pair vs. in-pair sexual attraction in natural cycles: implications for extended sexuality. Horm. Behav. 78, 211–219. https://doi.org/10.1016/j.yhbeh.2015.11.008.

Hampson, E., Finestone, J.M., Levy, N., 2005. Menstrual cycle effects on perceptual closure mediate changes in performance on a fragmented objects test of implicit memory. Brain Cogn. 57, 107–110. https://doi.org/10.1016/j.bandc.2004.08.028.

Harries, M.L., Walker, J.M., Williams, D.M., Hawkins, S., Hughes, I.A., 1997. Changes in the male voice at puberty. Arch. Dis. Child. 77, 445–447. https://doi.org/10.1136/adc.77.5.445.

Harries, M., Hawkins, S., Hacking, J., Hughes, I., 1998. Changes in the male voice at puberty: vocal fold length and its relationship to the fundamental frequency of the voice. J. Laryngol. Otol. 112, 451–454. https://doi.org/10.1017/S0022215100140757.

Harris, C.R., 2011. Menstrual cycle and facial preferences reconsidered. Sex Roles 64, 669–681. https://doi.org/10.1007/s11199-010-9772-8.

Harris, C.R., 2013. Shifts in masculinity preferences across the menstrual cycle: still not there. Sex Roles 69, 507–515. https://doi.org/10.1007/s11199-012-0229-0.

Haselton, M.G., Gangestad, S.W., 2006. Conditional expression of women's desires and men's mate guarding across the ovulatory cycle. Horm. Behav. 49, 509–518. https://doi.org/10.1016/j.yhbeh.2005.10.006.

Havlíček, J., Roberts, S.C., Flegr, J., 2005. Women's preference for dominant male odour: effects of menstrual cycle and relationship status. Biol. Lett. 1, 256–259. https://doi.org/10.1098/rsbl.2005.0332.

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. Behav. Ecol. 26, 1249–1260. https://doi.org/10.1093/beheco/arv064.

Herrera, A.Y., Nielsen, S.E., Mather, M., 2016. Stress-induced increases in progesterone and cortisol in naturally cycling women. Neurobiol. Stress 3, 96–104. https://doi.org/10.1016/j.ynstr.2016.02.006.

Hodges-Simeon, C.R., Gurven, M., Gaulin, S.J., 2015. The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. Evol. Hum. Behav. 36, 294–302. https://doi.org/10.1016/j.evolhumbehav.2015.01.002.

Jones, B.C., Hahn, A.C., Fisher, C.I., Wang, H., Kandrik, M., Han, C., ... DeBruine, L.M., 2018a. No compelling evidence that preferences for facial masculinity track changes in women's hormonal status. Psychol. Sci. https://doi.org/10.1177/0956797618760197.

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

Jünger, J., Gerlach, T.M., Penke, L., 2018a. No Evidence for Ovulatory Cycle Shifts in Women's Preferences for Men's Behaviors in a Pre-registered Study. https://doi.org/10.17605/OSF.IO/7G3XC. (Retrieved from https://psyarxiv.com/7g3xc).

Jünger, J., Kordsmeyer, T., Gerlach, T.M., Penke, L., 2018b. Fertile women evaluate male bodies as more attractive, regardless of masculinity. Evol. Hum. Behav. 39, 412–423. https://doi.org/10.1016/j.evolhumbehav.2018.03.007.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2015. Package 'lmerTest'. R Package Version, 2.0-33.

Little, A.C., Jones, B.C., Burriss, R.P., 2007. Preferences for masculinity in male bodies change across the menstrual cycle. Horm. Behav. 51, 633–639. https://doi.org/10.1016/j.yhbeh.2007.03.006.

Marcinkowska, U.M., Ellison, P.T., Galbarczyk, A., Milkowska, K., Pawlowski, B., Thune, I., Jasienska, G., 2016. Lack of support for relation between woman's masculinity preference, estradiol level and mating context. Horm. Behav. 78, 1–7. https://doi.org/10.1016/j.yhbeh.2015.10.012.

Marcinkowska, U.M., Galbarczyk, A., Jasienska, G., 2018a. 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. https://doi.org/10.106/j.psyneuen.2017.11.007.

Marcinkowska, U.M., Kaminski, G., Little, A.C., Jasienska, G., 2018b. Average ovarian hormone levels, rather than daily values and their fluctuations, are related to facial preferences among women. Horm. Behav. 102, 114–119. https://doi.org/10.1016/j.yhbeh.2018.05.013.

Michaelis, D., Fröhlich, M., Strube, H.W., 1998. Selection and combination of acoustic features for the description of pathologic voices. J. Acoust. Soc. Am. 103, 1628–1639. https://doi.org/10.1121/1.421305.

Motta-Mena, N.V., Puts, D.A., 2017. Endocrinology of human female sexuality, mating, and reproductive behavior. Horm. Behav. 91, 19–35. https://doi.org/10.1016/j.yhbeh.2016.11.012.

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. Ann. Psychol. 30, 667–675. https://doi.org/10.6018/analesps.30.2.145221.

Oinonen, K.A., Mazmanian, D., 2007. Facial symmetry detection ability changes across the menstrual cycle. Biol. Psychol. 75, 136–145. https://doi.org/10.1016/j.biopsycho.2007.01.003.

Penton-Voak, I.S., Perrett, D.I., 2000. Female preference for male faces changes cyclically: further evidence. Evol. Hum. Behav. 21, 39–48. https://doi.org/10.1016/S1090-5138(99)00033-1.

Penton-Voak, I.S., Perrett, D.I., Castles, D.L., Kobayashi, T., Burt, D.M., Murray, L.K., Minamisawa, R., 1999. Menstrual cycle alters face preference. Nature 399, 741–742. https://doi.org/10.1038/21557.

Peters, M., Simmons, L.W., Rhodes, G., 2009. Preferences across the menstrual cycle for masculinity and symmetry in photographs of male faces and bodies. PLoS One 4, e4138. https://doi.org/10.1371/journal.pone.0004138.

Pillsworth, E.G., Haselton, M.G., 2006. Women's sexual strategies: the evolution of long-term bonds and extrapair sex. Annu. Rev. Sex Res. 17, 59–100. https://doi.org/10.1080/10532528.2006.10559837.

Pisanski, K., Hahn, A.C., Fisher, C.I., DeBruine, L.M., Feinberg, D.R., Jones, B.C., 2014. Changes in salivary estradiol predict changes in women's preferences for vocal masculinity. Horm. Behav. 66, 493–497. https://doi.org/10.1016/j.yhbeh.2014.07.006.

Puts, D.A., 2005. Mating context and menstrual phase affect women's preferences for male voice pitch. Evol. Hum. Behav. 26, 388–397. https://doi.org/10.1016/j.evolhumbehav.2005.03.001.

Puts, D.A., 2006. Cyclic variation in women's preferences for masculine traits: potential hormonal causes. Hum. Nat. 17, 114–127. https://doi.org/10.1007/s12110-006-1023-x.

Puts, D.A., Gaulin, S.J., Verdolini, K., 2006. Dominance and the evolution of sexual dimorphism in human voice pitch. Evol. Hum. Behav. 27, 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003.

Puts, D.A., Hodges, C.R., Cárdenas, R.A., Gaulin, S.J., 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. Evol. Hum. Behav. 28, 340–344. https://doi.org/10.1016/j.evolhumbehav.2007.05.002.

Puts, D.A., Apicella, C.L., Cárdenas, R.A., 2012. Masculine voices signal men's threat potential in forager and industrial societies. Proc. R. Soc. B 279, 601–609. https://doi.org/10.1098/rspb.2011.0829.

Puts, D.A., Bailey, D.H., Cárdenas, R.A., Burriss, R.P., Welling, L.L., Wheatley, J.R., Dawood, K., 2013. Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle. Horm. Behav. 63, 13–19. https://doi.org/10.1016/j.yhbeh.2012.11.007.

Puts, D.A., Hill, A.K., Bailey, D.H., Walker, R.S., Rendall, D., Wheatley, J.R., ... Ramos-Fernandez, G., 2016. Sexual selection on male vocal fundamental frequency in humans and other anthropoids. Proc. R. Soc. B Biol. Sci. 283. https://doi.org/10.1098/rspb.2015.2830.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria Retrieved from. http://www.R-project.org/.

Revelle, W., 2016. Psych: Procedures for Personality and Psychological Research (Version 1.6.12) [Computer Software]. Retrieved from. https://CRAN.R-project.org/packagepsych.

Roney, J.R., Simmons, Z.L., 2013. Hormonal predictors of sexual motivation in natural menstrual cycles. Horm. Behav. 63, 636–645. https://doi.org/10.1016/j.yhbeh.2013.02.013.

Roney, J.R., Simmons, Z.L., 2015. Elevated psychological stress predicts reduced estradiol concentrations in young women. Adapt. Hum. Behav. Physiol. 1, 30–40. https://doi.org/10.1007/s40750-014-0004-2.

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. Horm. Behav. 81, 45–52. https://doi.org/10.1016/j.yhbeh.2016.03.008.

Roney, J.R., Simmons, Z.L., 2017. Ovarian hormone fluctuations predict within-cycle shifts in women's food intake. Horm. Behav. 90, 8–14. https://doi.org/10.1016/j.yhbeh.2017.01.009.

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

Scott, I.M., Clark, A.P., Boothroyd, L.G., Penton-Voak, I.S., 2012. Do men's faces really signal heritable immunocompetence? Behav. Ecol. 24, 579–589. https://doi.org/10.1093/beheco/ars092.

Scott, I.M., Clark, A.P., Josephson, S.C., Boyette, A.H., Cuthill, I.C., Fried, R.L., ... Penton-Voak, I.S., 2014. Human preferences for sexually dimorphic faces may be evolutionarily novel. Proc. Natl. Acad. Sci. 111, 14388–14393. https://doi.org/10.1073/pnas.1409643111.

Shimoda, R., Campbell, A., Barton, R.A., 2018. Women's emotional and sexual attraction to men across the menstrual cycle. Behav. Ecol. 29, 51–59. https://doi.org/10.1093/beheco/arx124.

Simmons, L.W., Peters, M., Rhodes, G., 2011. Low pitched voices are perceived as masculine and attractive but do they predict semen quality in men? PLoS One 6, e29271. https://doi.org/10.1371/journal.pone.0029271.

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. Horm. Behav. 65, 319–327. https://doi.org/10.1016/j.yhbeh.2014.01.006.

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

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

Thornhill, R., Chapman, J.F., Gangestad, S.W., 2013. Women's preferences for men's scents associated with testosterone and cortisol levels: patterns across the ovulatory cycle. Evol. Hum. Behav. 34, 216–221. https://doi.org/10.1016/j.evolhumbehav.2013.01.003.

Tusing, K.J., Dillard, J.P., 2000. The sounds of dominance: vocal precursors of perceived dominance during interpersonal influence. Hum. Commun. Res. 26, 148–172. https://doi.org/10.1111/j.1468-2958.2000.tb00754.x.

Wallen, K., Rupp, H.A., 2010. Women's interest in visual sexual stimuli varies with menstrual cycle phase at first exposure and predicts later interest. Horm. Behav. 57, 263–268. https://doi.org/10.1016/j.yhbeh.2009.12.005.

Welling, L.L., Jones, B.C., DeBruine, L.M., Conway, C.A., Smith, M.L., Little, A.C., ... Al-Dujaili, E.A., 2007. Raised salivary testosterone in women is associated with increased attraction to masculine faces. Horm. Behav. 52, 156–161. https://doi.org/10.1016/j.yhbeh.2007.01.010.

Wickham, H., 2011. The split-apply-combine strategy for data analysis. J. Stat. Softw. 40, 1–29. https://doi.org/10.18637/jss.v040.i01.

Wolff, S.E., Puts, D.A., 2010. Vocal masculinity is a robust dominance signal in men. Behav. Ecol. Sociobiol. 64, 1673–1683. https://doi.org/10.1007/s00265-010-0981-5.

Wood, W., Kressel, L., Joshi, P.D., Louie, B., 2014. Meta-analysis of menstrual cycle effects on women's mate preferences. Emot. Rev. 6, 229–249. https://doi.org/10.1177/1754073914523073.

Xue, S.A., Hao, G.J., 2003. Changes in the human vocal tract due to aging and the acoustic correlates of speech production: a pilot study. J. Speech Lang. Hear. Res. 46, 689–701. https://doi.org/10.1044/1092-4388(2003/054).