



Examining the role of ovarian hormones in the association between worry and working memory across the menstrual cycle

Lilianne M. Gloe^{a,*}, Deborah A. Kashy^a, Emily G. Jacobs^b, Kelly L. Klump^a, Jason S. Moser^a

^a Michigan State University, United States

^b University of California, Santa Barbara, United States

ARTICLE INFO

Keywords:

Anxiety
Worry
Working memory
Estradiol
Progesterone
Menstrual cycle

ABSTRACT

Previous research indicates that worry is associated with poorer working memory performance. Moreover, prior work demonstrates that estradiol relates to both worry and working memory performance. In the present study, we sought to further examine interrelations between worry, estradiol and working memory by testing whether estradiol moderates the association between worry and working memory in females. We hypothesized that worry would be associated with poorer working memory performance at higher levels of estradiol. We also conducted exploratory analyses to examine the role of progesterone as a moderator of the association between worry and working memory. Participants were 97 naturally-cycling females who attended four lab sessions across their menstrual cycles. Consistent with predictions, higher average levels of worry were associated with lower working memory accuracy on particularly difficult trials when average levels of estradiol were also high. The same association between higher worry and lower working memory accuracy emerged when average levels of progesterone were high. Findings highlight the importance of considering ovarian hormones in future studies and current theories of anxiety and cognition.

1. Introduction

Although it is well established that females¹ suffer from higher rates and longer courses of anxiety than males (Kessler et al., 2012; McLean et al., 2011), studies have largely failed to examine the specific impact of anxiety on cognitive function in females. Ovarian hormones affect both cognition and anxiety (e.g., Beltz and Moser, 2019; Li and Graham, 2017) and thus are critical to examine in associations between anxiety and cognitive function. The current study, therefore, evaluated the role of ovarian hormones in the association between worry – a component of anxiety comprised of negative, future-oriented verbal thought activity (Borkovec et al., 1998) – and working memory (WM; i.e., one's ability to store and manipulate information) in naturally-cycling females.

Meta-analytic evidence demonstrates that worry negatively impacts WM (Moran, 2016). WM resources that maintain goal-directed behaviors in the presence of competing distractors are thought to be usurped by worry, thereby resulting in poorer attentional control and cognitive deficits more broadly (Baddeley, 1996, 1986; Eysenck et al., 2007; Kane

et al., 2007). Although females report higher levels of worry than males (Gould and Edelstein, 2010; Hunt et al., 2003; Nitschke et al., 2001) and are twice as likely as males to receive a diagnosis of Generalized Anxiety Disorder (GAD) – a disorder defined by chronic worry (American Psychiatric Association, 2013) – research has failed to delineate how worry and WM relate in females, specifically.

Ovarian hormones are critical to consider since they have been shown to affect both cognition and anxiety (Beltz and Moser, 2019; Gasbarri et al., 2008; Jacobs and D'Esposito, 2011; Li and Graham, 2017; Maeng and Milad, 2015; Man et al., 1999). For example, recent findings suggest an association between menstrual phase and worry-related phenotypes (e.g., repetitive negative thinking; Li et al., 2020). In a series of preclinical studies, Shansky and colleagues demonstrated that estradiol moderated the association between WM performance and stress, an anxiety-related construct shown to relate to worry in humans (Brosschot et al., 2006; Shansky and Lipps, 2013; Shansky et al., 2004, 2006, 2009). They showed that WM performance in female rats suffers under acute stress (i.e., after brief restraint or

* Correspondence to: Department of Psychology, Michigan State University, 316 Physics Road, East Lansing, MI 48824, United States.

E-mail address: rothsch7@msu.edu (L.M. Gloe).

¹ Throughout the manuscript, we use the term “female” to refer to individuals whose sex assigned at birth is female. We use this term given the focus on biological processes (i.e., the menstrual cycle) that are specific to those who are biologically female but could apply across all gender identities. The term “woman” is reserved for those individuals who identify as women.

pharmacological stress induction), particularly when estradiol levels are elevated (for review: [Shansky and Lipps, 2013](#)). Hypothesized mechanisms by which high estradiol may negatively impact WM performance under stress include the intensification of the effects of glucocorticoid release, increases in the availability of dopamine, and disruption of the balance between dopaminergic and noradrenergic receptor activation ([Shansky and Lipps, 2013](#)).

Work in humans also demonstrates that stress has a greater deleterious impact on WM in females than males (for review: [Shields et al., 2016](#)). However, no studies have examined the role of ovarian hormones, and the one study that utilized menstrual cycle phase as a proxy for hormone action reported null effects of the menstrual cycle on the association between induced stress (via the Trier Social Stress Test) and WM (on the N-back task) in females ([Schoofs et al., 2013](#)). The [Schoofs et al. \(2013\)](#) study is significantly limited, however, by reliance on (1) self-reported menstrual cycle information, (2) comparison of the first half versus second half of the cycle as a rough estimate of hormone action, and (3) menstrual cycle phase as a between-person variable (i.e., 50% of participants were in the first half and 50% were in the second half of their cycles) thereby confounding between-person and within-person ovarian hormone differences. Given limitations of the extant human literature, the current study drew upon the work of [Shansky and colleagues](#) for hypotheses regarding the effects of estradiol. We addressed the gaps in the human literature by directly examining the role of ovarian hormones in the association between worry and WM in females.

We expected that worry would be related to reduced WM accuracy when estradiol was high, based on both human and animal work ([Moran, 2016](#); [Shansky and Lipps, 2013](#)). WM was indexed by performance on the verbal N-back task. Based on previous research and theory suggesting that anxiety negatively impacts accuracy when a task is particularly difficult (e.g., [Eysenck et al., 2007](#); [Eysenck and Derakshan, 2011](#)), we manipulated WM difficulty by (1) varying the number of trials back the target letter occurred in the sequence (i.e., N-back load) and (2) including lure trials. Lures are letters that match a recently shown, but non-target letter. They are thought to produce more interference than generic non-targets because they match recently presented letters, and thus place greater demands on attentional control to reject ([Gray et al., 2003](#)). Anxiety may impact lure trial performance, in particular, because anxiety increases a response bias towards familiar items ([Moran, 2016](#)). We therefore predicted that worry would be associated with reduced accuracy when the task was most difficult (i.e., at higher WM loads and/or on lure trials), but only when estradiol was high. We also conducted an exploratory analysis to examine progesterone as a potential moderator given previous work has not examined its role in the association between anxiety and WM.

We investigated both between-person effects that represent a person's average level of worry/estradiol and within-person effects that reflect deviations from these averages for a particular person. Levels of estradiol and worry differ both on average between females and fluctuate considerably within females during the menstrual cycle or day to day ([Hampson, 2020](#); [Joos et al., 2012](#); [Klump et al., 2016](#)). Given that prior work has often been cross-sectional, such that between- and within- person effects are confounded, it is difficult to discern if prior work speaks more to between- or within- person effects. Therefore, our hypotheses were the same for between- and within- person analyses, although it is possible between- and within- person effects are functionally distinct.

2. Methods

2.1. Participants

Participants were 109 female volunteers (ages 18–25 years old; $M = 20.86$, $SD = 1.73$). Participants were recruited from the community via mail, local media, flyers, and online advertisements. Of the

97 participants with useable data, 64% of the participants were Caucasian/White, 22% were Black/African American, 7% were more than one race and 6% were Asian. The sample was predominately heterosexual (81%) with notable diversity (Bisexual: 10%; Gay/Lesbian: 5%; Asexual: 1%; and Pansexual 1%). Ninety-seven percent of participants identified as women (1 participant identified as nonbinary/fluid queer/gender queer and 2 participants declined to answer). In terms of annual income, 44% of participants reported earning less than \$25,000, 10% between \$25,000 and \$50,000, 14% between \$50,000 and \$100,000, and 29% over \$100,000. Sixty-eight percent of participants reported being financially supported by someone else in the past year. The majority of our sample reported being full-time students (76%).

Participants were selected based on several inclusion/exclusion criteria. Specifically, participants had to be naturally menstruating (i.e. every 22–35 days) and not taking hormonal contraceptives, psychotropic medications, or steroid medications during the past eight weeks before study participation. Participants were selected to be between the ages of 18 and 25 (see Online Supplement for rationale). They must have had no history of genetic or medical conditions or recent medication use known to have an impact upon the endocrine system (e.g., thyroid disorders, metabolic disorders, and steroid medication usage). Additional exclusion criteria included: epilepsy; hearing, visual, or other physical or mental impairments that could interfere with data quality; head trauma that resulted in a loss of consciousness for longer than five minutes; and being a non-native English speaker. Eligibility was confirmed during each study visit. Because we were interested in studying worry as a dimensional/ transdiagnostic phenomenon, participants were included regardless of anxiety levels or diagnoses.

2.2. Procedure

2.2.1. Overview

The study involved an initial eligibility screening by phone, one intake visit where informed consent was obtained, daily questionnaires measuring worry, daily saliva sample collections to assay for ovarian hormones, four visits during which participants completed the N-back task, and a visit for structured diagnostic clinical interview administration. An overview of data collection is provided in [Fig. 1](#).

During the eligibility phone screen, menstrual cycle history was collected. A hybrid projection method was adapted from [Lester et al. \(2003\)](#) and then utilized to prospectively define four phases (i.e., early follicular, late follicular, ovulation and mid-luteal) for scheduling purposes to ensure the majority of the cycle, and adequate variability in ovarian hormones, would be captured. The next date of menstruation was estimated based on the average length of the past 3–5 cycles, which participants were required to have tracked on a calendar or through a phone application. The timing of the first of the four lab visits was selected based on their current menstrual cycle phase to ensure that a similar number of participants started in each of the four phases. Estimated phases were used for scheduling purposes only and were not utilized in analyses because our hypotheses centered on the direct effects of ovarian hormones. Following study completion, we determined the menstrual cycle phases in which each of the four lab visits occurred using hormone levels/patterning recorded during the study (see [Klump et al., 2015](#)).

2.2.2. Daily questionnaire procedure

Participants were asked to fill out a series of questionnaires between 5:00PM and 10:00PM each day via Qualtrics, an online assessment portal. Questionnaires were completed each day for 35 consecutive days of the study. The daily questionnaires consisted of the Penn State Worry Questionnaire (PSWQ; [Meyer et al., 1990](#)) and other psychopathology assessments not included in the current analysis. The PSWQ completed on in-person lab visit days was utilized in the current set of analyses. The daily PSWQ assessed worry symptoms on the day the questionnaire was completed (Cronbach's $\alpha = 0.93$, across four visits). Evidence for

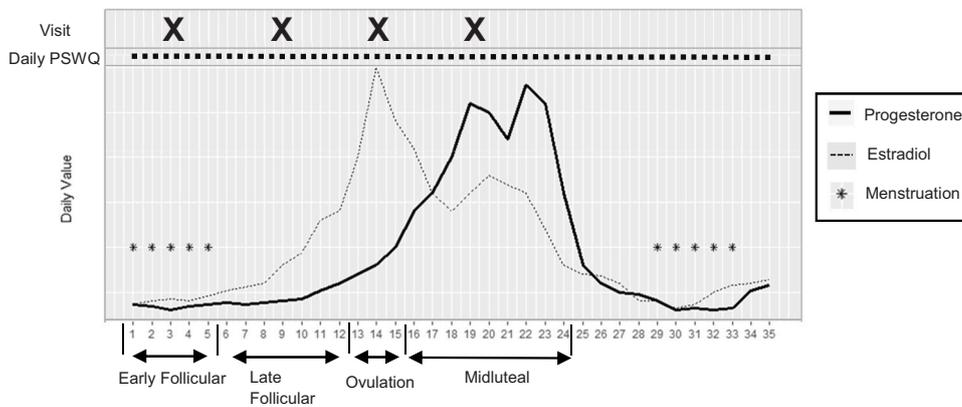


Fig. 1. Illustration of study data collection across the menstrual cycle. Day 1 is the first day of menstruation. The Penn State Worry Questionnaire (PSWQ) was completed and saliva samples were collected every day for 35 days. The N-back task was completed at each of the four visits illustrated by X's above. For each participant, visits were scheduled to fall in different phases using menstrual cycle projection based on the length of the past three menstrual cycles. For the purposes of the current analysis, only estradiol, progesterone and PSWQ scores on the day of visits will be utilized.

convergent validity of the daily version of the PSWQ can be found in the Online Supplement. Additional validity evidence for use of the PSWQ as a daily measure of worry is reported by Joos et al. (2012).

2.2.3. Daily saliva sample collection

Participants were instructed to provide their samples via the passive drool method in the proper tube (pre-labeled with study day and date) within thirty minutes of waking. Saliva samples were provided for 35 consecutive days to capture estradiol and progesterone levels across the menstrual cycle. Samples were logged at each in-person visit and stored in a lab – 80°F freezer until they were shipped to Salimetrics, LLC (State College, PA) for analysis. Saliva was analyzed using enzyme immunoassay kits for assaying estradiol, which have excellent intra- and inter-assay coefficients of variation (estradiol = 7.1% and 7.5%; progesterone 6.2% and 7.6%) and assay sensitivity (estradiol = 0.10 pg/mL; progesterone = 5 pg/mL). Although blood is also commonly used to assay for estradiol and progesterone, past studies have shown high correlations between ovarian hormones in saliva and blood (Dielen et al., 2017; Ellison, 1993; Shirtcliff et al., 2001). Saliva assays also reduce participant burden due to ease of collection compared to blood assays. Once collected, z-scored estradiol and progesterone values were visually examined and cycles were labeled anovulatory if they lacked the expected rise in progesterone and/or demonstrated a lack of patterning in ovarian hormones throughout the cycle as dictated by Klump et al. (2015). Any visit that occurred during an anovulatory cycle was removed from analyses.

2.2.4. Lab visit procedures

Lab visits lasted 2–3 h each. Participants completed a battery of cognitive assessments on the computer, which consisted of a Flanker task (not included here), followed by the verbal N-back task and automated O-Span and R-Span tasks (not included here because Span tasks do not systematically vary difficulty). Electroencephalogram (EEG) was also recorded during the N-back and Flanker tasks but will be reported on in future papers. As previously described, the N-back was the focus of the current investigation.

2.2.4.1. N-back task. Participants completed the verbal N-back task as described by Jacobs and D'Esposito (2011). Briefly, individual letters were displayed centrally on a computer screen for 1000 ms with an ITI of 1100 ms. Participants were asked to respond to each stimulus within 2000 ms by indicating whether it was a target (left button press) or non-target (right button press). Working memory load was manipulated across blocks of trials. Targets are defined as a letter that matches the letter shown N trials previously (where N = 0-back, 2-back or 3-back). The 0-back block consisted of a simple signal detection task in which participants were asked to respond to the target letter “X” whenever it appeared. The 2-back and 3-back trial blocks placed demands on working memory. For example, in the sequence A-T-R-T-C, the second T

in the sequence is a target in a 2-back block and all other letters are non-targets. Participants completed 16 blocks for a total of 320 trials (0-back = 160 trials, 2-back = 80 trials, 3-back = 80 trials). A special class of non-target trials, called lures, were also included in the paradigm. Lures are letters that match a recently shown, but non-target, letter. As an example, consider the sequence K-F-E-D-K in a 3-back trial block. The second “K” presented is a lure trial, because the first K appeared four letters back, not three letters back. Note that lures, by necessity, only occurred in 2- and 3-back blocks. In total, 20% of trials were targets, 15% were lures and 65% were generic non-targets. Participants were counterbalanced to one of four versions of the N-back using a Latin Square design. Participants completed one block of practice trials at each load to ensure instructions were understood. Only 2- and 3- back targets and lures were considered for the current analyses, as these are the levels at which we expected worry to relate to performance (i.e. levels at which the task was most difficult). Zero-back and non-targets were not analyzed, because (1) there was little variability in accuracy on 0-back ($SD = 8.85$) and non-target trials ($SD = 2.70$) to model compared to other trial types and (2) previous work with this version of the N-back task only examined 2- and 3-back targets and lures so that taking a similar approach increases robustness and replicability (Jacobs and D'Esposito, 2011).

Trials were included in calculations of reaction time and accuracy if the reaction time was greater than 200 ms. Accuracy and reaction time for each load-by-trial type level (i.e., 2-back targets, 2-back lures, 3-back targets, 3-back lures) were averaged across trials for each participant.

2.2.5. Analysis procedure

Multilevel modeling (MLM) with SAS Version 9.4 was used to examine the impact of worry, estradiol, and task difficulty (operationalized as a 2 by 2 design: 2- versus 3-back and lure versus target) on N-back task accuracy. Because participants attended four separate sessions across their menstrual cycle, the underlying data structure included two levels of repeated measurements: N-back trial types within visit and visits within person. If accuracy fell below 30% on any load-by-trial type level, accuracy from that trial type was treated as missing. Note that unlike repeated measures ANOVA that requires complete data for each participant, MLM allows for missing data and bases its estimates on all observations obtained.

2.2.5.1. Fixed effects model structure used to examine between- and within-person effects. To address between- and within- person effects of worry and estradiol, two fixed effects models were estimated for accuracy (models for reaction time are provided in the Online Supplement). All main effects and interactions between the four predictor variables (trial type, load, PSWQ score, and estradiol) were included in both models, but the models differed in how PSWQ scores and estradiol were centered. In the between-person model, the between-person effects of PSWQ and estradiol were tested using the individual's average PSWQ or

estradiol, averaged across the 4 visits and then grand-mean centered. In the within-person model, the within-person effects of PSWQ and estradiol were tested using the individual's four PSWQ scores or estradiol values and subtracting the individuals' mean on the variable (i.e. person-mean centered scores). To control for practice effects, a main effect of lab visit number was included in both models.

Two three-way interactions (i.e. Load x Estradiol x PSWQ score and Trial Type x Estradiol x PSWQ score) and the 4-way interaction (i.e. Load, Trial type, Estradiol x PSWQ score) were central to understanding the impact of estradiol and PSWQ scores on *N*-back accuracy at different levels of task difficulty. Examining these interactions allowed us to test the hypothesis that higher levels of worry (measured by PSWQ) would be related to decreased accuracy at higher levels of load (i.e. 3-back loads) and/or on more difficult trial types (i.e. lure trials; see Gray et al., 2003) when estradiol was high. A power analysis was performed using the G-power program with a repeated measures design as a proxy for MLM (number of repeated measurements = 32) and determined that 34 participants would be needed to achieve 80% power for the four-way interaction predicting accuracy ($f = 0.12$, mean correlation between repeated measures = 0.39). Data collected from the first 54 participants in our sample were used to estimate effect sizes and correlations prior to conducting analyses for this power analysis. Therefore, we were adequately powered to examine the four-way interaction and two three-way interactions in the models described.

Two additional models predicting accuracy were executed to examine the role of between- and within- person progesterone (models for RT are located in the Online Supplement). Progesterone models had the same fixed and random effects structures as those including estradiol.

Any significant interactions between continuous and categorical variables were broken down by simple slopes analyses in which separate intercepts and slopes for the continuous variable were computed at each level of the categorical variable(s) involved. Any interaction involving both estradiol or progesterone and PSWQ score (i.e. two, three, and four-way interactions involving estradiol or progesterone and PSWQ score) was followed-up using the procedures described by Aiken et al. (1991). Simple slopes analyses were conducted to examine the effects of one continuous variable at high (+1 SD) and low (-1 SD) values of the other continuous variable by recentering the variable. Lower-order interactions that were qualified by higher-order interactions were not interpreted in the text.

Estimated marginal means and estimates for ovarian hormones, PSWQ score and their interaction are presented in Table S4 and S5 in the Online Supplement. Significant two-way interactions are also interpreted in the Online Supplement.

2.2.5.2. Random effects structure to account for experimental design. A random intercept was included to account for nonindependence of multiple observations at the participant level. This intercept variance measures individual differences in accuracy, or the extent to which some participants were generally accurate, but others were generally inaccurate, across all trials and visits. Standard data analytic models (i.e., ANOVA, Regression, and many basic MLMs) assume homoskedasticity such that the residual variances and covariances are assumed to be equal across the levels of the predictors. However, preliminary analyses indicated that the accuracy data were highly heteroskedastic. Violations of homoskedasticity can bias Type I and Type II error rates, and so we specified models that allowed for separate residual variances and covariances for each of the four levels of the load-by-trial-type interaction (i.e., 2-back target, 2-back lure, 3-back target, 3-back lure) pooled across lab visits and participants. Thus, the random effects structure accounts for the experimental design and resulting heteroskedasticity at each load-by-trial type level. All assumptions of multilevel modeling were evaluated and found to be met for each model and no influential points were identified using a Cook's Distance cut-off of 0.5.

3. Results

3.1. Data retention

Of the 109 subjects enrolled (421 sessions), 97 subjects (89% of total sample) had useable data for the PSWQ, *N*-back behavioral data, and estradiol values on at least one visit day that did not occur during an anovulatory cycle. Ten participants had one cycle that was anovulatory and any visits that took place in these cycles were removed from analyses. Descriptive statistics for PSWQ scores, estradiol, and progesterone are provided in Table 1.

3.2. Model 1: testing the effects of between-subject variation in worry and estradiol

3.2.1. Basic task effects

Results of the model testing the effects of between-person analyses of worry and estradiol on accuracy (Model 1) are presented in Table 2. A significant main effect of load emerged confirming that participants were less accurate on 3-back ($M = 66.55\%$, $SE = 1.061$) compared to 2-back trials ($M = 78.038\%$, $SE = 1.088$). There was also a load by trial type interaction for accuracy. Using a Bonferroni adjustment for pairwise comparisons, accuracy scores significantly differed between loads (e.g., 2-back target accuracy was significantly greater than 3-back lures), but not by different trial types of the same load (e.g., 2-back targets and 2-back lures did not differ significantly). These effects of load and trial type are as expected based on previous work utilizing this version of the *N*-back task (Jacobs and D'Esposito, 2011). Additionally, the effect of visit number on *N*-back accuracy was significant, suggesting the presence of small practice effects: follow-up tests using the Bonferroni adjustment for pairwise comparisons, identified significant differences in *N*-back accuracy between all visits, except between visit 1 ($M = 69.015\%$, $SE = 1.217$) and visit 2 ($M = 70.877\%$, $SE = 1.209$) and between visit 3 ($M = 74.217\%$, $SE = 1.196$) and visit 4 ($M = 75.080\%$, $SE = 1.231$). These basic task effects are nearly identical for all other accuracy models for estradiol and progesterone (i.e., Models 2 - 4).

3.2.2. Results of predicted 3- and 4-way interactions

A significant interaction emerged between trial type, PSWQ score, and estradiol (see Table 2; $\eta^2 = 0.030$). In breaking down this significant interaction, we first examined the two-way interactions between estradiol and trial type for high and low levels of between-person worry (i.e., $PSWQ \pm 1$ SD). For high levels of worry there was evidence of an interaction between trial type and estradiol, $F(1, 302) = 16.52$, $p < 0.01$, but not at low levels of worry, $F(1, 294) = 0.30$, $p = 0.58$. For females with higher average PSWQ scores, higher average estradiol

Table 1
PSWQ score, estradiol, and progesterone means and standard deviations.

Measure	Mean (SD)	Minimum	Maximum
PSWQ Score	39.64 (14.94)	16	80
Within-Person PSWQ Score- Centered	0 (8.36)	-28.00	22.75
Between-Person PSWQ Score- Centered	0 (12.38)	-22.89	30.36
Estradiol (pg/mL)	1.54 (0.77)	0.17	5.48
Within-Person Estradiol- Centered (pg/mL)	0 (0.49)	-1.65	2.40
Between-Person Estradiol- Centered (pg/mL)	0 (0.60)	-1.19	1.68
Progesterone (pg/mL)	169.90 (136.38)	13.13	818.65
Within-Person Progesterone- Centered (pg/mL)	0 (99.51)	-244.59	453.98
Between-Person Progesterone- Centered (pg/mL)	0 (93.25)	-153.54	315.88

Note: PSWQ = Penn State Worry Questionnaire

Table 2

Models 1 & 2- test of type 3 fixed effect significance for multilevel model of N-back accuracy (%) related to trial type, load, PSWQ score, and estradiol controlling for visit number.

Effect	Model 1: Between-Person			Model 2: Within-Person		
	Den. Df	F-Value	p-value	Den. Df	F-Value	p-value
Trial Type	295	0.71	0.402	293	0.22	0.637
Load	311	233.12	< 0.001	308	229.74	< 0.001
PSWQ	90.4	1.49	0.226	198	0.41	0.524
Estradiol	91.1	0.10	0.758	200	0.00	0.966
Trial Type x Load	309	6.35	0.012	305	7.34	0.007
PSWQ x Trial Type	299	2.88	0.091	296	0.01	0.939
PSWQ x Load	320	0.11	0.738	316	0.06	0.806
Estradiol x Trial Type	294	6.34	0.012	298	0.21	0.649
Estradiol x Load	310	0.83	0.364	319	0.46	0.499
PSWQ x Estradiol	94.2	1.81	0.182	246	0.03	0.859
Estradiol x Load x Trial Type	308	0.17	0.678	316	0.13	0.718
PSWQ x Load x Trial Type	320	0.00	0.999	313	0.01	0.942
PSWQ x Estradiol x Trial Type	301	9.20	0.003	291	0.63	0.427
PSWQ x Estradiol x Load	323	0.33	0.565	303	4.09	0.044
PSWQ x Estradiol x Load x Trial Type	323	2.21	0.138	298	1.51	0.220
Visit Number	216	12.74	< 0.001	213	12.49	< 0.001

Notes: PSWQ = Penn State Worry Questionnaire; Between-Person Model: ICC = 0.269, $R^2 = 0.181$; Within-Person Model: ICC = 0.270, $R^2 = 0.150$; Numerator df for all tests besides Visit Number was 1. For Visit Number, it was 3.

predicted significantly lower accuracy on lure trials ($B = -7.342$, $SE = 2.82$, $t(150) = -2.60$, $p = 0.01$; see Fig. 2). This finding is consistent with our hypothesis that higher trait worry would relate to worse performance on more difficult trial types when estradiol levels were also high. No other effects of interest emerged.

3.3. Model 2: testing the effects of within-person variation in worry and estradiol

3.3.1. Results of predicted 3- and 4-way interactions

Results of the model testing the effects of within-person variation in worry and estradiol on accuracy (Model 2) are presented in Table 2. The only significant interaction of interest in Model 2 was between load, PSWQ score and estradiol (Table 2; $\eta^2 = 0.013$).

In breaking down this three-way interaction, we examined the two-way interactions between estradiol and load for when worry was high relative to the person's average, and when worry was low relative to the person's average (i.e., within-person PSWQ ± 1 SD). For days on which a person's PSWQ was high, relative to their average, there was marginal evidence of an interaction between load and estradiol, $F(1, 309) = 3.7$, $p = 0.06$, but not for days on which a person's PSWQ was low, $F(1, 313) = 0.84$, $p = 0.36$. When PSWQ was relatively high compared to a participant's average, neither simple slope was significant. However, effects were in opposite directions to each other, whereby increased estradiol was modestly and non-significantly related to increased accuracy on 3-back trials ($B = 2.21$, $SE = 1.60$, $t(252) = 1.38$, $p = 0.17$; Fig. S1 in the Online Supplement), but it was non-significantly related to decreased accuracy on 2-back trials ($B = -1.93$, $SE = 1.73$, $t(214) = -1.12$, $p = 0.27$). No other significant effects of interest emerged.

3.4. Model 3: testing the effects of between-person variation in worry and progesterone

3.4.1. Results of predicted 3- and 4-way interactions

Results of the model testing the effects of between-person variation in worry and progesterone on accuracy (Model 3) are presented in Table 3. Although progesterone analyses were exploratory, we focused on the same interactions that were of interest for the estradiol analyses. A significant three-way interaction between trial type, PSWQ score, and progesterone emerged (see Table 3; $\eta^2 = 0.019$). In breaking down this significant interaction, we first examined the two-way interactions between progesterone and trial type for high and low levels of between-person worry (i.e., PSWQ ± 1 SD). For high levels of worry there was evidence of an interaction between trial type and progesterone, $F(1, 316) = 8.48$, $p < 0.01$, but not at low levels of worry, $F(1, 301) = 0.62$, $p = 0.43$. For females with higher average PSWQ scores, higher average progesterone predicted significantly lower accuracy on lure trials ($B = -0.044$, $SE = 0.018$, $t(155) = -2.45$, $p = 0.02$; see Fig. 3).² No other effects of interest emerged.

3.5. Model 4: testing the effects of within-person variation in worry and progesterone

3.5.1. Results of predicted 3- and 4-way interactions

Results of the model testing the effects of within-person variation in worry and progesterone on accuracy (Model 4) are presented in Table 3. Three-way and 4-way interactions of interest were non-significant.

4. Discussion

The current study examined the relationship between worry and WM in females by considering the role of ovarian hormones across the menstrual cycle. Our findings provide mixed evidence for our hypotheses, as within- and between-person models demonstrated disparate results. Consistent with our prediction, we found that females higher in average worry and estradiol showed reduced accuracy on the difficult lure trials of the N-back task. Notably, females higher in average worry and progesterone likewise demonstrated decreased lure accuracy. However, within-person models were non-significant. Therefore, our results suggest that between-person differences in worry and ovarian hormones exert meaningful and disruptive effects on working memory, whereas within-person fluctuations may play a less critical role.

Significant between-person effects of worry and estradiol on WM are in line with findings by Shansky and colleagues (e.g., Shansky and Lipps, 2013). Thus, worry may be functionally similar to stress in that its negative relationship with WM is exacerbated by high levels of estradiol. Notably, our results revealed that average progesterone moderated the association between worry and working memory in the same manner as estradiol. Average, or basal, ovarian hormone levels have generally been under-studied (Eisenlohr-Moul and Owens, 2020). However, evidence suggests that average hormone levels are derived from interactions between genes and the environment and may not be stable throughout

² We conducted follow-up analyses to examine whether both estradiol and progesterone worked in tandem with worry to predict N-back accuracy. Given the constraints of our model (i.e., we are unable to interpret a five-way interaction), we used a median-split approach that yielded four groups of females ranked based on their average hormone levels (High Estradiol, High Progesterone; High Estradiol, Low Progesterone; Low Estradiol, High Progesterone; Low Estradiol, Low Progesterone). Results demonstrated a marginal interaction between ovarian hormone group, trial type and PSWQ ($F(3, 288) = 2.24$, $p = 0.084$; see Table S1). Further breaking down this marginal interaction showed that higher average PSWQ was associated with worse lure performance only in participants high in both average estradiol and progesterone ($B = -0.468$, $SE = 0.162$, $p = 0.005$; all other $p > 0.05$).

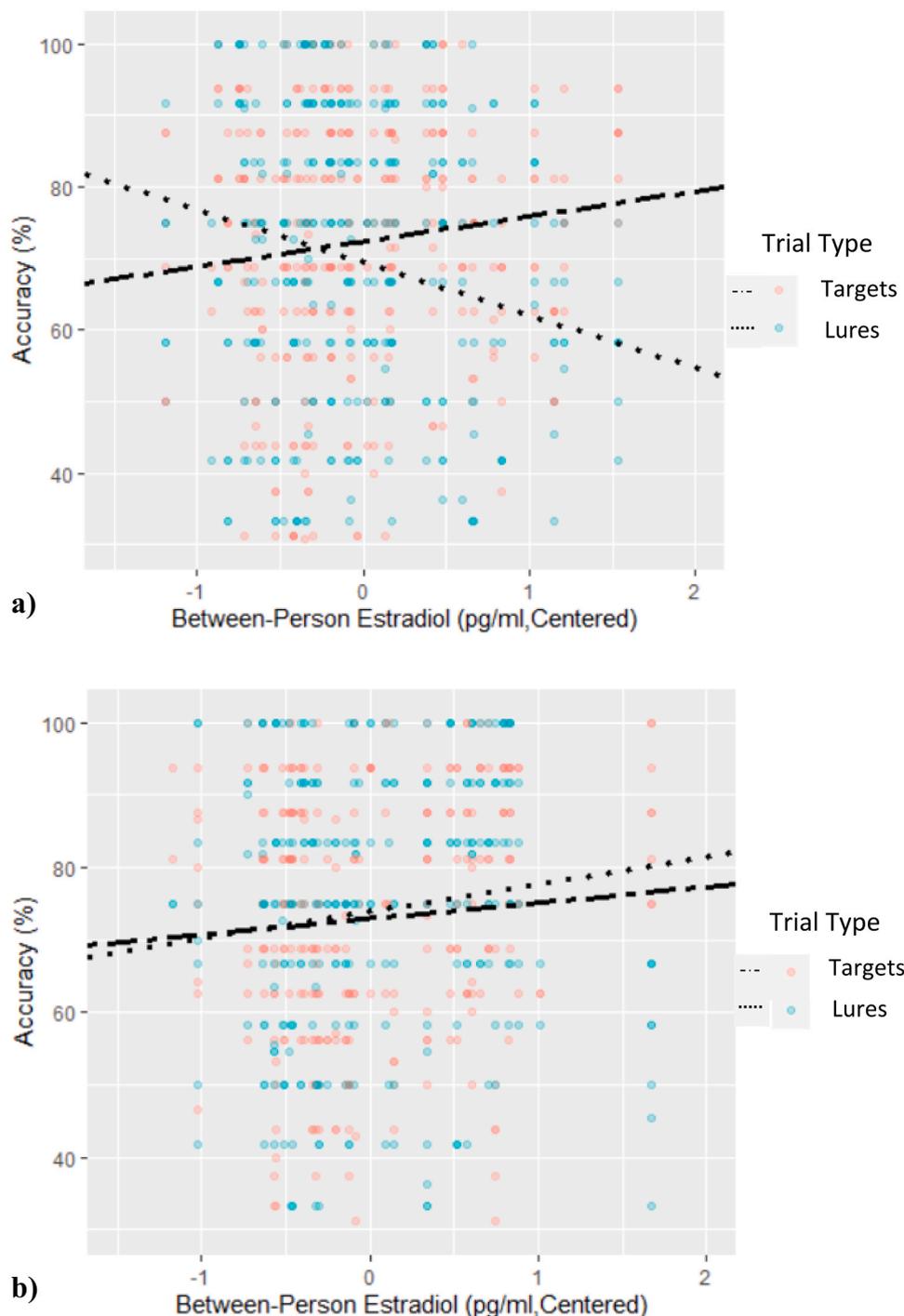


Fig. 2. Breaking down the 3-way between-person PSWQ score \times between-person estradiol \times trial type interaction identified in the multilevel model for accuracy. (a) For high average PSWQ Scores (+1 SD), there is a significant interaction between load and estradiol, such that increased average estradiol predicts decreased accuracy on lure trials. (b) For low average PSWQ Scores (-1 SD), there was no significant relationship between average estradiol and accuracy on any trial type. Each circle represents average accuracy achieved for the particular trial type on each visit per participant. Note that estradiol values shown represent average estradiol values that have been grand-mean centered. Uncentered estradiol values range from 0.17 to 5.48 pg/mL ($M = 1.54$ pg/mL, $SD = 0.77$).

one's lifetime (for review: Jasienska et al., 2017). Precedent for examining average levels of hormones can be found in the testosterone literature, in which testosterone has been shown to be heritable in females and a reliable and valid index of individual difference in personality (Harris et al., 1998; Sellers et al., 2007).

Although between-person ovarian hormones are only beginning to be explored, associations between the dopaminergic system and ovarian hormones provide an avenue for potential mechanisms underlying our findings. Strong evidence indicates dopamine is important for prefrontal cortex function (Cools and D'Esposito, 2011; Jacobs and D'Esposito, 2011). In humans, fluctuations of estradiol during the menstrual cycle are thought to exert their influence on WM through enhancing the

synthesis of dopamine (Becker, 1990; Jacobs and D'Esposito, 2011). While theory asserted by Shansky and Lipps (2013) and research by Jacobs and D'Esposito (2011) speak to within-person estradiol fluctuations (as opposed to between-person/average concentrations), emerging evidence implicates basal levels of ovarian hormones in dopaminergic pathways. Specifically, results from a recent resting-state functional magnetic resonance imaging study demonstrated that high absolute estradiol and progesterone levels were associated with enhanced connectivity between the dorsolateral PFC and substantia nigra, a region implicated in DA transfer and cognitive performance (Wang et al., 2020). Therefore, we speculate that progesterone may modulate or facilitate the action of estradiol, such that higher levels of both

Table 3

Models 3 & 4- test of type 3 fixed effect significance for multilevel model of N-back accuracy (%) related to trial type, load, PSWQ score, and progesterone controlling for visit number.

Effect	Model 3: Between-Person			Model 4: Within-Person		
	Den. Df	F-Value	p-value	Den. Df	F-Value	p-value
Trial Type	304	0.34	0.562	301	0.17	0.678
Load	320	233.63	< 0.001	319	238.00	< 0.001
PSWQ	90.7	1.36	0.246	207	0.26	0.612
Progesterone	92.3	0.28	0.598	210	0.01	0.907
Trial Type x Load	317	7.12	0.008	314	6.94	0.009
PSWQ x Trial Type	308	1.53	0.217	304	0.12	0.729
PSWQ x Load	329	0.18	0.670	325	0.04	0.840
Progesterone x Trial Type	303	2.21	0.138	305	3.65	0.057
Progesterone x Load	319	1.79	0.182	329	4.51	0.034
PSWQ x Progesterone	93.2	4.50	0.037	256	0.03	0.868
Progesterone x Load x Trial Type	315	0.11	0.745	325	0.83	0.363
PSWQ x Load x Trial Type	328	0.16	0.694	320	0.06	0.806
PSWQ x Progesterone x Trial Type	312	5.98	0.015	308	0.60	0.440
PSWQ x Progesterone x Load	337	0.15	0.700	334	0.90	0.343
PSWQ x Progesterone x Load x Trial Type	337	0.90	0.342	331	2.19	0.140
Visit Number	223	13.87	< 0.001	247	15.36	< 0.001

Notes: PSWQ = Penn State Worry Questionnaire; Between-Person Model: ICC = 0.297, $R^2 = 0.174$; Within-Person Model: ICC = 0.278, $R^2 = 0.172$; Numerator df for all tests besides Visit Number was 1. For Visit Number, it was 3.

hormones may increase DA activity within the PFC to an excessive level. Notably, work by Shanksy and Lipps (2013) and Jacobs and D'Esposito (2011) suggests that, indeed, excessive dopaminergic functioning would be expected to reduce WM performance in females with high worry.

More work is needed to explore the role of basal ovarian hormone levels in anxiety and WM, particularly with regard to progesterone. While progesterone receptors are present in brain areas implicated in cognitive processes in animals, their location in humans remains unclear (Beltz and Moser, 2019). Also, there are only three small-scale studies in the working memory literature that have considered progesterone (Baller et al., 2013; Berent-Spillson et al., 2015; Hidalgo-Lopez and Pletzer, 2017). One within-person study found participants demonstrated lower 2-back accuracy when receiving leuprolide, a GnRH agonist that suppresses the secretion of endogenous estradiol and progesterone, compared to when they were given exogenous estradiol or progesterone (Baller et al., 2013). The other two studies identified null effects of progesterone on WM (Berent-Spillson et al., 2015; Hidalgo-Lopez and Pletzer, 2017). Additionally, there have been mixed findings and differing theorized impacts of progesterone in anxiety and mood (for reviews: Li and Graham, 2017; Romans et al., 2012; Andreano et al., 2018). However, recent findings by Li et al. (2020) show that repetitive negative thinking is increased among females with GAD during the luteal phase, when estradiol and progesterone are both high. Further, two studies recently showed that higher between-person progesterone related to greater anxiety symptoms (Reynolds et al., 2018; Wang et al., 2020), with one study indicating that both high between-person progesterone and estradiol levels related to greater anxiety (Wang et al., 2020).

Within-person analyses involving estradiol resulted in smaller and less clear effects. Given the lack of robustness of these effects and non-

significant effects identified in follow-up completer analyses,³ we recommend results are replicated. Within-person analyses of progesterone also revealed no significant effects of interest.

The lack of within-person findings may seem surprising, especially given the theoretical importance of ovarian hormone fluctuations for anxiety (Li and Graham, 2017; Maeng and Milad, 2015). However, little extant work examining anxiety has directly assessed ovarian hormones over multiple observations across the menstrual cycle. Much of the work in this area has compared individuals in different phases of the menstrual cycle or ovariectomized rats to intact rats, conflating between- and within-person hormone effects (e.g., Gasbarri et al., 2008; Shanksy et al., 2004, 2006, 2009). Indeed, it is well known that data collected at a single time point can only speak to between-person effects (for review: Curran and Bauer, 2011). Therefore, the empirical evidence for within-person ovarian hormone effects in anxiety and working memory is unclear. It could also be that one must account for average ovarian hormone effects in order to determine the impact of within-person fluctuations in worry and WM. Further work is needed to tease apart within- and between- person influences of worry and hormones on working memory.

Importantly, our results suggest that between-person differences in estradiol and progesterone might be more important than within-person fluctuations. In other words, our results speak more to *who* is at risk for experiencing the negative cognitive effects of worry, rather than *when* risk will be greatest during the menstrual cycle for a particular female. Clinically, assaying for estradiol and progesterone may assist in identifying females experiencing anxiety-related cognitive impacts. Assays would likely need to be completed multiple times during the menstrual cycle to estimate average ovarian hormones levels for each individual. Normative ovarian hormone ranges for saliva have not yet been established, and ranges for blood would need to be altered to specify cut-offs for high average ovarian hormone levels (i.e., the ovarian hormone level at which anxiety would be expected to relate to WM). Thus, future research efforts should focus on identifying normative ranges for average hormones and examining the stability of average hormones over time. Understanding potential confounding factors (e.g., metabolic factors; Iversen et al., 2012) will also be important. Ultimately, the current findings suggest the possibility that interventions to reduce cognitive impacts of anxiety in females might be specifically tailored to females characterized by high average ovarian hormone levels.

Despite the novelty and strengths of the current study, it is not without its limitations. First, several characteristics of our sample reduce generalizability. Specifically, due to the need to exclude participants with conditions and medications impacting the endocrine system, the current sample was comprised of young, healthy naturally-cycling females. Further, the majority of our participants were full-time students, making it unclear if such results would apply to non-student or part-time student females.

Additionally, our conclusions are limited by several methodological challenges. Our methods for scheduling visits, while among the best available, could have been improved through use of ovulation kits to increase accuracy of scheduling visits during luteal phase. Also, although we provide evidence for the validity of the daily PSWQ in the Online Supplement, additional studies are needed to further establish measures of state worry. We also acknowledge that we assessed hormones in the morning and because there are diurnal changes in ovarian hormones, levels may have differed during the time of the visit compared with immediately after waking. Unfortunately, the time course of hormone effects on behavior (e.g., minutes, hours) is unknown

³ Follow-up analyses were conducted that included only participants with complete data from all four visits (Estradiol: $N = 49$, Progesterone: $N = 52$). While the two significant between-person three-way interactions reported in the full sample remained, the within-person estradiol x PSWQ x load interaction was no longer significant ($F(1, 187) = 1.16, p = 0.283$).

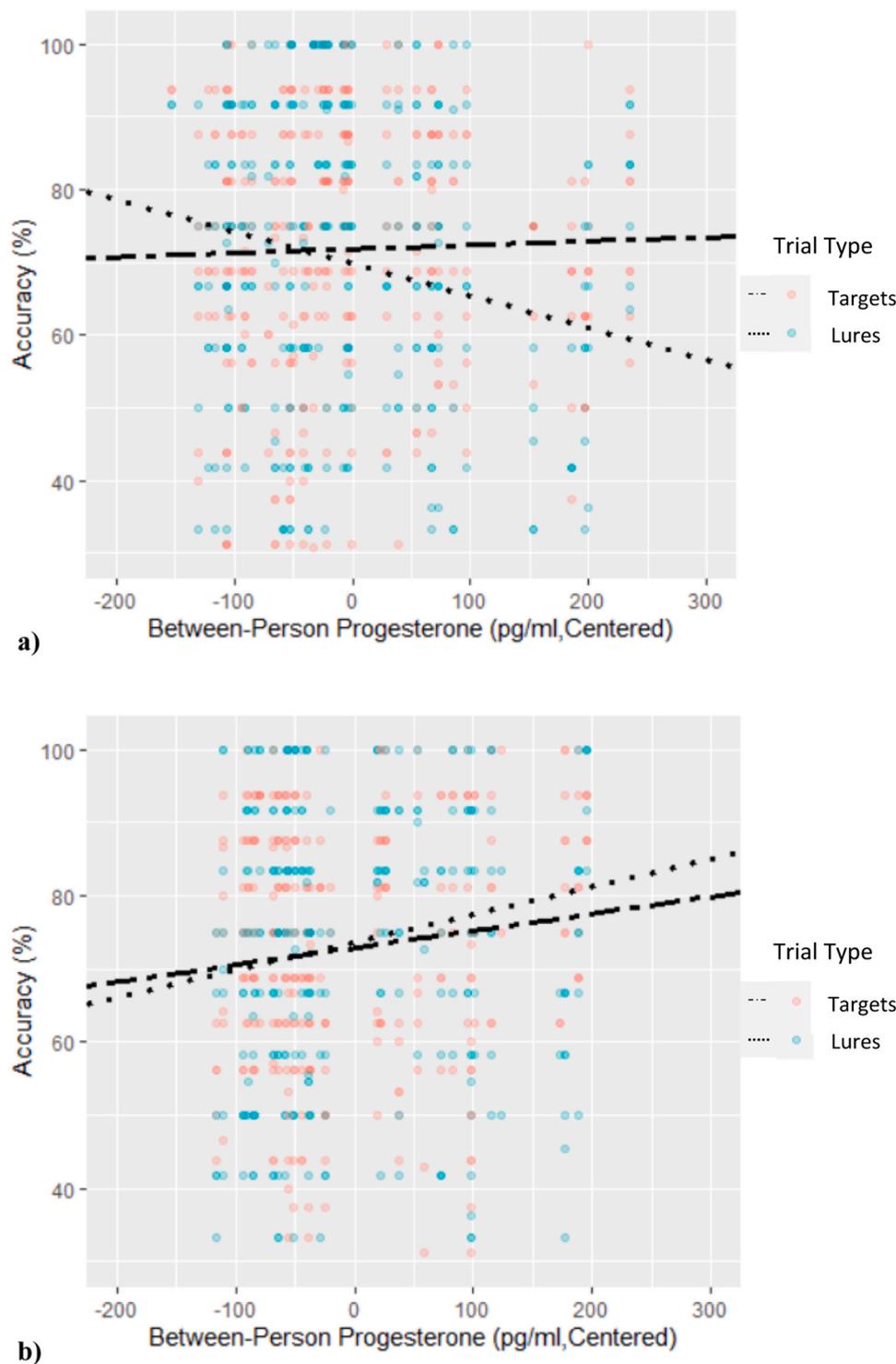


Fig. 3. Breaking down the 3-way between-person PSWQ score x between-person progesterone x trial type interaction identified in the multilevel model for accuracy. (a) For high average PSWQ Scores (+ 1SD), there is a significant interaction between trial type and progesterone, such that increased average progesterone predicts decreased accuracy on lure trials. (b) For low average PSWQ Scores (-1SD), there was no significant relationship between average progesterone and accuracy by trial type. Notably, for low average PSWQ scores, increased progesterone was related to marginally better performance regardless of load and trial type. Each circle represents average accuracy achieved for the particular trial type on each visit per participant. Note that progesterone values shown represent average progesterone values that have been grand-mean centered. Uncentered progesterone values range from 13.13 to 818.65 pg/mL ($M = 169.90$ pg/mL, $SD = 136.38$).

in humans (Nelson, 2011), making it difficult to identify the ideal time window in which to measure hormones and neurocognitive tasks. Future research is needed to examine whether collecting hormones at different times during the day affects hormone-behavior associations. Additionally, because of challenges with longitudinal designs and the number of measures collected, many of our participants were missing data for at least one of their four visits. Notably, multilevel modeling allows us to retain individuals who had at least one visit worth of data and over 85% of our participants had data from three or four visits available. Our results do not change when only participants with complete data were analyzed (see footnote ³).

Statistically, we were limited in the number of effects we could examine within a model, such that we could not examine an interaction involving more than four factors. Future work should consider combining within- and between-person effects in single models to examine interactive effects.

Despite these limitations, the present study revealed an important role for ovarian hormones in the association between worry and WM. Specifically, the current study points to higher average levels of ovarian hormones and worry as important predictors of lower WM in females. Unexpectedly, hormone fluctuations across the menstrual cycle were not involved in associations between worry and WM. Thus, our findings

speak more to *who* is likely to experience cognitive impacts as a function of worry and hormones rather than *when* such impacts might emerge during the menstrual cycle. Through future research, improved cognitive models of anxiety in females could be developed and, ultimately, leveraged to build targeted interventions in the context of average hormone levels, menstrual cycle fluctuations and hormonal transitions (e.g., puberty, menopause).

CRedit authorship contribution statement

J.S.M.: developed study concept and design with methodological contributions from E.G.J. and K.L.K. Data collection was performed by J.S.M. L.M.G. performed data analysis under supervision of D.A.K. L.M.G. drafted the paper. J.S.M., D.A.K., K.L.K. and E.G.J. provided critical revisions. All authors approved of the final version of the paper for submission.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The Brain Cycle Study is funded by the National Institute of Mental Health (NIMH, United States; grant number: 1R01MH108511-01). Lillian Gloe is supported by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP, United States, grant number: DGE-1848739).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.psyneuen.2021.105285](https://doi.org/10.1016/j.psyneuen.2021.105285).

References

- Aiken, L.S., West, S.G., Reno, R.R., 1991. *Multiple Regression: Testing and Interpreting Interactions*. Sage.
- American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.
- Andreano, J.M., Touroutoglou, A., Dickerson, B., Barrett, L.F., 2018. Hormonal cycles, brain network connectivity, and windows of vulnerability to affective disorder. *Trends Neurosci.* 41, 660–676.
- Baddeley, A., 1996. The fractionation of working memory. *Proc. Natl. Acad. Sci.* 93, 13468–13472.
- Baddeley, A.D., 1986. *Working Memory*. Elsevier.
- Baller, E.B., Wei, S.-M., Kohn, P.D., Rubiñow, D.R., Alarcón, G., Schmidt, P.J., Berman, K.F., 2013. Abnormalities of dorsolateral prefrontal function in women with premenstrual dysphoric disorder: a multimodal neuroimaging study. *AJP* 170, 305–314. <https://doi.org/10.1176/appi.ajp.2012.12030385>.
- Becker, J.B., 1990. Direct effect of 17 β -estradiol on striatum: sex differences in dopamine release. *Synapse* 5, 157–164.
- Beltz, A.M., Moser, J.S., 2019. Ovarian hormones: a long overlooked but critical contributor to cognitive brain structures and function. *Ann. N. Y. Acad. Sci.* 1464. <https://doi.org/10.1111/nyas.14255>.
- Berent-Spillon, A., Briceno, E., Pinsky, A., Simmen, A., Persad, C.C., Zubieta, J.-K., Smith, Y.R., 2015. Distinct cognitive effects of estrogen and progesterone in menopausal women. *Psychoneuroendocrinology* 59, 25–36. <https://doi.org/10.1016/j.psyneuen.2015.04.020>.
- Borkovec, T.D., Ray, W.J., Stober, J., 1998. Worry: a cognitive phenomenon intimately linked to affective, physiological, and interpersonal behavioral processes. *Cogn. Ther. Res.* 22, 561–576.
- Brosschot, J.F., Gerin, W., Thayer, J.F., 2006. The perseverative cognition hypothesis: a review of worry, prolonged stress-related physiological activation, and health. *J. Psychosom. Res.* 60, 113–124. <https://doi.org/10.1016/j.jpsychores.2005.06.074>.
- Cools, R., D'Esposito, M., 2011. Inverted-u-shaped dopamine actions on human working memory and cognitive control. *Biol. Psychiatry Prefrontal Cortical Circuits Regul. Atten. Behav. Emot.* 69, e113–e125. <https://doi.org/10.1016/j.biopsycho.2011.03.028>.
- Curran, P.J., Bauer, D.J., 2011. The disaggregation of within-person and between-person effects in longitudinal models of change. *Annu. Rev. Psychol.* 62, 583–619. <https://doi.org/10.1146/annurev.psych.093008.100356>.
- Dielen, C., Fiers, T., Somers, S., Deschepper, E., Gerris, J., 2017. Correlation between saliva and serum concentrations of estradiol in women undergoing ovarian hyperstimulation with gonadotropins for IVF/ICSI. *Facts Views Vis. ObGyn* 9, 85–91.
- Eisenlohr-Moul, T.A., Owens, S.A., 2020. Hormones and personality. *Encycl. Personal. Individ. Differ.* 2010–2032.
- Ellison, P.T., 1993. Measurements of salivary progesterone. *Ann. N. Y. Acad. Sci.* 694, 161–175.
- Eysenck, M.W., Derakshan, N., 2011. New perspectives in attentional control theory. *Personal. Individ. Differ.* 50, 955–960.
- Eysenck, M.W., Derakshan, N., Santos, R., Calvo, M.G., 2007. Anxiety and cognitive performance: attentional control theory. *Emotion* 7, 336–353.
- Gasbarri, A., Pompili, A., d'Onofrio, A., Cifariello, A., Tavares, M.C., Tomaz, C., 2008. Working memory for emotional facial expressions: role of the estrogen in young women. *Psychoneuroendocrinology* 33, 964–972.
- Gould, C.E., Edelstein, B.A., 2010. Worry, emotion control, and anxiety control in older and young adults. *J. Anxiety Disord.* 24, 759–766.
- Gray, J.R., Chabris, C.F., Braver, T.S., 2003. Neural mechanisms of general fluid intelligence. *Nat. Neurosci.* 6, 316–322.
- Hampson, E., 2020. A brief guide to the menstrual cycle and oral contraceptive use for researchers in behavioral endocrinology. *Horm. Behav.* 119, 104655.
- Harris, J.A., Vernon, P.A., Boomsma, D.I., 1998. The heritability of testosterone: a study of dutch adolescent twins and their parents. *Behav. Genet* 28, 165–171. <https://doi.org/10.1023/A:1021466929053>.
- Hidalgo-Lopez, E., Pletzer, B., 2017. Interactive effects of dopamine baseline levels and cycle phase on executive functions: the role of progesterone. *Front. Neurosci.* 11. <https://doi.org/10.3389/fnins.2017.00403>.
- Hunt, S., Wisocki, P., Yanko, J., 2003. Worry and use of coping strategies among older and younger adults. *J. Anxiety Disord.* 17, 547–560.
- Iversen, A., Thune, I., McTiernan, A., Makar, K.W., Wilsgaard, T., Ellison, P.T., Jasienska, G., Flote, V., Poole, E.M., Furberg, A.-S., 2012. Genetic polymorphism CYP17 rs2486758 and metabolic risk factors predict daily salivary 17 β -estradiol concentration in healthy premenopausal norwegian women. The EBBA-I Study. *The Journal of Clinical Endocrinology & Metabolism* 97, E852–E857. <https://doi.org/10.1210/jc.2011-2577>.
- Jacobs, E., D'Esposito, M., 2011. Estrogen shapes dopamine-dependent cognitive processes: implications for women's health. *J. Neurosci.* 31, 5286–5293. <https://doi.org/10.1523/JNEUROSCI.6394-10.2011>.
- Jasienska, G., Bribiescas, R.G., Furberg, A.-S., Helle, S., Núñez-de la Mora, A., 2017. Human reproduction and health: an evolutionary perspective. *Lancet* 390, 510–520. [https://doi.org/10.1016/S0140-6736\(17\)30573-1](https://doi.org/10.1016/S0140-6736(17)30573-1).
- Joos, E., Vansteenwegen, D., Brunfaut, E., Bastiaens, T., Demyttenaere, K., Pieters, G., Hermans, D., 2012. The Penn State worry questionnaire—past day: development and validation of a measure assessing daily levels of worry. *J. Psychopathol. Behav. Assess.* 34, 35–47. <https://doi.org/10.1007/s10862-011-9265-2>.
- Kane, M.J., Conway, A.R., Hambrick, D.Z., Engle, R.W., 2007. Variation in working memory capacity as variation in executive attention and control. *Var. Work. Mem.* 1, 21–48.
- Kessler, R.C., Petukhova, M., Sampson, N.A., Zaslavsky, A.M., Wittchen, H.-U., 2012. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int. J. Methods Psychiatr. Res.* 21, 169–184.
- Klump, K.L., Hildebrandt, B.A., O'Connor, S.M., Keel, P.K., Neale, M., Sisk, C.L., Boker, S., Burt, S.A., 2015. Changes in genetic risk for emotional eating across the menstrual cycle: a longitudinal study. *Psychol. Med.* 45, 3227–3237. <https://doi.org/10.1017/S0033291715001221>.
- Klump, K.L., O'Connor, S.M., Hildebrandt, B.A., Keel, P.K., Neale, M., Sisk, C.L., Boker, S., Burt, S.A., 2016. Differential effects of estrogen and progesterone on genetic and environmental risk for emotional eating in women. *Clin. Psychol. Sci.* 4, 895–908. <https://doi.org/10.1177/2167702616641637>.
- Lester, N.A., Keel, P.K., Lipson, S.F., 2003. Symptom fluctuation in bulimia nervosa: relation to menstrual-cycle phase and cortisol levels. *Psychol. Med.* 33, 51–60.
- Li, S.H., Denson, T.F., Graham, B.M., 2020. Women with generalized anxiety disorder show increased repetitive negative thinking during the luteal phase of the menstrual cycle. *Clin. Psychol. Sci.* 8, 1037–1045. <https://doi.org/10.1177/2167702620929635>.
- Li, S.H., Graham, B.M., 2017. Why are women so vulnerable to anxiety, trauma-related and stress-related disorders? The potential role of sex hormones. *Lancet Psychiatry* 4, 73–82.
- Maeng, L.Y., Milad, M.R., 2015. Sex differences in anxiety disorders: interactions between fear, stress, and gonadal hormones. *Horm. Behav.* 76, 106–117.
- Man, M.S., MacMillan, I., Scott, J., Young, A.H., 1999. Mood, neuropsychological function and cognitions in premenstrual dysphoric disorder. *Psychol. Med.* 29, 727–733.
- McLean, C.P., Asnaani, A., Litz, B.T., Hofmann, S.G., 2011. Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. *J. Psychiatr. Res.* 45, 1027–1035.
- Meyer, T.J., Miller, M.L., Metzger, R.L., Borkovec, T.D., 1990. Development and validation of the penn state worry questionnaire. *Behav. Res. Ther.* 28, 487–495.
- Moran, T.P., 2016. Anxiety and working memory capacity: a meta-analysis and narrative review. *Psychol. Bull.* 142, 831–864. <https://doi.org.proxy2.cl.msu.edu/10.1037/bul0000051>.
- Nelson, R., 2011. *An Introduction to Behavioral Endocrinology*, 4th ed. Sinauer Associates, Inc, Sunderland, MA.
- Nitschke, J.B., Heller, W., Imig, J.C., McDonald, R.P., Miller, G.A., 2001. Distinguishing dimensions of anxiety and depression. *Cogn. Ther. Res.* 25, 1–22. <https://doi.org/10.1023/A:1026485530405>.

- Reynolds, T.A., Makhanova, A., Marcinkowska, U.M., Jasienska, G., McNulty, J.K., Eckel, L.A., Nikonova, L., Maner, J.K., 2018. Progesterone and women's anxiety across the menstrual cycle. *Horm. Behav.* 102, 34–40. <https://doi.org/10.1016/j.yhbeh.2018.04.008>.
- Romans, S., Clarkson, R., Einstein, G., Petrovic, M., Stewart, D., 2012. Mood and the menstrual cycle: a review of prospective data studies. *Gend. Med.* 9, 361–384.
- Schoofs, D., Pabst, S., Brand, M., Wolf, O.T., 2013. Working memory is differentially affected by stress in men and women. *Behav. Brain Res.* 241, 144–153.
- Sellers, J.G., Mehl, M.R., Josephs, R.A., 2007. Hormones and personality: testosterone as a marker of individual differences. *J. Res. Personal.* 41, 126–138. <https://doi.org/10.1016/j.jrp.2006.02.004>.
- Shansky, R., Rubinow, K., Brennan, A., Arnsten, A.F., 2006. The effects of sex and hormonal status on restraint-stress-induced working memory impairment. *Behav. Brain Funct.* 2, 8.
- Shansky, R.M., Bender, G., Arnsten, A.F.T., 2009. Estrogen prevents norepinephrine alpha-2a receptor reversal of stress-induced working memory impairment. *Stress* 12, 457–463.
- Shansky, R.M., Glavis-Bloom, C., Lerman, D., McRae, P., Benson, C., Miller, K., Cosand, L., Horvath, T.L., Arnsten, A.F.T., 2004. Estrogen mediates sex differences in stress-induced prefrontal cortex dysfunction. *Mol. Psychiatry* 9, 531–538.
- Shansky, R.M., Lipps, J., 2013. Stress-induced cognitive dysfunction: hormone-neurotransmitter interactions in the prefrontal cortex. *Front. Hum. Neurosci.* 7, 123.
- Shields, G.S., Sazma, M.A., Yonelinas, A.P., 2016. The effects of acute stress on core executive functions: a meta-analysis and comparison with cortisol. *Neurosci. Biobehav. Rev.* 68, 651–668. <https://doi.org/10.1016/j.neubiorev.2016.06.038>.
- Shirtcliff, E.A., Reavis, R., Overman, W.H., Granger, D.A., 2001. Measurement of gonadal hormones in dried blood spots versus serum: verification of menstrual cycle phase. *Horm. Behav.* 39, 258–266.
- Wang, J.-X., Zhuang, J.-Y., Fu, L., Lei, Q., Fan, M., Zhang, W., 2020. How ovarian hormones influence the behavioral activation and inhibition system through the dopamine pathway. *PLOS ONE* 15, 0237032. <https://doi.org/10.1371/journal.pone.0237032>.