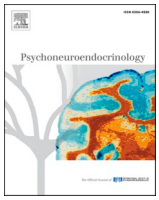
Psychoneuroendocrinology 148 (2023) 105977

Contents lists available at ScienceDirect

Psychoneuroendocrinology

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



Does emotion recognition change across phases of the ovulatory cycle?

Yasaman Rafiee [a,e,*], Julia Stern [d], Julia Ostner [b,e,f], Lars Penke [c,e], Anne Schacht [a,e]

[a] Department of Affective Neuroscience and Psychophysiology, Institute of Psychology, University of Goettingen, 37073 Goettingen, Germany
[b] Department of Behavioral Ecology, Institute for Zoology and Anthropology, University of Goettingen, 37077 Goettingen, Germany
[c] Department of Biological Personality Psychology, Institute of Psychology, University of Goettingen, 37073 Goettingen, Germany
[d] Department of Personality Psychology and Psychological Assessment, Institute of Psychology, University of Bremen, 28259 Bremen, Germany
[e] Leibniz ScienceCampus "Primate Cognition", 37077 Goettingen, Germany
[f] Primate Social Evolution Group, German Primate Center, Leibniz Institute for Primate Research, 37077 Goettingen, Germany

ARTICLE INFO

Keywords:
Estradiol
Progesterone
Ovulatory cycle
Multisensory emotion recognition
Expressions of emotions

ABSTRACT

Recognizing emotions is an essential ability for successful interpersonal interaction. Prior research indicates some links between the endocrine system and emotion recognition ability, but only a few studies focused on within-subject differences across distinct ovulatory cycle phases and this ability. These studies have demonstrated mixed results that might be potentially due to heterogeneity in experimental tasks, methodologies, and lacking ecological validity. In the current study, we investigated associations between within-subject differences in ovarian hormones levels and emotion recognition from auditory, visual, and audiovisual modalities in $N = 131$ naturally cycling participants across the late follicular and mid-luteal phase of the ovulatory cycle. We applied a within-subject design with sessions in the late follicular and mid-luteal cycle phase, and also assessed salivary progesterone and estradiol in these sessions. Our findings did not reveal any significant difference in emotion recognition ability across two cycle phases. Thus, they emphasize the necessity of employing large-scale replication studies with well-established study designs along with proper statistical analyses. Moreover, our findings indicate that the potential link between ovulatory cycle phases (late follicular and mid-luteal) and emotion recognition ability might have been overestimated in previous studies, and may contribute to theoretical and practical implications of socio-cognitive neuroendocrinology.

1. Introduction

Emotions evolved to help individuals to deal with various life tasks, including mating, resource finding, danger identification, and parenting (Al-Shawaf et al., 2015). One of their functions is to improve individuals' chances of survival and ultimately their reproductive success (Fischer and Manstead, 2009). In social contexts, emotional expressions carry important information supporting individuals to regulate their responses to environmental opportunities and risks (Keltner et al., 2016), also in the service of successful interpersonal interactions (Schlegel et al., 2016). The underlying mechanisms of emotion recognition are not completely understood; nevertheless, it is plausible to assume that this ability is influenced by neurochemicals, including hormones (Thagard, 2002). Some previous studies examined the link between endogenous or exogenous hormones and emotion recognition, including the specific role of female sex hormones fluctuating across women's menstrual cycle (e.g., Derntl et al., 2008, Gingnell et al., 2019; Lausen et al., 2020; Maner and Miller, 2014; Pahnke et al., 2019; Shirazi et al., 2020).

The menstrual cycle, due to its periodicity, provides a natural model to study relationships between female sex hormones, cognition, and emotion (Poromaa and Gingnell, 2014), and can roughly be divided into two main phases, namely follicular and luteal, across which the levels of ovarian hormones, i.e. estradiol and progesterone fluctuate in a cyclic fashion. The fluctuation of ovarian hormones, which is highly related to the reproductive state, could be associated with the processing of emotional expressions as an important component of reproductive success (Gingnell et al., 2019). Ovarian hormones could potentially yield alterations in women's recognition of emotions assumed to be involved in the facilitation of social interactions (Derntl et al., 2008) and flagging social threats (Maner and Miller, 2014), respectively. A high ability to recognize emotions might thus increase the chance of successful social

* Corresponding author at: Department of Affective Neuroscience and Psychophysiology, Institute of Psychology, University of Goettingen, 37073 Goettingen, Germany.
E-mail address: yasaman.rfe@psych.uni-goettingen.de (Y. Rafiee).

https://doi.org/10.1016/j.psyneuen.2022.105977
Received 29 June 2022; Received in revised form 15 November 2022; Accepted 15 November 2022
Available online 24 November 2022

interaction and, as a result, a higher chance for reproductive success (Derntl et al., 2008, Kamboj et al., 2015; Gingnell et al., 2019).

A recent review (Osório et al., 2018) and some studies (Derntl et al., 2008;, Pearson and Lewis, 2005; Rubin, 2012) suggested an improved emotion recognition accuracy (ERA) in the follicular compared to the luteal phase, presumably regulated by estradiol levels. Limited evidence indicated an overall impairment of emotion recognition by increased progesterone levels (see Osório et al., 2018). On the contrary, several studies found no evidence for a relationship between cycle phase or ovarian hormones with emotion recognition in healthy naturally cycling women (Di Tella et al., 2020; Kamboj et al., 2015; Pahnke et al., 2019; Shirazi et al., 2020; Zhang et al., 2013). In sum, the evidence concerning a possible association between cycle phases, ovarian hormones, and emotion recognition ability is inconsistent.

Regarding the interplay of ovarian hormones and specific emotions, previous findings are also mixed or even contradictory (for an overview see Gamsakhurdashvili et al., 2021; Osório et al., 2018). Sakaki and Mather (2012) suggested that increased levels of estradiol are related to a reduced reaction to negative stimuli, supported by reports of negative relationships between estradiol levels and accuracy in recognizing anger (Guapo et al., 2009; Kamboj et al., 2015) and disgust (Kamboj et al., 2015). Contradictory, Pearson and Lewis (2005) found a higher ability to recognize fear during the fertile phase, employing a between-subject study with a small sample ($N = 50$) and lacking hormone level measurement. Some studies reported higher levels of progesterone to be associated with an enhanced ability to recognize negative emotions, i.e., expressions of fear and disgust (Conway et al., 2007), and of anger, fear, disgust, and sadness (Maner and Miller, 2014), which was explained by an assumed behavioral defense mechanism to avoid physical danger or contamination (Conway et al., 2007). Other studies failed to demonstrate an association between ovarian hormones and recognizing specific emotions (e.g, Zhang et al., 2013).

Reasons for these inconsistencies might lie in the high variation in methodologies used for determining the cycle phases of interest (Allen et al., 2016), the absence of or differences in hormonal assessments, within vs. between-subject comparisons (Gingnell et al., 2019), the lack of statistical power to detect intraindividual differences (Schmalenberger et al., 2021), and finally, the employment of different experimental tasks (Gingnell et al., 2019; Gamsakhurdashvili et al., 2021). Recently, Shirazi et al. (2020) attempted to address these issues in a large-scale study with methodological rigor (e.g., high statistical power, proper analysis, direct hormonal measures), and did not find compelling evidence for a relationship between levels of ovarian hormones and the recognition of facially expressed complex emotions across two cycle phases (late follicular and mid-luteal). In spite of the methodological strength of their study, the ecological validity was limited: In a natural situation, emotions are usually expressed not only via faces but also via other modalites such as voices (Collignon et al., 2010) or in a bimodal context. So far, no study has investigated the association of ovarian hormone levels across two phases of the ovulatory cycle with emotion recognition from other modalities except faces.

In this preregistered study (https://osf.io/dkpf5/), we investigated within-subject differences in cycle phase and associated ovarian hormones levels and the recognition of different emotions expressed by faces, voices, and face-voice combinations across two cycle phases namely late follicular and mid-luteal. The reason to choose these phases across the menstrual cycle is to capture the peak of estradiol levels in the late follicular phase, and progesterone levels in the mid-luteal phase. We strictly followed recommendations for cycle studies by employing a within-subject design, direct hormone measurements, and luteinizing hormone (LH) tests to validate cycle phase estimates (Gangestad et al., 2016; Allen et al., 2016; Kiesner et al., 2020; Schmalenberger et al., 2021). Based on the existing literature, we hypothesized for the face domain H1) improved emotion recognition during the late follicular compared to the mid-luteal phase; that H2a) improved emotion recognition is related to increased estradiol levels, H2b) decreased emotion recognition is related to increased progesterone levels, and H2c) increased progesterone levels are related to a negativity bias, i.e. improved recognition of threat-related (angry, fearful, disgusted) compared to positive (happy) expressions. In an exploratory manner, we additionally examined whether the within-subject fluctuation of ovarian hormones across both cycle phases was differently associated with ERA in facial, vocal, and audiovisual expressions. We further tested for a potential association between the within-subject fluctuation of ovarian hormones across both cycle phases and ERA in response to stimuli expressed by male and female actors, as heterosexual women might pay more attention to male stimuli, potentially due to increased mating interest (e.g., Jünger et al., 2018). Finally, as previous studies suggest that between- rather than within-subjects hormone levels might be associated with psychological outcomes (e.g. Marcinkowska et al., 2018), we investigated relationships between averaged hormone levels and emotion recognition ability in an exploratory manner (following Shirazi et al., 2020).

2. Methods

The study was approved by the ethics committee of the Psychological Institute of the University of Goettingen. Each participant signed a consent form following the Declaration of Helsinki (DoH) ethical principles for human subjects. Before data collection, the study aims, hypotheses, and study design were preregistered at the open science framework (https://osf.io/dkpf5/). The open data and analysis script are available. Participants were compensated with either course credit or monetary rewards of 25€. As an additional incentive, participants who completed the study had the opportunity to win one of four Amazon gift vouchers with a maximum value of 50€.

2.1. Participants

Referring to Gangestad et al. (2016), achieving 80% power to detect a Cohen's d of 0.5 requires $N = 48$ participants in a within-subject study with two sessions each and LH validated fertile phase estimates. In the current study, our sample size substantially exceeded the mentioned recommendation to gain 80% statistical power to detect a medium-sized effect. Thus, the current study had sufficient power to detect even smaller effect sizes, also if we restricted our sample regarding women with LH-validated cycle phase estimates. Given the sample sizes and designs of previous studies, our study should at least have sufficient test power to detect previously reported effect sizes.

In total, $N = 131$ out of 180 females completed the study. Forty-nine participants withdrew from the study due to the following reasons: 1) having no more interest or time (22 subjects), 2) experiencing irregular cycle (cycle length less than 25 or more than 35 days) or intermenstrual bleeding through the course of study (16 subjects), 3) taking hormonal contraceptives (five subjects), 4) sickness (two subjects), 5) two participants were excluded by the decision of research team due to lack of laboratory capacity to prolong the course of the study, and 6) there was a mistake in sampling hormones for two participants, and therefore their samples and data were eliminated. Seventy-four participants observed a positive LH test in the estimated fertile phase (not more than three days before and two days after the late follicular session).

Included participants were German native speakers, self-reported being healthy, heterosexual, non-pregnant, naturally-cycling women with a cycle length of 25–35 days ($M_{Cycle\ length} = 28.77$, $SD_{Cycle\ length} = 2.03$), between 18 and 35 years old ($M_{Age} = 24.1$, $SD_{Age} = 3.5$), and having had a regular ovulatory cycle for at least three months before their first participation in the study. All participants reported normal or corrected-to-normal vision and normal hearing, without any history of psychiatric, neurological, metabolic, or hormonal disorders. In addition, they did not use any sort of hormonal medications such as contraceptives, nor did they breastfeed for at least three months before their first participation in the study. Of the 131 included participants, sixty

participants reported being single, sixty-three were in a relationship, two were engaged, three were married, and three reported to be in different forms of relationships.[1] One-hundred-nineteen participants were righthanded.

The range of salivary estradiol levels excluding outliers (± 3 SDs) in the late follicular phase was between 1.05 and 21.76 ($M_{Estradiol}$ = 6.76, SD = 4.47) pg/mL and the range of progesterone levels was between 14.60 and 218.16 ($M_{Progesterone}$ = 39.4, SD = 25.2) pg/mL. In the mid-luteal phase, the range of estradiol levels excluding outliers (± 3 SDs) was between 0.70 and 19.62 ($M_{Estradiol}$ = 6.12, SD = 3.82) and the range of progesterone levels was between 20.18 and 266.16 ($M_{Progesterone}$ = 83.8, SD = 45.8) pg/mL.

2.2. Procedure

The study consisted of an introductory session and two testing sessions that took place in the estimated late follicular and mid-luteal phases of each participants' ovulatory cycle, respectively. More precisely, we estimated the cycle days with the highest probability of being in the fertile or luteal phase based on backward counting from the expected next menstrual onset, as well as the average cycle length. These estimates were then validated with luteinizing hormone tests (see below for more details). To minimize potential carry-over effects, the order of testing sessions was counterbalanced across participants. Of the $N = 131$ participants, 63 started the first testing session in their late follicular phase and 68 started their first testing session in their mid-luteal phase. On average, intervals between the two testing sessions were 19.55 days (SD = 14.03, SEM = 1.23).

2.2.1. Introductory session

First, participants were screened according to the inclusion criteria. We estimated the onset of the next menstruation and used the backward-counting method to predict the ovulation date (Jünger et al., 2018). Moreover, to validate the fertile phase estimate, participants were asked to use highly sensitive (10mIU/mL) urine ovulation test strips from Runbio Biotech Co., Ltd., as soon as their menstruation ended and report to us whenever they saw a positive test. To standardize the influence of possible physiological factors, we asked participants to use LH strips between 10 am and 8 pm, preferably at the same time of the day. We also asked participants to send us photos of their LH tests on a voluntary basis.

2.2.2. Testing sessions

Sessions two and three took place in the late follicular and the mid-luteal phase of each participants' ovulatory cycle. Following Jünger et al. (2018) the late follicular phase was estimated as reverse cycle days[2] 16–18, with reverse cycle days 16 as the most ideal date. The mid-luteal phase was considered reverse cycle days 4–11, with reverse cycle days 6–8. In each testing session, participants first completed a computer-based screening questionnaire with regard to their health status and saliva sampling, adapted from Schultheiss and Stanton (2009) and Jünger et al. (2018), and the PANAS mood questionnaire[3] from Breyer and Bluemke (2016). Next, saliva samples were collected before participants performed the emotion recognition task.

2.3. Saliva sampling

To minimize the potential effect of the emotion recognition task on hormone levels, saliva samples were collected before the task. In each testing session, participants were asked to salivate into tubes via passive drool. Each sample was collected in tubes (max. 2 mL) from IBL SaliCap and kept frozen at − 80 °C until the delivery on dry ice to the laboratory for hormonal analysis. To reduce any risk of sample contamination such as blood or food debris, participants were asked to refrain from eating, drinking (except plain water), and teeth brushing for at least one hour before coming to the laboratory. After collecting the samples, a visual inspection was performed to lessen the risk of blood contamination in samples.

2.3.1. Hormone measures

Levels of estradiol (E2) and progesterone (P4) were measured via the Chemiluminescence Immunoassays method at the Endocrinology Laboratory at the Technical University of Dresden. Although our samples were analyzed as a single determination which is less accurate as compared to the duplicate determination; the lab still reported their procedure to determine the coefficient of variation below 10% and we furthermore found a highly significant association between cycle phase and estradiol to progesterone ratio (E/P) (β = 0.116, SE = 0.000, 95% CI = [0.11; 0.12], t = 133.1, $p < 0.001$) as an external validation for the hormonal measures. Additionally, two separate linear mixed models were performed to investigate whether levels of estradiol and progesterone differed across the two investigated cycle phases. In each model, one of the hormones (log-transformed) was included as the outcome variable, phase of the cycle as the predictor, and participant ID as the random effect. The model with estradiol as the outcome, showed a significant drop of estradiol levels in the mid-luteal phase compared to the late follicular phase (β = −0.09, SE = 0.002, 95% CI = [−0.09; −0.10], $p < 0.001$), and the model with progesterone as the outcome showed a significant rise of the progesterone levels in the mid-luteal phase compared to the late follicular phase (β = 0.84, SE = 0.002, 95% CI = 0.84; 0.85], $p < 0.001$). To minimize the possible diurnal fluctuations of hormones, all sessions were scheduled in the afternoon between 12.00 pm and 04.00 pm. Most of the participants were examined at the same time of the day for both sessions.

2.3.2. Handling hormonal data[4]

As preregistered and following previous studies (e.g., Jones et al., 2018; Stern et al., 2021), outliers of hormone measures ± 3 SDs from the sample mean were excluded. In total $N = 5$ including three measures of estradiol and two measures of progesterone were omitted from the data. Before including the variables in our statistical analysis, hormone values were visually inspected to see if they are distributed symmetrically. To check the distribution of estradiol and progesterone, a Shapiro-Wilk test was computed and showed that the distribution of both hormones significantly departed from normality (estradiol: W = 0.90, $p < 0.01$, progesterone: W = 0.83, $p < 0.01$).

To track the within-subject fluctuation of ovarian hormones across the ovulatory cycle, hormonal measures were subject mean-centered and scaled by being divided by a constant. The values varied from −.5 to .5, which eases the calculation in the linear mixed model (e.g. Jünger et al., 2018). Subject mean-centering distinguishes the effect of within-

[1] Two participants changed their relationship status from the first testing session to the second testing session, from "single" into "in an open relationship" (one participant), and from "other" into "in an open relationship" (one participant).

[2] Reverse cycle day or the backward-counting method is often used for estimating a woman's position in the menstrual cycle. This method counts days backward from the day one of the new cycle to the day of assessment (for an overview see Gangestad et al., 2016). Thus, reverse cycle day 16 means 16 days before the next menstrual onset.

[3] The mood questionnaire was part of a different study.

[4] In our preregistration, we wrote that we will log-transform hormone values to achieve normal distribution of hormone values. However, in the meantime, we learned that log-transformation might not be a good proxy for within-subjects hormonal mechanisms that may regulate fertility related cycle shifts (Roney, 2019). Importantly, Our results did not change when including log-transformed, subject mean-centered hormone measures or untransformed hormonal measures. Thus, no considerable difference was found in results regarding the applied procedures for handling hormonal data.

and between-subject variation of hormones, and therefore this method is recommended to track the influence of hormonal fluctuations across the ovulatory cycle (see Schmalenberger et al., 2021). To investigate the association between ovarian hormone variation between different individuals and their emotion recognition ability, hormonal measures were averaged across two sessions for each participant. Importantly, adding between-subject effects to our analyses did not affect any of the within-subjects results. We then log-transformed (base e) the average hormonal measures representing the between-subject levels of estradiol and progesterone. To facilitate the interpretation of model outcome and model convergence, we then z-transformed previously log-transformed between-subject levels of estradiol and progesterone.

2.4. Emotion recognition task

The emotion recognition task was adopted from Lausen et al. (2020) and included three separated blocks presenting facial, vocal, and audiovisual (combined facial and vocal) expressions of emotions. In each block, 144 randomized stimuli consisting of five basic emotions (angry, happy, sad, disgust, fear) and neutral expressions from female and male actors were presented. The order of blocks was randomized between participants, but constant for each participant within testing sessions. Participants received a message in the center of the screen at the end of each block asking whether they would like to take a break or whether they would like to continue. The experiment was resumed by pressing the Spacebar key. We measured emotion recognition accuracy and reaction times. However, as the emphasis of the task setup was on the accuracy, reaction times were not included in the inferential analysis.

At the beginning of the task, there were three practice trials to familiarize participants with the experimental procedure. Each trial started with a blank screen (1000 ms), followed by a fixation cross (1000 ms). The stimulus was presented after the fixation cross. The duration of stimulus presentation varied between 319 ms and 4821 ms (M = 1.84, SD = 1.12). After the presentation of the stimulus, a circular answer display containing all six categories of interest (i.e., anger, disgust, fear, happiness, neutral, and sadness) and the selection cursor (which appeared in the center of the display) was presented. Participants were asked to choose the correct emotion as accurately and quickly as possible. There was no time limitation to answer each trial (Fig. 1). The order of emotion labels was randomized across participants but was constant for each participant.

2.5. Stimuli

All stimuli were taken from the study by Lausen et al. (2020). The face stimuli were extracted from the Radboud face database (Langner et al., 2010). In total, 24 face identities including 12 females and 12 males were employed to create visual stimuli and were matched in their luminance. The auditory stimuli consisted of affect bursts from Montreal Affective Voices (Belin et al., 2008), pseudo-words from Magdeburg Prosody Corpus (Wendt and Scheich, 2002), and pseudo-sentences (Paulmann and Kotz, 2008) validated by Lausen and Hammerschmidt (2020). The loudness and background noises were adjusted by Adobe Audition CC (Version 8.1, Adobe 4 Systems, 2015, San Jose, CA). Audiovisual stimuli were created by combining visual and auditory stimuli, using Adobe Premiere Pro videos (see Lausen et al., 2020), with matched emotion category and sex of the actor.

2.6. Statistical analyses[5]

Data analyses were performed using R Software version 4.0.3 and R studio version 1.4.1106. Generalized Linear Mixed Models (GLMMs) with binomial error structure and logit link function was applied. To make inferences, standard p-value 0.05 was used as the cut-off criterion for two-tailed distributions. To preprocess the data, the following packages were used: tidyverse 1.3.1, knitr 1.33, dplyr 1.0.5. We used ggplot2 3.3.3, and sjplot 2.8.7 for data visualization, lme4 1.1.26 for computing models, and car 3.0.10 for assessing collinearity among predictors. Variance Inflation Factors (VIF) with a model lacking the interaction showed no collinearity issue in our models (maximum VIF: 1.044). To deal with the convergence issue we added "bobyqa" optimizer to fit the models (see supplementary document).

In each model, session number (first vs. second session) served as the variable to control for potential order effects. The outcome variable was emotion recognition accuracy (correct vs. incorrect). Since the same individuals were tested twice, we included subject ID as the random intercept in fitted models. To inspect the goodness of fit of the fitted model, we compared the log-likelihood function of the fitted model with the log-likelihood function of the minimal (reduced) model lacking the predictor or the interaction of interest, as recommended by Dobson (2002). In addition, model stability was estimated by dropping the levels of random effect one at a time and comparing the estimates derived from models fitted on the respective subsets with those obtained for the full data set. Model stability estimates revealed good stability for all models.

3. Results[6]

3.1. Descriptive statistics

Across both cycle phases, women had the highest performance in recognizing expressed emotions in the audiovisual modality ($M_{\text{proportion correct responses}}$ (PCR) = 0.96) and the lowest performance in recognizing auditory emotional expression (^{M}PCR =0.81) (see Table 1 for emotion recognition performance in each cycle phase). The most recognized emotion was the neutral expression (^{M}PCR =0.94) and the least recognized emotion was the disgust expression (^{M}PCR =0.81) across all phases of the ovulatory cycle. The recognition of emotions expressed by female actors was (^{M}PCR =0.91) and for male actors was (^{M}PCR =0.89).

3.2. Cycle shifts and facial emotion recognition (H1)

First, we investigated potential ovulatory cycle shifts in facial emotion recognition. We included the cycle phase as the fixed effect, session number as the control variable, and subject ID as the random intercept. The reference category for comparison was the mid-luteal phase. The included participants ($n = 74$) in this model observed a positive LH test during the optimal days (maximum of three days before and two days after their estimated day of ovulation) (see Blake et al., 2016). The model showed no significant differences in women's emotion recognition performance between the late follicular and mid-luteal phase of the ovulatory cycle ($\beta = -0.031$, SE = 0.028, 95% CI = [0.92; 1.02], z = −1.122, OR= 0.97, $p = 0.262$; Fig. 2, right panel). We further compared the log-likelihood of this model with a model lacking the ovulatory cycle phases to examine the goodness of fit of our model. The result showed no significant difference between the main model and the model lacking the ovulatory cycle phase ($\chi^2 = 1.25$, $df = 1$, $p = 0.262$). However, participants showed a better performance in the second session compared to the first session ($\beta = 0.315$, SE = 0.028, 95% CI = [1.30; 1.45], z = 11.388, OR= 1.37, $p < 0.001$), see Table 2. To control for the robustness of the findings, we fitted an additional model including all participants ($N = 131$). The results were consistent

[5] The analysis plan was preregistered as Generalized Linear Model (GLM); however, to avoid pseudoreplication caused by the repeated measure design, we applied Generalized Linear Mixed Model (GLMM) in this paper.

[6] An extra model was fitted including E/P, its interaction with emotion category and stimulus sex, and the maximal random slope with all possible interactions. The results were in line with above-mentioned models (see supplementary material).

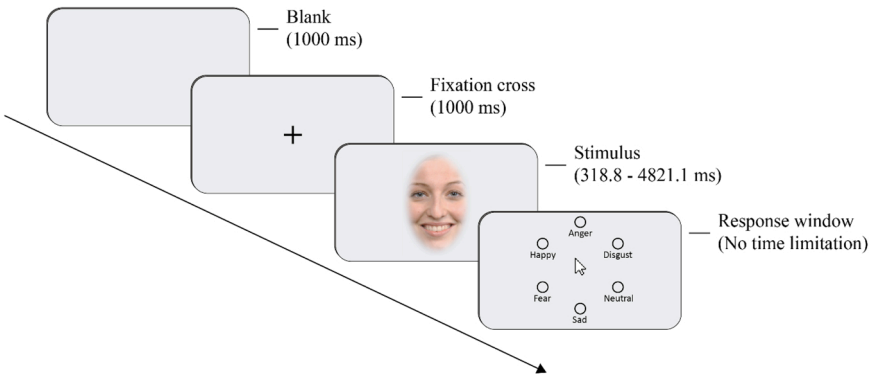


Fig. 1. Schematic overview of the trial scheme.

Table 1
Mean and standard deviation (SD) of emotion recognition accuracy (proportion of correct responses) and reaction times (seconds) for modality, emotion category, and sex of the actor across the ovulatory cycle, $N = 131$.

Late Follicular Phase				
Modality	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Audiovisual	0.962	0.037	1.275	0.348
Auditory	0.813	0.076	1.629	0.49
Visual	0.912	0.049	1.299	0.28
Emotion Category	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Anger	0.932	0.044	1.378	0.35
Disgust	0.81	0.105	1.489	0.427
Fear	0.905	0.072	1.551	0.373
Happy	0.93	0.047	1.206	0.262
Neutral	0.935	0.063	1.349	0.366
Sad	0.863	0.089	1.433	0.343
Sex of the actor	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Female	0.905	0.044	1.381	0.314
Male	0.886	0.05	1.421	0.316
Mid-Luteal Phase				
Modality	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Audiovisual	0.963	0.038	1.233	0.29
Auditory	0.816	0.079	1.612	0.563
Visual	0.914	0.053	1.282	0.316
Emotion Category	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Anger	0.933	0.059	1.326	0.342
Disgust	0.809	0.096	1.430	0.392
Fear	0.906	0.07	1.524	0.388
Happy	0.934	0.042	1.218	0.292
Neutral	0.938	0.07	1.355	0.71
Sad	0.865	0.088	1.402	0.332
Sex of the actor	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Female	0.908	0.044	1.360	0.367
Male	0.887	0.056	1.392	0.321

with the initial model (see supplementary document) Table 2.

3.3. Association between ovarian hormones levels and facial emotion recognition (H2a, 2b, 2c)

Next, we tested the association of within- and between-subjects variation of estradiol and progesterone levels and facial emotion recognition accuracy in the model including all participants ($N = 131$). In addition, the interplay of within-subject fluctuation of progesterone levels and threat-related emotions (progesterone × threat-related emotions) in facial emotion recognition accuracy was explored. Referring to our hypotheses the threat-related emotions including anger, disgust and fear were entered in the model and the happy expression was set as the reference category. Hence, sad and neutral expressions were dropped in this model, as they were not part of our hypotheses. Again, the session number served as a control variable. Subject ID was added as random intercept, and emotion category as random slope. The analysis revealed no significant association of within- or between-subject levels of estradiol or progesterone and facial emotion recognition (Fig. 2, panel left). Furthermore, the interaction between the within-subject fluctuation of progesterone levels and threat-related emotions was not significantly related to facial emotion recognition. The effect of session number was significant and suggests that women performed better in the second session (Table 3). The model outcome also showed that the emotion category contributes significantly to predicting the outcome variable and the happy expression had the highest recognition rate compared to threat-related emotions (Table 3).

The main model was compared with three different null models including all participants ($N = 131$). The first comparison was between the model and the null model lacking estradiol measures (within- and between-subject). The likelihood ratio test revealed no significant difference between the two models ($\chi^2 = 1.351$, $df = 2$, $p = 0.509$). Further, the main model was compared with a null model lacking progesterone measures (within- and between-subject). Again, the main model did not significantly explain the outcome better than the null model ($\chi^2 = 3.35$, $df = 5$, $p = 0.646$). Lastly, we compared the main model with a null model lacking the interaction between progesterone measures (within-subject) and the emotion category, which showed that the main and the null model are not significantly different ($\chi^2 = 2.07$, $df = 3$, $p = 0.556$). These comparisons imply that ovarian hormones and the interaction term did not predict emotion recognition considerably in our model.

We also fitted a separate model ($N = 131$) including within- and between-subject estradiol to progesterone ratio (E/P) rather than both hormones separately. The results showed no significant association of within- and between-subject E/P and facial emotion recognition. Again, participants showed a better performance in the second session compared to the first session. The main effect of the factor emotion category revealed that the happy expression was recognized significantly better than threat-related expressions namely anger, disgust, and fear The results of the likelihood ratio test for the model including E/P were consistent with the model tested ovarian hormones (see supplementary document).

3.4. Exploratory analysis

As preregistered, in $N = 131$ participants, we investigated if there was a moderating effect of stimulus modality on the association of within-subject fluctuations of ovarian hormones and emotion recognition accuracy. Session number was included as a control variable, subject ID was added as a random intercept, and stimulus modality as a random slope. Our analysis did not reveal a significant interaction

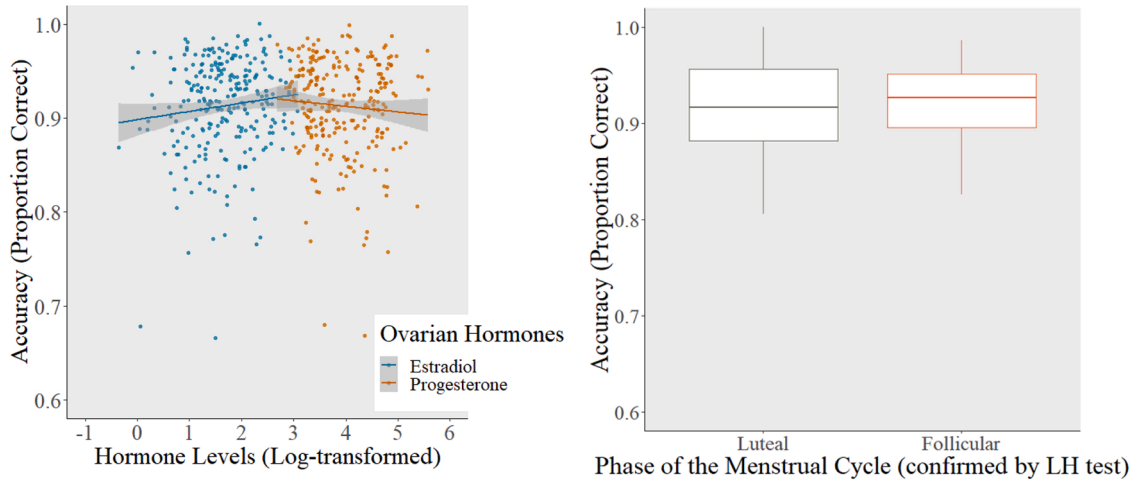


Fig. 2. Facial emotion recognition accuracy across the ovulatory cycle (right panel). The association of between-subject levels of ovarian hormones and facial emotion recognition accuracy (left panel).

Table 2
Results of the Generalized Linear Mixed Model testing ovulatory cycle shifts in facial emotion recognition.

	Estimates	*SE*	*z*	*p*	*OR*	*95% CI*
Model Phase with confirmed fertile phase ($n = 74$)						
Phase [late follicular]	-0.031	0.028	-1.122	0.262	0.97	0.92 – 1.02
Session	0.315	0.028	11.388	**< 0.001**	1.37	1.30 – 1.45
Model Phase ($N = 131$)						
Phase [late follicular]	-0.026	0.020	-1.330	0.184	0.97	0.94 – 1.01
Session	0.288	0.020	14.574	**< 0.001**	1.33	1.28 – 1.39

between ovarian hormone levels and stimulus modality. As in previous models, participants showed a better performance in the second session. The model showed the significant main effect of stimulus modality in which the audiovisual expression has the highest recognition rate compared to the other two modalities (see supplementary document).

Furthermore, we examined if the sex of the presented stimulus moderates the association of within-subject fluctuations of ovarian hormones and emotion recognition accuracy. As in previous models, the session number was entered as a control variable. Subject ID was entered as a random intercept. The results showed that hormone levels across the ovulatory cycle were not differentially associated with emotion recognition presented by male and female actors. The effect of session number was significant and participants showed a better performance in the second session. The results also showed significantly better recognition of emotions expressed by female actors (see supplementary document).

4. Discussion

The current study aimed at understanding the within-subject differences between ovulatory cycle phases (late follicular and mid-luteal), associated ovarian hormone levels and the recognition of emotions expressed in visual, auditory, and audiovisual modalities within a large-scale sample of healthy, naturally cycling females. We expected a higher accuracy of facial emotion recognition in the late follicular phase as compared to the mid-luteal phase, a positive relationship between levels of estradiol and facial emotion recognition accuracy, and a negative

Table 3
Results of Generalized Linear Mixed Models testing the association of ovarian hormones levels and facial emotion recognition, $N = 131$.

	Estimates	SE	z	p	OR	95% CI
Model Hormones						
Estradiol (within-subject)	0.058	0.183	0.315	0.753	1.06	0.74 – 1.52
Progesterone (within-subject)	-1.512	1.325	-1.141	0.254	0.22	0.02 – 2.96
Estradiol (between-subject)	0.064	0.057	1.129	0.259	1.07	0.95 – 1.19
Progesterone (between-subject)	0.023	0.058	0.400	0.689	1.02	0.91 – 1.15
Progesterone (within-subject) × Anger	1.777	1.353	1.313	0.189	5.91	0.42 – 83.84
Progesterone (within-subject) × Disgust	1.425	1.356	1.051	0.293	4.16	0.29 – 59.37
Progesterone (within-subject) × Fear	1.174	1.407	0.835	0.404	3.24	0.21 – 50.97
Emotion [Anger]	-3.625	0.257	14.095	**< 0.001**	0.03	0.02 – 0.04
Emotion [Disgust]	-3.861	0.257	15.016	**< 0.001**	0.02	0.01 – 0.03
Emotion [Fear]	-2.473	0.261	-9.460	**< 0.001**	0.08	0.05 – 0.14
Session	0.286	0.049	5.846	**< 0.001**	1.33	1.21 – 1.47

relationship between levels of progesterone and facial emotion recognition accuracy. We also predicted a positive association between levels of progesterone and the recognition of threat-related emotions (anger, disgust, and fear) presented in faces, known as negativity bias. In an exploratory manner, we investigated whether within-subject differences in ovarian hormones fluctuation across the late follicular and the mid-luteal phase of the cycle and emotion recognition differs among visual, auditory, and audiovisual modalities. Furthermore, we examined the interplay of within-subject ovarian hormone fluctuation and stimulus sex in emotion recognition.

4.1. No compelling evidence that women's emotion recognition ability shifts between the late follicular and mid-luteal phase

Contrary to our predictions and previous studies (for reviews see Gamsakhurdashvili et al., 2021; Osório et al., 2018) our analyses did not reveal any significant relationship of ovulatory cycle phase (late follicular vs. mid-luteal) or ovarian hormone levels across these two cycle phases with emotion recognition accuracy. In addition, our findings indicate that the modality or sex of the portrayer of the emotional expression did not moderate the assumed association between within-subject ovarian hormones fluctuation across the two ovulatory cycle phases and emotion recognition accuracy.

The lack of differences in emotion recognition between cycle phases, contrary to previous studies (e.g., Derntl et al., 2008, Rubin et al., 2012), might be explained by the methodological specificities of the studies. For instance, the current study employed a larger sample size as compared to most of the previous studies, used a within-subject design, and a large number of trials (144 trials per modality), which together resulted in higher statistical power. Moreover, the cycle phase estimation was confirmed via LH surge tests, and levels of ovarian hormones were directly measured in saliva. It should also be noted that different experimental setups could contribute to the diversity of findings in the field. As Gamsakhurdashvili et al. (2021) suggested, applying a standardized emotion recognition task could address this issue in the future. Another possible explanation for the heterogeneity of results in the existing literature could be publication bias which refers to publishing excessive significant results while non-significant results remain underreported (Francis, 2012). Our results, however, are in accordance with a recent large-scale study ($N = 192$) by Shirazi et al. (2020), examining the association of ovarian hormone levels in the fertile and mid-luteal phase, and recognizing complex emotions using the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 2001). In line with our results, the authors report no compelling evidence for a relationship between fertile and mid-luteal phase, ovarian hormones levels, and emotion recognition ability. Together, the findings of the current study and the study conducted by Shirazi et al. (2020) highlight the importance of employing a study design including high statistical power, within-subjects design, and direct hormonal measurements to study the association between ovulatory cycle phases and emotion recognition.

The apparent lack of association between ovulatory cycle phase or ovarian hormone levels measured across the ovulatory cycle phases (late follicular and mid-luteal) and emotion recognition ability in this study and the study by Shirazi et al. (2020) might also suggest that women's emotion recognition does not shift between the fertile and the mid-luteal phase. Many female ovulatory shifts with supposed adaptive benefits through increased reproductive success have been proposed in the literature, with very diverse empirical robustness (Stern and Penke, in press). The most robust one seems to be a higher sexual desire when fertile (e.g., Arslan et al., 2021; Jones et al., 2018). It might be that emotion recognition ability is among those ovulatory shifts that proof not replicable. In addition, there was no overt reproductive relevance in our stimuli, which could also be an explanation for the null findings. Another potential explanation is related to the broad debate on ecological validity and the gap between real-life experience and the abstract, artificial, and socially deprived environment of the laboratory (see Holleman et al., 2020). Although to bridge this gap we implemented visual, auditory, and audiovisual stimuli, it still might have been the case that women needed more sensory information (e.g., bodily expressions, or environmental cues) to assess the situation as relevant enough to make the extra effort which could show the difference in the performance. For instance, bodily expressions along with the moving facial expression of a talking person might create some boundary conditions to reveal the difference. Therefore, the assumed behavioral shift associated with the ovulatory cycle might be constrained by real-life experiences (e.g., the interacting effect of facial and bodily expression along with the attractiveness, intelligence, personality, and familiarity of the portrayer) and require enriched sensory stimulation.

Since the number of studies that investigated ovulatory cycle phases and emotion recognition ability is still limited, we encourage conducting replication studies with rigorous methods that will hopefully shed more light on the previously mixed findings and further our understanding regarding potential changes in cognitive and emotional capacities across the cycle that might manifest in behavioral adaptation.

4.2. Limitations

We only collected data in the estimated late follicular and mid-luteal cycle phases and assessed hormone levels therein; however, as recommended in a recent study by Stern et al. (2021) and a recent review by Gamsakhurdashvili et al. (2021), including more than two testing sessions (e.g. including the early follicular phase, or the premenstrual phase) might create a better contrast to show the possible effects of hormonal variation across the ovulatory cycle. Given that there is a second estradiol peak in the mid-luteal phase, a third session scheduled to collect data in a cycle phase characterized by low estradiol levels (e.g. early follicular or late luteal) might have provided a better insight into differential effects of estradiol levels. Furthermore, to provide a more reliable within-subjects measure for the random effect in the multilevel model, Schmalenberger et al. (2021) suggested to include at least three observations per cycle. In this study, however, due to practical concerns, we were only able to observe each participant twice per cycle. In addition, 43% of the participants ($n = 57$) did not observe positive LH tests during the ideal days and were therefore omitted from the main model that investigated the link between ovulatory cycle phase and (facial) emotion recognition (results remained virtually identical when including all participants in these analyses). Although the rate of observed cycles with negative LH tests only seems high, it is in a range of reported values from previous studies. Nevertheless, to ensure the detection of ovulation, future studies should rather employ more than 10 LH tests per participant to ensure captioning delayed ovulation and let participants provide pictures of the LH tests to the study team to avoid misinterpretation of positive results. It is also recommended to measure hormones on a daily basis rather than just tracking the fertile phase by LH test to identify false negative LH test results (Marcinkowska, 2020). A strong limitation to the current and previous studies is the common approach to measure salivary estradiol and progesterone with immunoassays. Although the immunoassays approach is an easy and accessible method to measure gonadal steroids in saliva, liquid chromatography-mass spectrometry (LC-MS/MS) provides more sensitivity, validity, and accuracy in measuring steroid hormone levels (Arslan et al., 2022). Thus, the results of our hormone models should be interpreted in light of this limitation and are in need for replication with a more valid analysis method. Moreover, to achieve reliable inter-and intra-assay CVs in hormonal samples, it is recommended to analyze hormone samples in duplicates (Stern et al., 2021). Further, we did not control for potentially confounding physiological factors associated with the menstruation such as headache, cramps, or other premenstrual symptoms which could be a threat to the internal validity of previous studies (Kiesner et al., 2020).

One task-related limitation in this study could be not implementing different intensities in emotional expressions that also explained the presence of the ceiling effect in our data which potentially explains very wide confidence intervals regarding some interaction effects (proportion correct responses = *.90*). Another limitation associated with the task might be the unbalanced number of positive and negative stimuli. One of our hypotheses particularly aimed at investigating the link between negativity bias (improved recognition of threat-related emotions) and the within-subject fluctuation of progesterone levels across the ovulatory cycle. To detect the negativity bias it is recommended to include a balanced number of positive and negative stimuli (Norris, 2019). Since we studied only a few emotions, the number of positive and negative emotions was not balanced in our design as happy expressions were the

only positive emotion. The main reason for using basic emotions in this study was due to previous studies on basic emotions that would allow us to compare our findings with the existing limited research literature. Secondly, validated auditory databases are mostly restricted to basic emotion expressions, and therefore, to create balanced modalities in the emotion recognition task we were limited to basic emotions. Nevertheless, this issue should be improved in future studies by including different emotions ranging from basic to complex expressions to provide a balanced set of stimuli in terms of valence (Gamsakhurdashvili et al., 2021). Moreover, the use of emotional prosody with still faces in the audiovisual condition might decrease the ecological validity of the study, as in the real environment we experience moving faces along with emotional prosody (Collignon et al., 2010).

One potential limitation concerning the study design is the presence of carry-over effect, as the natural shortcoming of within-subject designs (see Gangestad et al., 2016). Although, we randomized the order of stimuli, counter-balanced the testing sessions across the cycle phases (late follicular and mid-luteal), and controlled for testing session (first vs. second session), we still observed a significant carry-over effect in our findings that could be explained by using the same sets of stimuli in both testing sessions. Future studies should address this problem by implementing different sets of stimuli (Gangestad et al., 2016).

It is also worth noting that the current study counts as quasi-experimental, which means that the females' natural hormonal fluctuation was used (see Gingnell et al., 2019), and therefore drawing causal interpretation is not feasible from such a study. Studies employing hormonal administration may contribute more to our causal understanding of behavioral and cognitive changes moderated by hormones (Gingnell et al., 2019; see Gamsakhurdashvili et al., 2021).

4.3. Implications

Despite the above-mentioned limitations, the present study revealed a number of important implications. First, preregistered studies with well-established methodologies contribute to the growing body of literature on the underlying endocrinological correlates of emotion recognition. Given the mixed findings in the existing literature, preregistered studies may prevent biases in the literature by either decreasing false-positive findings or publication bias.

Second, considering the important role that ovarian hormones play across women's life span, it is worth investigating the possible association between these hormones, emotion, cognition, and behavior that would lead to improving women's health and well-being (Farage et al., 2008). The higher rate of affective disorders in women has been linked to ovarian hormones fluctuation (Van Wingen et al., 2011). Therefore, studies like the present one might contribute to an understanding of the mechanisms underlying this relationship in healthy and clinical populations. In some psychopathologies – e.g., borderline personality disorder – the ability to interpret facial expressions is impaired (e.g., Domes et al., 2009). Hence, it would be important to investigate whether the lack of association between cycle phase or ovarian hormones and emotion recognition ability would replicate in a clinical population.

Third, studies like the current one encourage the culture of publishing null findings which contributes to reducing the replication crises and publication bias. To be able to clearly define whether results are in favor of a null hypothesis or not, we recommend future studies to conduct Bayesian analyses with a priori defined regions of practical equivalence or smallest effect sizes of interest.

5. Conclusion

This study contributes to the limited existing literature on the link between the ovulatory cycle and emotion recognition ability. In conclusion, the current study did not find supporting evidence for the association between two different cycle phases (fertile and mid-luteal), fluctuations of ovarian hormones therein, and women's emotion recognition ability. Stimulus modality, stimulus sex, and emotion category did not significantly moderate the assumed association. We also found no support for shifts in facial emotion recognition ability across the ovulatory cycle in the subsample of participants with positive LH tests. The existence of such an association cannot be ruled out based on a single study; however, given the strength of the current study design, and given that our results are in line with another recent, well-designed study by Shirazi et al. (2020), we may consider that women's ability to recognize emotions might not shift between the fertile and mid-luteal phases of the ovulatory cycle.

Funding

This project was funded by Leibniz ScienceCampus "Primate Cognition", project number LSC-SF2018–10.

Author contributions

The study was conceptualized and designed by Y.R. and A.S., with input from J.O. and L.P. Y.R. carried out the experiment, analyzed the data (with support by J.S.), and drafted the manuscript. Tinterrerehe project was supervised by A.S. All authors contributed in commenting and revising the final manuscript.

Declaration of Competing Interest

All authors declare no competing interests.

Acknowledgment

We thank the German Academic Exchange Service (DAAD), the Research Training Group (RTG) 2070, and the Leibniz ScienceCampus "Primate Cognition" for the financial support of this project. We further thank our research assistants for data collection and programming. We also thank Francesco Grassi for helping with data preprocessing, Roger Mundry and Peter Zezula for providing statistical support, and all participants who took part in our study.

Author note

Parts of the data of the current study were presented at conferences and scientific meetings.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.psyneuen.2022.105977.

References

Allen, A.M., McRae-Clark, A.L., Carlson, S., Saladin, M.E., Gray, K.M., Wetherington, C. L., McKee, S.A., Allen, S.S., 2016. Determining menstrual phase in human biobehavioral research: a review with recommendations. Exp. Clin. Psychopharmacol. 24 (1), 1–11. https://doi.org/10.1037/pha0000057.

Al-Shawaf, L., Conroy-Beam, D., Asao, K., Buss, D., 2015. Human emotions: an evolutionary psychological perspective. Emot. Rev. https://doi.org/10.1177/1754073914565518.

Arslan, R.C., Schilling, K.M., Gerlach, T.M., Penke, L., 2021. Using 26 thousand diary entries to show ovulatory changes in sexual desire and behavior. J. Personal. Soc. Psychol. https://doi.org/10.1037/pspp0000208.

Arslan, R.C., Blake, K., Botzet, L.J., Bürkner, P.-C., DeBruine, L., Fiers, T., Grebe, N., Hahn, A., Jones, B.C., Marcinkowska, U.M., Mumford, S.L., Penke, L., Roney, J., Schisterman, E., & Stern, J. (2022). Not within spitting distance: salivary immunoassays of estradiol have subpar validity for cycle phase. https://doi.org/10.31234/osf.io/5r8mg.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I., 2001. The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. J. Child Psychol. Psychiatry Allied Discip. 42 (2), 241–251. https://doi.org/10.1017/S0021963001006643.

Belin, P., Fillion-Bilodeau, S., Gosselin, F., 2008. The montreal affective voices: a validated set of nonverbal affect bursts for research on auditory affective processing. Behav. Res. Methods 40 (2), 531–539. https://doi.org/10.3758/BRM.40.2.531.

Blake, K.R., Dixson, B.J.W., 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.

Breyer, B., Bluemke, M., 2016. Deutsche version der positive and negative affect schedulePANAS (GESIS panel). Zs. Soz. Item. Und Skalen. https://doi.org/10.6102/zis242.

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., Lassonde, M., 2010. Women process multisensory emotion expressions more efficiently than men. Neuropsychologia 48, 220–225. https://doi.org/10.1016/j.neuropsychologia.2009.09.007.

Conway, C., Jones, B., DeBruine, L., Welling, L., Law Smith, M., Perrett, D., Sharp, M., Al-Dujaili, E., 2007. Salience of emotional displays of danger and contagion in faces is enhanced when progesterone levels are raised. Horm. Behav. 51, 202–206. https://doi.org/10.1016/j.yhbeh.2006.10.002.

Derntl, B., Kryspin-Exner, I., Fernbach, E., Moser, E., Habel, U., 2008. Emotion recognition accuracy in healthy young females is associated with cycle phase. Horm. Behav. 53 (1), 90–95. https://doi.org/10.1016/j.yhbeh.2007.09.006.

Di Tella, M., Miti, F., Ardito, R.B., Adenzato, M., 2020. Social cognition and sex: are men and women really different. Personal. Individ. Differ. 162 (April) https://doi.org/10.1016/j.paid.2020.110045.

Dobson, A., 2002. An Introduction to Generalized Linear Models, second ed. Chapman & Hall/CRC, USA.

Domes, G., Schulze, L., Herpertz, S.C., 2009. Emotion recognition in borderline personality disorder-a review of the literature. J. Personal. Disord. 23 (*1*), 6–19.

Farage, M.A., Osborn, T.W., MacLean, A.B., 2008. Cognitive, sensory, and emotional changes associated with the menstrual cycle: a review. Arch. Gynecol. Obstet. 278 (4), 299–307. https://doi.org/10.1007/s00404-008-0708-2.

Fischer, A.H., Manstead, A.S.R., 2009. Social functions of emotions. In: Lewis, M., Haviland-Jones, J.M., Feldman Barrett, L. (Eds.), Handbook of Emotions. The Guilford Press, New York/ London.

Francis, G., 2012. Publication bias and the failure of replication in experimental psychology. Psychon. Bull. Rev. 19 (6), 975–991. https://doi.org/10.3758/s13423-012-0322-y.

Gamsakhurdashvili, D., Antov, M.I., Stockhorst, U., 2021. Facial emotion recognition and emotional memory from the ovarian-hormone perspective: a systematic review. Front. Psychol. 12 (641250), 1–14. https://doi.org/10.3389/fpsyg.2021.641250.

Gangestad, S.W., Haselton, M.G., Welling, L.L.M., 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 (2), 85–96 https://doi.org/10.1016/j.evolhumbehav.2015.09.001.

Gingnell, M., Hornung, J., Derntl, B., 2019. Emotional processing and sex hormones. Routledge Int. Handb. Soc. Neuroendocrinol. 30443, 403–419. https://doi.org/10.4324/9781315200439-24.

Guapo, V., Graeff, F., Zani, A., Labate, C., Dos Reis, R., Del-Ben, C., 2009. Effects of sex hormonal levels and phases of the menstrual cycle in the processing of emotional faces. Psychoneuroendocrinology 34, 1087–1094. https://doi.org/10.1016/j.psyneuen.2009.02.007.

Holleman, G.A., Hooge, I.T.C., Kemner, C., Hessels, R.S., 2020. The 'real-world approach' and its problems: a critique of the term ecological validity. Front. Psychol. 11, 721. https://doi.org/10.3389/fpsyg.2020.00721.

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

Jünger, J., Kordsmeyer, T., Gerlach, T.M., Penke, L., 2018. 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.

Kamboj, S.K., Krol, K.M., Curran, H.V., 2015. A specific association between facial disgust recognition and estradiol levels in naturally cycling women. PLoS ONE 10 (4), 1–12. https://doi.org/10.1371/journal.pone.0122311.

Keltner, D., Tracy, J., Sauter, D.A., Cordaro, D.C., Mcneil, G., 2016. Expression of emoitons. In: Feldman Barrett, L., Lewis, M., Haviland-Jones, J.M. (Eds.), Handbook of Emotions. The Guilford Press, New York/ London.

Kiesner, J., Eisenlohr-Moul, T., Mendle, J., 2020. Evolution, the menstrual cycle, and theoretical overreach. Perspect. Psychol. Sci. 15 (4), 1113–1130. https://doi.org/10.1177/1745691620906440.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., van Knippenberg, A., 2010. Presentation and validation of the radboud faces database. Cogn. Emot. 24 (8), 1377–1388. https://doi.org/10.1080/02699930903485076.

Lausen, A., Broering, C., Penke, L., Schacht, A., 2020. Hormonal and modality specific effects on males' emotion recognition ability. Psychoneuroendocrinology 119 (June 2019). https://doi.org/10.1016/j.psyneuen.2020.104719.

Lausen, A., Hammerschmidt, K., 2020. Emotion recognition and confidence ratings predicted by vocal stimulus type and prosodic parameters. Humanit. Soc. Sci. Commun. 7. https://doi.org/10.1057/s41599-020-0499-z.

Maner, J.K., Miller, S.L., 2014. Hormones and social monitoring: menstrual cycle shifts in progesterone underlie women's sensitivity to social information. Evol. Hum. Behav. 35 (1), 9–16. https://doi.org/10.1016/j.evolhumbehav.2013.09.001.

Marcinkowska, U.M., 2020. Importance of daily sex hormone measurements within the menstrual cycle for fertility estimates in cyclical shifts studies. Evolut. Psychol. 18 (1), 1–8. https://doi.org/10.1177/1474704919897913.

Marcinkowska, U.M., Kaminski, G., Little, A.C., Jasienska, G., 2018. Average ovarian hormone levels, rather than daily values and their fluctuations, are related to facial preferences among women. Horm. Behav. 102, 114–119.

Norris, C.J., 2019. The negativity bias, revisited: Evidence from neuroscience measures and an individual differences approach. Soc. Neurosci. 0 (0), 1. https://doi.org/10.1080/17470919.2019.1696225.

Osório, F.L., Cassis, J.M., de Sousa, J.P.M., Poli-Neto, O., Martín-Santos, R., 2018. Sex hormones and processing of facial expressions of emotion: a systematic literature review. Front. Psychol. 9 (APR) https://doi.org/10.3389/fpsyg.2018.00529.

Pahnke, R., Mau-Moeller, A., Junge, M., Wendt, J., Weymar, M., Hamm, A.O., Lischke, A., 2019. Oral contraceptives impair complex emotion recognition in healthy women. Front. Neurosci. 13 (FEB), 1–9. https://doi.org/10.3389/fnins.2018.01041.

Paulmann, S., Kotz, S.A., 2008. An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. Brain Lang. 105 (1), 59–69. https://doi.org/10.1016/j.bandl.2007.11.005.

Pearson, R., Lewis, M.B., 2005. Fear recognition across the menstrual cycle. Horm. Behav. 47 (3), 267–271. https://doi.org/10.1016/j.yhbeh.2004.11.003.

Poromaa, I.S., Gingnell, M., 2014. Menstrual cycle influence on cognitive function and emotion processing from a reproductive perspective. Front. Neurosci. 8 (Nov), 1–16. https://doi.org/10.3389/fnins.2014.00380.

Roney, J.R., 2019. On the use of log transformations when testing hormonal predictors of cycle phase shifts: commentary on. Evol. Hum. Behav. 40 (6), 526–530.

Rubin, L., 2012. Sex-specific associations between peripheral oxytocin, symptoms, and emotion perception in schizophrenia. Schizophr. Res. 136 (February), S70. https://doi.org/10.1016/s0920-9964(12)70261-x.

Sakaki, M., Mather, M., 2012. How reward and emotional stimuli induce different reactions across the menstrual cycle. Soc. Personal. Psychol. Compass 6, 1–17. https://doi.org/10.1111/j.1751-9004.2011.00415.x.

Schlegel, K., Fontaine, J., Scherer, K., 2016. The nomological network of emotion recognition ability: evidence from the Geneva emotion recognition test. Eur. J. Psychol. Assess. 35. https://doi.org/10.1027/1015-5759/a000396.

Schmalenberger, K.M., Tauseef, H.A., Barone, J.C., Owens, S.A., Lieberman, L., Jarczok, M.N., Girdler, S.S., Kiesner, J., Ditzen, B., Eisenlohr-Moul, T.A., 2021. How to study the menstrual cycle: Practical tools and recommendations. Psychoneuroendocrinology 123 (June 2020), 104895. https://doi.org/10.1016/j.psyneuen.2020.104895.

Schultheiss, O.C.O., Stanton, S.S.J., 2009. Assessment of salivary hormones. Methods Soc. Neurosci. 17–44.

Shirazi, T.N., Rosenfield, K.A., Cárdenas, R.A., Breedlove, S.M., Puts, D.A., 2020. No evidence that hormonal contraceptive use or circulating sex steroids predict complex emotion recognition. Horm. Behav. 119 (November 2019) https://doi.org/10.1016/j.yhbeh.2019.104647.

Stern, J., & Penke, L. (in press). Ovulatory cycle effects and hormonal influences on women's mating psychology. In D. M. Buss & P. Durkee (Eds.): Handbook of Human Mating. New York: Oxford University Press.

Stern, J., Kordsmeyer, T.L., Penke, L., 2021. A longitudinal evaluation of ovulatory cycle shifts in women's mate attraction and preferences. Horm. Behav. 128, 104916 https://doi.org/10.1016/j.yhbeh.2020.104916.

Thagard, P., 2002. How molecules matter to mental computation. Philos. Sci. 69 (3), 429–446. https://doi.org/10.1086/342454.

Van Wingen, G.A., Ossewaarde, L., Bäckström, T., Hermans, E.J., Fernández, G., 2011. Gonadal hormone regulation of the emotion circuitry in humans. Neuroscience 191, 38–45. https://doi.org/10.1016/j.neuroscience.2011.04.042.

Wendt, B., Scheich, H., 2002. The "Magdeburger Prosodie-Korpus". Speech Prosody 2002, 3.

Zhang, W., Zhou, R., Ye, M., 2013. Menstrual cycle modulation of the late positive potential evoked by emotional faces. Percept. Mot. Skills 116 (3), 707–723. https://doi.org/10.2466/22.27.PMS.116.3.707-723.