Mood and Hormonal Effects on Facial Emotion Detection:

The Role of Depressive Symptoms, Oral Contraceptive Use, and Premenstrual Symptoms

Bianca Boboc, HBSc

Department of Psychology

Lakehead University

Submitted in partial fulfillment of the degree of

Master of Arts in Clinical Psychology

Supervisor: Dr. K. Oinonen

Second Reader: Dr. J. Tan

External Examiner: Dr. A. Maranzan

September 8, 2023

Abstract

Past research has suggested altered facial emotion detection (FED) with depressive symptoms, oral contraceptive (OC) use, and premenstrual symptoms. Altered FED may contribute to and maintain negative mood associated with OC side effects, premenstrual dysphoric disorder (PMDD), and other depressive disorders. The purpose of this study was to investigate the effects of depressive symptoms, OC use, and premenstrual symptoms on FED using a novel task that examines detection thresholds. A sample of 163 participants (37 OC using women, 72 free-cycling (FC) women, 35 men, 19 other hormonal contraceptive using women) completed the Facial Emotion Detection Task. The task used neutral to emotional facial expression morphs (15 images per morph), and participants indicated what emotion they detected for each image within the progressive intensity morph. For all six basic emotions (happy, sad, angry, fearful, surprise, disgust) two types of scores were calculated: accuracy of responses and the intensity within the morph at which the correct emotion was detected (image number). Higher depression symptoms were associated with earlier (i.e., lower intensity) detection across overall emotions. Conversely, OC use was associated with later detection across overall emotions, and worse overall FEDT performance, with specific effects for happy and disgust. Also, women taking androgenic OC formulations detected happy emotions later than FC women. Premenstrual symptoms were associated with earlier detection of sad emotions (independent of cycle phase). An interaction with cycle phase also emerged: women with PMDD were more accurate at detecting emotions during the premenstrual phase, with a large effect size for sad emotions. This is the first study to investigate emotional intensity at detection based on OC use and premenstrual symptoms. Potential etiology, mechanisms, and implications are discussed.

Acknowledgements

I would like to first express my gratitude to my supervisor, Dr. Kirsten Oinonen, for her support in exploring my curiosities, and her continuous guidance and mentorship throughout this project. Thank you also to my committee members, Dr. Josephine Tan and Dr. Amanda Maranzan, whose thoughtful suggestions have helped improved this work. I am grateful to Chyenne Panetta and Nandini Parekh for their help with the data.

Finally, words cannot express my gratitude toward my support network, without which I would not have made it this far. To my parents, for undertaking difficult changes, endlessly believing in me, and proudly following along on my academic endeavours. To my cohort sisters, for the solidarity and uplifting each other during the lows, while curating many of the highs of graduate school. And to my friends at a distance, for staying immersed in my journey and reminding me about the bigger picture.

Table of Contents

List of Appendices

Mood and Hormonal Effects on Facial Emotion Detection:

The Role of Depressive Symptoms, Oral Contraceptive Use, and Premenstrual Symptoms

Oral contraceptives (OCs) are among the most commonly used contraceptive methods worldwide (Chae et al., 2021; Duesenberg et al., 2016). Research has indicated that OCs can influence mood, and OC use has been associated with greater emotional intensity and consequently, increased rates of depression and reduced positive affect variability (Jarva & Oinonen, 2007; Skovlund et al., 2016). Facial emotion detection (FED) is a component of emotional processing, that has also been found to differ in women based on OC use, and with cyclical hormonal changes across the menstrual cycle (Gamsakhurdashvili et al., 2021; Gingnell et al., 2013; Guapo et al., 2009; Osório et al., 2018). FED has also been widely studied in relation to depression. Research suggests a consensus that depression is associated with a negativity bias in FED, whereby individuals with depression detect negative emotions faster and more accurately compared to other emotions (Bourke et al., 2010). One study has pointed to a similar processing bias associated with Premenstrual Dysphoric Disorder (PMDD), a mood disorder linked to hormonal fluctuations (Rubinow et al., 2007). FED is an important aspect of reciprocal social interactions (e.g., it affects how we perceive others and how we behave in response), which highlights the importance of studying this process.

Presently, there is strong evidence that facial emotion processing changes with hormonal fluctuations, both induced by OC use and occurring with a natural menstrual cycle, and mood disorders, such as depression and PMDD. However due to limitations in the methodology of previous studies and a lack of integration of all of these factors within individual studies, consensus regarding the specific pattern of facial emotion processing based on hormonal changes, and the relationship between hormones and mood, has not been reached. For example,

hormonal changes are often found to affect the perception of negative emotions more than they do positive emotions (Gültekin et al., 2017; Hamstra et al., 2014; Maner & Miller, 2014). Research on differences in facial emotion processing as a function of OC use, hormonal changes, and depression, and the intercorrelation between these factors is important. This information can help guide women to make informed choices regarding contraceptive use and their health.

Depression

Depression is among the most prevalent mental health disorders worldwide. In 2015 4.4% of the world population was diagnosed with Major Depressive Disorder (MDD) (World Health Organization, 2017). In addition, rates of depression are approximately double in women compared to men (Jenkins et al., 2018). This ratio can be seen cross-culturally, indicating that, aside from just socio-cultural influences, there may be intrinsic biological factors contributing to this difference (Bromet et al., 2011). There is also ample evidence that indicates hormones, specifically disturbances in the hypothalamic-pituitary-adrenal (HPA), hypothalamic-pituitary thyroid (HPT), and hypothalamic-pituitary-gonadal (HPG) axes, contribute to the pathophysiology of MDD (Dwyer et al., 2020). On top of its high prevalence, MDD is one of the most debilitating health conditions worldwide (Lépine & Briley, 2011). With rates on the rise, these statistics point to the importance of understanding all etiological factors contributing to depression and of the gender difference within depression.

Within current literature, a biopsychosocial model of MDD is generally accepted, suggesting that an interaction between biological, psychological, and social predisposing factors contribute to the risk of developing MDD (Kupferberg et al., 2016). Not only is it a predisposing factor, but deficits in social functioning are considered a key feature of the presentation of MDD in many people, and contribute to the maintenance and relapse of the disorder (Kupferberg et al.,

2016). Within this context, social functioning refers to both the productive aspects, such as interpersonal behaviours, and receptive aspects, such as emotion recognition in others, associated with interpersonal interaction (Kupferberg et al., 2016). For example, depression may occur when repeated misperceiving (or failing to detect) emotions leads to social misunderstandings, conflict, and relationship dissolution. Emotional recognition has been studied experimentally by assessing FED. Such tasks can give insight into distinct socially relevant cognitive processes and social symptoms within depression (Bourke et al., 2010). It has also been suggested FED tasks may also be used as a predictive and monitoring measure of depression (Bourke et al., 2010).

Depression and Facial Emotion Detection

Facial emotion detection within depression is an important area of research, as this process often acts as a compass, guiding the perceiver's own emotional experience and subsequent behaviour. Past research has persistently identified that individuals with MDD display a distinct pattern of facial emotion perception. Specifically, meta-analysis has found that MDD is associated with a negative response bias, meaning that individuals with MDD are more likely to appraise positive and ambiguous emotions as more negative, compared to euthymic individuals (Bourke et al., 2010). Additionally, those with MDD show a hypervigilance and attentional bias toward negative emotions, meaning that when presented with both positive and negative facial emotions, their attention will be drawn to, and they will therefore process, the negative emotion more readily (Bourke et al., 2010). Another meta-analysis has found that those with MDD show decreased FED accuracy for all facial emotions except for sadness (Dalili et al., 2015). The most recent meta-analysis found a similar relationship, indicating that as MDD severity increases, FED accuracy decreases proportionally (Krause et al., 2021). This effect was observed for all facial emotions, however the decrease in accuracy corresponding to MDD, was

smaller when detecting the negative emotions sadness and anger, compared to other facial emotions (Krause et al., 2021). They also found that MDD was specifically associated with poorer detection of happy and neutral emotions compared to negative emotions (Krause et al., 2021). That is, those with MDD, compared to those without, were more likely to make accuracy errors in the detection of happy emotions, and were more likely to categorize neutral expressions as negative, which is commonly referred to as a negativity bias (Krause et al., 2021). Other research has also found that these observed patterns are robust and persist even within studies that used a female-only sample, and that depressed women were able to accurately detect sad facial emotions at lower intensities compared to control participants (Bento de Souza et al., 2014).

Taken together, the research suggests a mood-congruent bias within depression, indicating a bidirectional relationship between negative affect and emotional processing, contributing to the development and maintenance of depression. That is, due to their negative mood, individuals with depression may show enhanced processing and attention towards negative facial emotions compared to positive ones and a negativity bias in perception, which may contribute to negative social interactions and subsequently circle back to exacerbate negative mood.

Depression and Sex Differences in Facial Emotion Detection

Consistent with the sex differences in the rates of MDD, there is some evidence to suggest that there are also sex differences in emotional processing within depression. One study found that women with MDD made more errors in identifying facial emotions compared to euthymic women and both MDD and euthymic men (Wright et al., 2009). The same study also identified that women with MDD were more likely to make errors in the identification of sad

emotions compared to euthymic women and both MDD and euthymic men, and when making these errors were more likely to incorrectly identify emotions as angry compared to all other groups (Wright et al., 2009). These findings suggest that depressed women are more affected by facial emotion processing biases characteristic of depression, compared to men.

These differences may have a neurobiological source. One neuroimaging study found distinct activation differences within the superior frontal gyrus (a brain area involved in emotion regulation) between women and men with depression (Jenkins et al., 2018). Additionally, it is well recognized that the neurotransmitter dopamine is a major contributor to the symptoms associated with depression, such as anhedonia (Williams et al., 2021). Aside from just differences between individuals with and without depression, research has also shown that dopamine activity differs between men and women with depression (Williams et al., 2021). This may be in part contributed to by sex hormone effects on dopamine synthesis, reuptake, and transport (Barth et al., 2015). Specifically, estrogen and progesterone have global effects on dopamine, but also have a role in modulating dopamine across the menstrual cycle (Barth et al., 2015; Hidalgo-Lopez & Pletzer, 2017). Most notably, women with depression show a greater dopamine active transporter binding affinity, meaning that dopamine more readily binds to reuptake transporters in synapses and gets recycled faster (Williams et al., 2021). This difference is pronounced in the caudate nucleus, which is a brain region affecting reward, motivation, and emotion systems (Williams et al., 2021). Dopamine has also recently been linked to facial emotion processing, further suggesting the link between sex differences in MDD and facial emotion detection (Rypma et al., 2015). Despite the interest in facial emotion processing and sex differences in depression, other researchers (Jenkins et al., 2018), have noted that studies

investigating facial emotion processing and MDD rarely consider the sex of participants within their analyses.

Facial Emotion Detection Methods

Research on facial emotion processing normally uses some variation of a facial emotion detection task. An important methodological variable within facial emotion research is the nature of the task used. Some tasks rely on a simple detection process by asking participants to identify the emotion presented in a facial emotion stimulus (i.e., out of six basic emotion options, identifying what emotion is presented) (e.g., Derntl, Kryspin-Exner, et al., 2008; Maner & Miller, 2014). Many of such tasks use the six basic emotions (happy, sad, angry, fearful, surprise, disgust) as possible options, however some studies include only a subset of these emotions (e.g., Maner & Miller, 2014), and others include a broader range of more complex emotions as well (e.g., Shirazi et al., 2020). Other tasks use a matching paradigm, and ask participants to select out of a larger pool of facial emotion pictures the one that is emotionally congruent with the facial emotion stimulus (i.e., identifying if an image of an angry expression or sad expression matches the stimulus image) (e.g., Gingnell et al., 2012). The nature of facial emotion stimuli can also vary greatly. Some tasks use static images, in which an image of an emotion at 100% intensity is presented (e.g., Maner & Miller, 2014; Weisenbach et al., 2014). Some tasks use static images with varying levels of intensity such that some emotions are presented at less than 100% intensity, such as 10% intensity instead (e.g., Kamboj et al., 2015; Sasson et al., 2010). Less intense emotional expressions more closely resemble a neutral expression and are therefore more difficult to detect. Using such stimuli may be useful as they may reflect emotional expressions that tend to be misperceived in common social situations, leading to greater effect sizes (e.g., Hamstra et al., 2014, 2015, 2016, 2017). Finally, some tasks use dynamic stimuli, either by using

videos of facial expressions or facial morphs created by presenting several still photos (e.g., Wingenbach et al., 2018).

Sex Differences in Facial Emotion Detection

Independent of depression, past research has also pointed to a sex difference in emotional processing in general, including the processing of facial emotions. Empirical studies have consistently found that women are more accurate and faster at detecting facial emotions than men (Guapo et al., 2009; Saylik et al., 2018). Furthermore, some research has suggested an emotion-specific sex difference. Women may be particularly adept at identifying negative emotions, such as fear and sadness (Hampson et al., 2006). Meanwhile, while some older studies replicated the female advantage for negative emotions, they also indicated that men may be better at identifying angry expressions, especially anger in male faces (e.g., Rotter & Rotter, 1988). The female advantage persists across the lifespan (Sasson et al., 2010). Some research also suggests that this pattern is persistent even in the identification of sub-conscious facial stimuli (i.e., stimuli presented for <200ms) (Hall & Matsumoto, 2004).

Within a static facial matching task with no variation in emotion intensity, both female and male participants typically exhibit almost 100% accuracy, however females show significantly faster reaction times, especially for negative emotions (Hampson et al., 2006). This difference did not persist on a task testing facial processing in emotionally neutral faces, indicating that the observed difference cannot be accounted for simply by a facial processing speed difference (Hampson et al., 2006).

The female advantage was detected even within a more complex task. The stimuli used in this task were 10-second long videos of 10 different emotions at three different intensity levels (Wingenbach et al., 2018). Participants were asked to identify what emotion they perceived.

Women showed higher accuracy in the detection of all facial emotions at all intensity levels. Additionally, women showed faster reaction times for all six basic emotions (i.e., happiness, sadness, anger, fear, surprise, disgust) across all intensity levels. These differences did not persist in the identification of neutral faces, indicating they are specific to FED. These findings indicate that the observed higher accuracy and speed of FED in women is not dependent on a particular methodology but rather is robust.

Speculation about the source of this sex difference has prompted two related theories: the primary caretaker hypothesis and the fitness threat hypothesis. The primary caretaker hypothesis posits that since women have been the main caretakers of children throughout evolution, they have developed an evolutionary advantage to better identify emotions, as this allows them to optimally care for offspring (Wingenbach et al., 2018). As an alternative perspective of this theory, the fitness threat hypothesis states that women are superior at identifying *negative* emotions, as this indicates potential threat to the offspring and prompts caretaking action (Hampson et al., 2006, 2021). A study by Hampson et al. (2021) supports this theory with the finding that the female advantage in emotion identification extends to the identification of facial emotions in infant faces.

These differences have also been considered from a neurological perspective. The localization of emotional processing in the brain is vast and lateralized. Emotions are broadly processed in the right hemisphere, though negative emotion processing is generally more lateralized to the left hemisphere (Blom et al., 2020). Additionally, the patterns of activation in response to certain types of emotional stimuli differ in women vs. men. Meta-analysis has indicated that women have greater brain activation when processing negative emotional stimuli (referring to both facial stimuli and other stimuli such as emotional videos, sounds, and smells),

while men show greater activation when processing positive emotions (Stevens & Hamann, 2012). It has also been found that men recruit more widespread cortical and limbic brain areas for the identification of angry, happy, and sad emotions (Weisenbach et al., 2014). As women are better at identifying facial emotions, this suggests that men activate more brain areas in order to perform the same task (Weisenbach et al., 2014). Overall, women and men show differing intensity and localization of brain activity when processing emotional stimuli and localization depends on the valence of the emotional stimuli.

In terms of specific brain regions, the amygdala is an important structure contributing to emotional processing, especially in processing fearful or threatening stimuli. When processing facial emotions, in particular angry faces, men show higher right amygdala activation compared to women (Schneider et al., 2011). Men also show greater amygdala activation when viewing other types of threatening emotional stimuli, such as videos of animal and human attacks (Schienle et al., 2005). Aside from the amygdala, Kret and De Gelder's (2012) review of the research indicated that increased activation in response to emotional stimuli is also observed within the frontal and temporal cortex, the cingulate cortex, frontal gyrus, fusiform gyrus, and thalamus in men. Overall, when completing facial emotion detection tasks, men recruit greater brain areas compared to women, referring to both location and intensity of activation, yet women tend to outperform men on these tasks. These brain activation studies provide further evidence that facial emotion processing may be a more optimized process in women. One hypothesis is that hormones may play a role in the sex differences.

Hormones and Emotion in Women and Men

One of the factors contributing to the cognitive and emotional differences between women and men are sex hormones. The hormones progesterone, estrogen, and testosterone cross

the blood-brain barrier and bind to receptors within the cerebral cortex and areas associated with emotional processing such as the hypothalamus and the limbic system (Sundström Poromaa & Gingnell, 2014). Within women, the hormones estrogen and progesterone have a major impact on mood and related processes (van Wingen et al., 2011). In terms of emotional processing, testosterone is generally associated with increased amygdala activity in response to threatening faces (van Wingen et al., 2011). One study found that exogenous administration of testosterone increased amygdala reactivity in response to angry facial emotions in both men and women (Hermans et al., 2008). Within men, testosterone affects social behaviours, aggression, and is associated with substance use (van Wingen et al., 2011). Testosterone may also protect against depressive symptoms, while estrogen and progesterone have been found to moderate mood disorders (van Wingen et al., 2011).

Hormonal fluctuations occur in women across the menstrual cycle. The typical menstrual cycle lasts 25 to 35 days, with 28 days being the average length (Reed & Carr, 2000). The cycle can be characterized by two main phases; the follicular phase and the luteal phase (Reed & Carr, 2000). Within a typical 28 day cycle, the follicular phase encompasses days 1 through 14. The first day of the cycle corresponds to the beginning of menstruation, during which the uterine lining is shed. During menstruation, both estradiol and progesterone levels are low (Hampson & Young, 2007). Following menstruation, a preovulatory follicle begins to develop into a mature oocyte, which is accompanied by a steady increase in estradiol. Estradiol peaks around day 14 triggering ovulation, which is the release of the mature oocyte from the ovary. Ovulation represents the end of the follicular phase. The luteal phase follows and encompasses days 15 through 28. During the luteal phase the uterine lining thickens in preparation for possible pregnancy. The early luteal phase is characterized by a sharp drop in estradiol from its peak

during the periovulatory phase. Over the course of the luteal phase estradiol and progesterone levels gradually increase. During the mid-luteal phase, around days 20 to 24, progesterone peaks and estradiol reaches moderate levels (Hampson & Young, 2007). If the released oocyte has not been fertilized by this point, progesterone and estradiol levels drop and trigger the beginning of the next menstrual cycle (Reed & Carr, 2000).

It is common to refer to days 1 to 7 of the menstrual cycle as the early follicular phase, days 8 to 14 as the late follicular phase, days 15 to 21 as the early luteal phase, and days 22 to 28 as the late luteal phase (Sundström Poromaa & Gingnell, 2014). However, there is variability in how researchers may choose to categorize the phases of the menstrual cycle. Alternatively, it is also possible to distinguish shorter cycle phases more specifically as follows; days 1 to 5 (menstruation/early follicular), days 6 to 10 (mid follicular), days 11 to 14 (late follicular), days 15 to 19 (early luteal), days 20 to 24 (mid- luteal), days 25 to 28 (late luteal), and days 24 to 28 (late luteal) phase (Hawkins & Matzuk, 2008). These differences contribute to a source of variability among menstrual cycle studies.

Studies examining the effects of hormones on cognition often compare cognitive processes across menstrual cycle days or phases. This is because the hormones estrogen and progesterone are able to pass through the blood-brain barrier and bind to receptors in the brain, as well as influence other signaling pathways, such as those governing synaptic formation (Le et al., 2020). To assess cognition over the menstrual cycle measurements are often taken in phases of the menstrual cycle during which estrogen and progesterone are highest (e.g., days 11 to 14 for estradiol; days 20 to 24 for progesterone). The early follicular phase, specifically days 2 to 5 is generally taken as the point of lowest estrogen and progesterone levels, and may be used as a baseline measure (Mordecai et al., 2008). To study the effects of estrogen, the late follicular

phase represents the time of peak estrogen levels (Sundström Poromaa & Gingnell, 2014). Meanwhile to study progesterone, the mid luteal phase, is the period of high progesterone levels, with progesterone generally peaking at day 21 of the cycle (Sundström Poromaa & Gingnell, 2014). Finally, to study the premenstrual phase, such as to examine premenstrual symptoms, days 25 to 2 of the cycle generally coincide with the low levels of estrogen and progesterone, and the greatest levels of symptoms (Schmalenberger et al., 2021). However, menstrual symptoms generally also peak during the first days of the menstrual period, so studying premenstrual symptoms during menstruation is generally discouraged (Schmalenberger et al., 2021). Using this methodology allows for a convenient assessment of cognitive performance across the menstrual cycle, however it does pose certain issues. Specifically, when hormone assays are not used, individual differences in cyclical hormone release and anovulatory cycles, can increase error variance (López et al., 2010). Thus, cycle phase comparisons involve an inference about hormonal status (Le et al., 2020). Despite these limitations, studies employing this methodology have found evidence that hormones affect cognitive processes, including emotional recognition.

Hormones and Facial Emotion Recognition

Estrogen and Progesterone

The chronobiological changes in hormones across the menstrual cycle affect cognitive and emotional processes as well as physical ones. When testing performance across the menstrual cycle, three studies have shown that women exhibit increased accuracy in the detection of all facial emotions across the entire follicular phase compared to the luteal phase (Derntl, Kryspin-Exner, et al., 2008; Derntl, Windischberger, et al., 2008; Rubin et al., 2011). The cycle days examined in the follicular phase (i.e., two studies looked at days 1 to 14; one study looked at days 2 to 4) suggest that the follicular phase advantage occurred even when

looking at days where both estradiol and progesterone were very low (Rubin et al., 2011). Consistent with these findings, low progesterone levels, which occur during the follicular phase, were also associated with greater accuracy in detection of all emotions (Derntl, Kryspin-Exner, et al., 2008). Low progesterone has also been associated with enhanced detection of facial symmetry in men's faces (Oinonen & Mazmanian, 2007). Another study found no difference in accuracy of FED across the menstrual cycle at all (Zhang et al., 2013). Conversely, one other study reported a nonsignificant trend towards an increase in accuracy of FED during the follicular phase, but the lack of significance may have been due to a very small sample size (Gingnell et al., 2012). Other studies did not find a general follicular phase accuracy effect, but instead found evidence that accuracy is increased only for certain emotions during the follicular phase. Specifically accuracy was greater for angry and sad expressions during the early follicular phase compared to ovulatory and luteal phases (Guapo et al., 2009), and for fearful expressions during the late follicular phase compared to the early luteal phase (R. Pearson & Lewis, 2005). Evolutionary explanations suggest that the follicular phase is the period during which women would look for a mate, and having heightened social/emotional sensitivity would enhance mating chances (Derntl, Windischberger, et al., 2008). These findings may also extend to facial processing in general. One study identified that women in the follicular phase were more accurate in detecting facial symmetry (Oinonen & Mazmanian, 2007). Other studies have found evidence that visuospatial processing more broadly is enhanced during the follicular phase (Hampson, 1990; Maki et al., 2002), though these results tend not to be reproduced in more recent studies (Halari et al., 2005; Leeners et al., 2017). These findings have been conceptualized as a visuospatial social sensitivity advantage during the follicular phase, whereby enhanced facial processing would contribute to this advantage.

Lower estradiol, which occurs in the early follicular phase, has also been associated with greater accuracy in the detection of angry emotions (Guapo et al., 2009). Estrogen levels have also been differentially correlated with accuracy in the detection of other specific emotions. One study found that fear was detected more accurately during the late follicular phase when estrogen is at its peak (Pearson & Lewis, 2005) but high estrogen was associated with lower accuracy in the detection of disgust (Kamboj et al., 2015). However, these results have not been replicated so specific conclusions can not be drawn.

Progesterone levels are also correlated with accuracy in the detection of negative emotions such as anger, fear, and disgust. The direction of this relationship also differs depending on the study and results are generally not consistent (Conway et al., 2007; Derntl, Kryspin-Exner, et al., 2008; Kamboj et al., 2015). Overall, the sex hormones estradiol and progesterone are both associated with facial emotion processing, although presently the role of estradiol is better understood.

Androgens

Androgens such as testosterone also have an effect on FED. As testosterone is higher in men, it is possible that high levels partially contribute to the lower accuracy in FED that is observed within men compared to women. Indeed, research has indicated a negative correlation between testosterone levels in men and the ability to accurately detect expressions of fear (Rukavina et al., 2018; van Honk & Schutter, 2007). One of the same studies also identified that when comparing groups of individuals with high versus low testosterone, those with high testosterone showed less accurate general recognition of all facial emotions (Rukavina et al., 2018). While testosterone likely contributes to some of the variability seen between the sexes, it is likely not independent. The decreased accuracy in emotion detection seen in men is likely

attributed to testosterone and other factors such as differences in brain structure (Kret & De Gelder, 2012) and social factors (Hampson et al., 2021) that were discussed previously.

Testosterone also modulates responses to social aggression or threat. Detection of angry facial emotions is an element of social aggression detection. One study looking at endogenous salivary testosterone levels measured an increase in testosterone in response to viewing images of angry facial emotions (Zilioli et al., 2014). Interestingly, while testosterone increases activity in the amygdala, some studies have found that it decreases recognition of facial expressions of anger (van Honk & Schutter, 2007). There is also evidence suggesting that the testosterone effect may be mediated by cortisol. In one study, participants with a high testosterone to cortisol ratio showed greater brain activity in response to angry faces (Hermans et al., 2008). The hormonal combination of high testosterone and low cortisol is also generally associated with more aggressive behaviour and less threat in response to angry emotions in others (Hermans et al., 2008). Consistent with this, administration of endogenous testosterone to women decreased their avoidance of angry faces (Romero-Martínez et al., 2021). It may be the case that testosterone decreases recognition of anger in others, and therefore decreases feelings of threat, and in combination with low stress (corroborated by low cortisol) prompts aggressive behaviours. This suggests a need for further investigation of the independent effects of cortisol and stress on facial emotion processing.

Cortisol

There is evidence that the immediate effects of elevating endogenous cortisol include effects on attention and reaction time to certain emotions. Some studies induce a high cortisol condition through a stress test that causes psychosocial stress. Such studies tend to find mixed effects of cortisol on FED. For example, one study found that high cortisol induced by stress was

associated with detection of disgust at a higher intensity (later detection) and surprise at a lower intensity (earlier detection) (Daudelin-Peltier et al., 2017). Alternatively, another study found that those who had high cortisol induced by stress had greater accuracy in the detection of happy expressions and lower accuracy in the detection of angry expressions compared to controls (von Dawans et al., 2020). The variability in these findings has been attributed to potential confounds due to the manipulation of stress by inducing psychosocial stress (Daudelin-Peltier et al., 2017).

To mitigate these confounds, the majority of studies looking at cortisol often induce a high cortisol condition through the administration of exogenous hydrocortisone (Romero-Martínez et al., 2021). When studied in this way, a clearer pattern of cortisol's effect of facial emotion processing is seen. While women normally show a heightened ability to discriminate facial emotions compared to men, administration of cortisol decreased this effect, and lead to no gender difference in the accuracy of identification of angry and sad expressions (Romero-Martínez et al., 2021). Additionally, cortisol administration leads to faster reaction times in identifying angry and sad expressions (Romero-Martínez et al., 2021). Similarly, administration of cortisol also increased memory for angry expressions, while decreasing memory for fearful expressions (Putman et al., 2007). This has been conceptualized as an adaptive mechanism, whereby short term cortisol spikes facilitate an approach motivation, leading to increased aggression and bias toward anger emotions. In contrast, chronic maladaptive elevated cortisol levels contribute to an avoidant motivation, facilitating detection of fearful emotions (Putman et al., 2007). Consistent with this finding, a review by Bérubé et al. (2021) found that those with a history of adverse childhood experiences (ACEs) were able to detect anger and fear faster and at lower intensities, compared to those with no ACEs. While cortisol seems to affect attentional processes relating to FED, the exact pattern of action is not yet clear in the literature.

Oral Contraceptives

OCs are pills taken by women in order to prevent unwanted pregnancy. They are also taken for various hormonal conditions (Słopień et al., 2018). Worldwide there are 151 million users of OCs, making it one of the most used contraceptive methods (Chae et al., 2021; Duesenberg et al., 2016). OC pills contain either a combination of synthetic estrogen (ethinyl estradiol [EE]) and progesterone (progestin), or a progestin only. In addition to this there are four classes of OCs that mainly differ in the type of synthetic progesterone they administer (Batur et al., 2003). The three classes are first generation, second generation, third generation, and new (fourth) generation (Batur et al., 2003). New generation OCs, such as Yasmin™ and Zarah™, are monophasic, meaning they deliver the same dose of hormones throughout the entire menstrual cycle. These OCs contain the progestin Drospirenone which is the most anti-androgenic, making them ideal for women that suffer with certain conditions linked to androgens, such as acne and hirsutism (Machado et al., 2011). Second generation OCs, such as Alesse™ and Min-Ovral™, are also typically monophasic, but normally contain the progestin Levonorgestrel. This progestin replicates the actions of androgens and binds to androgen receptors upregulating the action of these receptors, making it the most androgenic formulation (Darney, 1995). Third generation OCs, such as Tri-Cyclen™, are often triphasic, meaning they deliver different doses of hormones throughout the cycle (Batur et al., 2003). Third generation OCs contain a lower dose of estrogen and the progestin Norgestimate or Desogestrel, which have lower androgenic activity compared to Levonorgestrel, but are still androgenic (Kaunitz, 2004). These different classes ultimately contribute to different hormone profiles and side effects. For example, for women struggling with acne or polycystic ovary syndrome, the new-generation or third generation OC formulations may be more appropriate since they have a higher anti-

androgenic effect, that can help manage these conditions (Powell, 2017). Other research has also suggested that second generation OCs are associated with a greater increase in negative mood side effects compared to third generation OCs (Shahnazi et al., 2014). To summarize, OCs can differ based on whether they are monophasic vs. triphasic, androgenic vs. antiandrogenic, the type of progesterone they contain, and whether they deliver estrogen and progesterone or just progesterone. All of these factors can contribute to different side effect profiles.

Regardless of OC type, all OCs generally have the same mechanism of action. Since OCs typically deliver consistent levels of progesterone and estrogen, depending on the type, they increase levels of exogenous hormones. These hormones contribute to a negative feedback loop, inhibiting the HPG axis. This suppresses the release of FSH and LH from the anterior pituitary which normally results in a downregulation of the production of endogenous hormones (Hamstra et al., 2015; Marečková et al., 2014). This means that while OC use increases exogenous estrogen and progesterone, it leads to reduced production of endogenous estrogen and progesterone. Hence, in pill-taking women the endogenous levels of both estrogen and progesterone are lower compared with free-cycling (FC) women (Fleischman et al., 2010). Due to the decrease in endogenous female hormones, it has been suggested that all OCs exert a defeminizing effect. Defeminization refers to the suppression of anatomical and behavioural characteristics that are typical and unique to females (Wallen, 2017). Additionally, androgenic formulations may provide a larger defeminizing effect than anti-androgenic ones.

The effects of OCs also extend to other hormones as well. One meta-analysis found that OCs containing both estrogen and progesterone decrease levels of testosterone by inhibiting androgen synthesis (Zimmerman et al., 2014). OCs also affect other steroid hormones, such as cortisol. Progesterone binds to the mineralocorticoid receptor with greater affinity than cortisol,

however upon binding of progesterone the mineralocorticoid receptor produces a weaker response (Quinkler et al., 2002). This may suggest that progesterone may decrease cortisol's efficacy. Interestingly, endogenous estrogen was found to increase unbound cortisol levels, but only minimally increase cortisol metabolism (Burke, 1969). Consistent with these findings, the relationship between cortisol and experienced stress is weakened in women using OCs (Lewis et al., 2019; Nielsen et al., 2013). Additionally, a meta-analysis has identified that women taking OCs have a decreased cortisol response compared to FC women in response to a psychological stress task (Gervasio et al., 2022). Nielsen and colleagues failed to find that OCs affected the peak level of cortisol, but rather suggest that OCs may lead to a faster decrease in cortisol levels. It has also been found that higher baseline cortisol levels were associated with less stress response in OC-using women, but not in FC women. (Nielsen et al., 2013). This suggests there may be a subset of women, with an intrinsically higher baseline cortisol level, that are more sensitive to the various stress-related side effects associated with OCs.

OC use is associated with various adverse side effects, including mood-related side effects. Mood side effects, including depressive symptoms, are particularly an issue for those taking combined estrogen and progesterone OCs, and are one of the major reasons for OC use discontinuation (Rosenberg & Waugh, 1998). While some women do experience improved mood while taking OCs, approximately 10% of OC users experience depressive mood (Poromaa & Segebladh, 2012). The incidence rates of depression and antidepressant use is also higher among OC users compared to non-users, and those that use OCs in adolescence are particularly prone to developing depressive symptoms (Skovlund et al., 2016).

Although these mood side effects are often referenced in the literature, relatively little is known about the mechanisms contributing to these effects (Poromaa & Segebladh, 2012). The

combined alteration of endogenous estrogen, progesterone, testosterone, and cortisol in OC users likely contributes to these observed negative effects, however, vitamin B deficiencies, history of depression, anxiety and eating disorders, risk use of alcohol (Bengtsdotter et al., 2018), family history of negative mood symptoms related to OC-use (Kendler et al., 1988) have also been suggested as contributors (see Oinonen & Mazmanian, 2002 for a review). To better understand mood effects associated with OCs, it is also beneficial to consider the other processes that OCs affect. It has been found that OC users experience less mood variability across the menstrual cycle (Oinonen & Mazmanian, 2002), and reduced positive affect in response to emotional mood primes (Jarva & Oinonen, 2007). In addition to these processes, there is ample evidence that OCs affect facial emotion processing.

OC Use and Facial Emotion Detection

One of the goals of this thesis was to conduct a scoping review on OCs and FED. There have been two reviews conducted to date outlining the effects of OCs on the processing of facial emotions (Gamsakhurdashvili et al., 2021; Osório et al., 2018). In addition to the papers included in these reviews, searches were performed in order to attempt to identify all articles examining OCs and facial emotion detection. A full description of databases searched and search terms used can be found in Appendix A.

After full article review, and the incorporation of additional studies that were outlined in the review articles, 12 relevant original empirical studies were identified in addition to the two review papers. Four articles did not appear in any of the reviews. Table 1 provides a summary of the relevant studies.

The Osório et al. (2018) review concluded that OCs affect facial emotion processing. While the seven studies they included showed mixed results, their main finding was that OCs are

associated with lower accuracy in the detection of negative emotions. Alternatively, the Gamsakhurdashvili et al. (2021) review concluded that OC use is associated with impaired FED in general, that it is not restricted to negative emotions. However, there is variability in the outcomes of the studies included above and in the reviews discussed. Overall, of the 12 studies examined in the present review, 6 (50%) studies found that OC users had lower accuracy when detecting facial emotions, compared to FC women (Gurvich et al., 2020; Hamstra et al., 2014, 2015, 2016, 2017; Pahnke et al., 2019). Two of these studies found that OC use was associated with lower accuracy in the detection of *all* emotions (Gurvich et al., 2020; Pahnke et al., 2019). Three of the studies, which were all from the same lab, found that this effect was observed only in the processing of negative emotions (Hamstra et al., 2014, 2015, 2017). This negativity bias, whereby OC users show enhanced negative emotion processing, also persists on other emotional tasks such as emotional categorization and memory tasks (Hamstra et al., 2014, 2015). Interestingly, while OC users make more mistakes in categorizing negative emotions, some studies have found that in terms of recognition speed, OC users are faster at identifying negative emotions compared to non-users (Hamstra et al., 2014, 2017).

Other studies failed to replicate these findings. Out of the 12 studies, six (50%) studies found no evidence that facial emotion processing differs between OC users and non-users (Gingnell et al., 2013; Kimmig et al., 2022; Menting-Henry et al., 2022; Radke & Derntl, 2016; Shirazi et al., 2020). It is noteworthy that no effect of OC use was found in the largest study to date on OCs and FED (Shirazi et al., 2020). However, it is also noteworthy that 0 of 12 (0%) studies found that OC use was associated with increased accuracy of FED.

The discrepancies in the findings may be related to the different facial emotional detection tasks used in the studies. Hamstra and colleagues (2014, 2015, 2016, 2017)

Table 1

Summary of Studies Assessing Facial Emotion Detection and Oral Contraceptive (OC) Use

Authors	*n*	Facial Expressions [a]							Task Design [b]	OCs	Main Findings
		N	H	S	A	F	S_p	D			
Gingnell et al. (2013)	15 placebo/ 15 OC				X	X			no intensity, static, select image matching emotion of image presented	combination (30 µg ethinyl estradiol (EE) & 0.15 mg levonorgestrel)	No OC effect
Hamstra et al. (2014)	14 non-OC / 26 OC	X	X	X	X	X		X	10% intensity, static	no details	OC ↓ accuracy for sad, angry, disgust OC ↑ speed for sad, disgust
Maner & Miller (2014)	23 non-OC/ 21 OC			X	X	X		X	no intensity, static	no details	No OC effect P ↑ accuracy all negative emotions
Hamstra et al. (2015)	41 non-OC/ 44 OC	X	X	X	X	X		X	10 % intensity, static	39 monophasic combination/ 5 no detail	OC ↓ accuracy for sad MR 1/3 ↑ accuracy for angry, disgust (OC & MR2) ↓ speed for disgust
Hamstra et al. (2016)	44 non-OC/ 49 OC	X	X	X	X	X		X	10 % intensity, static	combination (30 µg ethinyl estradiol (EE) & 0.15 mg levonorgestrel)	(OC & MR 1/3) ↓ accuracy for all emotions
Radke & Derntl (2016)	43 non-OC/ 30 OC	X	X	X	X	X		X	VERT: no intensity, static	monophasic combination	No OC effect
Hamstra et al. (2017)	39 non-OC/ 57 OC	X	X	X	X	X		X	10 % intensity, static	combination (30 µg ethinyl estradiol (EE) & 0.15 mg levonorgestrel)	OC ↓accuracy for sad E ↑ accuracy for happy OC ↑ speed for angry
Pahnke et al. (2019)	53 non-OC/ 42 OC								RMET (no intensity, static images of only eye region)	21 androgenic/ 21 anti-androgenic (exact formulations included in article)	OC (any) ↓ accuracy for all emotions
Shirazi et al. (2020)	192 non-OC/ 203 OC								RMET (no intensity, static images of only eye region)	no details	No OC effect
Gurvich et al. (2020)	35 OC								CogState: Pick 1 odd emotion/intensity out of 4 - digital faces	18 androgenic/ 17 anti-androgenic (exact formulations included in article)	Anti-androgenic OC ↓ accuracy compared to androgenic OC
Kimmig et al. (2022)	56 non-OC/ 30 OC	X	X	X	X	X		X	no intensity, static	mean exoE: 72.7 pmol/L mean exoP: 33.6 nmol/L	No OC effect
Menting-Henry et al. (2022)	20 non-OC/ 32 OC	X	X	X	X	X		X	No intensity, static	16 androgenic/ 16 anti-androgenic	No OC effect

Note. [a] Tasks indicated the inclusion of the following expressions: ; A = angry; D = disgust; F = fear; H = happy; N = neutral; S = sad; S_p = surprise;. [b] CogState = Cognitive Evaluation Social - Emotional Task; no intensity = the task used only 100% intensity stimuli; RMET = Reading the Mind in the Eyes Test (no intensity, static images of only eye region); static = the task used static picture stimuli. VERT = Vienna Emotion Recognition Task; 10% intensity = the task used stimuli varying in 10% intensity intervals. E = estrogen; exoE = exogenous estrogen; exoP = exogenous progesterone; MR 1/3 = mineralocorticoid receptor haplotype 1 and 3; MR 2 = mineralocorticoid receptor haplotype 2; P = progesterone.

consistently used the same Facial Emotion Recognition Task (FERT) in their studies and consistently found oral contraceptive effects. Within this task, participants are randomly presented with either a happy, sad, angry, fearful or disgust expression at varying intensities (in 10% intensity increments) and neutral expressions, and are asked to indicate what emotion they perceive. Compared to just using 100% intensity emotions, the presentation of faces varying in emotion intensities allows tasks to be more sensitive and increases the likelihood of identifying subtle detection differences (Montagne et al., 2007). Some other studies used the Vienna Emotion Recognition Task (VERT) or a comparable task, which tests recognition of static and only high intensity facial emotions (Gingnell et al., 2013; Kimmig et al., 2022; Radke & Derntl, 2016). Consistent with the previous statement, studies using these tasks found no differences between OC users and non-users in emotional detection, which may be due to the lower sensitivity of the task used. Meanwhile, Shirazi et al. (2020) used the Reading the Mind in the Eyes Task (RMET), which presents participants with a picture of a set of eyes, and prompts them to make a choice between various complex emotions (i.e., options were not limited to the basic six emotions). Within this task participants do not process whole face emotions, which may have skewed results. For example, features of the face such as the mouth, not the eyes, are most fixated on when processing negative emotions (Duran & Atkinson, 2021). Shirazi et al. (2020) found no OC effect when testing facial emotion detection using the RMET. Since the RMET involves detecting facial emotions in the eyes as opposed to whole face emotions, it may not have been appropriate to replicate the negative emotion bias observed in OC users in other studies. Additionally, the VERT and comparable tasks represent a simplified test of FED. Differences in detection abilities are more robust within tasks that are more challenging, such as the FERT, which tests emotion detection at varying intensities (e.g., 10%, 20%,... 100%

intensity). Some evidence also suggests that processing differences between OC users and non-users can be observed in the detection of difficult-to-detect expressions, but not easy ones (Pahnke et al., 2019). Taken together, this points to a potential methodological reason for the discrepancies in the literature thus far (i.e., a ceiling effect due to task simplicity in those studies that did not find OC effects) and suggests the need for more challenging tasks in future studies.

It has also been suggested that OC effects on facial emotion processing are not unidimensional and are instead affected by additional state and trait factors, which provides another potential explanation for discrepancies between the above 12 studies. In their study, Hamsta et al. (2016) found that OC users compared to non-users displayed poorer accuracy when processing all facial emotions only if they carried a specific mineralocorticoid receptor allele. The mineralocorticoid receptor binds cortisol and mediates stress reactions, such as attention and emotional memory (Hamstra et al., 2016). Poorer accuracy was observed only in OC using participants carrying the MR 1 or 3 haplotype, but not the MR 2 haplotype (Hamstra et al., 2016). OC users carrying the MR 1 or 3 haplotypes also exhibited increased accuracy for sad and disgust emotions, while this was not observed within the MR 2 haplotype carriers (Hamstra et al., 2015). Kimmig et al. (2022) found that participant emotional state also affected facial processing. OC users in a negative affective state exhibited a negative processing bias, and were more likely to misclassify neutral expressions as negative (Kimmig et al., 2022). Finally, OC type may also affect processing. Gurvich et al. (2020) found decreased emotional recognition accuracy in users taking anti-androgenic OCs, but not within those taking androgenic OCs. All of these individual difference and OC-related factors deserve further study as mediators of OC effects on FED.

Premenstrual Syndrome (PMS) and Premenstrual Dysphoric Disorder (PMDD)

Premenstrual syndrome refers to negative physical and emotional symptoms that some women experience during the late luteal phase, leading up to the beginning of menstruation. The most common emotional symptoms associated with PMS include irritability, depressed mood, nervousness, and tension (Angst et al., 2001). Worldwide it is estimated that 47% of women experience some PMS symptoms (Direkvand-Moghadam et al., 2014).

Relatively mild presentation of these symptoms are quite prevalent in the population, however about 5-8% of women experience more severe premenstrual symptoms that cause significant distress and functional impairment (Mishra et al., 2022). To recognize these severe symptoms, premenstrual dysphoric disorder (PMDD) was included in the *DSM-5* as a depressive disorder. Compared to PMS, PMDD is characterized by more severe emotional symptoms, including irritability or anger, and depression symptoms accompanied by feelings of worthlessness and/or hopelessness (APA, 2013; Hantsoo & Epperson, 2015). The DSM-5 criteria for PMDD requires the presence of at least five such symptoms, at least one of which must be: decreased interest in usual activities; subjective difficulty in concentration; lethargy, easy fatigability, or marked lack of energy; marked change in appetite; overeating; or specific food cravings; hypersomnia or insomnia; a sense of being overwhelmed or out of control; or, physical symptoms such as breast tenderness or swelling, joint or muscle pain, a sensation of "bloating," or weight gain (APA, 2013). Consistent with many other DSM-5 diagnoses, in order for PMDD to be diagnosed the outlined symptoms must be severe enough to interfere significantly with social, occupational, sexual, or scholastic functioning (APA, 2013; Mishra et al., 2022). While severe PMS and PMDD have high symptom overlap, women with clinically significant PMD may not meet PMDD criteria as they may not show five distinct symptoms (Yonkers et al.,

2008). Within scientific reports, women with clinically severe PMS generally correspond to women with PMDD (Yonkers et al., 2008). This speaks to the value of using a dimensional approach to measure PMS, which is consistent with the approach adopted by the DSM-5 (APA, 2013). A dimensional approach allows one to look at subclinical symptoms experienced by many women and examine the full range of symptoms experienced by women (Richards & Oinonen, 2021). This can help increase the scientific and psychosocial understanding of women's experience. For this reason, within the current study, symptoms may be considered on a continuum, with PMDD representing the severe end of the spectrum. Premenstrual symptoms will be considered in this way, and PMS and PMDD will be referred to concurrently.

Although the exact mechanisms leading to PMS and PMDD are not entirely understood, there is evidence that a hormonal sensitivity may contribute to these symptoms (Pope et al., 2017; Steiner et al., 2003). There is mixed literature regarding the hormone profiles of women with PMS/PMDD. Some researchers found no differences in plasma levels of estrogen and progesterone between women with PMS/PMDD and those with no symptoms (Schmidt et al., 1994; Steiner et al., 2003). Alternatively, other research has indicated that women with PMDD have lower estrogen levels and higher progesterone levels during the early luteal phase compared to women with no symptoms (Yen et al., 2019). There is also some evidence that testosterone may contribute to certain features of PMS/PMDD, in particular the irritability symptom (Steiner et al., 2003). Some research has identified elevated testosterone levels during the luteal phase in women with PMS/PMDD, whereas other studies have failed to replicate these findings (Steiner et al., 2003). Notably, one study actually detected lower overall testosterone in PMS/PMDD prone women, however this study did not measure testosterone at a particular phase of the menstrual cycle (Hashemi et al., 2016). It may be the case that PMS/PMDD is not necessarily

induced by dysregulated hormone levels, but rather by a hypersensitivity to abrupt hormonal changes (Payne et al., 2009). Consistent with this theory, the administration of a monophasic OC for 24 consecutive weeks, that is, without taking the placebo pills during what would normally be the withdrawal bleeding period, provided a consistent level of hormones and alleviated premenstrual symptoms (Coffee et al., 2006). Given no evidence of consistent hormonal differences between women with and without PMDD, the hormonal sensitivity hypothesis remains a compelling one.

PMS and Facial Emotion Detection

Due to its effect on emotions, it is worthwhile to investigate how PMS/PMDD may affect other emotional processes such as emotion detection. One of the goals of this thesis was to conduct a scoping review on PMS/PMDD and FED. Searches were performed in order to attempt to identify all published articles examining PMS/PMDD and facial emotion detection. A full description of databases searched and search terms used can be found in Appendix A. One study found that during the luteal phase women with PMDD were more likely to judge emotions as more negative (displayed a negativity bias) and displayed difficulties in discriminating between happy and sad emotions and neutral expressions by being more likely to report emotional expressions in general as being neutral (Rubinow et al., 2007). Similarly, another study found that, compared to women without PMS symptoms, women with PMS were less accurate at detecting expressions of sadness and surprise during the luteal phase (Gültekin et al., 2017). Finally, one study found that women with PMDD were more accurate detecting sad emotions in male faces compared to female faces (Ramos-Loyo & Sanz-Martin, 2017).

Another study that looked at FED across the menstrual cycle considered PMS symptoms as one of various factors that may discriminate between phase groups (Kamboj et al., 2015).

They found that PMS symptoms at the time of testing were not a predictor of FED, however they did not directly examine differences between a PMS versus non-PMS group (Kamboj et al., 2015).

The findings of the Gültekin et al. (2017) study were conceptualized as a potential difficulty in social interactions during the luteal phase, which is the time that PMS symptoms generally occur. This is consistent with the literature on the mood-congruent bias within depression. Both depression and PMS are characterized by negative affect and inferior facial emotion detection. Additionally, those with PMDD show the same negativity bias in judging facial emotions as those with depression (Rubinow et al., 2007). This indicates that further research on the interrelation between hormonal sensitivity and mood expression and FED is warranted.

Limitations of Past Studies

There are a few limitations within past studies examining hormonal effects on facial emotion detection. As mentioned previously, the type of FED tasks used show a lot of variability across studies. Some studies used a simplified task, including static images and only one level of emotion intensity (e.g., Radke & Derntl, 2016), while other studies employed more complex tasks, either ones including a greater variety of emotions (Wingenbach et al., 2018), varying emotional intensities (Hamstra et al., 2014, 2015, 2016, 2017), or some dynamic component (e.g., Wingenbach et al., 2018). Generally, emotion recognition abilities were more challenged by the complex studies, such as those conducted by Hamstra et al. (2014, 2015), compared to simpler ones (Pahnke et al., 2019). The majority of studies also fail to report descriptive statistics, such as the percentage of correct trials within facial emotions tasks, which further makes it difficult to assess the appropriateness, the complexity, or the power of tasks used.

Additionally, within the Pahnke et al. (2019) study, it is noteworthy that the distinction between OC users and non-users was more evident when processing expressions that are more difficult to recognize. Between-group differences in accuracy and speed of detection are also more generalizable and robust when the FED task used has greater external validity, such as by considering emotion intensity and including a dynamic component (Krumhuber et al., 2013; Wingenbach et al., 2018). The studies that did present varying intensities of facial emotions were more likely to find an effect of OCs on performance (i.e., Hamstra et al., 2014, 2015, 2016, 2017). Finally, the type of emotions displayed may also play a role in study outcomes. Several studies failed to include both positive and negative emotions in their tasks. For example, the Gingnell et al. (2013) and Maner and Miller (2014) studies only included negative facial emotions, meaning that it is impossible to draw conclusions regarding differences in processing positive versus negative emotions. Since a negativity bias is present within depression and PMS/PMDD research, it is essential that tasks include both positive and negative emotion types, as well as neutral expressions that may act as a control and can also be examined to determine if biases (e.g., the negativity bias) persists in neutral faces. It would also be ideal if studies measure participant affect/mood and examine this variable as a factor or covariate.

The RMET task is also often used in emotion detection research, however this task does not display the entire face. There is evidence that entire face processing is distinct from just eye processing. For example, one imaging study identified a brain region, the right inferior frontal junction, that is activated primarily by eye processing (Chan & Downing, 2011). This indicates that results from face-emotion versus eye-emotion detection tasks may not be interchangeable.

Aside from task type, the type of data that is collected and reported is also inconsistent across studies. Some studies report enhanced detection of certain emotions but fail to specify

whether such a statement is based on the accuracy or the speed of detection (e.g., Putman et al., 2007). This may be because many studies do not measure both accuracy and speed (e.g., Gurvich et al., 2020; Maner & Miller, 2014; Radke & Derntl, 2016). Additionally, it is also atypical to find reports of other descriptive statistics such as the percentage of correct responses for each type of emotion (e.g., the only studies that provided descriptive statistics were Kimmig et al., 2022; Menting-Henry et al., 2022; Radke & Derntl, 2016) or the types of errors that are being made (e.g., none of the studies on OC use and FED examined the types of errors being made). Those that do report errors often fail to specify whether incorrect responses are included within statistical analyses for reaction times. Failure to report this information makes it impossible to draw conclusions about potential confounds affecting outcomes.

The Current Study

To summarize, many previous studies looking at OC use and FED have likely not utilized methodology that is comprehensive, complex, or sensitive enough in order to draw conclusions regarding the role of OCs in FED. Additionally, the literature is lacking studies looking at the influence of PMS/ PMDD on FED. Finally, given the link between hormonal fluctuations and mood, including the co-occurrence of mood disorders, it is valuable to investigate the effects of negative mood, OCs, and PMS all within one study.

The current study used a dynamic FED task, that assessed the accuracy, and intensity at which individuals were able to detect basic facial emotions. This methodology was used in order to increase the likelihood of determining even subtle differences in FED. Individual differences in performance on this task were primarily examined as a function of their level of depressive symptoms, OC use, and level of PMDD symptoms. The current study has the largest sample size of all studies investigating OCs and whole face emotion detection (i.e., not including the Shirazi

et al. (2020) study, which had a larger sample size but used the RMET). This is also the first study to investigate PMS/PMDD and FED accuracy and intensity across all six basic emotions.

Previous literature points toward three main hypotheses. Hypothesis 1: (a) Women with high depression symptoms will be faster and more accurate at detecting negative emotions compared to participants with low depression symptoms; (b) and this effect will be strongest for sad expressions. Hypothesis 2: OC users will be slower and less accurate at detecting negative emotions compared to FC women and men. Hypothesis 3: Women with a provisional PMDD diagnosis will be faster and more accurate at detecting negative emotions compared to participants with no PMDD diagnosis; (b) and this effect will be strongest for participants in the premenstrual period.

Method

Participants

The final sample consisted of 163 participants (37 OC using women, 72 FC women, 35 men, 19 other hormonal contraceptive using women). The age of participants ranged from 18 to 47 (M = 22.69, SD = 5.40), and 72.0% of participants were White/European. In terms of education, 96.3% of participants were current university students. Additionally, 19.5% of participants self-reported they have a current depression diagnosis. Finally, 48.4% of female participants reported having significant PMS symptoms.

A total of 335 participants were initially recruited and participated in a study on "*Hormones and Facial Emotion Detection*" from Lakehead University through the Psychology Department research recruitment platform, SONA, and from the general university community, the Thunder Bay community, and the online internet community (See Appendix B for sample recruitment material). There were no initial exclusion criteria other than age (16 and older for

Lakehead students; 18 and older for community participants) and having access to a computer with a keyboard. However, participants were screened for inclusion in data analyses using the Demographic and General Background Questionnaire. Participants were excluded if they met any of the following criteria: (a) had consumed alcohol or other cognition-altering substance in the past 5 hours (n = 15), (b) pregnant or lactating (n = 7), (c) menopausal or have had a hysterectomy (n = 4), (d) intersex (n = 0), (e) had changed their OC use status in the past 2 months (n = 1), or (f) were taking any hormonal medication (e.g., thyroid medication, polycystic ovary syndrome medication, hormone therapy for transgender care, hormone replacement therapy) (n = 16). Participants were also excluded if they had more than 10% missing data on the Facial Emotion Detection Task (i.e., had more than two invalid trials) (n = 22). Finally, due to the FEDT being hosted on a separate platform, some participants had difficulties accessing the task, some participant codes were incongruent and could not be linked, and some participants did not complete the task at all (n = 126).

The total number of participants excluded based on the exclusion criteria was 43, based on task performance was 150, and a total number of 172 participants were excluded for one or both reasons. The final sample for hypothesis 1 included only women and consisted of 101 participants (49 low depression symptom women, 52 high depression symptom women). The final sample for hypothesis 2 consisted of 144 participants (37 OC-users, 72 FC women, 35 men), and excluded participants using other hormonal contraceptives. The type of OCs that participants used is presented in Table 2. The number of participants using androgenic (1st and 2nd generation) formulations (n = 24) was greater than the number using anti-androgenic (3rd and 4th generation) OC formulations (n = 12). The final sample for hypothesis 3 included only FC women and consisted of 72 participants (30 no/minimal PMDD, 34 mild PMDD, 8 severe

PMDD). The study was approved by The Lakehead University Research Ethics Board (see approval in Appendix C) and participants gave informed consent prior to participating. Participants recruited through SONA received 1.5 bonus points as Psychology course credit for participating.

Measures

The study was completed online and contained three sections: A Demographic and General Background Questionnaire, a Facial Emotion Detection Task, and a Final Questionnaire.

Demographic And General Background Questionnaire

The Demographic and General Background Questionnaire consisted of a demographics questionnaire and assessed for depressive symptomology and PMS symptoms using the self-report measures. The demographic questionnaire (see Appendix D) collected information on demographics (e.g., age, sex, ethnicity), health history (e.g., history of psychological disorders and hormonal disorders), substance use, and sleep, and also included questions for women about contraceptive use history and their menstrual cycle. Self-report measures used include the 16-item Quick Inventory of Depressive Symptomatology ($QIDS_{16}$; Rush et al., 2003), the DSM-5-Based Screening for Premenstrual Symptoms (DSPMS; (Richards & Oinonen, 2021), and the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988). These are described in more detail below.

16-item Quick Inventory of Depressive Symptomatology ($QIDS_{16}$). The $QIDS_{16}$ was developed by Rush and colleagues (2003) as a 16-item shortened version of the full-scale 30-item Inventory of Depressive Symptomatology (IDS). The scale assesses depressive symptoms across nine domains including sad mood, concentration, self-criticism, suicidal ideation, interest, energy/fatigue, sleep disturbance, decrease/increase in appetite/weight, psychomotor

Table 2

Type of Oral Contraceptives (OCs) Used by Participants

Generation of Progestin	Phase Type	Frequency per Type (%)	Brand Name	*n*
1st Generation	Monophasic	13.9	Aviane	2
			Lolo	3
	Biphasic	2.8	Synphasic	1
2nd Generation	Monophasic	50.0	Alesse	2
			Alysena	12
			Indayo	2
			Ovima	2
3rd Generation	Monophasic	5.6	Apri	1
			Freya	1
	Triphasic	13.9	Linessa	1
			Tri-Cyclen	1
			Tricira Lo	3
4th Generation	Monophasic	13.9	Mya	1
			Slinda	1
			Yaz	1
			Zamine	2

Note. Progestins used within formulations are as follows: 1st generation: Norethindrone; 2nd Generation: Levonorgestrel; 3rd Generation: Desogestrel; 4th Generation: Drospirenone. 2nd generation progestins represent androgenic formulations, while 4th generation progestins are the most anti-androgenic.

agitation/retardation (Rush et al., 2003). Two domains are measured by only taking the highest score amongst a few items. For example, the appetite/weight domain is assessed based on the highest score on any of four items asking about increased appetite, decreased appetite, increased weight or decreased weight. The question assessing suicidal ideation was deemed inappropriate for the present study for ethical reasons, and was replaced with a new question asking about hopelessness. The $QIDS_{16}$ includes all criterion items to diagnose MDD based on the DSM-IV (Rush et al., 2003) and the DSM-5 (APA, 2013). The 16 items are rated on a four-point Likert scale, yielding a total score between 0 and 27 (Rush et al., 2003). Recommended cut-offs for depression severity are as follows: normal (0-5), mild (6-10), moderate (11-15), severe (16-20), very severe (21+). The $QIDS_{16}$ has a high internal consistency of .86, and scores are highly correlated with the IDS (.96) (Rush et al., 2003). The $QIDS_{16}$ is also highly correlated with the Hamilton Rating Scale for Depression (HRSD) (.86), and it has been suggested that clinician administration of the IDS or HRSD is comparable to the self-report version of the $QIDS_{16}$ (Rush et al., 2006). In the present study, the $QIDS_{16}$ score was used to delineate depression groups. The participants with scores in the bottom 40% (scores ≤ 8) were assigned to the low depression group, and participants with scores in the top 40% (scores ≥ 12) were assigned to the high depression group.

Positive and Negative Affect Schedule (PANAS). The PANAS is a scale developed by Watson et al. (1988), designed to measure affective states. It contains 20 adjectives describing 10 positive affect (PA) states and 10 negative affect (NA) states (Watson et al., 1988). For each item participants indicate the extent to which they feel that way on a Likert scale ranging from *very slightly* or *not at all* (1) to *extremely* (5). Participants initially completed only the NA scale and were asked to indicate the extent to which they felt that way over the past 2 weeks. This was

used to supplement information about long term negative/depressive affect obtained from the $QIDS_{16}$. Prior to the Facial Emotions Task, the NA and PA scales were both administered using a shorter time frame. Participants were asked to respond based on how they currently feel. The scale has been validated for both the two-week and the current time frames. The scale has high internal consistencies of .89 for the PA subscale and .87 for the NA subscale (Watson et al., 1988).

DSM-5-Based Screening for Premenstrual Symptoms (DSPMS). The DSPMS is an 11-item scale developed by Richards and Oinonen (2021). The scale was created to assess for each criterion in the DSM-5 corresponding to PMDD, and allows for both a dimensional and categorical/diagnostic measure of PMS/PMDD. Each item outlines 1 of 11 criteria and inquires about the severity of the symptom, the impairment it causes and the frequency with which it has occurred during the week prior to menstruation over the past year (Richards & Oinonen, 2021). The first two questions for each item are rated on a 5-point Likert scale (0 = *not at all* to 4 = *severe/extremely*), and the final question is rated based on number of months, ranging from 0 to 12. Possible scores range from 0 to 44, with higher scores indicating greater severity of pre-menstrual symptoms. This measure can be used to assess the severity of PMS symptoms, which was its use in the present study. The DSPMS has a high internal consistency of .92. It also has a high convergent and predictive validity demonstrated by strong correlations with another PMS measure, the Menstrual Distress Questionnaire (Moos, 1968) (r = .70), and with prospective ratings of negative affect during the premenstrual phase (r = .70) (Richards & Oinonen, 2021).

As is described in Richards and Oinonen (2021), the DSPMS can be used to provide a provisional diagnosis of PMDD based on DSM-5 criteria (APA, 2013). Criterion A was met if participants endorsed experiencing any five symptoms for more than six months. Criterion B was

met if participants endorsed experiencing at least one of items 1 to 4 on the DSPMS for more than six months. Criterion C was met if participants endorsed experiencing at least one of items 5 to 11 on the DSPMS for more than six months. Criterion D was determined by the total scores on the Intensity and Severity scales. If either score was between 0-11, Criterion D was not met. Criterion D was met if the product was 12 or greater. Additionally, distress was categorized as mild for scores of 12 to 32, and moderate-severe for scores of 33 or higher.

Participants meeting all criteria were assigned to the Mild PMDD group, if their criterion D score was in the mild range, or the Moderate-Severe PMDD group, if their criterion D score was in the moderate-severe range. Participants that did not meet all criteria were assigned to the No/Minimal PMDD group.

Facial Emotion Detection Task

The facial emotion detection task (FEDT) is a novel task developed for this project. The task measures intensity and accuracy of facial emotion detection as participants view and respond to facial emotion images within an intensity morph (a morph from a neutral expression to an emotional expression). Morphing tasks, similar to this one, have been shown to be effective in differentiating perceptual differences across various settings and populations (Stottinger et al., 2016). This methodology was employed to increase the likelihood of identifying subtle differences in facial emotion detection abilities across groups. Additionally, the task utilizes full face stimuli and includes all six basic emotions, in order to address the limitations of past studies.

Stimuli. The stimuli in the task included images of 24 models expressing neutral, and emotional facial expressions, retrieved from the RADIATE face database (Conley et al., 2018), and one model expressing anger retrieved from the NimStim database which was used for the practice trial only (Tottenham et al., 2009). To reduce possible gender or race perceptual biases,

an equal number of female and male faces, and an equal number of Black, White, Asian and Hispanic faces were selected for the morphs (the database does not have South Asian, Latino or Indigenous models in sufficient numbers for us to include here; Tottenham et al., 2009). Validity ratings of the models' emotional presentations are available in Conley et al. (2018). Conley and colleagues had approximately 50 participants report what emotion they perceived for each image (each model and each emotion type) in the database and a validly score was reported reflecting the percentage of correct responses for each image (Conley et al., 2018). For the present task, these percentages were used in order to choose the stimuli with the most valid emotion presentations. This was done to maximize agreement/validity for the emotional expressions used. A composite score was calculated for each emotion each model expressed by taking the mean of the neutral validity score and the emotional validity score. In calculating the composite score the emotional validity score was given a weight that was double the neutral validity score[1]. The models with the most unambiguous presentations (i.e., the largest validity composite scores for a particular neutral-emotion pair) were chosen to be utilized in the task. The models used in the task had an average composite score of 91.4%. The images were morphed using Psychomorph software following the procedure described by Sutherland (2015). Each model was morphed from neutral to its target emotion in 15 steps. For each full morph the 15 levels of intensity correlated with an increase of 6.6% intensity of the emotion from each image to the next, such that the first image within the morph was neutral and 0% emotion, the second image was 6.6% emotion, the third 13.2% emotion, and each subsequent increased such that the 15th image displayed 100% of the emotion. The task design and similar stimuli were previously used in Boboc et al. (2021). A sample morph is shown in Figure 1.

[1] Since the participants in the present study were told that the first neutral image was neutral, this was done in order to place more weight on the validity of the emotional image.

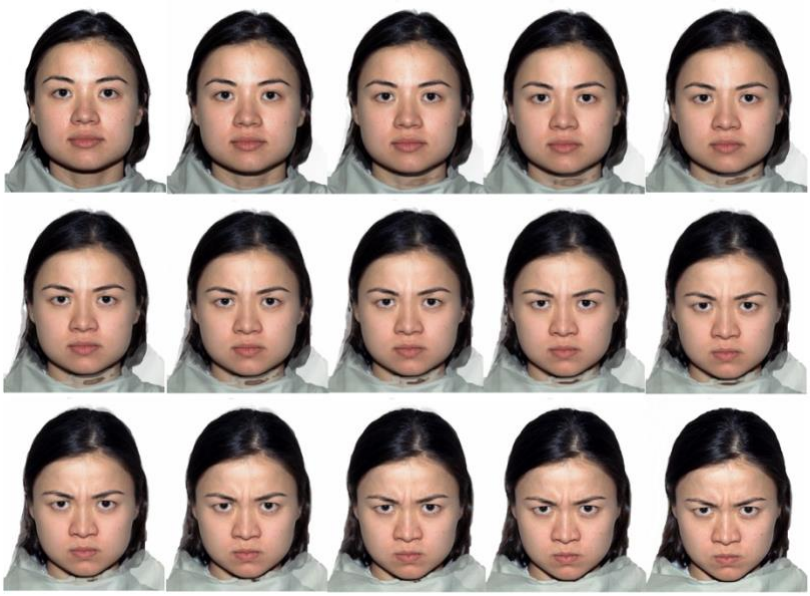
Training Session. The task commenced with a Training Session, to acquaint the participants with the mapping of responses on the keyboard. The emotion response mapping is presented in Figure 2. First participants were shown images with the seven facial emotions used in the task at 100% intensity and asked to respond to them using the appropriate keys. While each face image was presented, instructions relating to the mapping of response options on the keyboard also remained on the screen. See Figure 3 for a sample of what participants saw on the screen during each trial. The facial emotions were presented in the order in which they were mapped on the keyboard from left to right (i.e., disgust, fear, sadness, neutral, anger, happiness, surprise). Only once participants pressed the appropriate key response was the next image shown.

Participants were then told that they would be shown several more images of facial emotions, and to respond to them by pressing the appropriate key response as quickly as they can. Two images of each facial emotion type at 100% intensity (14 images total) were shown in random order, and participant accuracy and reaction time was measured. For each image, once participants indicated the emotion they perceived, only then was the next image presented.

Practice Trial and Measurable Trials. Following the training session to orientate participants to the Facial Emotions Task, they completed a Practice Trial identical to the task. Participants viewed an image of a neutral facial expression, that over 15 steps morphed into a distinct emotion, either anger, disgust, fear, happiness, sadness or surprise. Participants were informed that all trials will commence with an image of an emotionally neutral face, which over the course of 15 images will morph into a detectable emotion, either anger, disgust, fear, happiness, sadness, or surprise. Participants were shown one image at a time and were instructed to indicate what emotion they perceived using the appropriate keys on the keyboard that were

Figure 1

Sample 15-Step Facial Emotion Morph



Note. A sample 15-step facial emotion morph (from neutral in the upper left image to angry in the lower right image (Boboc et al., 2021). Note, participants see only one image at a time within the Facial Emotions Task and are asked to report what emotion they perceive. Unmorphed images were retrieved from http://www.macbrain.org/resources.htm. Printed with permission from Tottenham et al. (2009).

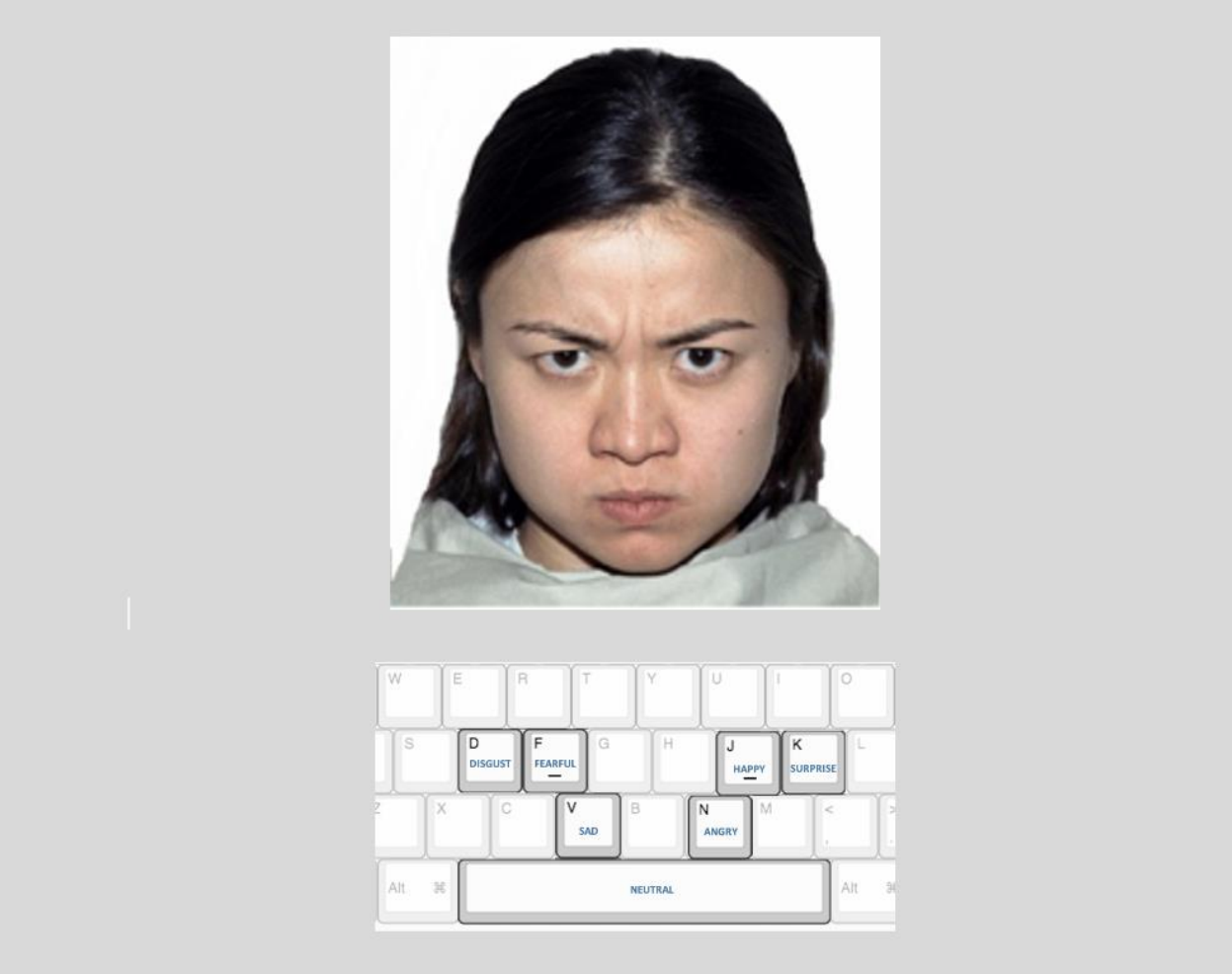
Figure 2

Keyboard Emotion Response Mapping

Note. Mapping of facial emotion response options on the keyboard. Participants were instructed to press the key reflecting the appropriate facial emotion as soon as they detected it.

Figure 3

Sample FEDT Screen



Note. Sample of what participants saw on their screen for every image presented within a trial.

learned in the Training Session (see Figure 2 and 3). The exact instructions participants were given are, "You will be shown an image of a neutral face which, over the span of 15 images, will gradually morph/change into one of the following emotions: disgust, fear, sad, angry, happy, or surprise. For every image in the morph indicate what emotion you see. We are interested in how fast and accurately you can identify the emotion. The first face of each morph will appear neutral. Begin by pressing the SPACE BAR to indicate this. Do this for all of the following faces that appear neutral. Once you see an emotion in the face, press the key corresponding to the emotion that you see. If you see an emotion but are unsure what it is, please only guess the emotion you see when you are reasonably confident. For each image try to respond as quickly as possible."

For each image, once participants indicated the emotion they perceived, only then was the next image in the morph presented, and participants were required to respond to each image in the morph, regardless of whether or not they responded correctly. As in the Training Session, the mapping of the keyboard response options remained on the screen (See Figure 3) while each face image was presented. Following the Practice Trials, participants completed 24 trials of the Facial Emotions Task (four of each emotion). Each trial proceeded identically to the Practice Trial.

Error Minimization. Several measures were put in place to minimize errors on the task. First, participants were only permitted to respond with "neutral" for the first image of each morph. If participants attempt to give a different response, they were reminded of the task instructions. Secondly, trials in which participants appeared to give invalid responses were not included in analyses. Invalid responses included three patterns of responding: (a) participants did not change response options (i.e., persisting in responding neutral to all faces in the trial), (b)

randomly oscillated between response options (i.e., if participants oscillated between responses at least four times in a row or if participants reported four or more emotions, not including neutral), and (c) incomplete trials (i.e., participants stopped responding halfway through the trials, or did not start the trial). Finally, responses were examined to identify any single key mistakes, where participants indicated an incongruent emotion for *one* image in the trial. The incongruent response had to follow at least two congruent responses (e.g., responding "happy" for at least two images, then randomly responding "sad" on one image, and then responding "happy" again for the subsequent image). In this case the singular incongruent response was considered a mistake and was modified to the emotion reported in the adjacent images. All trials were manually inspected for invalidity and single key mistakes by two raters.

Outcome Measures To determine overall performance at the trial level, trials were labeled based on if participants made errors and if they correctly identified the emotion. Trials in which participants identified the correct emotion directly (i.e., switched from reporting a neutral emotion to the correct emotion without reporting incorrect emotions in between) were coded as *primary* trials. If participants reported an incorrect emotion prior to correctly identifying the emotion for the trial, the trial was coded as a *secondary* trial. Conversely, trials in which participants never identified the correct emotion were coded as *error* trials. Two primary outcome measures were calculated: measures of intensity and measures of accuracy.

Intensity Measure. Intensity level of detection is represented by the image number at which participants reported the correct emotion for each trial. However, error trials do not have an intensity level because participants never identify the correct emotion. To prevent a loss of data due to missing values for the error trials, an Image Number at Detection score was computed for each trial. For primary and secondary trials, the Image Number at Detection score

equalled the image number at which participants reported the correct emotion, which could range from 2-15. Lower scores mean better performance. All error trials were assigned an intensity level of 17, which corresponds to two units larger than the highest possible value on primary/secondary trials. This value was determined to produce the highest sensitivity to error trials, while maintaining normal distribution of the variable. The Image Number at Detection for each of the six emotion types was computed as the mean Image Number at Detection across the four trials of that emotion (i.e., six scores).

Accuracy Measure and Overall FEDT Performance Measure. Two scores were calculated to reflect the percentage of correct and incorrect responses that participants make across emotions. Accuracy influences both Percentage of Correct Responses and the Percentage of Incorrect Responses at the emotion level (i.e., not computed at the trial level). The Percentage of Correct Responses for each emotion was computed as the number of correct responses divided by the total number of possible responses across trials for that emotion type. The total number of possible responses depended on the number of valid trials, where if the participant had no invalid trials their total of possible responses was 60 per emotion (i.e., 4 trials * 15 possible responses per trial), and the number decreased by 15 for every invalid trial. For example:

$$\textit{Happy Percentage of Correct Responses} = \frac{\#\ Happy\ Responses}{15\ (4 - \#\ Invalid\ Happy\ Trials)}$$

Conversely, the Percentage of Incorrect Responses was computed as the number of emotional responses that were not the correct emotion divided by the total number of possible responses. For example:

$$\textit{Happy Percentage of Incorrect Responses} = \frac{\#\ Disgust,Fear,Sad,Angry,Surprise\ Responses}{15\ (4 - \#Invalid\ Happy\ Trials)}$$

Accuracy generally refers to whether participants incorrectly detected one emotion as a different one (e.g., being presented a sad emotion and reporting detecting anger). Therefore, the

Percentage of Incorrect Responses was used as the main accuracy outcome variable. The Percentage of Correct Responses is dependent on both accuracy and intensity, because if participants detect emotions earlier they will have a higher number of correct responses. For this reason, Percentage of Correct Responses was not used as an accuracy variable for the main analyses, but was used for supplemental analyses as a measure of overall FEDT performance. Neutral responses were not counted as incorrect, and thus the number of neutral responses explains why there isn't a prefect inverse correlation between the percentage correct and incorrect.

Error Bias Measure. Secondary and error trials were examined in order to identify potential error biases. The type of incorrect responses (i.e., what incorrect emotion the presented emotion was mistaken for) were analysed for each emotion type as a function of group. That is, for the six emotion trial types (disgust, fear, sad, angry, happy, surprise), the percentage of each type of incorrect response was calculated out of the total possible valid responses (i.e., excluding invalid trials).

Pilot Testing. A similar facial emotions task was previously administered to a comparable sample of undergraduate students ($N = 126$; mean age = 19.9) (Boboc et al., 2021). Facial stimuli for this task were retrieved from the NimStim face database, which is an older database developed by the same researchers that developed the RADIATE database {Citation}. The NimStim database contains less models than the RADIATE database, but models are posed and evaluated identically (i.e., same positions, emotions and validation of emotions) within both databases (Tottenham et al., 2009). Within this first administration of the task only happy, angry, and disgust emotional morphs were included. The task and instruction wording were optimized to reduce common errors. The mean image of correct detection was between 5.76 – 7.52 (out of a

possible 15) across the different emotions presented. Across all emotions assessed, the percentage of responses identifying the correct emotion ranged from 10.9% on image 4, to 66.6% on image 7, to 94.0% by image 10. All participants were able to correctly identify the emotion by the last image. These values may reflect earlier detection than what is expected within the current study since participants in pilot testing were asked to select only among three emotions, and within the current task they will have six options.

Final Questionnaire

The Final Questionnaire (see Appendix C) consisted of several self-report measures designed to provide supplementary information that is not integral to the main hypotheses. The self-report measures used were the Behavioural Inhibition System/Behavioural Activation System (BIS/BAS; Carver & White, 1994), the OC Side Effect Questionnaire (OCQ; Oinonen, 2009), the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994), the Adverse Childhood Experiences Questionnaire (ACE-Q; Felitti et al., 1998), and the Infrequency (INF), Negative Impression Management (NIM), and Positive Impression Management (PIM) Scales from the Personality Assessment Inventory (PAI; Morey, 2007). Information from these measures was used for supplementary analyses and as potential covariates.

Behavioural Inhibition System/Behavioural Activation System (BIS/BAS). The BIS/BAS scale assesses the behavioral inhibition system (BIS), characterized by the motivation to avoid aversive outcomes, and the behavioral activation system (BAS), characterized by the motivation to approach goal-oriented outcomes (Carver & White, 1994). The BIS/BAS scale contains 24 items that participants rate on a 4-point Likert scale (1 = *very true for me* to 4 = *very false for me*). Sample items on this scale include *"When I get something I want, I feel excited and energized,"* and "*I often act on the spur of the moment"* (Carver & White, 1994). Carver and

White (1994) report internal consistencies of 0.74 and 0.71, and test-retest reliabilities of 0.66 and 0.64, on the BIS and BAS scale respectively.

OC Side Effects Scale. The OC Side Effect Scale is a questionnaire adapted for this study to assess the severity of adverse physical, emotional, and sexual side effects of OC use. Current and past OC users were asked to complete this questionnaire. The scale was designed based on other scales that have been developed within the Health Hormones and Behaviour lab in past studies (Oinonen, 2009). The original scale contains 50 symptoms, the frequency of which were tested in at least two studies and other theses. Of the 50 symptoms only the symptoms that were endorsed by at least 3% of women in both studies were included. For brevity, this was done to exclude a couple of outlier symptoms that were only endorsed by one or two participants. The questionnaire being used in this study contains 43 symptoms covering the most common physical, emotional and sexual side effects of OCs. There are also certain items that are known to be associated with estrogenic or androgenic side effect profiles (Dickey, 2000; Nelson, 2007). For each item participants rate to what extent they have experienced each symptom when taking oral contraceptives that they believe may be due to oral contraceptives, on a 5-point Likert scale (0 = *Not at all* to 4 = *Extreme*). Participants also indicate whether the experience of each side effect was positive or negative. Performance on the facial emotion detection task may be examined post-thesis based on estrogenic versus androgenic OC side effect profiles (Dickey, 2000; Nelson, 2007).

Toronto Alexithymia Scale (TAS-20). The TAS-20 is a scale designed by Bagby et al. (1994) to measure alexithymia, which refers to difficulties in identifying and describing emotions that one feels. The TAS-20 contains 20 items and is comprised of three subscales: Difficulty Describing Feelings, Difficulty Identifying Feelings and Externally-Oriented

Thinking. Each item is rated on a 5-point Likert scare ranging from 1 (strongly disagree) to 5 (strongly agree). Total scores range from 20 to 100, with greater scores indicating greater difficulties in emotional processing. Score between 52-60 indicate possible alexithymia, and scores equal to or greater than 61 indicate alexithymia. Group differences in TAS-20 may be used as additional evidence of group effects.

Adverse Childhood Experiences Questionnaire (ACE-Q). The ACE-Q is a scale designed to assess abuse and household dysfunction experienced in childhood (Felitti et al., 1998). The ACE-Q contains 10 items describing different types of abuse or dysfunction that participants indicate whether they have experienced (yes or no) prior to the age of 18. The scale has an acceptable internal consistency, with a Cronbach alpha ranging from .66 to .74, and a test-retest reliability of .71 (Zanotti et al., 2017). In the present study the ACE-Q was used as a covariate in exploratory analyses to determine whether childhood stress affects facial emotion detection. It was examined as a potential covariate for analyses.

Response Bias Measure. The Infrequency (INF; 8 items), Negative Impression Management (NIM; 9 items), and Positive Impression Management (PIM; 9 items) subscales from the Personality Assessment Inventory (PAI) were used in order to assess the validity of participants' responses. The IN*F* scale assesses whether participants endorse items that are infrequently endorsed by others. Over-endorsement of these items may indicate that the participant is not responding accurately to questionnaire items due to random responding or other sources of error (Morey & Lowmaster, 2010). The NIM scale items are designed to detect if participants are attempting to purposely present themselves in a more negative manner, meanwhile the PIM scales are designed to detect if they are purposely presenting themselves in a more positive manner (Morey & Lowmaster, 2010).

All scales are widely used and have good psychometric properties. For each item participants are asked to indicate whether the statement is accurate for them on a 4-point Likert scale (1 = *False, not at all true* to 4 = *Very true*). The IN*F* scale has a Cronbach's alpha ranging from .22 to .52 and a test-retest reliability ranging from .42 to .55 (Morey, 2007). The NIM scale has a Cronbach's alpha ranging from .63 to .74, and a test-retest reliability ranging from .71 to .80 (Morey, 2007). The PIM scale has a Cronbach's alpha ranging from .71 to .77, and a test-retest reliability ranging from .75 to .81 (Morey, 2007). In the present study, elevated scores on the INF scale, exceeding the cut-off of 74, identified participants who may not have been responding truthfully to study questions, and these participants' responses were screened to determine if they should be excluded from analysis (Morey, 2007). However, no participants were excluded on the basis of this criteria.

Procedure

Following recruitment, participants were directed to a link online in Survey Monkey to complete a study on *Hormones and Facial Emotion Detection*. Participants were told they are participating in a study investigating the relationship between social processes, such as facial emotion detection, and aspects of health, such as mood and hormones. Participants provided informed consent (see Appendix E for the full Letter of Information and Consent Form). They then completed a virtual questionnaire and cognitive task, consisting of the three sections: demographic and general background questionnaire, the FEDT and the final questionnaire, completed in that order. The FEDT was hosted on the website Pavlovia. Participant codes were used to link the data from Survey Monkey with the FEDT data. After completing the study, participants were given a debriefing form (see Appendix E). Lakehead University student were given one bonus credit towards their final grade if they were in a participating undergraduate

psychology class. Additionally, all participants were invited to follow a separate link to enter their name into a draw for a $50 prepaid Visa gift card.

Data Analysis

Hypothesis 1 was that (a) women with elevated depression scores are earlier and more accurate at detecting negative emotions compared to participants with low depression scores; and (b) this effect is strongest for sad expressions. To test hypothesis 1, two separate two-group (low vs. high depression) MANCOVAs were conducted with the following dependent variables (DVs): (a) three intensity dependent variables: Image Number at Detection for fearful, sad, and angry emotions, and (b) Percentage of Incorrect Responses for fearful, sad, and angry emotion trials. Significant results (Pillai's Trace *F*-statistic, $\alpha < .05$) were followed-up with univariate ANCOVAs. Exploratory MANCOVAs were also done to determine how the two groups (low vs. high depression scores) perform when detecting all emotions (i.e., fear, sad, angry, disgust, happy, and surprise).

Hypothesis 2 was that female participants taking OCs are later and less accurate at detecting negative emotions compared to FC participants and men. To test hypothesis 2, two separate two-group (OC users vs. FC women) MANCOVAs (with follow-up ANCOVAs) were conducted with the following DVs: (a) Image Number at Detection for fearful, sad, and angry emotions, and (b) Percentage of Incorrect Responses for fearful, sad, and angry emotions. To maximize power, these analyses were initially run with just the primary OC and FC groups. Then, to explore sex effects, they were repeated with three groups (OC users, FC women, men). Significant results (Pillai's Trace *F*-statistic, $\alpha = .05$) were followed-up with univariate ANCOVAs. Exploratory MANCOVAs were also done to determine how the two groups (OC

users vs. FC women) and three groups (OC users, FC women, men) perform when detecting all emotions (i.e., fear, sad, angry, disgust, happy, and surprise).

Hypothesis 3 was that (a) participants with a high PMDD symptoms are earlier and more accurate at detecting negative emotions compared to participants with no/minimal PMDD symptoms, and (b) especially when participants with PMDD are in the premenstrual phase. To test hypothesis 3, two separate three-group (no/minimal PMDD, mild PMDD, moderate-severe PMDD) MANCOVAs were conducted with the following DVs: (a) Image Number at Detection for fearful, sad, and angry emotions, and (b) Percentage of Incorrect Responses for fearful, sad, and angry emotions. Significant results (Pillai's Trace *F*-statistic, $\alpha = .05$) were followed-up with univariate ANCOVAs. Exploratory MANCOVAs were also done to determine how the three groups perform when detecting all emotions (i.e., fear, sad, angry, disgust, happy, and surprise).

Then, two separate between-subjects 2 x 2 MANCOVAs were conducted to examine the effects of PMDD group (no/minimal PMDD vs. mild-severe PMDD) and the premenstrual phase (weeks 2-3 vs. week 4), on (a) Image Number at Detections for fearful, sad, and angry emotions, and (b) Percentage of Incorrect Responses for fearful, sad, and angry emotions. Significant results (Pillai's Trace *F*-statistic, $\alpha = .05$) were followed-up with univariate ANCOVAs. Exploratory 2 x 2 MANCOVAs were also done to examine the effects of PMDD group (no/minimal PMDD vs. mild-severe PMDD) and the premenstrual phase (weeks 2-3 vs. week 4), on detection of all emotions (i.e., fear, sad, angry, disgust, happy, and surprise).

Age, BMI, typical alcohol consumption, hours of sleep last night, typical hours of sleep, ethnicity, education, typing skills, ADHD diagnosis, caffeine withdrawal, nicotine withdrawal, typical THC use, menstrual cycle week, and contraceptive use (where applicable), were all considered as potential covariates.

Results

Data Screening and Statistical Considerations

IBM SPSS Statistics was used to perform statistical analyses. All data entry was manually checked for accuracy. The main hypotheses were examined using MANCOVAs and ANCOVAs. For all analyses, a significance level of $< .05$ was chosen. A significance level of $< .10$ was chosen to represent nonsignificant trends. Pillai's trace criterion was used to evaluate multivariate significance. Significant MANCOVAs were followed-up with univariate ANCOVAs. The Bonferroni adjustment was used for follow-up comparisons to reduce Type I errors. All means reported are untransformed unadjusted means, unless otherwise indicated (e.g., figures represent adjusted means and their standard errors).

Missing Data

It is recommended that if less than 10% of data is missing, mean imputation may be used to replace missing data (Tabachnick et al., 2019; Tsikriktsis, 2005). Thus, mean imputation based on individual item scores was used for its conservativeness. Missing values (and maximum number of items replaced) were replaced for items comprising the DSPMS Severity and Intensity subscales (1 item each), QIDS-16 (1 item), PANAS NA and PA subscales (1 item each), BAS subscale (1 item), TAS-20 (2 items), ACES-Q (1 item), and OC Side Effect Physical and Emotional subscales (1 item each). In cases in which missing values comprised more than 10% of the data, no data were replaced, and participants were excluded from the relevant analyses. For the FEDT data, participants with more than 10% missing data (i.e., more than two invalid trials out of a total of 24 trials) were excluded from analyses. For the remaining participants, missing performance scores at the trial level were replaced by mean imputation.

Assessing Statistical Assumptions

Prior to running analyses to test the main hypotheses, the data were examined to ensure that statistical assumptions were met. The main outcome variables from the FEDT, performance scores and accuracy variables, were tested for outliers and normality. Outliers were identified by looking for z-scores larger than an absolute value of 3.29 (Tabachnick et al., 2019). The distribution of scores for all outcome variables were also examined for normality as a function of the groups utilized within each analysis. Normality was examined using the following criteria: skewness divided by the standard error of skewness <3; kurtosis divided by the standard error of kurtosis <3 (Tabachnick et al., 2019).

Among the FEDT Image Number at Detection variables, one outlier was identified within the surprise performance score. Instead, the outlier was replaced with a raw score one unit larger than the next more extreme value (Tabachnick et al., 2019). The Image Number at Detection variables all met normality tests across all groups.

The Percentage of Incorrect Responses and Percentage of Correct Responses variables had eight outliers overall. Each participant's overall performance on the relevant variable was examined for validity. In particular, happy and sad Percentage of Incorrect Response variables had a large number of outliers. Since these two emotions tend to be detected with fewer errors than other emotions, any incorrect responses made with these emotions are more likely to appear as extreme values although they represent actual performance on the test. For this reason, outliers were considered to be representative of actual variance within the sample, and were not modified. The majority of the Percentage of Incorrect Responses and Percentage of Correct Responses variables met normality criteria across all groups. However, there were some variables that exceeded skewness and kurtosis values of three. Upon visual inspection some distributions appeared visibly skewed, especially for happy and sad variables. As these

distributions are expected, as described above, no corrections were made. As a check to ensure that the outliers and non-normality were not responsible for any findings, follow-up non-parametric tests (Kruskal-Wallis analysis of ranks) were run for the Percentage of Incorrect Responses variables after the main analyses.

All variables examined as potential covariates were also scrutinized for normality and outliers. The variables had no outliers and met normality criteria, except for the ACES scale. The ACES scale had 3 outliers and did not meet normality criteria. However, the distribution for this scale is expected to be positively skewed (Centers for Disease Control and Prevention (CDC), 2010), so no corrections were made, and it was not used as a covariate

Levene's Test for Equality of Variances and Box's Test for Equivalence of Covariance Matrices were conducted with all multivariate analyses. The homogeneity assumption was met for all analyses.

Group Equivalency: Identification of Covariates and Examining Validity of Groups

To determine potential covariates for the main analyses and to identify any expected group differences, ANOVAs, *t*-tests, and chi-squares were run to identify group differences. The groups associate with each hypothesis (H1: low depression, high depression; H2: OC using women, FC women, men; H3: no PMDD, mild PMDD, moderate-severe PMDD) were examined for equivalency.

The results for the low and high depression groups are presented in Tables 3 and 4. Groups differed in negative affect over the past two weeks [F (1, 98) = 54.174, p <.001], inhibition behaviour [F (1, 96) = 9.306, p = .003], alexithymia scores [F (1, 97) = 39.326, p <.001], history of adverse childhood experiences [F (1, 97) = 8.933, p = .004], current positive affect [F (1, 99) = 10.934, p = .001], current negative affect [F (1, 99) = 28.435, p < .001],

PMDD symptoms [F (1, 99) = 42.446, p < .001], current depression diagnosis status [X^2 (1, N = 92) = 11.379, p < .001], and current other psychological disorder diagnosis status [X^2 (1, N = 92) = 7.906, p = .005]. The high depression group had higher rates of negative affect over the past two weeks, current negative affect, and current depression diagnoses, and lower current positive affect. All of these results were expected and confirmed the validity of group characteristics. Similarly, rates of depression are approximately double in women compared to men, which was observed within the sex differences of the sample (Jenkins et al., 2018). Depression is also associated with higher rates of inhibition behaviour (as measured by the BIS) (Johnson et al., 2003), alexithymia (Sagar et al., 2021), ACEs (Tsehay et al., 2020), comorbidity with other psychological disorders (Thaipisuttikul et al., 2014), and PMS symptoms (Padhy et al., 2015). As these factors may be integral to mediating the relationship between depression and FED they are considered sources of variability that should not be controlled for, so they were not used as covariates (Tabachnick et al., 2019). No necessary covariates were identified based on the depression group differences.

The comparisons of OC users, FC women, and men are presented in Tables 5 and 6. Groups differed in the number of hours of sleep the night prior to testing [F (2, 137) = 3.523, p = .032], inhibition behaviour [F (2, 138) = 13.821, p < .001], history of adverse childhood experiences [F (2, 138) = 5.763, p =.004], current positive affect [F (2, 141) = 4.606, p = .012], typing skills [X^2 (4, N = 142) = 12.349, p = .015], current depression diagnosis status [X^2 (2, N = 134) = 12.636, p = .002], and current other psychological disorder diagnosis status [X^2 (2, N = 134) = 6.431, p = .040]. Follow-up t-tests revealed that OC users slept less the night prior to testing than FC women (p = .014), but all other differences were primarily expected sex differences, and no other significant differences between OC users and non-users were found.

Table 3

Examination of Group Equivalency Between Low and High Female Depression Score Groups (t-tests): Means (SDs).

Variable	Low depression (n = 49)	High depression (n = 52)
Age (years)	23.63 (5.98)	21.98 (4.14)
BMI (kg/m^2)	25.56 (7.35)	25.03 (4.74)
Typical alcohol use score[a]	4.33 (3.21)	5.58 (3.24)
Typical hours of sleep	7.06 (0.92)	6.75 (1.31)
Hours of sleep last night	7.41 (1.29)	6.96 (1.74)
Negative affect (past 2 weeks) ***	18.29 (5.32)	27.88 (7.49)
BIS score**	18.4 (3.12)	20.22 (2.76)
BAS score	31.17 (3.63)	30.52 (3.96)
Alexithymia score ***	55.68 (7.79)	65.53 (7.83)
ACEs score **	1.54 (1.95)	2.9 (2.53)
Positive affect (now) **	23.67 (8.42)	18.69 (6.66)
Negative affect (now) ***	13.46 (4.45)	19.28 (6.29)
DSPMS score[b] ***	-0.53 (0.86)	0.51 (0.69)

Note: Variables above the dotted line were examined as potential covariates and variables below the line reflect ones expected to differ or that were of theoretical interest. [a] Variable represents alcohol consumption frequency and amount over the past 6 months; [b] Mean of the Impairment, Severity, and Frequency DSM-5-Based Screening for Premenstrual Symptoms (DSPMS) subscale z-scores. BMI = Body Mass Index; BIS = Behavioural Inhibition System; BAS = Behavioural Activation System; ACEs = Adverse Childhood Experiences.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4

Examination of Group Equivalency Between Low and High Female Depression Score Groups (Chi-Square Tests): Frequencies (%)

Variable	Low Depression (n = 49)	High Depression (n = 52)
Ethnicity		
White	33 (67.3)	39 (75.0)
Other	16 (32.7)	13 (25.0)
Highest Education		
High school	34 (69.4)	39 (75.0)
Diploma/associate degree	8 (16.3)	8 (15.4)
Undergraduate bachelor's degree	5 (10.2)	4 (7.7)
Post-graduate degree	2 (4.1)	1 (1.9)
Typing skills		
Weak	2 (4.3)	3 (5.8)
Neither weak nor strong	8 (17.0)	11 (21.2)
Strong	37 (78.7)	38 (73.1)
Diagnosed with ADHD		
No	44 (89.8)	45 (86.5)
Yes	5 (10.2)	7 (13.5)
Currently in caffeine withdrawal		
No	48 (98.0)	48 (92.3)
Yes	1 (2.0)	4 (7.70)
Currently in nicotine withdrawal		
No	49 (100.0)	49 (94.2)
Yes	0 (0.0)	3 (5.8)
Typical THC monthly use		
<1 time	37 (77.1)	35 (67.3)
1-4 times	4 (8.3)	6 (11.5)
5+ times	7 (14.6)	11 (21.2)

Variable	Low Depression (n = 49)	High Depression (n = 52)
Menstrual cycle week[a]		
Week 1	13 (28.3)	20 (40.0)
Week 2	10 (21.7)	6 (12.0)
Week 3	17 (37.0)	11 (22.0)
Week 4	6 (13.0)	13 (26.0)
Contraceptive use[a]		
OC user	14 (28.6)	16 (30.8)
Other HC user	8 (16.3)	8 (15.4)
Non HC user	27 (55.1)	28 (53.8)
Diagnosed with depression ***		
No	41 (87.2)	25 (55.6)
Yes	6 (12.8)	20 (44.4)
Diagnosed with other psychological disorder **		
No	34 (75.6)	22 (46.8)
Yes	11 (24.4)	25 (53.2)

Note: Variables above the dotted line were examined as potential covariates and variables below the line reflect ones expecting to differ or are of theoretical interest.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Past research predicts that women have higher rates of inhibition behaviour (as measured by the BIS) (Jung et al., 2022), ACEs (Jones et al., 2022), and depression and other psychological disorder diagnoses (Jenkins et al., 2018). Women are also expected to have lower positive affect (Speed et al., 2017). Additionally, OC users have less positive affect variability compared to non-users and men (Jarva & Oinonen, 2007). While men reported overall weaker typing skills, typing skills were not correlated with any of the outcome variables, and was therefore not used as a covariate. Having more sleep the night before was associated with earlier detection of surprise [surprise Image Number at Detection, $r(160) = -.221$, $p = .005$] and more errors when detecting disgust [disgust Percentage of Incorrect Responses, $r(160) = .174$, $p = .037$]. There is also other evidence that a previous night's sleep can affect FED (van der Helm et al., 2010). For these reasons, sleep was included as a covariate.

It is also noteworthy that the length of time that OC users were taking OCs correlated with angry Image Number at Detection, $r(38) = -.527$, $p < .001$, and sad Percentage of Incorrect Responses, $r(38) = .364$, $p = .025$. Shorter duration of OC use was associated with taking longer to detect angry emotions, but with making fewer errors when detecting sad emotions, while long-time users detected anger earlier but made more errors with sad emotions.

Group comparisons for the PMDD groups are presented in Tables 7 and 8. Groups differed in typical alcohol use over the past 6 months [F (2, 69) = 4.087, $p = .021$], negative affect over the past two weeks [F (2, 67) = 9.511, $p < .001$], inhibition behaviour [F (2, 67) = 5.446, $p = .006$], alexithymia [F (2, 68) = 6.886, $p = .002$], current positive affect [F (2, 69) = 4.090, $p = .021$], current negative affect [F (2, 69) = 3.588, $p = .033$], DSPMS score [F (2, 69) = 61.742, $p < .001$], nicotine withdrawal [X^2 (2, $N = 72$) = 16.457, $p < .001$], and self-reported PMS symptoms [X^2 (2, $N = 72$) = 8.576, $p = .014$]. The moderate-severe PMDD group reported

Table 5

Examination of Group Equivalency Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men (ANOVAs): Means (SDs)

Variable	OC users (n = 37)	FC women (n = 72)	Men (n = 35)
Age (years)	21.38 (3.3)	23.31 (6.09)	23.34 (6.53)
BMI (kg/m^2)	24.78 (4.73)	25.55 (6.97)	25.05 (4.99)
Typical alcohol use score[a]	5.78 (3.51)	4.60 (2.92)	5.34 (3.73)
Typical hours of sleep	6.70 (0.88)	7.00 (1.1)	7.15 (0.99)
Hours of sleep last night *	6.73 (1.56)[y]	7.49 (1.47)[y]	6.88 (1.77)
Negative affect (past 2 weeks)	22.95 (7.34)	23.15 (8.2)	19.4 (7.89)
BIS score ***	19.30 (3.01)	19.74 (2.75)	16.62 (3.07)[x]
BAS score	30.94 (3.38)	31.29 (4.1)	31.82 (3.75)
Alexithymia score	60.18 (8.46)	61.35 (8.74)	57.65 (7.52)
ACEs score **	1.80 (2.01)	2.21 (2.36)[y]	0.78 (1.21)[y]
Positive affect now *	20.43 (8.4)	20.5 (8.09)	25.23 (7.61)[x]
Negative affect now	15.93 (6.15)	16.65 (6.32)	14.37 (5.88)
DSPMS score[b]	35.93 (18.16)	39.32 (13.69)	

Note: Variables above the dotted line were examined as potential covariates and variables below the line reflect ones expected to differ or that were of theoretical interest. [a] Variable represents alcohol consumption frequency and amount over the past 6 months; [b] Mean of the Impairment, Severity, and Frequency DSM-5-Based Screening for Premenstrual Symptoms (DSPMS) subscale z-scores. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups. BMI = Body Mass Index; BIS = Behavioural Inhibition System; BAS = Behavioural Activation System; ACEs = Adverse Childhood Experiences. * $p < .05$. ** $p < .01$. *** $p < .001$

Table 6

Examination of Group Equivalency Between Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men (Chi-Square Tests): Frequencies (%)

Variable	OC users (n = 37)	FC women (n = 72)	Men (n = 35)
Ethnicity			
White	29 (78.4)	48 (66.7)	25 (71.4)
Other	8 (21.6)	24 (33.3)	10 (28.6)
Highest Education			
High school	27 (73.0)	52 (72.2)	22 (62.9)
Diploma/associate degree	5 (13.5)	10 (13.9)	6 (17.1)
Undergraduate bachelor's degree	4 (10.8)	7 (9.7)	5 (14.3)
Post-graduate degree	1 (2.7)	3 (4.2)	2 (5.7)
Typing skills *			
Weak	0 (0.0)	6 (8.3)	2 (5.7)
Neither weak nor strong	12 (34.3)	14 (19.4)[y]	17 (48.6)[y]
Strong	23 (65.7)	52 (72.2)[y]	16 (45.7)[y]
Diagnosed with ADHD			
No	33 (89.2)	65 (90.3)	33 (94.3)
Yes	4 (10.8)	7 (9.7)	2 (5.7)
Currently in caffeine withdrawal			
No	35 (94.6)	69 (95.8)	32 (91.4)
Yes	2 (5.4)	3 (4.2)	3 (8.6)
Currently in nicotine withdrawal			
No	37 (100.0)	70 (97.2)	34 (97.1)
Yes	0 (0.00)	2 (2.80)	1 (2.90)
Typical THC monthly use			
<1 time	31 (83.8)	49 (69.0)	28 (80.0)
1-4 times	2 (5.4)	9 (12.7)	6 (17.1)
5+ times	4 (10.8)	13 (18.3)	1 (02.9)

Variable	OC users (n = 37)	FC women (n = 72)	Men (n = 35)
Menstrual cycle week [a]			
Week 1	14 (37.8)	20 (28.2)	
Week 2	5 (13.5)	15 (21.1)	
Week 3	10 (27.0)	20 (28.2)	
Week 4	8 (21.6)	16 (22.5)	
Current depression diagnosis **			
No	23 (69.7)	48 (72.7)	35 (100.0)[x]
Yes	10 (30.3)	18 (27.3)	0 (0.00)[x]
Current other psychological disorder diagnosis *			
No	21 (61.8)[y]	46 (69.7)	30 (88.2)[y]
Yes	13 (38.2)[y]	20 (30.3)	4 (11.8)[y]

Note: Variables above the dotted line were examined as potential covariates and variables below the line reflect ones expected to differ or of theoretical interest. [a] Reported for female participants only. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

* $p < .05$. ** $p < .01$. *** $p < .001$.

more PMS symptoms than the combined no/minimal and mild PMDD groups (p = .028), while the no/minimal PMDD group reported less PMS symptoms than the combined mild and moderate-severe groups (p = .004), confirming the validity of groups. PMDD is associated with higher negative affect, less positive affect, and inhibition behaviour (Petersen et al., 2016), and higher rates of alexithymia (De Berardis et al., 2005), so these group differences were not unexpected. Given group differences in typical alcohol use and nicotine withdrawal, their correlations with the outcome variables were examined. None of the outcome variables were associated with nicotine withdrawal, therefore it was not used as a covariate. However, typical alcohol use was correlated with Image Number at Detection for surprise emotions within the sample of FC women, $r(72) = -.241$, $p = .041$. While people with higher typical alcohol use had earlier surprise detection, there is other evidence that alcohol use can cause impairments to FED (Donadon & Osório, 2014). For these reasons, typical alcohol use was included as a covariate in the analyses pertaining to the PMDD symptom groups.

Based on the group equivalency analyses outlined above, potential covariates included hours of sleep last night (i.e., the night before testing), and typical alcohol consumption. Due to the immediate effects of sleep the night before testing, hours of sleep last night was used as a covariate for all analyses pertaining to hypotheses 1, 2, and 3. Typical alcohol use was also used as a covariate for the analyses pertaining to hypothesis 3 due to group differences in that sample.

Table 7

Examination of Group Equivalency Between No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (ANOVAs): Means (SDs)

Variable	No/Minimal PMDD (n = 30)	Mild PMDD (n = 34)	Moderate-Severe PMDD (n = 8)
Age (years)	24.23 (6.55)	22.91 (5.83)	21.50 (5.5)
BMI (kg/m^2)	24.75 (8.16)	25.97 (5.74)	27.01 (7.82)
Typical alcohol use score[a] *	3.50 (2.45)[y]	5.26 (3.04)[y]	5.88 (3.04)
Typical hours of sleep	7.23 (1.10)	6.88 (1.05)	6.63 (1.19)
Hours of sleep last night	7.70 (1.32)	7.45 (1.68)	6.88 (0.99)
Negative affect (past 2 weeks) ***	18.94 (7.65)[x]	25.18 (6.77)	30.00 (8.50)
BIS score **	18.63 (3.05)[x]	20.34 (2.15)	21.5 (2.33)
BAS score	31.49 (4.38)	31.26 (3.93)	30.63 (4.14)
Alexithymia score **	57.26 (8.16)[x]	63.91 (8.12)	66.13 (7.57)
ACEs score	1.90 (2.52)	2.45 (2.41)	2.38 (1.51)
Positive affect (now) *	23.16 (9.00)[y]	19.45 (7.29)	15.00 (2.78)[y]
Negative affect (now) *	14.57 (5.98)[y]	17.64 (6.22)	20.25 (5.99)[y]
DSPMS score[b] ***	-0.89 (0.62)[y]	0.46 (0.58)[y]	1.21 (0.44)[y]

Note: Variables above the dotted line were examined as potential covariates and variables below the line reflect ones expected to differ or that were of theoretical interest. [a] Variable represents alcohol consumption frequency and amount over the past 6 months; [b] Mean of the Impairment, Severity, and Frequency DSM-5-Based Screening for Premenstrual Symptoms (DSPMS) subscale z-scores. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups. BMI = Body Mass Index; BIS = Behavioural Inhibition System; BAS = Behavioural Activation System; ACEs = Adverse Childhood Experiences. * $p < .05$. ** $p < .01$. *** $p < .001$

Table 8

Examination of Group Equivalency Between No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (Chi-Square Tests): Frequencies (%)

Variable	No/Minimal PMDD (n = 30)	Mild PMDD (n = 34)	Moderate-Severe PMDD (n = 8)
Ethnicity			
White	15 (50.0)	26 (76.5)	7 (87.5)
Other	15 (50.0)	8 (23.5)	1 (12.5)
Highest Education			
High school	19 (63.3)	25 (73.5)	8 (100.0)
Diploma/associate degree	6 (20.0)	4 (11.8)	0 (0.0)
Undergraduate bachelor's degree	3 (10.0)	4 (11.8)	0 (0.0)
Post-graduate degree	2 (6.7)	1 (2.9)	0 (0.0)
Typing skills			
Weak	3 (10.0)	3 (08.8)	0 (0.0)
Neither weak nor strong	4 (13.3)	9 (26.5)	1 (12.5)
Strong	23 (76.7)	22 (64.7)	7 (87.5)
Diagnosed with ADHD			
No	27 (90.0)	32 (94.1)	6 (75.0)
Yes	3 (10.0)	2 (5.9)	2 (25.0)
Currently in caffeine withdrawal			
No	2 (6.7)	1 (2.9)	0 (0.0)
Yes	28 (93.3)	33 (97.1)	8 (100.0)
Currently in nicotine withdrawal***			
No	30 (100.0)	34 (100.0)	6 (75.0)[x]
Yes	0 (0.0)	0 (0.0)	2 (25.0)[x]

Variable	No/Minimal PMDD (n = 30)	Mild PMDD (n = 34)	Moderate-Severe PMDD (n = 8)
Typical THC monthly use			
<1 time	22 (75.9)	24 (70.6)	3 (37.5)
1-4 times	2 (6.9)	5 (14.7)	2 (25.0)
5+ times	5 (17.2)	5 (14.7)	3 (37.5)
Menstrual cycle week			
Week 1	8 (26.7)	7 (21.2)	5 (62.5)
Week 2	7 (23.3)	8 (24.2)	0 (0.0)
Week 3	10 (33.3)	10 (30.3)	0 (0.0)
Week 4	5 (16.7)	8 (24.2)	3 (37.5)
Current depression diagnosis			
No	23 (85.2)	21 (65.6)	4 (57.1)
Yes	4 (14.8)	11 (34.4)	3 (42.9)
Current other psychological disorder diagnosis			
No	21 (75.0)	21 (67.7)	4 (57.1)
Yes	7 (25.0)	10 (32.3)	3 (42.9)
PMS symptoms[a] *			
Never – Mild	19 (67.9)[y]	13 (39.4)	1 (14.3)[y]
Mild – Severe	9 (32.1)[y]	20 (60.6)	6 (85.7)[y]

Note: Variables above the dotted line were examined as potential covariates and variables below the line reflect ones expecting to differ or of theoretical interest. [a] One-item retrospective self-report measure of PMS symptoms. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Descriptive Supplementary Data

Descriptive data for supplementary variables that were not used for the thesis analyses can be found in Appendix F. The unadjusted means and *SD*s of raw image number at detection for primary and secondary trials (i.e., error trials are excluded instead of being assigned an intensity level of 17 as was done for the Image Number at Detection score used within analyses) per groups are presented. Additionally, figures outlining the number of primary, secondary, error, and invalid trials per emotion per group are also included in Appendix F. These figures show the number of trials in which the correct emotion was detected directly, in which incorrect responses were provided prior to correctly detecting, and in which the correct emotion was never detected.

Main Analyses

Hypothesis 1

Hypothesis 1 was that female participants with high depression symptoms are earlier and more accurate at detecting negative (sad, angry, and fearful) emotions compared to female participants with low depression symptoms, with the strongest effect for sad expressions. Hours of sleep the night before testing was included as a covariate for all MANCOVAs.

Image Number at Detection MANCOVAs. Table 9 contains the unadjusted means and *SD*s of all Image Number at Detection scores. Visual examination of the Image Number at Detection means revealed that they were all in the same direction such that the high depression group had lower Image Number at Detection (i.e., earlier detection) than the low depression group. Despite this, the MANCOVA testing Image Number at Detection for negative (fear, sad, angry) emotions was non-significant, $F(3, 101) = 0.651$, $p = .584$, $\eta^2 = .020$, suggesting that the women with low vs. high depression symptoms did not differ in their intensity at detection for

negative emotions. The MANCOVA testing Image Number at Detection for all six emotions found a significant multivariate effect, $F(6, 101) = 2.428$, $p = .032$, $\eta^2 = .135$, suggesting that women with high depression scores detected overall emotions at lower intensities (earlier) than women with low depression scores.

Table 10 contains the results of follow-up ANCOVAs. Figure 4 displays the group differences in Image Number at Detection, and all significant follow-up ANCOVA results. Follow-up ANCOVAs revealed that the high depression group had a significantly lower Image Number at Detection than the low depression group on surprise emotions, with a medium-large effect size ($p = <.001$, $\eta^2 = .108$). This indicates that the high depression group detected surprise earlier (i.e., at lower intensity).

Percentage of Incorrect Responses MANCOVA. Table 11 contains the unadjusted means and *SD*s of Percentage of Incorrect Responses scores. The MANCOVA testing group differences in Percentage of Incorrect Responses for negative emotions, was non-significant, $F(3, 101) = 0.516$, $p = .673$, $\eta^2 = .016$. The MANCOVA testing Percentage of Incorrect Responses across all six emotions, was also non-significant, $F(6, 101) = 1.053$, $p = .396$, $\eta^2 = .064$. Non-parametric Kruskal-Wallis tests on the individual emotions also yielded no significant group differences. These findings suggest that the women with low vs. high depression symptoms did not differ in their accuracy of detection for negative emotions, or across all emotions.

Depression Diagnosis IV MANCOVA. Given that previous research has fairly consistently indicated that depression is associated with lower accuracy of detection, to ensure that results are robust, the MANCOVAs were repeated with self-reported current depression diagnosis (no vs. yes) as the IV. The means and *SD*s for the MANCOVAs, as well as the follow-

Table 9

Hypothesis 1: Unadjusted Means (SDs) of Image Number at Detection per Emotion for Low and High Female Depression Score Groups

Emotion	Mean (*SD*) of Image Number at Detection	
	Low depression (*n* = 49)	High depression (*n* = 52)
Fear	10.95 (2.27)	10.42 (2.42)
Sad	9.25 (1.80)	8.98 (1.76)
Angry	10.23 (2.03)	10.02 (2.12)
Disgust	11.45 (2.15)	10.79 (2.19)
Happy	6.98 (1.89)	6.77 (1.78)
Surprise ***	7.69 (1.71)	6.62 (1.67)

Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. Variables used in the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The means are unadjusted for covariates, but all analyses controlled for hours of sleep last night.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 10

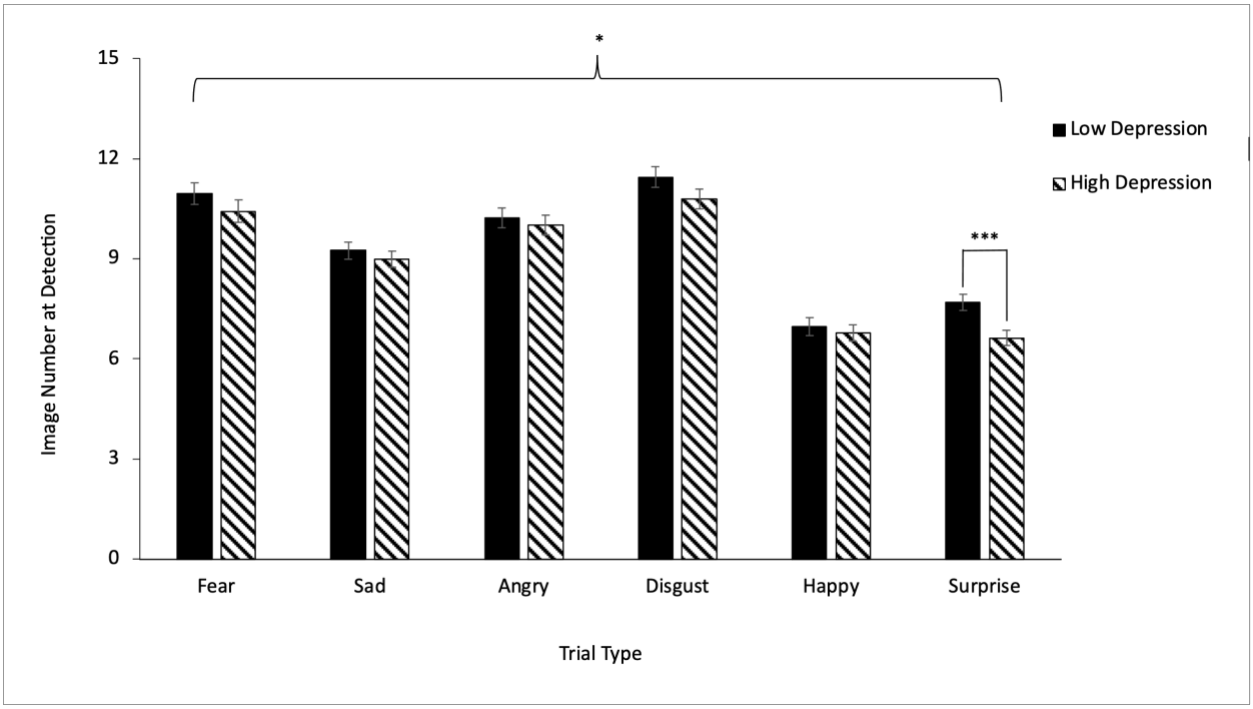
Hypothesis 1: ANCOVA Results for Image Number at Detection per Emotion for Low vs. High Female Depression Score Groups

Emotion	*df*	*F*	*p*	η^2
Fear	2, 98	1.866	.175	.019
Sad	2, 98	0.763	.384	.008
Angry	2, 98	0.344	.559	.003
Disgust	2, 98	2.23	.139	.022
Happy	2, 98	0.497	.482	.005
Surprise	2, 98	11.832	<.001***	.108

Note. Results of the follow-up univariate ANCOVAs testing group differences in Image Number at Detection for individual emotions. All analyses controlled for hours of sleep last night.
$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 4

Image Number at Detection per Emotion for Low and High Female Depression Score Groups



Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. All analyses controlled for hours of sleep last night. There was a significant multivariate effect for depression group indicating that, across all the emotions, women with high depression symptoms responded correctly earlier (i.e., lower Image Number at Detection), $F(6, 101) = 2.428$, $p = .032$, $\eta^2 = .135$. The high depression group also responded earlier to surprise than the low depression group, ($p < .001$, $\eta^2 = .108$). Error bars represent standard error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 11

Hypothesis 1: Unadjusted Means (SDs) for Percentage of Incorrect Responses per Emotion for Low and High Female Depression Score Groups

Emotion	Mean (*SD*) of Percentage of Incorrect Responses	
	Low depression (*n* = 49)	High depression (*n* = 52)
Fear	14.91 (11.72)	15.49 (11.8)
Sad	2.46 (5.82)	2.79 (5.00)
Angry	5.11 (8.09)	7.33 (11.32)
Disgust	22.98 (13.86)	22.23 (17.26)
Happy	0.94 (2.93)	0.93 (3.39)
Surprise [t]	6.12 (7.94)	3.75 (6.23)

Note. Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. Variables pertaining to the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The data here are unadjusted for covariates, but all analyses controlled for hours of sleep last night.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

up ANCOVAs are presented in Appendix G. The MANCOVA assessing Image Number at Detection was not significant for negative emotions, F(3, 112) = 0.554, p = .646, η^2 = .015, or across all the emotions, F(6, 109) = 0.625, p = .710, η^2 = .033. However, visual examination of the means revealed that they were in the same direction as the initial analysis (i.e., women with depression had lower Image Numbers at Detection for all emotions). The MANCOVA assessing Percentage of Incorrect Responses was also not significant for negative emotions, F(3, 112) = 0.565, p = .639, η^2 = .015, or across all emotions, F(6, 109) = 0.418, p = .866, η^2 = .022. None of the follow-up ANCOVAS for either MANCOVA were significant. Thus, no new significant results emerged, and the direction of the findings was consistent with both measures of depression. This provides validity for the $QIDS_{16}$.

Hypothesis 2

Hypothesis 2 was that women taking OCs are slower and less accurate at detecting negative emotions compared to FC participants and men. All MANCOVAs and ANCOVAs included hours of sleep last night as a covariate.

Image Number at Detection MANCOVAs. Table 12 contains the unadjusted means and *SD*s of all Image Number at Detection scores. Visual examination of the means show that OC users took longer to detect all emotions (i.e., had higher Image Number at Detection) than FC women. The two-group (OC users, FC women) MANCOVA testing Image Number at Detection for negative emotions was non-significant, $F(3, 108) = 0.629$, $p = .598$, $\eta^2 = .018$. The MANCOVA examining differences between the three groups (OC users, FC women, men) on negative emotions also did not find a multivariate effect, $F(6, 140) = 0.773$, $p = .587$, $\eta^2 = .017$. However, the two-group (OC users, FC women) MANCOVA testing group differences in Image Number at Detection across all the emotions was significant, $F(6, 108) = 2.749$, $p = .016$, η^2 =

.142, with OC users taking longer to detect emotions than FC women. The three group (OC users, FC women, men) MANCOVA testing Image Number at Detection across all six emotions was not significant, $F(12, 140) = 1.758$, $p = .055$, $\eta^2 = .074$, although a non-significant trend did emerge. These findings suggest that OC users and FC women differed in their intensity of detection across all emotions, with OC users correctly detecting emotions later (i.e., at higher intensity).

Table 13 contains all follow-up ANCOVA results. Figure 5 displays the group differences in Image Number at Detection, and reflects all significant follow-up ANCOVA results. The two-group univariate ANCOVAs revealed that OC users had a significantly higher Image Number at Detection than FC women with happy emotions, with a medium effect size ($p = .002$, $\eta^2 = .085$), and with disgust emotions ($p = .031$, $\eta^2 = .044$). The three-group ANCOVAs revealed that OC users, FC women, and men also differed in their Image Number at Detection for happy emotions *(*$p = .007$, $\eta^2 = .070$), and disgust emotions ($p = .039$, $\eta^2 = .047$). Pairwise comparisons simply reflected the above findings. None of the pairwise comparisons between men and the other groups were significant.

Percentage of Incorrect Responses. Table 14 contains the unadjusted means and *SD*s of all Percentage of Incorrect Responses scores. The two-group MANCOVA testing Percentage of Incorrect Responses for negative emotions was non-significant, $F(3, 108) = 0.448$, $p = .719$, $\eta^2 = .013$. Similarly, the three-group MANCOVA for negative emotions was not significant, $F(6, 140) = 0.816$, $p = .558$, $\eta^2 = .018$. The two-group MANCOVA testing Percentage of Incorrect Responses across all emotions, was not significant, $F(6, 108) = 1.222$, $p = .301$, $\eta^2 = .068$. The three-group MANCOVA for all six emotions was also not significant, $F(12, 140) = 1.288$, $p =$

Table 12

Hypothesis 2: Unadjusted Means (SDs) of Image Number at Detection per Emotion for Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men

Emotion	Mean (*SD*) of Image Number at Detection		
	OC users (*n* = 37)	FC women (*n* = 72)	Men (*n* = 35)
Fear	10.67 (2.40)	10.28 (2.36)	11.28 (2.50)
Sad [t]	9.31 (1.56)	8.83 (1.81)	9.14 (1.52)
Angry	10.42 (2.14)	9.87 (2.04)	10.31 (2.11)
Disgust *	11.60 (2.39)[y]	10.64 (2.02)[y]	11.44 (2.49)
Happy**	7.69 (2.11)[y]	6.40 (1.68)[y]	6.81 (1.95)
Surprise	7.05 (1.76)	6.91 (1.80)	7.32 (1.58)

Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. Variables used in the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The means here are unadjusted for covariates, but all analyses controlled for hours of sleep last night. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.
[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 13

Hypothesis 2: ANCOVA Results for Image Number at Detections per Emotion for Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men

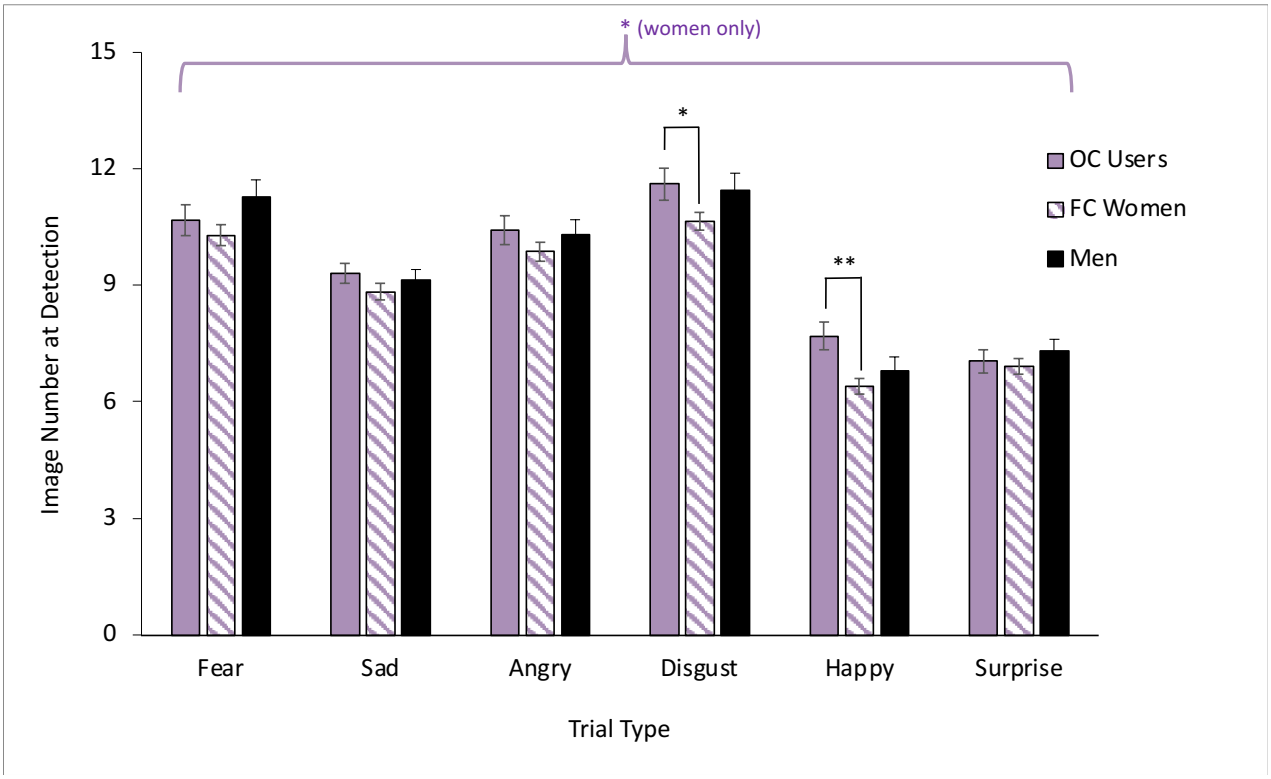
Image Number at Detection								
	2 groups (OC users vs. FC women)				3 groups (OC users vs. FC women vs. Men)			
Emotion	*df*	*F*	*p*	η^2	*df*	*F*	*p*	η^2
Fear	2, 105	0.245	.622	.002	3, 136	1.280	.281	.018
Sad	2, 105	1.189	.278	.011	3, 136	0.778	.461	.011
Angry	2, 105	1.087	.299	.010	3, 136	0.675	.511	.010
Disgust	2, 105	4.791	.031*	.044	3, 136	3.328	.039*	.047
Happy	2, 105	9.806	.002**	.085	3, 136	5.095	.007**	.070
Surprise	2, 105	0.010	.920	.000	3, 136	0.220	.803	.003

Note. Results of the follow-up univariate ANCOVAs testing group differences in Image Number at Detection for individual emotions. The ANCOVA results for the 2-group (OC users vs. FC women) analyses are presented on the left, and for the 3-group analyses are presented on the right. All analyses controlled for hours of sleep last night.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 5

Image Number at Detection per Emotion for Oral Contraceptive (OC) Users, Free Cycling (FC) Women, and Men



Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. All analyses controlled for hours of sleep last night. There was a significant multivariate effect for group when comparing OC users and FC women. OC users correctly detected emotions later (i.e., higher Image Number at Detection) across all the emotions, $F(6, 108) = 2.749$, $p = .016$, $\eta^2 = .142$. OC users detected emotions later/slower than FC women when viewing happy emotions ($p = .002$, $\eta^2 = .085$), and disgust emotions ($p = .031$, $\eta^2 = .044$). Error bars represent standard error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

.225, $\eta^2 = .055$. Non-parametric Kruskal-Wallis tests also yielded no significant results. This suggests that the three groups (OC users, FC women, and men) did not differ in their accuracy of detection across negative emotions or all emotions.

OC Androgenicity as the IV. Given that some past studies have identified differences in detection based on OC androgenicity (Gurvich et al., 2020; Menting-Henry et al., 2022), OC type was examined in the present study. To explore the effect of OC type, specifically the difference between androgenic and anti-androgenic formulations, the significant two-group (OC users, FC women) MANCOVA testing group differences in Image Number at Detection across all the emotions was repeated with three groups (Androgenic OC users, Anti-androgenic OC users, FC women).

Table 15 contains the unadjusted means and *SD*s of scores for the three groups. The MANCOVA testing Image Number at Detection across all six emotions presented a nonsignificant trend, $F(12, 108) = 1.713$, $p = .066$, $\eta^2 = .094$. The results of follow-up univariate ANCOVAs are presented in Table 16. The three groups differed in their Image Number at Detection for happy emotions. Pairwise comparisons indicated that androgenic OC users took longer than FC women to detect happy emotions ($p = .004$). This suggests that OC formulation may impact detection of happy emotions.

Hypothesis 3

Hypothesis 3 was that participants with a provisional PMDD diagnosis are earlier and more accurate at detecting negative emotions compared to participants with no PMDD diagnosis, and that this effect will be strongest for participants in the premenstrual period. Hours of sleep the night before testing and typical alcohol use were included as covariates for all MANCOVAs.

Table 14

Hypothesis 2: Unadjusted Means and SDs of Percentage of Incorrect Responses per Emotion for Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men

	Mean (*SD*) of Percentage of Incorrect Responses		
Emotion	OC users (*n* = 37)	FC women (*n* = 72)	Men (*n* = 35)
Fear	13.39 (10.08)	15.20 (12.11)	19.19 (12.71)
Sad	2.36 (4.64)	2.79 (5.07)	2.30 (4.86)
Angry	7.82 (10.28)	6.40 (9.90)	6.84 (8.24)
Disgust	24.02 (17.63)	21.03 (14.03)	25.83 (16.99)
Happy	1.71 (4.96)	0.79 (3.15)	0.76 (3.11)
Surprise	3.38 (6.07)	4.81 (7.45)	7.21 (8.93)

Note. Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. Variables pertaining to the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 15

Unadjusted Means (SDs) of Image Number at Detection per Emotion for Androgenic and Anti-androgenic Oral Contraceptive (OC) Users and Free-cycling (FC) Women

Emotion	Mean (*SD*) of Image Number at Detection		
	Androgenic OC users (*n* = 24)	Anti-androgenic OC users (*n* = 12)	FC women (*n* = 72)
Fear	10.69 (2.63)	10.53 (2.05)	10.28 (2.36)
Sad	9.31 (1.53)	9.19 (1.67)	8.83 (1.81)
Angry	10.37 (1.75)	10.95 (2.47)	9.87 (2.04)
Disgust	11.21 (2.27)	12.13 (2.54)	10.64 (2.02)
Happy**	7.94 (2.14)[y]	6.96 (1.91)	6.40 (1.68)[y]
Surprise	6.94 (2.00)	7.22 (1.28)	6.91 (1.80)

Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. Significance values reflect the results of follow-up ANCOVAs. The means are unadjusted for covariates, but all analyses controlled for hours of sleep last night. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 16

ANCOVA Results for Image Number at Detection per Emotion for Androgenic and Anti-androgenic Oral Contraceptive (OC) Users and Free-cycling (FC) Women

Emotion	*df*	*F*	*p*	η^2
Fear	2, 108	0.109	.897	.002
Sad	2, 108	0.451	.639	.009
Angry	2, 108	1.281	.282	.024
Disgust	2, 108	2.675	.074	.049
Happy	2, 108	5.474	.006**	.096
Surprise	2, 108	0.059	.943	.001

Note. Results of the follow-up univariate ANCOVAs testing group differences in Image Number at Detection for individual emotions. All analyses controlled for hours of sleep last night.
$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Image Number at Detection MANCOVAs. Table 17 contains the unadjusted means and *SD*s of Image Number at Detection scores. The three-group (No/Minimal PMDD, Mild PMDD, Moderate-Severe PMDD) MANCOVA comparing groups on Image Number at Detection for negative emotions was non-significant, $F(6, 71) = 1.412$, $p = .215$, $\eta^2 = .061$. Similarly, the three-group MANCOVA testing across all the emotions was non-significant, $F(12, 71) = 1.740$, $p = .066$, $\eta^2 = .144$, although a non-significant trend did present. These findings suggest that PMDD groups do not differ in their intensity at detection across negative emotions. However, the non-significant trend suggests that the PMDD groups may differ in their intensity at detection across all emotions. Follow-up ANCOVAs for intensity at detection across all emotions were performed given that there was a medium-large effect size.

Table 18 contains all follow-up ANCOVA results. ANCOVAs revealed that PMDD groups differed in their Image Number at Detection for disgust emotions ($p = .024$, $\eta^2 = .107$), and a non-significant trend emerged for sad emotions ($p = .081$, $\eta^2 = .073$). Pairwise comparisons determined that the moderate-severe PMDD group detected disgust earlier than the mild PMDD group ($p = .038$). Figure 6a displays the group differences in intensity for disgust.

Percentage of Incorrect Responses MANCOVAs. Table 19 contains the unadjusted means and *SD*s of Percentage of Incorrect Responses scores. The three-group MANCOVA testing Percentage of Incorrect Responses for negative emotions was non-significant $F(6, 71) = 1.082$, $p = .376$, $\eta^2 = .048$. The three-group MANCOVA testing Percentage of Incorrect Responses across all emotions, was also not significant, $F(12, 71) = 1.376$, $p = .186$, $\eta^2 = .117$. Given that there was a medium-large effect size, follow-up ANCOVAs for intensity at detection for each of the emotions was performed.

Table 17

Hypothesis 3: Unadjusted Means and SDs of Image Number at Detections per Emotion for No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups

Emotion	Mean (*SD*) of Image Number at Detection		
	No/Minimal PMDD (*n* = 30)	Mild PMDD (*n* = 34)	Moderate-Severe PMDD (*n* = 8)
Fear	10.67 (2.36)	9.89 (2.41)	10.53 (2.2)
Sad [t]	9.17 (2.04)	8.65 (1.68)	8.31 (1.28)
Angry	10.06 (1.89)	9.54 (2.21)	10.53 (1.75)
Disgust *	10.21 (1.75)	11.32 (2.13)[y]	9.38 (1.59)[y]
Happy	6.38 (1.64)	6.34 (1.81)	6.78 (1.4)
Surprise	7.16 (1.73)	6.66 (1.98)	7.03 (1.24)

Note. Lower scores indicate earlier (i.e., lower intensity of emotion) detection. Variables used in the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The means are unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol consumption. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 18

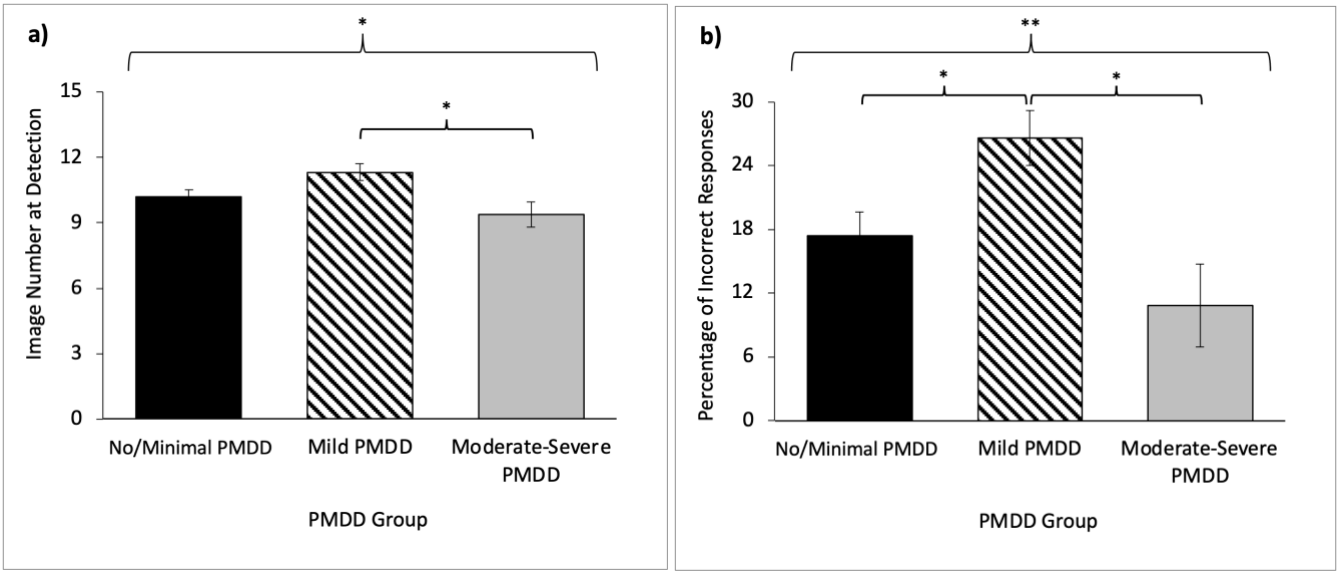
Hypothesis 3: ANCOVA Results for Image Number at Detections per Emotion for No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups

	Image Number at Detection			
Emotion	*df*	*F*	*p*	η^2
Fear	2, 66	0.869	.424	.026
Sad	2, 66	2.606	.081[t]	.073
Angry	2, 66	0.858	.429	.025
Disgust	2, 66	3.969	.024*	.107
Happy	2, 66	0.229	.796	.007
Surprise	2, 66	0.648	.526	.019

Note. Results of the follow-up univariate ANCOVAs testing group differences in Image Number at Detection for individual emotions. All analyses controlled for hours of sleep last night.
[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 6

Disgust Detection: (a) Image Number at Detection and (b) Percentage of Incorrect Responses as a function of Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (No/Minimal, Mild, and Moderate-Severe)



Note. Women with moderate-severe provisional PMDD detected disgust earlier and more accurately. a) Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. There was a univariate effect for group in Image Number at Detection for disgust, $F(12, 71) = 3.969$, $p = .024$, $\eta^2 = .107$. The moderate-severe PMDD group detected disgust emotions earlier than the mild PMDD group ($p = .038$). b) Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. There was a univariate effect for group in Percentage of Incorrect Responses for disgust, $F(12, 71) = 5.971$, $p = .004$, $\eta^2 = .153$. The no/minimal PMDD group ($p = .043$), and the moderate-severe PMDD group ($p = .014$), made fewer incorrect responses than the mild PMDD group. All analyses controlled for hours of sleep last night and typical alcohol use. Error bars represent standard error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 19

Hypothesis 3: Unadjusted Means and SDs of Percentage of Incorrect Responses per Emotion for No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD)Groups

Emotion	Mean (*SD*) of Percentage of Incorrect Responses		
	No/Minimal PMDD (*n* = 30)	Mild PMDD (*n* = 34)	Moderate-Severe PMDD (*n* = 8)
Fear	16.78 (12.98)	14.4 (11.52)	12.71 (11.98)
Sad [t]	1.59 (3.66)	4.36 (6.15)	0.63 (1.77)
Angry	6.11 (8.83)	6.76 (11.11)	5.97 (9.35)
Disgust *	17.44 (11.80)	26.59 (14.39)[x]	10.83 (10.98)
Happy	1.33 (4.68)	0.34 (0.99)	0.63 (1.77)
Surprise	4.94 (7.85)	5.74 (7.70)	0.42 (1.18)

Note. Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. Variables pertaining to the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 20 contains all follow-up ANCOVA results. The ANCOVA revealed that PMDD groups differed in their Percentage of Incorrect Responses for disgust emotions (p = .004, η^2 = .153), and a non-significant trend emerged for sad emotions (p = .094, η^2 = .069). Pairwise comparisons determined that the mild PMDD group made more incorrect responses on disgust than the no/minimal PMDD group (p = .043), and the moderate-severe PMDD group (p = .014). Non-parametric Kruskal-Wallis tests also found that the groups differed in their accuracy for disgust (p = .002. Figure 6b displays the group differences in accuracy for disgust.

Menstrual Cycle Phase as a Covariate. To ensure that the above effects were not due to potential group differences in cycle phase at the time of testing, any significant univariate analyses were re-run with cycle phase (follicular [weeks 1-2] vs. luteal [weeks 3-4]) as an additional covariate. Within the first set of ANCOVAs testing Image Number at Detection for all emotions, PMDD groups remained significantly different in Image Number at Detection for disgust emotions, $F(2, 64) = 3.634$, p = .032, η^2 = .102, but the non-significant trend disappeared for sad emotions (p = .104, η^2 = .068). Within the second set of ANCOVAs testing Percentage of Incorrect Responses for all emotions, the group differences between PMDD groups in their Percentage of Incorrect Responses for disgust emotions persisted, $F(2, 64) = 6.287$, p = .003, η^2 = .164, and the non-significant trend for sad emotions also persisted (p = .079, η^2 = .076). Controlling for cycle phase did not change the outcomes, suggesting that the findings, especially for disgust emotions, are due to group characteristics and not caused by cycle phase.

No-PMDD vs. PMDD Analyses. Given the small sample size in the moderate-severe PMDD group (n = 8), the ANCOVAS for Image Number at Detection and Percentage of Incorrect Responses for disgust and sad emotions were re-run using two groups: no-PMDD

Table 20

Hypothesis 3: ANCOVA Results for Percentage of Incorrect Responses per Emotion for No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups

Percentage of Incorrect Responses				
Emotion	*df*	*F*	*p*	η^2
Fear	2, 66	0.772	.466	.023
Sad	2, 66	2.446	.094 [t]	.069
Angry	2, 66	0.124	.883	.004
Disgust	2, 66	5.971	.004*	.153
Happy	2, 66	0.545	.582	.016
Surprise	2, 66	1.445	.243	.042

Note. Results of the follow-up univariate ANCOVAs testing group differences in Percentage of Incorrect Responses for individual emotions. All analyses controlled for hours of sleep last night and typical alcohol consumption.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

(comprising the no/minimal PMDD group) and PMDD (comprising the mild PMDD and moderate-severe PMDD groups). Table 21 contains the unadjusted means and *SD*s of Image Number at Detection and Percentage of Incorrect Responses, and the ANCOVA results. Image Number at Detection for disgust between the no-PMDD and the PMDD groups was no longer significant. However, a nonsignificant trend suggested the PMDD group had a higher Percentage of Incorrect Responses for disgust emotions, and were therefore less accurate, than the no PMDD group, (p = .089, η^2 = .043). Also, the PMDD group detected sad emotions earlier than the no-PMDD group (p = .032, η^2 = .067). Percentage of Incorrect Responses for sad did not differ between the groups. This suggests sad emotions are detected earlier within the PMDD group than those with no/minimal symptoms. Accuracy for disgust may be worse within the PMDD group.

Image Number at Detection 2 x 2 MANCOVAs. Table 22 contains the unadjusted means and *SD*s of Image Number at Detection scores pertaining to the 2 between (PMDD group) x 2 between (Cycle Phase) MANCOVAs. Group by phase interaction effects were examined to determine how women with PMDD performed during the premenstrual phase. For the 2 x 2 MANCOVA testing Image Number at Detection for negative emotions, there was no significant Group x Phase effect, $F(3, 42) = 0.517$, p = .673, η^2 = .036. For the 2 x 2 MANCOVA testing Image Number at Detection for all emotions, there was also no significant interaction effect, $F(6, 39) = 0.692$, p = .657, η^2 = .096. These findings suggest that PMDD groups did not differ in their Image Number at Detection (i.e., intensity) for negative emotions or across all emotions based on where they are in their cycle.

Table 21

Unadjusted Means, SDs and ANCOVA Results for Image Number at Detection and Percentage of Incorrect Responses per Emotion for the Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (No/Yes)

Emotion	No PMDD (n = 30)	PMDD (n = 42)	ANCOVA Results			
	Image Number at Detection		*df*	*F*	*p*	η^2
Sad *	9.17 (2.04)	8.50 (1.53)	1, 67	4.803	.032	.067
Disgust	10.21 (1.75)	10.89 (2.17)	1, 67	1.254	.267	.018
	Percentage of Incorrect Responses		*df*	*F*	*p*	η^2
Sad	1.59 (3.66)	3.21 (5.07)	1, 67	1.490	.227	.022
Disgust [t]	17.44 (11.8)	23.59 (15.23)	1, 67	2.985	.089	.043

Note. Lower Image Number at Detection scores indicate earlier (i.e., requiring lower intensity of emotion) detection. Lower Percentage of Incorrect Responses scores indicate a lower percentage of errors when detecting trials with the identified emotion. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 22

Hypothesis 3: Unadjusted Means and SDs of Image Number at Detection per Emotion for the Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (No/Yes) as a Function of Cycle Phase

	Mean (*SD*) of Image Number at Detection			
	Non-premenstrual Phase		Premenstrual Phase	
Emotion	No PMDD (*n* = 17)	PMDD (*n* = 18)	No PMDD (*n* = 5)	PMDD (*n* = 10)
Fear	10.85 (2.91)	9.63 (2.73)	9.9 (1.04)	10.15 (1.94)
Sad	9.25 (2.33)	8.46 (1.58)	8.75 (1.98)	8.78 (1.59)
Angry	10.51 (1.86)	9.10 (1.92)	9.45 (2)	9.60 (2.18)
Disgust	10.47 (1.64)	11.14 (2.58)	10.35 (1.97)	10.13 (1.8)
Happy	6.35 (1.86)	6.31 (2.07)	5.95 (1.44)	5.83 (1.39)
Surprise	7.53 (1.88)	6.78 (2.14)	7.10 (2.05)	6.13 (1.47)

Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. The non-premenstrual phase encompasses those in weeks 2-3 at the time of testing, and the premenstrual phase encompasses those in week 4. Variables used in the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Variables pertaining to the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the MANCOVA on all emotions. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Percentage of Incorrect Responses 2 x 2 MANCOVAs. Table 23 contains the unadjusted means and *SD*s for Percentage of Incorrect Responses scores pertaining to the 2 between (PMDD group) x 2 between (Menstrual Cycle Phase) MANCOVAs. Visual examination of the means suggested that, compared to the other three groups, the women with PMDD in the premenstrual phase had the lowest or second lowest number of incorrect responses for all six emotions. For the 2 x 2 MANCOVA testing Percentage of Incorrect Responses for negative emotions, the Group x Phase interaction effect was significant $F(3, 42) = 5.966$, $p = .002$, $\eta^2 = .299$. For the 2 x 2 MANCOVA testing Percentage of Incorrect Responses across all emotions, the interaction effect was also significant $F(6, 39) = 3.492$, $p = .007$, $\eta^2 = .349$.

Table 24 contains all follow-up ANCOVA results. Figure 7 displays the group differences in Percentage of Incorrect Responses and all significant pairwise results. The ANCOVA testing Percentage of Incorrect Responses for individual emotions revealed significant Group x Phase interactions for the detection of sad ($p < .001$, $\eta^2 = .288$), and surprise emotions, ($p = .030$, $\eta^2 = .102$). Pairwise comparisons indicated that the premenstrual PMDD group was more accurate at detecting sad emotions than the premenstrual no-PMDD group ($p = .005$), and the non-premenstrual PMDD group ($p = .004$). Conversely, the non-premenstrual no-PMDD group was more accurate at detecting sad emotions than the non-premenstrual PMDD group ($p = .005$), and the premenstrual no PMDD group ($p = .005$). See Figure 8 for an illustration of this interaction. Finally, the PMDD group was significantly more accurate at detecting disgust emotions in the premenstrual phase compared to the non-premenstrual phase ($p = .017$).

Supplemental Analyses

Supplemental analyses were run to determine whether there are group differences in Percentage of Correct Responses. These analyses were done given that there were differences in

Table 23

Hypothesis 3: Unadjusted Means (SDs) for Percentage of Incorrect Responses per Emotion for Provisional Premenstrual Dysphoric Disorder (PMDD) Groups

	Means (*SD*) of Percentage of Incorrect Responses			
	Non-premenstrual Phase		Premenstrual Phase	
Emotion	No PMDD (*n* = 17)	PMDD (*n* = 18)	No PMDD (*n* = 5)	PMDD (*n* = 10)
Fear	15.78 (13.91)	15.15 (14.29)	20.00 (13.99)	14.67 (8.45)
Sad ***	0.36 (1.13)$^{x, y}$	5.09 (5.61)$^{x, w}$	6.67 (6.12)$^{z, y}$	0.50 (1.58)$^{z, w}$
Angry	5.20 (7.52)	5.93 (11.00)	8.00 (8.37)	3.06 (4.88)
Disgust *	18.04 (10.04)	27.59 (16.85)^{y}	16.67 (17.00)	14.17 (10.04)^{y}
Happy	0.20 (0.81)	0.19 (0.54)	0 (0.00)	0.50 (1.58)
Surprise	3.73 (6.42)	6.20 (7.64)	9.67 (10.95)	2.33 (3.26)

Note. Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. The non-premenstrual phase encompasses those in weeks 2-3 at the time of testing, and the premenstrual phase encompasses those in week 4. Variables pertaining to the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the MANCOVA on all emotions. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use. ^{w,x,y,z} Group differences between the two indicated groups.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 24

Hypothesis 3: ANCOVA Interaction Results for Percentage of Incorrect Responses per Emotion: Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (No PMDD vs. PMDD) x Cycle Phase (Non-premenstrual vs. Premenstrual)

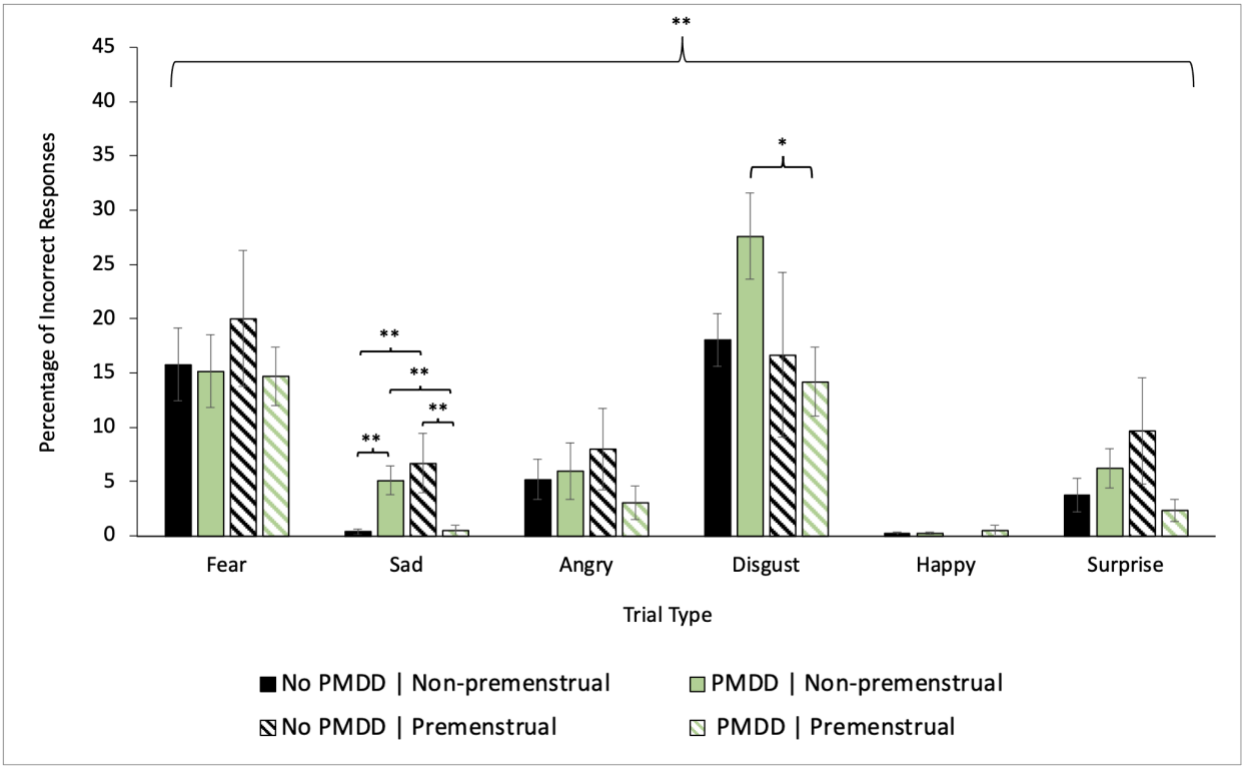
Percentage of Incorrect Responses				
Emotion	*df*	*F*	*p*	η^2
Fear	1, 63	0.292	.592	.007
Sad	1, 63	17.768	<.001***	.288
Angry	1, 63	0.996	.324	.022
Disgust	1, 63	1.847	.181	.040
Happy	1, 63	0.739	.395	.017
Surprise	1, 63	5.004	.030*	.102

Note. Results of the follow-up univariate ANCOVAs testing the interaction effect of PMDD group (no PMDD vs. PMDD) and cycle phase (weeks 2-3 vs. week 4 [premenstrual phase]) in Percentage of Incorrect Responses for individual emotions. All analyses controlled for hours of sleep last night and typical alcohol consumption.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 7

Percentage of Incorrect Responses per Emotion for Provisional Premenstrual Dysphoric Disorder (PMDD) Groups (Yes/No) as a Function of Cycle Phase

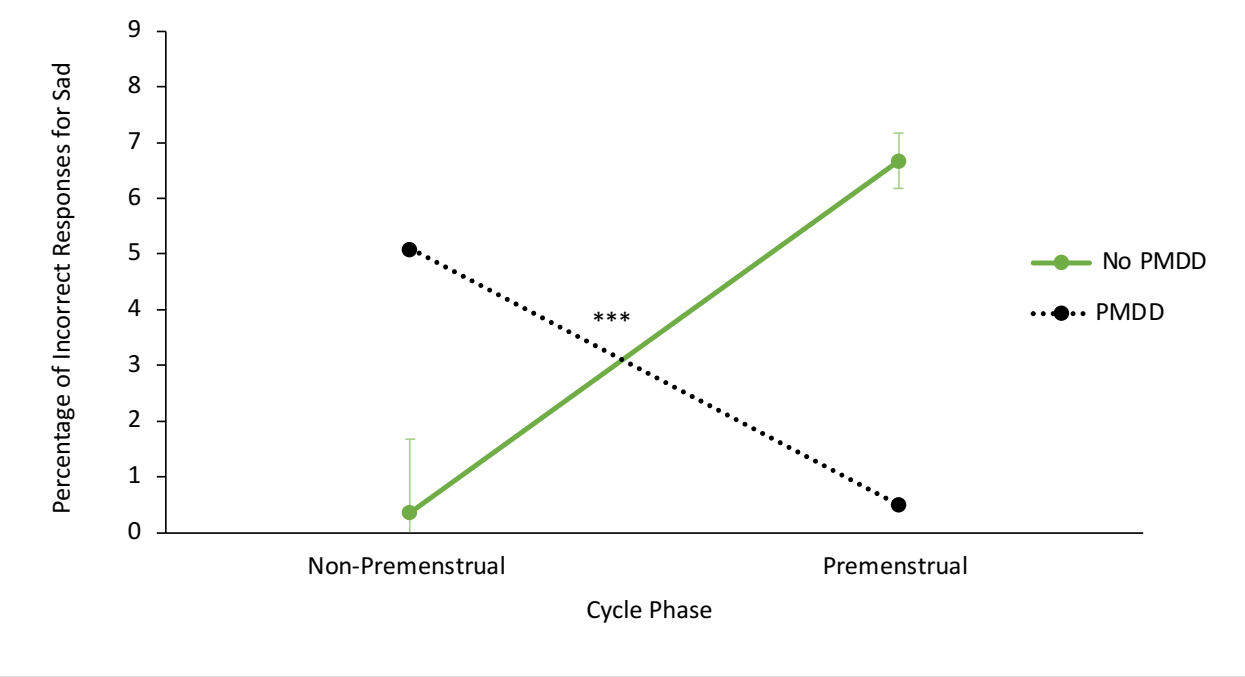


Note. Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. Non-premenstrual phase = in weeks 2-3 at the time of testing. Premenstrual phase = in week 4 at the time of testing. All analyses controlled for hours of sleep last night and typical alcohol use. There is a significant interaction effect between PMDD Group and Cycle Phase on Percentage of Incorrect Responses across the emotions, $F(6, 39 = 3.492$, $p = .007$, $\eta^2 = .349$. Error bars represent standard error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 8

Interaction Between Provisional Premenstrual Dysphoric Disorder (PMDD) Group (Yes, No) and Cycle Phase on Percentage of Incorrect Responses for Sad Emotions



Note. Lower scores indicate a lower percentage of errors when detecting trials with sad emotions. Non-premenstrual phase = weeks 2-3 at the time of testing. Premenstrual phase = week 4 at the time of testing. There was a significant PMDD Group x Phase interaction effect for Percentage of Incorrect Responses for sad emotions, $F(1, 63) = 17.768$, $p < .001$, $\eta^2 = .288$. Women with PMDD tested during the premenstrual phase made fewer errors than PMDD women during the non-premenstrual phase ($p = .004$) and than women without PMDD during the premenstrual phase ($p = .005$). Error bars represent standard error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

the calculation of correct and incorrect response percentages (see Methods for details). The main findings are outlined below, while a detailed description of these analyses can be found in Appendix G.

Succinctly, the two-group (OC users, FC women) MANCOVA testing Percentage of Correct Responses across all emotions was significant, $F(6, 108) = 2.564$, $p = .024$, $\eta^2 = .133$, suggesting OC users had overall worse performance on the task (i.e., had a lower Percentage of Correct Responses) than FC women. Additionally, the three-group (no/minimal PMDD, mild PMDD, moderate-severe PMDD) MANCOVA examining Percentage of Correct Responses across all emotions was also significant, $F(12, 124) = 1.963$, $p = .033$, $\eta^2 = .160$. This finding suggests that PMDD groups differ in their overall FEDT performance across all emotions, and univariate ANCOVAs indicated that these results were driven by the moderate-severe PMDD group providing more correct response on disgust trials than the mild PMDD group ($p = .058$). These results are consistent with the main findings for Image Number at Detection and Percentage of Incorrect Responses.

Error Biases

Error biases were examined at the trial type level by looking at the kind of incorrect responses participants were making. That is, for the six emotion trial types (disgust, fear, sad, angry, happy, surprise), the percentage of each type of incorrect response was calculated out of the total possible valid responses (i.e., excluding invalid trials). Repeated-measures MANOVAS were run for each emotion trial type to examine interaction, and multivariate effects.
The main findings are outlined below, while a detailed description of these analyses can be found in Appendix H.

Low and high depression score groups did not significantly differ in the type of errors made. Three non-significant small effect size nonsignificant trends suggest the possibility that OC users may provide more angry responses to disgust trials, $F(1, 107) = 3.615$, $p = .060$, $\eta^2 = .033$, less sad responses to fear trials, $F(1, 107) = 3.513$, $p = .064$, $\eta^2 = .032$, and more sad responses to happy trials, $F(1, 107) = 2.925$, $p = .090$, $\eta^2 = .027$, than FC women. Additionally, three non-significant small-medium effect size trends suggest that the PMDD group provides more sad responses to fear trials, $F(1, 70) = 3.649$, $p = .060$, $\eta^2 = .050$; less sad responses to angry trials, $F(1, 70) = 2.854$, $p = .096$, $\eta^2 = .039$; and more happy responses to angry trials, $F(1, 70) = 3.828$, $p = .054$, $\eta^2 = .052$, than the no-PMDD group.

Discussion

The purpose of the present study was to examine whether FED intensity and accuracy and are altered within women with depression symptoms, OC users, and women with PMS/PMDD symptoms. The first hypothesis that high depression symptoms would be associated with earlier and more accurate detection of negative emotions, in particular sad emotions, was not supported. However, women with high depression symptoms did have earlier detection across all of the emotions (largely driven by surprise). The second hypothesis that OC users would be slower and less accurate at detecting negative emotions compared to NC women and men was not entirely supported, although a similar finding presented. OC users exhibited overall later/slower detection (i.e., detection at higher intensity) across all emotions than FC women (largely driven by happy and disgust emotions). The final hypothesis was that women with high PMDD symptoms would be earlier and more accurate at detecting negative emotions, and that this effect would be strongest for participants in the premenstrual period. This hypothesis was partially supported. The women with and without provisional PMDD differed in their accuracy of detection across

negative emotions and across all six emotions based cycle phase (premenstrual, non-premenstrual), and this interaction was strongest for sad emotions. The women with PMDD were better at detecting sad emotions during the premenstrual phase.

Supplementary analyses indicated that OC users performed overall better on the FEDT than FC women (largely driven by happy and disgust emotions). Additionally, PMDD severity was associated with overall performance on the FEDT, with the moderate-severe PMDD group detecting disgust emotions better than the mild PMDD group. Error bias analyses did not reveal any significant differences between any of the groups.

Depression

Hypothesis 1 was not supported, as the women with high (vs. low) depression symptoms were not significantly earlier or more accurate at detecting the negative facial emotions. Additionally, no significant differences between low and high depression score groups emerged in intensity or accuracy for detection of sad emotions.

Women With High Depression Scores Detect Emotions Earlier

While significant group differences in accuracy and intensity at which *negative* emotions are detected did not emerge, groups did differ in timing of detection across all of the emotions. Specifically, women with high depression symptom scores detected emotions at significantly lower intensity scores (i.e., earlier) than those with low symptoms. The fact that this effect size was large, provides strong evidence that women experiencing depression are earlier to detect emotions across the board, and that this effect is not specific to negative emotions. Looking at individual emotions, there was a medium-large effect size for surprise, a small-medium effect size for disgust, and small effect size for fear, driving this relationship. However, only the surprise effect was significant.

While there is extensive literature examining FED in depression (e.g., Bourke et al., 2010; Dalili et al., 2015; Krause et al., 2021), few studies have examined the effects of emotional intensity on detection in depression (instead, most studies look at accuracy). Of those studies that have examined intensity, some have measured emotion detection across different intensities as we have done, while others have asked participants to rate the intensity of emotions. Consistent with our findings, Bomfim et al. (2019) found that depressed men and women provided more correct responses compared to controls to both low and high intensity presentations of the same six emotions we tested. One other study that looked at intensity suggests that sadness may be more accurately detected at low intensities within individuals with MDD compared to controls (Milders et al., 2010).

Two previous studies are inconsistent with the current findings as they found that depressed individuals required significantly more emotion intensity in order to detect happy emotions compared to controls, but not for sad or angry emotions (Joormann & Gotlib, 2006; LeMoult et al., 2009). Similar to these findings, Yoon et al. (2009) found that participants with MDD tended to judge happy emotions as less intense, which may explain why they require a higher intensity level to detect them. While our study did not find a significant effect for happy emotions, the effect size we did find for happy emotions was in the opposite direction to these studies but was also very small compared to the other emotions. Conversely, another study did not find evidence that low and higher depression groups differ in intensity of detection (Bediou et al., 2005). Notably, no study was identified that found that depression was associated with slower detection. Thus, there are no previous studies that have found the opposite of the current findings.

One potential explanation for the current finding that depressed women detected emotions at lower intensities is their higher BIS scores. High BIS scores are generally indicative of inhibitory behaviour and priming toward the fight or flight response (Carver & White, 1994). However, more current discussions surrounding the role of the BIS suggests that it is primarily activated when one is faced with conflict or uncertainty (Berkman et al., 2009). In this uncertain state (i.e., when attempting to detect emotions at the threshold of detection), elevated BIS levels are conducive to faster and better behavioural responses. Within the current task, there is uncertainty around the point of detection, when the face is no longer neutral but the expression of the emotion is still sub-threshold. Performance at this point in the task is what determines the intensity score. The high depression symptoms group exhibited significantly higher BIS scores than the low depression group. Consequentially, it may be this tendency toward activating the BIS system that enhanced their intensity at detection for all emotions. Overall, since this FED intensity effect presented with a large effect size, it should be followed up in future studies.

Women With High Depression Scores Detect Surprise Earlier and More Accurately

When looking at individual emotions, women with high depression scores were significantly earlier at detecting surprise emotions, with a medium-large effect size. Additionally, a small-medium effect size nonsignificant trend emerged suggesting that more depressed women were more accurate at detecting surprise emotions. Consistent with these results, women with high depression scores also had a higher Percentage of Correct Responses for surprise emotions, with a medium effect size, which is a variable that encompasses both intensity and accuracy. This suggests that women with high depression scores had overall better performance on the FEDT for surprise emotions.

The finding of a depression advantage in detecting surprise is not consistent with past research. In a review and meta-analyses on FED in depression, surprise stimuli were used in only 33% (Bourke et al., 2010), 32% (Dalili et al. 2015), and 35% (Krause et al., 2021) of included studies. Thus, there is limited past research on detecting this emotion. Two meta-analyses found worse detection for surprise among individuals with depression, with small effect sizes (Dalili et al. 2015; Krause et al., 2021). Out of 13 studies examining surprise, Bourke et al. (2010) identified only 1 study that found group differences, which suggested depression symptoms were associated with poorer recognition of surprise. Thus, past findings are not in line with ours.

It should be noted that within the present study, the surprise emotion was the only one where the actors had an open mouth. Participants may pick up on this, and potentially may observe an open mouth developing within a morph and be able to detect surprise based on this feature alone, as opposed to taking longer to integrate other facial features as well. Examination of mean intensity scores indicated that, while surprise was detected fairly early compared to other emotions, happy emotions were consistently detected earlier, suggesting surprise intensities do not represent potential outliers. However, it may be the case that women with high depression symptoms picked up on this open mouth feature more readily than the low depression group. Consistent with this, there is evidence that depressed individuals spend more time looking at the mouth region compared to the nose regions when processing fear, sad, angry, and neutral emotions, compared to controls (Hunter et al., 2020). There is also some evidence that higher BIS scores are associated with more right hemispherical activity when viewing surprise emotions (Balconi & Mazza, 2009), and the right hemisphere plays a role in processing uncertainty and negative facial emotion detection (Marinsek et al., 2014). It may be the case that detection of surprise involved a lot of uncertainty, which enhanced the performance of the higher BIS

depression group. Due to limited past research, and the potential effect of the open mouth, it is difficult to draw conclusions as to why the high depression group detected surprise emotions better. Future research is needed to replicate the finding for surprise and investigating the source of this relationship.

Facial Emotion Detection Accuracy Not Associated with Depression

There was no evidence of differences between the low and high depression groups on any measures of accuracy (e.g., errors) used within this study. Analyses looking at accuracy for negative emotions and across all six emotions were non-significant at the multivariate level, and for individual emotions. Only a non-significant trend emerged for accuracy of surprise emotions, suggesting more depressed women may be more accurate (see previous section). Additionally, Percentage of Correct Responses across all of the emotions did not differ based on group.

The lack of association between depressive symptoms and emotion detection accuracy was unexpected given the large literature base linking depression with decreased FED accuracy (See Dalili et al., 2015; Krause et al., 2021 for meta-analysis reviews). Four potential reasons for the present finding are considered below.

One difference between the current study and past ones is the criteria used for depression groupings. Dalili et al. (2015) only included studies that assessed depression diagnosis using DSM or International Classification of Diseases criteria. However, Krause et al. (2021) included several studies that used semi-structured measures to determine depression groups, including the measure we used, the $QIDS_{16}$. To determine whether the current results may be dependent on grouping criteria, all analyses were re-run using self-reported current depression diagnosis as the grouping variable. No new significant results emerged, and the direction of the means for the

intensity scores remained the same, favouring the depression group. This provided reassurance that the results are robust for grouping criteria.

A second potential explanation is that antidepressant use may have influenced the observed results. Some research suggests that antidepressant use attenuates the negativity bias in depression (Harmer et al., 2009), and meta-analysis has determined it may eliminate differences in FED (Bourke et al., 2010). However, Dalili et al. (2015) found no differences in FED between medicated and unmedicated depressed individuals. Antidepressant use was not asked about in the current study, and therefore it could not be used as a covariate or exclusion criterion to determine its influence. Another study collecting data from the same sample pool at the same time identified that out of a total sample of 160 participants, 16.5% were using antidepressants (Venkateshan, 2023). Additionally, of 71 depressed participants, 45.7% were using antidepressants, with the depressed group being significantly more likely to use antidepressants than the non-depressed group. The high rates of use in the population sampled suggest that antidepressant use may have been a factor in the current study, and should be investigated in future studies.

Third, given the vast literature outlining sex differences in FED (Rypma et al., 2015), and suggesting sex differences in the influence of depression on FED (Jenkins et al., 2018; Rypma et al., 2015; Wright et al., 2009), it is surprising that there have been few studies that have controlled for effects of sex. Within the meta-analysis by Krause et al. (2021), only 6/23 studies considered the effect of gender. Within our study we only included women in the depression analyses, to eliminate the interaction between sex differences in FED and the higher rates of depression in women. If the majority of past studies have included both men and women, and have not corrected for sex, their findings for depression may be confounded by sex differences

(i.e., women tend to have higher rates of depression and better FED, suggesting that better FED in depressed groups may reflect the sex difference). This is another factor that should be considered when interpreting the findings of this study in the context of past research.

However, the most likely explanation for our lack of association between depression and FED accuracy is that the nature of the task may have influenced accuracy. Indeed, Joormann and Gotlib (2006) also failed to find differences in accuracy between depressed and non-depressed groups when using a similar intensity morph task. Further, many other tasks that have tested FED using stimuli displaying different emotional intensities have presented the varying intensities in random order. This only provides insight on detection differences in the final intensity of emotions. In our study we presented the images of increasing intensities in order, essentially slowing down the process of how emotions develop on a face and measuring differences in the threshold of detection during this process. Our task simulates the in-vivo experience of engaging in a social interaction and observing an individual's face as they are developing an emotion. Interestingly, final emotions are rated as more intense when they are shown within an intensity morph, similar to what was used within the current task, than when simply displayed directly at 100% intensity (Biele & Grabowska, 2006). Within the present task this morph may have helped facilitate accurate detection, minimizing potential group differences. This may also account for the differences in the current findings compared to past studies, which have used only 100% intensity images or faces with lower emotion intensities that are presented in random order.

Error Biases Not Associated with Depression

The exploratory analysis of error biases did not find any differences between women with low and high depression symptoms. These analyses examined the cases in which participants

provided incorrect responses to determine whether the kind of incorrect emotional response per emotion type trial differed based on depression group. Depression symptoms did not influence the overall pattern in the number of errors that were made for each trial type (e.g., the distribution in number of errors on fear, sad, angry, happy, surprise errors or disgust trials), and also did not influence the rate of unique error types (e.g., the percentage of responses that were fear on disgust trials). Given the literature on negativity biases, suggesting that individuals with depression appraise neutral and positive stimuli as more negative (Bourke et al., 2010), it was expected that the high depression group would exhibit a bias toward overreporting negative, or sad, emotions. Indeed, one study found that MDD participants made more mistakes than a control group in discriminating between the following emotion pairs: disgust-anger, sadness-anger, fear-anger, sadness-disgust, fear-disgust, surprise-anger, surprise-disgust, surprise-happiness, and happiness-anger (Mo et al., 2021). However, they did not find that MDD participants overreported negative emotions, but rather that they were more likely to confuse anger with other emotions. This was the only other study found that examined error biases in depression.

It is also worthwhile to look at overall error biases across the entire sample. Within the current study, across all participants, disgust was most likely to be mis-detected as anger, fear as surprise, sad as anger, angry as disgust, happy as surprise, and surprise as happy. This was fairly consistent with Mo et al. (2021), who also reported which emotions were most likely to be confused for other ones. They found all the same error pairs, except that their happy and surprise trials were most likely to be mis-detected as fear. They used a very different task, including only 100% intensity emotions, and providing participants with only two possible response options per image. No other studies were found that calculated error biases for all six emotions similarly to

how we have. However, these findings provide some validity for the error biases calculated in the present study.

Implications

The present study found that individuals with high depression symptoms detected emotions earlier, overall, but particularly so for surprise emotions. This finding may be conceptualized within the lens of depressive realism, a theory regarding the mechanisms of depression (Moore & Fresco, 2012). This theory suggests that depressed individuals process stimuli more realistically and make more realistic inferences. It is therefore proposed that depressed individuals do not exhibit negative biases, but rather that non-depressed individuals favour positively-biased perceptions, which are adaptive and protective against depression (Moore & Fresco, 2012). Depression may then characterize a tendency to attend to subtle indicators of emotions, and perceive them in a more realistic, yet comparatively more negative way, which may maintain negative mood (i.e., a greater tendency to quickly perceive actual emotions). Our findings are congruent with this theory, as the high depression symptom group was earlier to detect emotions, with no differences in accuracy compared to the low depression group. It should also be noted that while evidence for the negativity bias did not present within the depression group, the current results can still stand alongside the negativity bias theory (Bourke et al., 2010) in that it may be the case that the negativity bias is more salient across different types of FED tasks.

Oral Contraceptives

Hypothesis 2 was partially supported. Women taking OCs were not significantly slower or less accurate at detecting negative emotions compared to FC women and men, when fear, sad, and angry faces were examined together (i.e., there were no significant group differences in

intensity or accuracy). However, consistent with hypothesis 2, OC users were slower than FC women to detect emotions when *all facial emotions* (i.e., not just negative emotions) were considered together. There was also evidence that OC users were slower to identify happy and disgust emotions. Supplemental analyses (i.e., Percentage of Correct Responses) also revealed that OC users' overall performance on the FEDT was worse than FC women, with specific effects for happy and disgust.

OC Users Detect Emotions Later

The expected group differences across negative emotions did not present, however OC users were slower to detect across overall emotions. Specifically, OC users required a higher Image Number at Detection than FC women to accurately detect the facial emotions, and this effect was of a large effect size. Further examination of these differences found that OC users were later than FC women to detect happy emotions, with a medium effect size, and disgust emotions, with a small-medium effect size. No significant differences at the individual emotion level presented between the three groups (OC users, FC women, and men).

While other studies have examined a variety of emotion intensities, they presented the different emotional intensities (e.g., 40%, 20%, 100%) in random order (Hamstra et al., 2014, 2015, 2016, 2017). The present study is the first to assess detection at varying intensities within a sequential morph. Only one previous study reported the effect of OCs on intensity, and found that fear and sad emotions were more accurately recognized at low intensity levels by OC users, however they did not examine effects for happy or disgust emotions (Hamstra et al., 2015). These findings seem inconsistent with ours. Another study found that OC users rated the valence of facial emotions as more extreme (Spalek et al., 2019). That is, OC users rated negative facial emotions as more negative, positive as more positive, and neutral as more neutral, compared to

FC women. Interestingly, FC women rated neutral emotions as more positive than OC users (Spalek et al., 2019), and OC users rated negative emotions as more aversive (Gamsakhurdashvili et al., 2021). If replicable, the tendency for OC women to evaluate neutral facial emotions as more neutral may have made it more difficult for OC users to see the emerging emotions in all the morphs given that the images morphed form neutral to the emotion. Further, FC women might have been quicker to see the happy emotion within the neutral first part of the morph, given the finding that they tend to view neutral emotions as more positive. However, while OC users may have been slower to detect disgust due to the tendency to view neutral emotions as more neutral, the tendency for OC users to view negative emotions as more negative may have reduced this effect size. These two previous studies (Spalek et al., 2019; Gamsakhurdashvili et al., 2021) might help explain the medium-large effect size for happy and small-medium for disgust, but do not explain why these particular two emotions show the group differences over the other four.

As expected, OC users had a lower Percentage of Correct Responses (indicating worse overall FEDT performance) across all emotions, compared to FC women with a large effect size. Additionally, OC users provided significantly less correct responses on happy and disgust emotions than FC women. This finding is consistent with other findings in the literature suggesting that OC users are worse than non-users at detecting emotions. Specifically, Hamstra et al. (2014) also found that OC users are worse at detecting disgust emotions, and several other studies have found that OC users are worse at detecting all emotions (Hamstra et al., 2016; Pahnke et al., 2019).

Cortisol levels may help explain the slower and overall worse detection of happy and disgust in OC users. Meta-analysis has found that women taking OCs have a decreased cortisol

response compared to FC women (Gervasio et al., 2022). Lower cortisol response has been associated with lower accuracy in detecting happy emotions (von Dawans et al., 2020), and later (higher intensity) detection of disgust emotions (Daudelin-Peltier et al., 2017), similar to what we have found. Thus, the blunted cortisol response in OC users might help explain slower and less accurate detection of happy and disgust facial emotions.

Androgenic OC Users Detect Happy Emotions Later

Further examination of the differences in intensity at FED between OC users and non-users found that using androgenic compared to anti-androgenic OC formulations affects detection tendencies. Androgenic OC users, anti-androgenic OC users, and FC women significantly differed in the intensity at which they detected overall emotions, with a large effect size. Specifically, androgenic (but not anti-androgenic) OC users were slower than FC women at detecting happy emotions, with a medium-large effect size. No differences were found when comparing the androgenic and anti-androgenic users. Only two other studies have investigated the effect of OC androgenicity on FED. One study found no difference between androgenic and anti-androgenic OCs on FED accuracy (Pahnke et al., 2019). The second found that anti-androgenic OC users were less accurate than androgenic OC users across all emotions tested (Gurvich et al., 2020). Thus, the three studies to date do not show consistent findings.

There are a number of differences between the three studies that have examined OC androgenicity and FED. It should be noted that the categorization of androgenic versus anti-androgenic OC formulations can be considered on a spectrum and is not perfectly dichotomous. While categorization within the present study was based on the criteria used in these two past studies (i.e., 1st and 2nd generation OC are androgenic, and 3rd and 4th generation are anti-androgenic), there are still differences in the overall formulations (i.e., the type of progestins and

concentration of estradiol and progestin) that groups were taking across the two past studies and the present one. Additionally, Pahnke and colleagues (2019) did not assess detection of full faces (i.e., they used the RMET task), and Gurvich and colleagues (2020) used a discrimination task requiring participants to look at four images and select the odd one out, as opposed to a detection task where they look at one image and state what they detect. Our study tested threshold of detection within intensity morphs, while other studies have only used individual images. These methodological differences may contribute to the different findings.

These results suggest that our overall finding that OC users detect happy emotions slower than FC women may apply to only those taking androgenic OC formulations. This may contribute to the higher rates of negative mood side effects that are associated with starting 2nd generation formulations, which are the most androgenic, compared to 3rd generation formulations (Shahnazi et al., 2014), and the higher rates of mood swings, depression, and fatigue among 2nd generation users compared to non-users (Gingnell et al., 2013). Replication of this androgenic OC happy effect is important for better understanding this relationship.

Accuracy Not Associated With OC Use

There were no significant differences in accuracy across OC users, FC women and men. That is, Percentage of Incorrect Responses did not differ between any of the groups when assessed across negative emotions, across overall emotions, or for individual emotions. Additionally, examination of the mean scores did not indicate that means were distributed in any reliable direction among the groups. This finding was surprising, as the majority of other studies that reported effects of OCs on FED found differences in accuracy (Gurvich et al., 2020; Hamstra et al., 2014, 2015a, 2016, 2017; Pahnke et al., 2019). Similarly, the literature suggests that sex differences in accuracy, with women performing better, are expected (Guapo et al.,

2009; Saylik et al., 2018; Wingenbach et al., 2018). It should be noted that, as mentioned above, FC women did have a better Percentage of Correct Responses across overall emotions, however this variable represents overall FEDT performance (i.e., a composite of intensity and accuracy, but does not represent accuracy alone). The lack of group differences specific to accuracy within the present study may be attributed to the design of the FEDT used, as other studies have presented intensities in random order, or simply not included different emotional intensities.

Error Biases Not Significantly Associated With OC Use

Exploratory examination of error biases did not find significant differences between OC users and FC women. That is, the type of errors that were made for each trial type, and the rate of unique error types did not differ between groups. Some non-significant trends in unique error types did present. Namely, OC users may provide more incorrect angry responses on disgust trials, less incorrect sad responses on fear trials, and more incorrect sad responses on happy trials, all with a small-medium effect size. This is the first study to look at OC use and error biases in FED, therefore these analyses were exploratory in nature and the trends are only mentioned to inform future research. Several past studies have pointed toward a greater rate of errors for sad emotions among OC users (but have not investigated the type of errors) (Hamstra et al., 2014, 2015b, 2017), however this was not observed in the present study, nor did any error biases for sad emotions present. However, the nonsignificant trends for OC users to provide more sad responses on happy trials, and less sad responses on fear trials may be consistent with the differences in processing sad emotions that were found in past studies.

Implications

Consistent with past research, the present study continues to point toward the finding that OC use is associated with worse FED. We found that OC users detected emotions slower, and

had a worse overall FEDT performance. This was the first study to investigate how OC use may affect intensity at detection, and we had the second largest sample size of all OC and FED studies, second to Shirazi et al. (2020). Interestingly, our strongest emotion-specific finding, that OC users detect happy emotions slower, persisted only for androgenic OC users. This highlights how different OC formulations may differentially affect FED which may then contribute to mood side effects. Specifically, if androgenic OC users do not detect happy emotions as quickly as other emotions, the tendency to perceive other emotions at a normal speed would contribute to a slight negativity bias in their experience. In this case, a similar negativity bias as seen in depression may present (Bourke et al., 2010). Perceiving negative stimuli earlier may influence their subsequent mood to reinforce feeling a negative emotional state. More generally, perceiving overall emotions slower could contribute to OC users experiencing social difficulties as they may not be able to process non-verbal cues as quickly as FC women. This may cause misinterpretations within social interactions and contribute to negative mood. Understanding the effects of different OC formulations is important in the informed decision-making process, and in aiding women and providers to select OC types based on individual needs. Replication of the current findings and further research is needed.

PMS/PMDD

Hypothesis 3 was partially supported. In line with the hypothesis, while in the premenstrual phase, women with provisional PMDD were more accurate in detecting sad emotions compared to women without PMDD. Also, an overall Group x Phase interaction with a large effect size suggested that, within the premenstrual phase, women with PMDD symptoms were more accurate in detecting negative emotions (when all types were examined together). Additionally, while in the non-premenstrual phase, women without PMDD more accurately

detected sad emotions than women with PMDD. However, PMDD symptoms were not related to intensity of detection across all negative emotions, neither across the menstrual cycle nor at specific cycle phases as hypothesized.

Women With PMDD: More Accurate Emotion Detection During the Premenstrual Phase

As was mentioned above, a significant interaction between PMDD and cycle phase on accuracy of detection emerged. This interaction was present across the negative emotions and across all emotions, with a large effect size for both. The finding that women with PMDD were more accurate at detecting sad emotions when in the premenstrual phase also presented with a very large effect size. The premenstrual PMDD group was significantly more accurate at detecting sad emotions than the premenstrual non-PMDD group and the non-premenstrual PMDD group (Conversely, the non-premenstrual non-PMDD group was significantly more accurate at detecting sad emotions than the non-premenstrual PMDD group and the premenstrual non-PMDD group). A similar significant interaction for surprise emotions also presented with a medium-large effect size. A final significant group difference indicated that the PMDD group was more accurate in detecting disgust within the premenstrual compared to the non-premenstrual phase.

Only three previous studies have examined FED in PMS/PMDD, and only two examined PMDD-cycle phase interactions. In their study, Rubinow et al. (2007) found an effect that may be congruent with the present one, namely that women with PMDD were more likely to rate neutral emotions as sad in the luteal phase. This may suggest a tendency toward processing sad during the luteal/premenstrual phase. However, a second study using a within-subjects design found that women with PMS were less accurate at detecting sad emotions during the luteal phase compared to the follicular phase, and found no differences between women with and without

PMS at any point (Gultekin et al., 2017). One final study found that women with PMDD were better at detecting sadness in male compared to female faces, but no other differences in emotion detection (Ramos-Loyo & Sanz-Martin, 2017). It should be noted that the present study defined the premenstrual phase as week four of the menstrual cycle, while the non-premenstrual phase was defined as weeks two through three. Week one, the week of menstruation, was excluded from these analyses to avoid any influence of menstrual discomfort on FED, and to isolate the influence of premenstrual symptoms. This approach differs from past studies which have included the entire cycle and compared between luteal and follicular phase (i.e., not specifically examining the premenstrual phase), suggesting a possible source for the difference in results. Nevertheless, our findings are consistent with one of two previous studies that examined FED as a function of PMDD and cycle phase.

The present findings suggest that women with PMDD experience a tendency toward negative processing during their premenstrual phase, which facilitates the detection of negative, and especially sad, emotions. PMDD is characterized as a depressive disorder, and mood lability is one of the essential features of the disorder (APA, 2013). Similar to the negativity bias within depression, this negativity bias during the premenstrual phase may prime individuals to exhibit a hypervigilance to negative emotional stimuli. This can cause them to attend to negative stimuli, and perceive other kinds of emotional stimuli as more negative, which can reinforce a negative affective state (Bourke et al., 2010).

Women With PMDD Detect Sad Emotions Earlier. When women with and without PMDD were compared on FED independent of cycle phase, significant group differences were found in the intensity of detection for sad emotions. The PMDD group was earlier to detect sad emotions compared to the no-PMDD group across the cycle, with a medium effect size. This

finding indicates that enhanced detection of sad emotions among women with PMDD may also exist across the entire cycle, and supports the notion of a general negativity bias within PMDD, outlined below.

While the current study is only the second to find evidence of greater negative processing in PMDD, this tendency persists when producing and processing other kinds of emotional stimuli as well. Mass et al. (2008) found that women with PMS symptoms produced more sad facial expressions when viewing emotionally charged stimuli during the luteal phase. Meanwhile, Gonda et al. (2010) asked women about their experience of negative, positive, and neutral life events over different timelines (ranging from within the past month to the past three years). When tested in the late-follicular phases, PMS symptoms were positively associated with reporting having experienced more negative life events and less positive life events. Similarly, when led to believe they had failed at a task, women with PMS/PMDD in the luteal phase described themselves more negatively and experienced more sadness and irritation (Śliwerski & Bielawska-Batorowicz, 2019). However, Śliwerski and Bielawska-Batorowicz (2019), did not find the same effect in women taking OCs, suggesting hormonal fluctuations are essential to this relationship. Thus, the present findings fit well with previous evidence of a negative bias in women with PMS/PMDD in the premenstrual phase and extend the findings to suggest an enhanced ability to detect sad emotions.

Severity of PMDD May Affect Intensity and Accuracy

When removing the influence of cycle phase, the severity of PMDD symptoms, characterized as no/minimal, mild, or moderate-severe, did not significantly affect intensity or accuracy of detection across negative emotions or across overall emotions. However, a non-significant trend with a medium-large effect size was found for intensity, and a medium-large

effect size was found for accuracy, across overall emotions. Group means were not distributed in any consistent direction across the groups. Shown below, examination of individual emotions did yield significant findings.

Severity of PMDD Associated With Detection of Disgust

Level of PMDD symptoms influenced the intensity and accuracy of detection for disgust emotions, when detecting across the menstrual cycle. The three PMDD level groups (i.e., no/minimal, mild, and moderate-severe) differed in their intensity of detection, with a medium-large effect size, and accuracy, with a large effect size, for disgust emotions. These findings persisted when menstrual cycle phase was statistically controlled (i.e., used as a covariate). The mild PMDD group displayed the worst detection, being slower (i.e., required a higher intensity) to detect disgust than the moderate-severe PMDD group, and less accurate at detecting disgust than both the no/minimal and moderate-severe PMDD groups. A similar trend appeared for Percentage of Correct Responses (representing overall FEDT performance), with the mild PMDD group performing worse than the no/minimal and moderate-severe groups. It should be noted that the subsample of moderate-severe PMDD was small (n = 8). When severity of PMDD was removed (assessing only no PMDD vs. PMDD), a small-medium effect size trend toward the PMDD group exhibiting lower accuracy did persist, however the intensity findings did not. This suggests the most robust findings are that PMDD symptoms may be related to less accurate detection of disgust, while women with more severe PMDD symptoms appear to detect emotions earlier than women with mild symptoms. However, the results do not point toward a linear relationship between PMDD and disgust detection.

In considering possible explanations for this finding, past research has linked the detection of disgust emotions to estrogen and progesterone levels. Some studies have found that

high estrogen levels are associated with lower accuracy in detecting disgust emotions (Kamboj et al., 2015; Gasbarri et al., 2008). Conversely, high progesterone levels have been associated with better intensity ratings for disgust emotions (Conway et al., 2007). As expected, women also experience weaker feelings of disgust during high estrogen and low progesterone periods, such as around the ovulatory period, which is proposed to be conducive toward mating behaviour (Liu et al., 2023). It has also been suggested that women with PMDD exhibit lower estrogen and higher progesterone levels (Roomruangwong et al., 2019; Yen et al., 2019). The luteal phase is associated with a rapid rise in estrogen and progesterone, with progesterone reaching a higher peak, and then a rapid fall of both hormones in the premenstrual phase (Reed & Carr, 2000). These rapid hormone changes likely contribute to our findings.

The postpartum period is also characterized by a rapid drop in estrogen and progesterone. Additionally, many women experience postpartum depression (PPD) which is postulated to be in part caused by this hormone change and has many overlapping symptoms with PMDD (Schiller et al., 2015). Consistent with FED in depression, women with PPD are better at detecting negative emotions, and tend to view neutral infant facial expressions as more negative (Gil et al., 2011). This has been conceptualized as an adaptation toward perceiving their infant's distress consistent with the primary caretaker hypothesis and the fitness threat hypothesis (Hampson et al., 2006; Wingenbach et al., 2018; Hampson et al., 2021). Interestingly, they also found that high anxiety in the postpartum sample was associated with perceiving disgust as more intense (Gil et al., 2011). However, another study found that PPD was associated with decreased accuracy in detecting disgust (Flanagan et al., 2011). These findings may be relevant as the similar drop in hormones during postpartum and the premenstrual phase may suggest similar mechanisms affecting FED in PPD and PMDD.

There is also evidence toward a Hormonal Sensitivity Syndrome (HSS), suggesting that a subset of women are more negatively affected by hormonal changes (Pope et al., 2017). The subset of women with HSS are more likely to have PMS, higher rates of postpartum symptoms, a history of antidepressant use, and report lower sex drive, which is relevant in the context of decreased disgust processing around ovulation being adaptive for mating. Other studies have also found that postpartum depression and PMDD symptoms tend to be experienced in the same women (Lee et al., 2015; Takayama et al., 2020).

These factors may explain why women with PMDD may be more affected by these hormonal changes and exhibit differences in disgust detection. The PMDD group likely comprises more women with HSS. Consistent with the findings from Flanagan et al. (2011), our PMDD group had worse accuracy in detecting disgust. When considering severity of PMDD, it may be the case that the moderate-severe PMDD group may comprise women who are more sensitive to drops in progesterone and have more anxiety symptoms. Indeed, our group equivalency did indicate the moderate-severe PMDD group had higher rates of other psychological disorders (the majority of which were anxiety), although this difference was not significant. Then, consistent with the findings from Conway et al. (2007), and Gil et al. (2011), it may be a higher progesterone sensitivity and higher anxiety in the moderate-severe PMDD group that is contributing to the observed lower intensity at detection for disgust. These findings must be replicated, and more research is needed on the dimensional characteristics of PMDD, before a definitive conclusion can be drawn.

Error Biases Not Significantly Associated With PMDD

Exploratory examination of error biases found trends, but no significant differences, between women with and without provisional PMDD. There were trends suggesting that the

PMDD group may provide less incorrect sad responses on fear trials, more incorrect sad responses on angry trials, and less incorrect happy responses on angry trials. This was the first study to look at PMS/PMDD and error biases in FED, therefore these trends in the data are exploratory.

Implications

The findings from the present study may help provide insight into the etiology of PMDD or some of the negative symptoms in PMDD. Past research has suggested that women with high PMS/PMDD symptoms may exhibit both a hormonal sensitivity (Pope et al., 2017; Steiner et al., 2003), and trait-like negative cognitions (Gonda et al., 2010). Then, during hormonal changes, they experience of a negativity bias (Bourke et al., 2010), or depressive realism (Moore & Fresco, 2012), similar to what is seen in depression, may have influenced the enhanced sad processing within PMDD women, especially in the premenstrual phase. This may cause the PMDD group to also appraise these hormonal changes as more negative and impairing relative to those without symptoms, and may contribute to the clinically significant, especially mood-related, symptoms that characterize PMDD. The rates within the present sample are consistent with past studies in suggesting that about half of women experience some amount of PMS symptoms (Direkvand-Moghadam et al., 2014). However, the combination of a high hormonal sensitivity and tendency toward trait negative cognitions may characterize women with severe PMS/PMDD, which are even more sensitive to hormonal changes in the premenstrual period. This could contribute to the observed differential detection of disgust emotions within the moderate-severe PMDD group. The hormonal sensitivity is also associated with higher postpartum depression symptoms, and may explain why women with PMDD exhibit similar disgust FED as women with PPD (Flanagan et al., 2011; Lee et al., 2015; Takayama et al., 2020).

This was the first study to investigate intensity at detection and PMDD, and has the largest sample size of all PMS/PMDD and FED studies. While the etiology of PMDD is still not clearly understood, these findings point toward these women experiencing a hormonal sensitivity that is activated within the hormonal milieu of the late luteal phase which contributes to the differences in FED found and the mood symptoms associated with the disorder.

Strengths, Limitations, and Future Research

There are many strengths to the current study, in part because the limitations of past studies were used to guide our current approach. One main strength relates to the design of the FEDT. The task was comprehensive in testing all six basic emotions, with full face stimuli, and across different intensity levels. As emotion recognition abilities are more challenged in complex studies, this may have allowed for the detection of subtle group differences. Additionally, the stimuli included in the task were validated (in terms of the emotion being presented), and sex and race of stimuli photos were counterbalanced across emotions, which reduces the chance of external biases in detection. The task is also comprehensive in that we have reported both raw descriptive data (See Appendix F), which many past studies have failed to do, and calculated separate intensity and accuracy scores, and error biases, allowing us to pinpoint which process of FED is different across groups.

Additional strengths pertain to our group variables. Stringent exclusion criteria aimed to eliminate external confounds that may affect detection (e.g., exclusion of people that changed their OC status in the past 2 months or were taking hormonal medication). Some examples include the PMDD groups which included only FC women. Additionally, the OC analyses reported on and investigated the effects of androgenic vs. anti-androgenic OC formulations,

which most past studies have not done. Our overall sample size was large, making it the second largest sample size of the OC studies, and the largest of all the PMS/PMDD studies.

There are three main study limitations. First, all groups were created based on participant self-report/self-selection. Participants may have been biased in reporting their rate and severity of depression and PMDD symptoms. We used symptom severity as our grouping variable, however using more stringent diagnosis criteria for depression and PMDD (i.e, semi-structured interview based on ICD or DSM) may have yielded different results. Also, OC use was self-selected, and participants that choose to use OCs may differ in other characteristics from FC women. To fully examine the effect of grouping variables on detection, future studies can use diagnostic criteria (i.e., semi-structured interviews) to create depression/PMDD groups. Additionally, there is no longitudinal research on the causal relationship between OCs and FED. Future research may recruit women interested in using OCs and assess their FED before starting, and at different time points after starting to take OCs.

Second, some data from recruited participants could not be used. In the future, exclusion criteria could be outlined prior to recruitment, such that certain scenarios (e.g., substance use before participating) could be prevented. Additionally, many participants could not or had difficulties accessing the FEDT due to technical issues with Pavlovia, the site that hosted the task, suggesting a different hosting software could be considered. Similarly, the linking of data from Survey Money and Pavlovia via the participant code was prone to errors (e.g., participants did not input the same code on both platforms), and fifteen responses could not be linked. The sample for some of the groups ended up being small, and we found many results with large effect sizes that were only trends or non-significant. Future studies should aim to address these limitations in sample size in order to increase the power of analyses.

Finally, as this was the first use of the FEDT, there are several areas of improvement for the task. First, there were only 4 trials per emotion type, however the demand per trial was high as participants had to provide 15 responses. More than 2 missing trials resulted in excess missing data such that the participant's overall data could not be used. Future iterations of the task may consider increasing the number of trials per emotion, in order to increase the power of analyses, but researchers will also have to consider the length of the task as this will also affect drop-out rates. Additionally, as this was the first time that error biases were calculated for the task, the methodology used (i.e., percentage of all incorrect response types per emotion trial) was selected due to its simplicity. However, we found no significant differences in error biases although group means did appear to differ. Future iterations of the task may consider using more applied error bias calculations, such as the approach used by Lee et al. (2022), which looked at the differences in error types to emotional and neutral expressions. In the present task "neutral" expressions could represent the images in the first half of the emotional morph, and emotional expressions could represent expressions in the second half of the morph. Such an approach may be more sensitive to group differences.

Summary and Conclusions

The present study examined the effects of depression symptoms, OC use, and premenstrual symptoms on FED. The findings provide support for the role of hormones and mood in emotional processing. Women with high depression symptoms detected emotions earlier. Conversely, OC users detected emotions later. However, OC type was relevant and androgenic formulations were associated with later detection of happy emotions. PMDD was associated with earlier detection of sad emotions, more accurate detection of emotions while in

the premenstrual phase, particularly sad emotions, and differential detection of disgust emotions (i.e., PMDD group was more accurate and moderate-severe PMDD group was earlier).

These findings point to a relationship between hormones, mood, and facial emotion processing that is strong yet complex. Emotional processing in depression may not follow a straightforward negativity bias as past studies have suggested, and instead depression may involve a tendency to pay attention to subtle facial emotions but then perceive them in more realistic, yet ultimately more negative way compared to non-depressed individuals. This tendency toward negative perceptions, combined with a hormonal sensitivity, may differentiate PMDD from other depressive disorders. While individuals with PMDD are prone to exhibit similar negative perceptions as in depression, they may be more sensitive to them when large hormonal changes occur, specifically during the premenstrual period. Across the rest of the cycle, hormonal sensitivity experienced by some women may still contribute to negative mood, which may be one reason why rates of depression are higher in women compared to men. Conversely, OCs disrupt the natural cyclicity of hormones across the course of the cycle. Instead, OC use allowed us to observe the role of altered steady hormone levels on emotional processing, specifically that androgenic formulations increased the threshold of detection for happy emotions. This may be helpful in understanding different types of side effects associated with different OC formulations. In conclusion these findings begin to highlight the complex interplay between hormones, mood, and women's overall well-being. They provide novel insight into how FED may be relevant to mechanisms involved in depression and PMDD, and the effects of OCs on mood.

References

Angst, J., Sellaro, R., Stolar, M., Merikangas, K. R., & Endicott, J. (2001). The epidemiology of perimenstrual psychological symptoms. *Acta Psychiatrica Scandinavica*, *104*(2), 110–116. https://doi.org/10.1034/j.1600-0447.2001.00412

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Bagby, R., Parker, J., & Taylor, G. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*, 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Balconi, M., & Mazza, G. (2009). Brain oscillations and BIS/BAS (behavioral inhibition/activation system) effects on processing masked emotional cues. ERS/ERD and coherence measures of alpha band. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *74*(2), 158–165. https://doi.org/10.1016/j.ijpsycho.2009.08.006

Barth, C., Villringer, A., & Sacher, J. (2015). Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. *Frontiers in Neuroscience*, *9*, 37. https://doi.org/10.3389/fnins.2015.00037

Batur, P., Elder, J., & Mayer, M. (2003). Update on contraception: Benefits and risks of the new formulations. *Cleveland Clinic Journal of Medicine*, *70*(8), 681–682, 685–686, 668-690 passim. https://doi.org/10.3949/ccjm.70.8.681

Bediou, B., Krolak-Salmon, P., Saoud, M., Henaff, M.-A., Burt, M., Dalery, J., & D'Amato, T. (2005). Facial expression and sex recognition in schizophrenia and depression. *Canadian*

Journal of Psychiatry. Revue Canadienne De Psychiatrie, *50*(9), 525–533. https://doi.org/10.1177/070674370505000905

Bengtsdotter, H., Lundin, C., Gemzell Danielsson, K., Bixo, M., Baumgart, J., Marions, L., Brynhildsen, J., Malmborg, A., Lindh, I., & Sundström Poromaa, I. (2018). Ongoing or previous mental disorders predispose to adverse mood reporting during combined oral contraceptive use. *The European Journal of Contraception & Reproductive Health Care*, *23*(1), 45–51. https://doi.org/10.1080/13625187.2017.1422239

Bento de Souza, I. B. M., Barbosa, F. F., Lacerda, A. M., dos Santos, N. A., & Torro-Alves, N. (2014). Evaluation of facial expressions in women with major depression: Is there a negative bias? *Psychology & Neuroscience*, *7*(4), 513–519. http://dx.doi.org.ezproxy.lakeheadu.ca/10.3922/j.psns.2014.4.10

Berkman, E. T., Lieberman, M. D., & Gable, S. L. (2009). BIS, BAS, and response conflict: Testing predictions of the revised reinforcement sensitivity theory. *Personality and Individual Differences*, *46*(5–6), 586–591. https://doi.org/10.1016/j.paid.2008.12.015

Bérubé, A., Turgeon, J., Blais, C., & Fiset, D. (2021). Emotion recognition in adults with a history of childhood naltreatment: A systematic review. *Trauma, Violence, & Abuse*, 15248380211029404. https://doi.org/10.1177/15248380211029403

Biele, C., & Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Experimental Brain Research*, *171*(1), 1–6. https://doi.org/10.1007/s00221-005-0254-0

Blom, S. S. A. H., Aarts, H., & Semin, G. R. (2020). Lateralization of facial emotion processing and facial mimicry. *Laterality*, *25*(3), 259–274. https://doi.org/10.1080/1357650X.2019.1657127

Bomfim, A. J. de L., Ribeiro, R. A. dos S., & Chagas, M. H. N. (2019). Recognition of dynamic and static facial expressions of emotion among older adults with major depression. *Trends in Psychiatry and Psychotherapy*, *41*, 159–166. https://doi.org/10.1590/2237-6089-2018-0054

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: A review. *Australian & New Zealand Journal of Psychiatry*, *44*(8), 681–696. https://doi.org/10.3109/00048674.2010.496359

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., de Graaf, R., Demyttenaere, K., Hu, C., Iwata, N., Karam, A. N., Kaur, J., Kostyuchenko, S., Lépine, J.-P., Levinson, D., Matschinger, H., Mora, M. E. M., Browne, M. O., Posada-Villa, J., … Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, *9*(1), 90. https://doi.org/10.1186/1741-7015-9-90

Burke, C. W. (1969). The effect of oral contraceptives on cortisol metabolism. *Journal of Clinical Pathology. Supplement (Ass. Clin. Path.)*, *3*, 11–18.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Centers for Disease Control and Prevention (CDC). (2010). Adverse childhood experiences reported by adults—Five states, 2009. *MMWR. Morbidity and Mortality Weekly Report*, *59*(49), 1609–1613.

Chan, A. W.-, & Downing, P. (2011). Faces and eyes in human lateral prefrontal cortex. *Frontiers in Human Neuroscience*, *5*, 51.

Coffee, A. L., Kuehl, T. J., Willis, S., & Sulak, P. J. (2006). Oral contraceptives and premenstrual symptoms: Comparison of a 21/7 and extended regimen. *American Journal of Obstetrics and Gynecology*, *195*(5), 1311–1319. https://doi.org/10.1016/j.ajog.2006.05.012

Conley, M. I., Dellarco, D. V., Rubien-Thomas, E., Cohen, A. O., Cervera, A., Tottenham, N., & Casey, B. (2018). The racially diverse affective expression (RADIATE) face stimulus set. *Psychiatry Research*, *270*, 1059–1067. https://doi.org/10.1016/j.psychres.2018.04.066

Conway, C. A., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Law Smith, M. J., Perrett, D. I., Sharp, M. A., & Al-Dujaili, E. A. S. (2007). Salience of emotional displays of danger and contagion in faces is enhanced when progesterone levels are raised. *Hormones and Behavior*, *51*(2), 202–206. https://doi.org/10.1016/j.yhbeh.2006.10.002

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafò, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine*, *45*(6), 1135–1144. https://doi.org/10.1017/S0033291714002591

Darney, P. D. (1995). The androgenicity of progestins. *The American Journal of Medicine*, *98*(1, Supplement 1), S104–S110. https://doi.org/10.1016/S0002-9343(99)80067-9

Daudelin-Peltier, C., Forget, H., Blais, C., Deschênes, A., & Fiset, D. (2017). The effect of acute social stress on the recognition of facial expression of emotions. *Scientific Reports*, *7*(1), 1–13. https://doi.org/10.1038/s41598-017-01053-3

De Berardis, D., Campanella, D., Gambi, F., Sepede, G., Carano, A., Pelusi, L., La Rovere, R., Di Matteo, D., Salini, G., Cotellessa, C., Salerno, R. M., & Ferro, F. M. (2005). Alexithymia and body image disturbances in women with Premenstrual Dysphoric

Disorder. *Journal of Psychosomatic Obstetrics and Gynaecology*, *26*(4), 257–264. https://doi.org/10.1080/01674820500109081

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

Derntl, B., Windischberger, C., Robinson, S., Lamplmayr, E., Kryspin-Exner, I., Gur, R. C., Moser, E., & Habel, U. (2008). Facial emotion recognition and amygdala activation are associated with menstrual cycle phase. *Psychoneuroendocrinology*, *33*(8), 1031–1040. https://doi.org/10.1016/j.psyneuen.2008.04.014

Direkvand-Moghadam, A., Sayehmiri, K., Delpisheh, A., & Kaikhavandi, S. (2014). Epidemiology of Premenstrual Syndrome (PMS)-A systematic review and meta-analysis study. *Journal of Clinical and Diagnostic Research: JCDR*, *8*(2), 106–109. https://doi.org/10.7860/JCDR/2014/8024.4021

Donadon, M. F., & Osório, F. de L. (2014). Recognition of facial expressions by alcoholic patients: A systematic literature review. *Neuropsychiatric Disease and Treatment*, *10*, 1655–1663. https://doi.org/10.2147/NDT.S65376

Duran, N., & Atkinson, A. P. (2021). Foveal processing of emotion-informative facial features. *PLOS ONE*, *16*(12), e0260814. https://doi.org/10.1371/journal.pone.0260814

Dwyer, J. B., Aftab, A., Radhakrishnan, R., Widge, A., Rodriguez, C. I., Carpenter, L. L., Nemeroff, C. B., McDonald, W. M., & Kalin, N. H. (2020). Hormonal treatments for major depressive disorder: State of the art. *American Journal of Psychiatry*, *177*(8), 686–705. https://doi.org/10.1176/appi.ajp.2020.19080848

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine*, *14*(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Flanagan, T. J., White, H., & Carter, B. G. (2011). Differential impairments in emotion face recognition in postpartum and nonpostpartum depressed women. *Journal of Affective Disorders*, *128*(3), 314–318. https://doi.org/10.1016/j.jad.2010.07.021

Fleischman, D. S., Navarrete, C. D., & Fessler, D. M. T. (2010). Oral contraceptives suppress ovarian hormone production. *Psychological Science*, *21*(5), 750–752.

Gamsakhurdashvili, D., Antov, M. I., & Stockhorst, U. (2021). Facial emotion recognition and emotional memory from the ovarian-hormone perspective: A systematic review. *Frontiers in Psychology*, *12*, 641250. https://doi.org/10.3389/fpsyg.2021.641250

Gervasio, J., Zheng, S., Skrotzki, C., & Pachete, A. (2022). The effect of oral contraceptive use on cortisol reactivity to the Trier Social Stress Test: A meta-analysis. *Psychoneuroendocrinology*, *136*, 105626. https://doi.org/10.1016/j.psyneuen.2021.105626

Gil, S., Teissèdre, F., Chambres, P., & Droit-Volet, S. (2011). The evaluation of emotional facial expressions in early postpartum depression mood: A difference between adult and baby faces? *Psychiatry Research*, *186*(2), 281–286. https://doi.org/10.1016/j.psychres.2010.06.015

Gingnell, M., Engman, J., Frick, A., Moby, L., Wikström, J., Fredrikson, M., & Sundström-Poromaa, I. (2013). Oral contraceptive use changes brain activity and mood in women

with previous negative affect on the pill—A double-blinded, placebo-controlled randomized trial of a levonorgestrel-containing combined oral contraceptive. *Psychoneuroendocrinology*, *38*(7), 1133–1144. https://doi.org/10.1016/j.psyneuen.2012.11.006

Gingnell, M., Morell, A., Bannbers, E., Wikström, J., & Sundström Poromaa, I. (2012). Menstrual cycle effects on amygdala reactivity to emotional stimulation in premenstrual dysphoric disorder. *Hormones and Behavior*, *62*(4), 400–406. https://doi.org/10.1016/j.yhbeh.2012.07.005

Gonda, X., Fountoulakis, K. N., Csukly, G., Telek, T., Pap, D., Rihmer, Z., & Bagdy, G. (2010). Association of a trait-like bias towards the perception of negative subjective life events with risk of developing premenstrual symptoms. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *34*(3), 500–505. https://doi.org/10.1016/j.pnpbp.2010.02.004

Guapo, V. G., Graeff, F. G., Zani, A. C. T., Labate, C. M., dos Reis, R. M., & Del-Ben, C. M. (2009). Effects of sex hormonal levels and phases of the menstrual cycle in the processing of emotional faces. *Psychoneuroendocrinology*, *34*(7), 1087–1094. https://doi.org/10.1016/j.psyneuen.2009.02.007

Gultekin, G., Uludag, C., Cetinkaya, S., Altun, I., Ozan, E., Acikgoz, S., Dalcik, E., & Emul, M. (2017). The comparison of facial emotion recognition ability in women with and without Premenstrual Syndrome. *Turkish Journal of Psychiatry*, *28*(4). https://doi.org/10.5080/u20494

Gültekin, G., Uludağ, C., Çetinkaya, S., Altun, İ., Ozan, E., Açıkgöz, S., Dalcık, E. N., & Emül, M. (2017). The Comparison of Facial Emotion Recognition Ability in Women with and

without Premenstrual Syndrome. *Turk Psikiyatri Dergisi = Turkish Journal of Psychiatry*, *28*(4), 234–239.

Gurvich, C., Warren, A. M., Worsley, R., Hudaib, A.-R., Thomas, N., & Kulkarni, J. (2020). Effects of oral contraceptive androgenicity on visuospatial and social-emotional cognition: A prospective observational trial. *Brain Sciences*, *10*(4). https://doi.org/10.3390/brainsci10040194

Halari, R., Hines, M., Kumari, V., Mehrotra, R., Wheeler, M., Ng, V., & Sharma, T. (2005). Sex differences and individual differences in cognitive performance and their relationship to endogenous gonadal hormones and gonadotropins. *Behavioral Neuroscience*, *119*(1), 104–117. https://doi.org/10.1037/0735-7044.119.1.104

Hall, J., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion (Washington, D.C.)*, *4*, 201–206. https://doi.org/10.1037/1528-3542.4.2.201

Hampson, E. (1990). Estrogen-related variations in human spatial and articulatory-motor skills. *Psychoneuroendocrinology*, *15*(2), 97–111. https://doi.org/10.1016/0306-4530(90)90018-5

Hampson, E., Istasy, P., Owais, S., Chow, J., Howidi, B., & Ouellette, S. (2021). Sex differences in the recognition of children's emotional Expressions: A test of the fitness threat hypothesis. *Evolutionary Psychological Science*, *7*(1), 45–60. https://doi.org/10.1007/s40806-020-00254-w

Hampson, E., van Anders, S. M., & Mullin, L. I. (2006). A female advantage in the recognition of emotional facial expressions: Test of an evolutionary hypothesis. *Evolution and Human Behavior*, *27*(6), 401–416. https://doi.org/10.1016/j.evolhumbehav.2006.05.002

Hampson, E., & Young, E. (2007). Methodological issues in the study of hormone-behavior relations in humans: Understanding and monitoring the menstrual cycle. *Sex Differences in the Brain: From Genes to Behavior*. https://doi.org/10.1093/acprof:oso/9780195311587.003.0004

Hamstra, D. A., de Kloet, E. R., Quataert, I., Jansen, M., & Van der Does, W. (2017). Mineralocorticoid receptor haplotype, estradiol, progesterone and emotional information processing. *Psychoneuroendocrinology*, *76*, 162–173. https://doi.org/10.1016/j.psyneuen.2016.11.037

Hamstra, D. A., de Kloet, E. R., Tollenaar, M., Verkuil, B., Manai, M., Putman, P., & Van der Does, W. (2016). Mineralocorticoid receptor haplotype moderates the effects of oral contraceptives and menstrual cycle on emotional information processing. *Journal of Psychopharmacology*, *30*(10), 1054–1061. https://doi.org/10.1177/0269881116647504

Hamstra, D. A., de Kloet, E. R., van Hemert, A. M., de Rijk, R. H., & Van der Does, A. J. W. (2015). Mineralocorticoid receptor haplotype, oral contraceptives and emotional information processing. *Neuroscience*, *286*, 412–422. https://doi.org/10.1016/j.neuroscience.2014.12.004

Hamstra, D. A., De Rover, M., De Rijk, R. H., & Van der Does, W. (2014). Oral contraceptives may alter the detection of emotions in facial expressions. *European Neuropsychopharmacology*, *24*(11), 1855–1859. https://doi.org/10.1016/j.euroneuro.2014.08.015

Hantsoo, L., & Epperson, C. N. (2015). Premenstrual Dysphoric Disorder: Epidemiology and treatment. *Current Psychiatry Reports*, *17*(11), 87. https://doi.org/10.1007/s11920-015-0628-3

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., Goodwin, G. M., & Cowen, P. J. (2009). Effect of acute antidepressant administration on negative affective bias in depressed patients. *American Journal of Psychiatry*, *166*(10), 1178–1184. https://doi.org/10.1176/appi.ajp.2009.09020149

Hashemi, S., Ramezani Tehrani, F., Mohammadi, N., Rostami Dovom, M., Torkestani, F., Simbar, M., & Azizi, F. (2016). Comparison of metabolic and hormonal profiles of women with and without premenstrual syndrome: A community based cross-sectional study. *International Journal of Endocrinology and Metabolism*, *14*(2), e28422. https://doi.org/10.5812/ijem.28422

Hawkins, S. M., & Matzuk, M. M. (2008). The menstrual cycle. *Annals of the New York Academy of Sciences*, *1135*(1), 10–18. https://doi.org/10.1196/annals.1429.018

Hermans, E. J., Ramsey, N. F., & van Honk, J. (2008). Exogenous testosterone enhances responsiveness to social threat in the neural circuitry of social aggression in humans. *Biological Psychiatry*, *63*(3), 263–270. https://doi.org/10.1016/j.biopsych.2007.05.013

Hidalgo-Lopez, E., & Pletzer, B. (2017). Interactive effects of dopamine baseline levels and cycle phase on executive functions: The role of progesterone. *Frontiers in Neuroscience*, *11*, 403.

Hudgens, G. A., Fatkin, L. T., Billingsley, P. A., & Mazurczak, J. (1988). Hand steadiness: Effects of sex, menstrual phase, oral contraceptives, practice, and handgun weight. *Human Factors*, *30*(1), 51–60.

Hunter, L., Roland, L., & Ferozpuri, A. (2020). Emotional expression processing and depressive symptomatology: Eye-tracking reveals differential importance of lower and middle facial

areas of interest. *Depression Research and Treatment*, *2020*, e1049851. https://doi.org/10.1155/2020/1049851

Jarva, J. A., & Oinonen, K. A. (2007). Do oral contraceptives act as mood stabilizers? Evidence of positive affect stabilization. *Archives of Women's Mental Health*, *10*(5), 225–234. https://doi.org/10.1007/s00737-007-0197-5

Jenkins, L. M., Kendall, A. D., Kassel, M. T., Patrón, V. G., Gowins, J. R., Dion, C., Shankman, S. A., Weisenbach, S. L., Maki, P., & Langenecker, S. A. (2018). Considering sex differences clarifies the effects of depression on facial emotion processing during fMRI. *Journal of Affective Disorders*, *225*, 129–136. https://doi.org/10.1016/j.jad.2017.08.027

Johnson, S. L., Turner, R. J., & Iwata, N. (2003). BIS/BAS levels and psychiatric disorder: An epidemiological study. *Journal of Psychopathology and Behavioral Assessment*, *25*(1), 25–36. https://doi.org/10.1023/A:1022247919288

Jones, M. S., Pierce, H., & Shafer, K. (2022). Gender differences in early adverse childhood experiences and youth psychological distress. *Journal of Criminal Justice*, *83*, 101925. https://doi.org/10.1016/j.jcrimjus.2022.101925

Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology*, *115*(4), 705–714. https://doi.org/10.1037/0021-843X.115.4.705

Jung, W. H., Lee, T. Y., Kim, M., Lee, J., Oh, S., Lho, S. K., Moon, S.-Y., & Kwon, J. S. (2022). Sex differences in the behavioral inhibition system and ventromedial prefrontal cortex connectivity. *Social Cognitive and Affective Neuroscience*, *17*(6), 571–578. https://doi.org/10.1093/scan/nsab118

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), e0122311. https://doi.org/10.1371/journal.pone.0122311

Kaunitz, A. M. (2004). Enhancing oral contraceptive success: The potential of new formulations. *American Journal of Obstetrics and Gynecology*, *190*(4, Supplement), S23–S29. https://doi.org/10.1016/j.ajog.2004.01.062

Kendler, K. S., Martin, N. G., Heath, A. C., Handelsman, D., & Eaves, L. J. (1988). A twin study of the psychiatric side effects of oral contraceptives. *The Journal of Nervous and Mental Disease*, *176*(3), 153–160. https://doi.org/10.1097/00005053-198803000-00003

Kimmig, A.-C. S., Bischofberger, J. A., Birrenbach, A. D., Drotleff, B., Lämmerhofer, M., Sundström-Poromaa, I., & Derntl, B. (2022). No evidence for a role of oral contraceptive-Use in emotion recognition but higher negativity bias in early follicular women. *Frontiers in Behavioral Neuroscience*, *15*, 773961. https://doi.org/10.3389/fnbeh.2021.773961

Krause, F. C., Linardatos, E., Fresco, D. M., & Moore, M. T. (2021). Facial emotion recognition in major depressive disorder: A meta-analytic review. *Journal of Affective Disorders*, *293*, 320–328. https://doi.org/10.1016/j.jad.2021.06.053

Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*(7), 1211–1221. https://doi.org/10.1016/j.neuropsychologia.2011.12.022

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of Dynamic Aspects of Facial Expressions: A Review. *Emotion Review*, *5*(1), 41–46. https://doi.org/10.1177/1754073912451349

Kupferberg, A., Bicks, L., & Hasler, G. (2016). Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, *69*, 313–332. https://doi.org/10.1016/j.neubiorev.2016.07.002

Le, J., Thomas, N., & Gurvich, C. (2020). Cognition, The Menstrual Cycle, and Premenstrual Disorders: A Review. *Brain Sciences*, *10*(4), Article 4. https://doi.org/10.3390/brainsci10040198

Lee, S.-C., Lin, G.-H., Shih, C.-L., Chen, K.-W., Liu, C.-C., Kuo, C.-J., & Hsieh, C.-L. (2022). Error patterns of facial emotion recognition in patients with schizophrenia. *Journal of Affective Disorders*, *300*, 441–448. https://doi.org/10.1016/j.jad.2021.12.130

Lee, Y.-J., Yi, S.-W., Ju, D.-H., Lee, S.-S., Sohn, W.-S., & Kim, I.-J. (2015). Correlation between postpartum depression and premenstrual dysphoric disorder: Single center study. *Obstetrics & Gynecology Science*, *58*(5), 353–358. https://doi.org/10.5468/ogs.2015.58.5.353

Leeners, B., Kruger, T. H. C., Geraedts, K., Tronci, E., Mancini, T., Ille, F., Egli, M., Röblitz, S., Saleh, L., Spanaus, K., Schippert, C., Zhang, Y., & Hengartner, M. P. (2017). Lack of associations between female hormone levels and visuospatial working memory, divided attention and cognitive bias across two consecutive menstrual cycles. *Frontiers in Behavioral Neuroscience*, *11*, 120.

LeMoult, J., Joormann, J., Sherdell, L., Wright, Y., & Gotlib, I. H. (2009). Identification of emotional facial expressions following recovery from depression. *Journal of Abnormal Psychology*, *118*(4), 828–833. https://doi.org/10.1037/a0016944

Lépine, J.-P., & Briley, M. (2011). The increasing burden of depression. *Neuropsychiatric Disease and Treatment*, *7*(Suppl 1), 3–7. https://doi.org/10.2147/NDT.S19617

Lewis, C. A., Kimmig, A.-C. S., Zsido, R. G., Jank, A., Derntl, B., & Sacher, J. (2019). Effects of hormonal contraceptives on mood: A focus on emotion recognition and reactivity, reward processing, and stress response. *Current Psychiatry Reports*, *21*(11), 115. https://doi.org/10.1007/s11920-019-1095-z

Liu, M., Zhang, X., He, Z., Liang, Y., Zou, B., Ma, X., Gu, S., & Wang, F. (2023). Opposite effects of estradiol and progesterone on woman's disgust processing. *Frontiers in Psychiatry*, *14*. https://www.frontiersin.org/articles/10.3389/fpsyt.2023.1161488

López, L. E., Verdejo, E. C., Javier, F. G., Martín, J. R. O., & Gómez-Amor, J. (2010). Incidence of anovulatory menstrual cycles among dysmenorrheic and non-dismenorrheic women: Effects on symptomatology and mood. *Psicothema*, *22*, 654–658.

Machado, R. B., Pompei, L. de M., Giribela, A. G., & Giribela, C. G. (2011). Drospirenone/ethinylestradiol: A review on efficacy and noncontraceptive benefits. *Women's Health*, *7*(1), 19–30. https://doi.org/10.2217/WHE.10.84

Maki, P. M., Rich, J. B., & Rosenbaum, R. S. (2002). Implicit memory varies across the menstrual cycle: Estrogen effects in young women. *Neuropsychologia*, *40*(5), 518–529. https://doi.org/10.1016/s0028-3932(01)00126-9

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

Marečková, K., Perrin, J. S., Nawaz Khan, I., Lawrence, C., Dickie, E., McQuiggan, D. A., Paus, T., & the IMAGEN Consortium. (2014). Hormonal contraceptives, menstrual cycle and brain response to faces. *Social Cognitive and Affective Neuroscience*, *9*(2), 191–200. https://doi.org/10.1093/scan/nss128

Marinsek, N., Turner, B. O., Gazzaniga, M., & Miller, M. B. (2014). Divergent hemispheric reasoning strategies: Reducing uncertainty versus resolving inconsistency. *Frontiers in Human Neuroscience*, *8*. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00839

Mass, R., Moll, B., Hölldorfer, M., Wiedemann, K., Richter-Appelt, H., Dahme, B., & Wolf, K. (2008). Effects of the premenstrual syndrome on facial expressions of sadness. *Scandinavian Journal of Psychology*, *49*(3), 293–298. https://doi.org/10.1111/j.1467-9450.2008.00645.x

Menting-Henry, S., Hidalgo-Lopez, E., Aichhorn, M., Kronbichler, M., Kerschbaum, H., & Pletzer, B. (2022). Oral contraceptives modulate the relationship between resting brain activity, amygdala connectivity and emotion recognition – A resting state fMRI Study. *Frontiers in Behavioral Neuroscience*, *16*, 775796. https://doi.org/10.3389/fnbeh.2022.775796

Milders, M., Bell, S., Platt, J., Serrano, R., & Runcie, O. (2010). Stable expression recognition abnormalities in unipolar depression. *Psychiatry Research*, *179*(1), 38–42. https://doi.org/10.1016/j.psychres.2009.05.015

Mishra, S., Elliott, H., & Marwaha, R. (2022). Premenstrual Dysphoric Disorder. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK532307/

Mo, F., Gu, J., Zhao, K., & Fu, X. (2021). Confusion effects of facial expression recognition in patients with major depressive disorder and healthy controls. *Frontiers in Psychology*, *12*, 703888. https://doi.org/10.3389/fpsyg.2021.703888

Montagne, B., Kessels, R. P. C., De Haan, E. H. F., & Perrett, D. I. (2007). The emotion recognition task: A paradigm to measure the perception of facial emotional expressions at

different intensities. *Perceptual and Motor Skills*, *104*(2), 589–598. https://doi.org/10.2466/pms.104.2.589-598

Moore, M. T., & Fresco, D. M. (2012). Depressive realism: A meta-analytic review. *Clinical Psychology Review*, *32*(6), 496–509. https://doi.org/10.1016/j.cpr.2012.05.004

Moos, R. H. (1968). The development of a menstrual distress questionnaire. *Psychosomatic Medicine*, *30*(6), 853–867. https://doi.org/10.1097/00006842-196811000-00006

Mordecai, K. L., Rubin, L. H., & Maki, P. M. (2008). Effects of menstrual cycle phase and oral contraceptive use on verbal memory. *Hormones and Behavior*, *54*(2), 286–293. https://doi.org/10.1016/j.yhbeh.2008.03.006

Morey, L. C., & Lowmaster, S. E. (2010). Personality Assessment Inventory. In *The Corsini Encyclopedia of Psychology* (pp. 1–4). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470479216.corpsy0663

Nielsen, S. E., Segal, S. K., Worden, I. V., Yim, I. S., & Cahill, L. (2013). Hormonal contraception use alters stress responses and emotional memory. *Biological Psychology*, *92*(2), 257–266. https://doi.org/10.1016/j.biopsycho.2012.10.007

Oinonen, K. A. (2009). Putting a finger on potential predictors of oral contraceptive side effects: 2D:4D and middle-phalangeal hair. *Psychoneuroendocrinology*, *34*(5), 713–726. https://doi.org/10.1016/j.psyneuen.2008.11.009

Oinonen, K. A., & Mazmanian, D. (2002). To what extent do oral contraceptives influence mood and affect? *Journal of Affective Disorders*, *70*(3), 229–240. https://doi.org/10.1016/S0165-0327(01)00356-1

Oinonen, K. A., & Mazmanian, D. (2007). Facial symmetry detection ability changes across the menstrual cycle. *Biological Psychology*, *75*(2), 136–145. https://doi.org/10.1016/j.biopsycho.2007.01.003

Osório, F. L., de Paula Cassis, J. M., Machado de Sousa, J. P., Poli-Neto, O., & Martín-Santos, R. (2018). Sex hormones and processing of facial expressions of emotion: A systematic literature review. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00529

Padhy, S. K., Sarkar, S., Beherre, P. B., Rathi, R., Panigrahi, M., & Patil, P. S. (2015). Relationship of Premenstrual Syndrome and Premenstrual Dysphoric Disorder with major depression: Relevance to clinical practice. *Indian Journal of Psychological Medicine*, *37*(2), 159–164. https://doi.org/10.4103/0253-7176.155614

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. *Frontiers in Neuroscience*, *12*, 1041.

Payne, J. L., Palmer, J. T., & Joffe, H. (2009). A reproductive subtype of depression: Conceptualizing models and moving toward etiology. *Harvard Review of Psychiatry*, *17*(2), 72–86. https://doi.org/10.1080/10673220902899706

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

Petersen, N., London, E. D., Liang, L., Ghahremani, D. G., Gerards, R., Goldman, L., & Rapkin, A. J. (2016). Emotion regulation in women with premenstrual dysphoric disorder. *Archives of Women's Mental Health*, *19*(5), 891–898. https://doi.org/10.1007/s00737-016-0634-4

Poromaa, I. S., & Segebladh, B. (2012). Adverse mood symptoms with oral contraceptives. *Acta Obstetricia et Gynecologica Scandinavica*, *91*(4), 420–427. https://doi.org/10.1111/j.1600-0412.2011.01333.x

Powell, A. (2017). Choosing the right oral contraceptive pill for teens. *Pediatric Clinics of North America*, *64*(2), 343–358. https://doi.org/10.1016/j.pcl.2016.11.005

Putman, P., Hermans, E. J., & van Honk, J. (2007). Exogenous cortisol shifts a motivated bias from fear to anger in spatial working memory for facial expressions. *Psychoneuroendocrinology*, *32*(1), 14–21. https://doi.org/10.1016/j.psyneuen.2006.09.010

Quinkler, M., Meyer, B., Bumke-Vogt, C., Grossmann, C., Gruber, U., Oelkers, W., Diederich, S., & Bahr, V. (2002). Agonistic and antagonistic properties of progesterone metabolites at the human mineralocorticoid receptor. *European Journal of Endocrinology*, *146*(6), 789–799. https://doi.org/10.1530/eje.0.1460789

Radke, S., & Derntl, B. (2016). Affective responsiveness is influenced by intake of oral contraceptives. *European Neuropsychopharmacology*, *26*(6), 1014–1019. https://doi.org/10.1016/j.euroneuro.2016.03.004

Ramos-Loyo, J., & Sanz-Martin, A. (2017). Emotional experience and recognition across menstrual cycle and in premenstrual disorder. *International Journal of Psychological Studies*, *9*(4), 33. https://doi.org/10.5539/ijps.v9n4p33

Reed, B. G., & Carr, B. R. (2000). The Normal Menstrual Cycle and the Control of Ovulation. In K. R. Feingold, B. Anawalt, A. Boyce, G. Chrousos, W. W. de Herder, K. Dhatariya, K. Dungan, J. M. Hershman, J. Hofland, S. Kalra, G. Kaltsas, C. Koch, P. Kopp, M. Korbonits, C. S. Kovacs, W. Kuohung, B. Laferrère, M. Levy, E. A. McGee, … D. P.

Wilson (Eds.), *Endotext*. MDText.com, Inc. http://www.ncbi.nlm.nih.gov/books/NBK279054/

Richards, M. A., & Oinonen, K. A. (2021). Psychometric properties of a DSM-5-based screening tool for women's perceptions of premenstrual symptoms. *Psychological Reports*, *125*(2), 1–32. https://doi.org/10.1177/0033294120979696

Romero-Martínez, Á., Sarrate-Costa, C., & Moya-Albiol, L. (2021). A systematic review of the role of oxytocin, cortisol, and testosterone in facial emotional processing. *Biology*, *10*(12), Article 12. https://doi.org/10.3390/biology10121334

Roomruangwong, C., Carvalho, A. F., Comhaire, F., & Maes, M. (2019). Lowered plasma steady-state levels of progesterone combined with declining progesterone levels during the luteal phase predict peri-menstrual syndrome and its major subdomains. *Frontiers in Psychology*, *10*, 2446. https://doi.org/10.3389/fpsyg.2019.02446

Rosenberg, M. J., & Waugh, M. S. (1998). Oral contraceptive discontinuation: A prospective evaluation of frequency and reasons. *American Journal of Obstetrics and Gynecology*, *179*(3), 577–582. https://doi.org/10.1016/S0002-9378(98)70047-X

Rotter, N. G., & Rotter, G. S. (1988). Sex differences in the encoding and decoding of negative facial emotions. *Journal of Nonverbal Behavior*, *12*(2), 139–148. https://doi.org/10.1007/BF00986931

Rubin, L. H., Carter, C. S., Drogos, L., Jamadar, R., Pournajafi-Nazarloo, H., Sweeney, J. A., & Maki, P. M. (2011). Sex-specific associations between peripheral oxytocin and emotion perception in schizophrenia. *Schizophrenia Research*, *130*(1), 266–270. https://doi.org/10.1016/j.schres.2011.06.002

Rubinow, D. R., Smith, M. J., Schenkel, L. A., Schmidt, P. J., & Dancer, K. (2007). Facial emotion discrimination across the menstrual cycle in women with Premenstrual Dysphoric Disorder (PMDD) and controls. *Journal of Affective Disorders*, *104*(1–3), 37–44. https://doi.org/10.1016/j.jad.2007.01.031

Rukavina, S., Sachsenweger, F., Jerg-Bretzke, L., Daucher, A. E., Traue, H. C., Walter, S., & Hoffmann, H. (2018). Testosterone and its influence on emotion recognition in young, healthy males. *Psychology*, *9*(7), Article 7. https://doi.org/10.4236/psych.2018.97106

Rush, A. J., Bernstein, I. H., Trivedi, M. H., Carmody, T. J., Wisniewski, S., Mundt, J. C., Shores-Wilson, K., Biggs, M. M., Woo, A., Nierenberg, A. A., & Fava, M. (2006). An evaluation of the Quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: A sequenced treatment alternatives to relieve depression trial report. *Biological Psychiatry*, *59*(6), 493–501. https://doi.org/10.1016/j.biopsych.2005.08.022

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., & Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic Major Depression. *Biological Psychiatry*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(02)01866-8

Rypma, B., Fischer, H., Rieckmann, A., Hubbard, N. A., Nyberg, L., & Bäckman, L. (2015). Dopamine D1 binding potential predicts fusiform BOLD activity during face-recognition performance. *Journal of Neuroscience*, *35*(44), 14702–14707. https://doi.org/10.1523/JNEUROSCI.1298-15.2015

Sagar, R., Talwar, S., Desai, G., & Chaturvedi, S. K. (2021). Relationship between alexithymia and depression: A narrative review. *Indian Journal of Psychiatry*, *63*(2), 127–133. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_738_19

Sasson, N. J., Pinkham, A. E., Richard, J., Hughett, P., Gur, R. E., & Gur, R. C. (2010). Controlling for response biases clarifies sex and age differences in facial affect recognition. *Journal of Nonverbal Behavior*, *34*(4), 207–221. https://doi.org/10.1007/s10919-010-0092-z

Saylik, R., Raman, E., & Szameitat, A. J. (2018). Sex differences in emotion recognition and working memory tasks. *Frontiers in Psychology*, *9*, 1072.

Schienle, A., Schäfer, A., Stark, R., Walter, B., & Vaitl, D. (2005). Relationship between disgust sensitivity, trait anxiety and brain activity during disgust induction. *Neuropsychobiology*, *51*(2), 86–92. https://doi.org/10.1159/000084165

Schiller, C. E., Meltzer-Brody, S., & Rubinow, D. R. (2015). The role of reproductive hormones in postpartum depression. *CNS Spectrums*, *20*(1), 48–59. https://doi.org/10.1017/S1092852914000480

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*, 104895. https://doi.org/10.1016/j.psyneuen.2020.104895

Schneider, S., Peters, J., Bromberg, U., Brassen, S., Menz, M., Miedl, S., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Conrod, P., Dalley, J., Flor, H., Gallinat, J., Garavan, H., Heinz, A., Itterman, B., Mallik, C., Mann, K., & Büchel, C. (2011). Boys do it the right way: Sex-dependent amygdala lateralization during face processing in

adolescents. *NeuroImage*, *56*, 1847–1853. https://doi.org/10.1016/j.neuroimage.2011.02.019

Shahnazi, M., Farshbaf Khalili, A., Ranjbar Kochaksaraei, F., Asghari Jafarabadi, M., Gaza Banoi, K., Nahaee, J., & Bayati Payan, S. (2014). A comparison of second and third generations combined oral contraceptive pills' effect on mood. *Iranian Red Crescent Medical Journal*, *16*(8), e13628. https://doi.org/10.5812/ircmj.13628

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. *Hormones and Behavior*, *119*, 104647. https://doi.org/10.1016/j.yhbeh.2019.104647

Skovlund, C. W., Mørch, L. S., Kessing, L. V., & Lidegaard, Ø. (2016). Association of hormonal contraception with depression. *JAMA Psychiatry*, *73*(11), 1154–1162. https://doi.org/10.1001/jamapsychiatry.2016.2387

Śliwerski, A., & Bielawska-Batorowicz, E. (2019). Negative cognitive styles as risk factors for the occurrence of PMS and PMDD. *Journal of Reproductive and Infant Psychology*, *37*(3), 322–337. https://doi.org/10.1080/02646838.2018.1543943

Słopień, R., Milewska, E., Rynio, P., & Męczekalski, B. (2018). Use of oral contraceptives for management of acne vulgaris and hirsutism in women of reproductive and late reproductive age. *Przegląd Menopauzalny = Menopause Review*, *17*(1), 1–4. https://doi.org/10.5114/pm.2018.74895

Spalek, K., Loos, E., Schicktanz, N., Hartmann, F., de Quervain, D., Stier, C., & Milnik, A. (2019). Women using hormonal contraceptives show increased valence ratings and

memory performance for emotional information. *Neuropsychopharmacology*, *44*(7), 1258–1264. https://doi.org/10.1038/s41386-019-0362-3

Speed, T. J., Richards, J. M., Finan, P. H., & Smith, M. T. (2017). Sex moderates the effects of positive and negative affect on clinical pain in patients with knee osteoarthritis. *Scandinavian Journal of Pain*, *16*, 66–73. https://doi.org/10.1016/j.sjpain.2017.03.005

Steiner, M., Dunn, E., & Born, L. (2003). Hormones and mood: From menarche to menopause and beyond. *Journal of Affective Disorders*, *74*(1), 67–83. https://doi.org/10.1016/S0165-0327(02)00432-9

Stevens, J. S., & Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. *Neuropsychologia*, *50*(7), 1578–1593. https://doi.org/10.1016/j.neuropsychologia.2012.03.011

Stottinger, E., Rafetseder, E., Anderson, B., & Danckert, J. (2016). The picture morphing task – an efficient and quick means to measure updating. *Journal of Vision*, *16*(12), 172. https://doi.org/10.1167/16.12.172

Sundström Poromaa, I., & Gingnell, M. (2014). Menstrual cycle influence on cognitive function and emotion processing—From a reproductive perspective. *Frontiers in Neuroscience*, *8*, 380.

Sutherland, C. (2015). *A basic guide to Psychomorph*. https://aura.abdn.ac.uk/handle/2164/12696

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2019). *Using multivariate statistics* (Seventh edition). Pearson.

Takayama, E., Tanaka, H., Kamimoto, Y., Sugiyama, T., Okano, T., Kondo, E., & Ikeda, T. (2020). Relationship between a high Edinburgh Postnatal Depression Scale score and

premenstrual syndrome: A prospective, observational study. *Taiwanese Journal of Obstetrics and Gynecology*, *59*(3), 356–360. https://doi.org/10.1016/j.tjog.2020.03.003

Thaipisuttikul, P., Ittasakul, P., Waleeprakhon, P., Wisajun, P., & Jullagate, S. (2014). Psychiatric comorbidities in patients with major depressive disorder. *Neuropsychiatric Disease and Treatment*, *10*, 2097–2103. https://doi.org/10.2147/NDT.S72026

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Tsehay, M., Necho, M., & Mekonnen, W. (2020). The role of adverse childhood experience on depression symptom, prevalence, and severity among school going adolescents. *Depression Research and Treatment*, *2020*, e5951792. https://doi.org/10.1155/2020/5951792

Tsikriktsis, N. (2005). A review of techniques for treating missing data in OM survey research. *Journal of Operations Management*, *24*, 53–62. https://doi.org/10.1016/j.jom.2005.03.001

United Nations, & Social Affairs. (2019). *Contraceptive use by method 2019: Data booklet*. Department of Economic and Social Affairs, Population Division (2019). https://doi.org/10.18356/1bd58a10-en

van der Helm, E., Gujar, N., & Walker, M. P. (2010). Sleep deprivation impairs the accurate recognition of human emotions. *Sleep*, *33*(3), 335–342.

van Honk, J., & Schutter, D. J. L. G. (2007). Testosterone reduces conscious detection of signals serving social correction: Implications for antisocial behavior. *Psychological Science*, *18*(8), 663–667.

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

Venkateshan, S. (2023). Facial emotion recognition in women with symptoms of polycystic ovary syndrome [Unpublished masters thesis]. Lakehead University

von Dawans, B., Spenthof, I., Zimmer, P., & Domes, G. (2020). Acute psychosocial stress modulates the detection sensitivity for facial emotions. *Experimental Psychology*, *67*(2), 140–149. https://doi.org/10.1027/1618-3169/a000473

Wallen, K. (2017). Sexual differentiation of behavior in nonhuman primates. In D. W. Pfaff & M. Joëls (Eds.), *Hormones, brain and behavior (Third edition)* (pp. 225–245). Academic Press. https://doi.org/10.1016/B978-0-12-803592-4.00102-4

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037//0022-3514.54.6.1063

Weisenbach, S. L., Rapport, L. J., Briceno, E. M., Haase, B. D., Vederman, A. C., Bieliauskas, L. A., Welsh, R. C., Starkman, M. N., McInnis, M. G., Zubieta, J.-K., & Langenecker, S. A. (2014). Reduced emotion processing efficiency in healthy males relative to females. *Social Cognitive and Affective Neuroscience*, *9*(3), 316–325. https://doi.org/10.1093/scan/nss137

Williams, O. O. F., Coppolino, M., George, S. R., & Perreault, M. L. (2021). Sex differences in dopamine receptors and relevance to neuropsychiatric disorders. *Brain Sciences*, *11*(9), 1199. https://doi.org/10.3390/brainsci11091199

Wingenbach, T. S. H., Ashwin, C., & Brosnan, M. (2018). Sex differences in facial emotion recognition across varying expression intensity levels from videos. *PLoS ONE*, *13*(1), e0190634. https://doi.org/10.1371/journal.pone.0190634

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. https://scholar.google.com/scholar_lookup?title=Depression%20and%20Other%20Common%20Mental%20Disorders%3A%20Global%20Health%20Estimates&publication_year=2017&author=World%20Health%20Organization

Wright, S. L., Langenecker, S. A., Deldin, P. J., Rapport, L. J., Nielson, K. A., Kade, A. A., Own, L. S., Akil, H., Young, E. A., & Zubieta, J.-K. (2009). Gender specific disruptions in emotion processing in younger adults with depression. *Depression and Anxiety*, *26*(2), 182–189. https://doi.org/10.1002/da.20502

Yen, J.Y., Lin, H.C., Lin, P.C., Liu, T.L., Long, C.Y., & Ko, C.H. (2019). Early- and late-luteal-phase estrogen and progesterone levels of women with premenstrual dysphoric disorder. *International Journal of Environmental Research and Public Health*, *16*(22), 4352. https://doi.org/10.3390/ijerph16224352

Yonkers, K. A., O'Brien, P. M. S., & Eriksson, E. (2008). Premenstrual syndrome. *Lancet*, *371*(9619), 1200–1210. https://doi.org/10.1016/S0140-6736(08)60527-9

Yoon, K. L., Joormann, J., & Gotlib, I. H. (2009). Judging the intensity of facial expressions of emotion: Depression-related biases in the processing of positive affect. *Journal of Abnormal Psychology*, *118*(1), 223–228. https://doi.org/10.1037/a0014658

Zanotti, D., Kaier, E., Vanasse, R., Davis, J., Strunk, K., & Cromer, L. (2017). An examination of the test-retest reliability of the ACE-SQ in a sample of college athletes. *Psychological Trauma: Theory, Research, Practice, and Policy*, *10*. https://doi.org/10.1037/tra0000299

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

Zilioli, S., Caldbick, E., & Watson, N. V. (2014). Testosterone reactivity to facial display of emotions in men and women. *Hormones and Behavior*, *65*(5), 461–468. https://doi.org/10.1016/j.yhbeh.2014.04.006

Zimmerman, Y., Eijkemans, M. J. C., Coelingh Bennink, H. J. T., Blankenstein, M. A., & Fauser, B. C. J. M. (2014). The effect of combined oral contraception on testosterone levels in healthy women: A systematic review and meta-analysis. *Human Reproduction Update*, *20*(1), 76–105. https://doi.org/10.1093/humupd/dmt038

Appendix A

OCs and Facial Emotion Detection Search Terms

For the literature search attempting to identify all articles examining OCs and facial emotion detection, a search was conducted within the electronic databases *PubMed* and *PsychINFO.*

Within *PsychINFO* the search terms used were ("oral contracept*" OR "hormonal contracep*" OR "the pill" OR "contracep* pill") AND ("facial expression" OR "facial emotion" OR "emotion*" OR "emotion* face" OR "facial affect") AND ("detection" OR "recognition" OR "perception" OR "indetif*" OR "decoding" OR "discrimination" OR "processing").

Within *PubMed* the search terms used were ("oral contracept*" OR "hormonal contracep*" OR "the pill" OR "contracep* pill") AND ("facial expression" OR "facial emotion" OR "emotion*" OR "emotion* face" OR "facial affect") AND ("detection" OR "recognition" OR "perception" OR "indetif*" OR "decoding" OR "discrimination" OR "processing").

The *PsychINFO* search yielded 55 results, and the *PubMed* search yielded 82 results. After duplicate studies were removed, 103 results remained. Following abstract review, 90 of these studies were deemed not relevant, 2 were relevant review articles, and 11 were relevant empirical studies. One additional empirical study was found through one of the review articles. The total number of relevant empirical studies included in qualitative synthesis was 12.

PMS/PMDD and Facial Emotion Detection Search Terms

For the literature search attempting to identify all articles examining OCs and facial emotion detection, a search was conducted within the electronic databases *PubMed* and *PsychINFO.*

Within *PsychINFO* the search terms used were (mainsubject.Exact("facial expressions" OR “emotion recognition" OR “face perception") AND (mainsubject.exact("Premenstrual Dysphoric Disorder") OR mainsubject.exact("Premenstrual Syndrome"))

Within *PubMed* the search terms used were (("facial emotion"[Title] OR "emotion*"[Title]) AND ("processing"[Title] OR "detection"[Title] OR "recognition"[Title] OR "perception"[Title] OR “discrimination”[Title]])) AND ("oral contracept*"[Title] OR "contracept*"[Title] OR "hormonal contracept*"[Title]).

The *PsychINFO and PubMed* searches both yielded 5 results (10 in total). After duplicate studies were removed and abstract review, 3 relevant empirical articles were identified.

Appendix B

Online and Community Poster Recruitment Advertisement

HORMONES AND FACIAL EMOTION DETECTION STUDY

Researchers in the Department of Psychology at Lakehead University are conducting a study investigating mood and hormonal effects on facial emotion detection. We are looking for men and women who are 18 years of age or older, to participate in an online study which will take approximately 60 minutes to complete. The study will involve completing short questionnaires and participating in an interesting facial emotion detection task. Responses will collected through the Health Hormones and Behaviour Laboratory (HHABLAB) in the department of Psychology at Lakehead University. All responses will be anonymous and will be kept confidential.

Additionally, all participants can choose to be entered in a draw for the chance to win a $50 prepaid Visa gift card!

This study has been approved by Lakehead University's Research Ethics Board, (807) 343-8283 or research@lakeheadu.ca


Lakehead UNIVERSITY
HHAB Lab
PARTICIPANTS NEEDED
FOR A STUDY INVESTIGATING MOOD, HORMONES, AND FACIAL EMOTION DETECTION
We are looking for men and women to participate in an online study. The study will involve completing short questionnaires and participating in an interesting facial emotion detection task.
COMPENSATION
Lakehead Students will receive 1.5
SONA SYSTEMS
bonus marks towards an eligible Psychology course.
All participants will be entered into a draw for a $50
VISA
giftcard.
You are eligible if you:
Are 18+ years old (Or 16+ for LU students)
Own a laptop/desktop computer
Have 45min - 1h15min of time
TO PARTICIPATE:
(and more information)
For Lakehead Students:
For Non-Students:
Contacts:
Bianca Boboc - bboboc@lakeheadu.ca
Dr. Kirsten Oinonen
This study has been approved by Lakehead University's Research Ethics Board (File No. 1469489). (807) 343-8283/research@lakeheadu.ca

Appendix C

REB Approval

Research Ethics Board
t: (807) 343-8283
research@lakeheadu.ca

November 25, 2022

Principal Investigator: Dr. Kirsten Oinonen
Student Investigator: Bianca Boboc
Health and Behavioural Sciences\Psychology
Lakehead University
955 Oliver Road
Thunder Bay, ON P7B 5E1

Dear Dr. Oinonen and Bianca:

Re: Romeo File No: 1469489
Granting Agency: N/A
Agency Reference #: N/A

On behalf of the Research Ethics Board, I am pleased to grant ethical approval to your research project titled, "Mood and Hormonal Effects on Facial Emotion Detection: The Role of Depressive Symptoms, Oral Contraceptive Use, and Premenstrual Symptoms".

Ethics approval is valid until November 25, 2023. Please submit a Request for Renewal to the Office of Research Services via the Romeo Research Portal by October 25, 2023 if your research involving human participants will continue for longer than one year. A Final Report must be submitted promptly upon completion of the project. Access the Romeo Research Portal by logging into myInfo at:

https://erpwp.lakeheadu.ca/

During the course of the study, any modifications to the protocol or forms must not be initiated without prior written approval from the REB. You must promptly notify the REB of any adverse events that may occur.

Best wishes for a successful research project.

Sincerely,

Dr. Claudio Pousa
Chair, Research Ethics Board

/sw

Appendix D

Demographic and General Background Questionnaire

Demographics Questionnaire

Participant Code

The following questions will be used to make a unique participant code for you. This code will be used to link your answers from this survey to your answers on the Facial Emotion Detection Task, should you choose to complete both. This will ensure anonymity of your answers.

On what DAY were you born? (July 16, 1985 – 16 is the DAY) _____
What are the FIRST THREE letters of you mother's FIRST name? _____
What is the FIRST letter of your middle name? (if you do not have a middle name please put the letter 'X') ____

General

1. Age: ____
2. Biological Sex: Male ; Female ; Other
3. Gender: Male ; Female ; Indigenous or other cultural gender minority (e.g., two-spirit); Something else (e.g., gender fluid, non-binary), please specify: _____
4. Sexual Orientation: Straight; Gay; Bisexual: Asexual; Pansexual; Demi-sexual; Other: _____
5. What is your current relationship status? (Choose the one that best describes you)
   [ ] Married
   [ ] Common-law or living together
   [ ] One steady dating partner
   [ ] More than one dating partner
   [ ] One regular sexual partner
   [ ] More than one sexual partner
   [ ] Single
   [ ] Other (please specify): _____
6. Please choose the response that best represents your ethnic background.
   [ ] White, or Euro-American/Canadian
   [ ] South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)
   [ ] Chinese
   [ ] Black, Afro-Caribbean, or African-American/Canadian
   [ ] Filipino
   [ ] First Nations (North American Indian), Métis or Inuk (Inuit)
   [ ] Latin American
   [ ] Arab
   [ ] Southeast Asian (e.g., Vietnamese, Cambodian, Laotian, Thai, etc.)
   [ ] West Asian (e.g., Iranian, Afghan, etc.)
   [ ] Korean
   [ ] Japanese
   [ ] Other (please specify)
7. What is the highest level of education you have completed or are currently undertaking: High School; Diploma/Associates degree; Bachelor's degree; Post-graduate degree
8. What is the highest level of education your mother/parent has completed? High School; Diploma/Associates degree; Bachelor's degree; Post-graduate degree.

9. What is the highest level of education your father/parent has completed? High School; Diploma/Associates degree; Bachelor's degree; Post-graduate degree.
10. Is English your native language (the language you learned to speak first)? Yes; No, if not, please specify the language:____
11. Is English your primary language (the language you currently speak most fluently)? Yes; No, if not, please specify the language:____
12. What country do you call your home?
    [ ] Canada
    [ ] USA
    [ ] Brazil
    [ ] China
    [ ] France
    [ ] Korea
    [ ] India
    [ ] Iran
    [ ] Mexico
    [ ] Nigeria
    [ ] Pakistan
    [ ] Philippines
    [ ] Syria
    [ ] Vietnam
    [ ] Other, please specify: ___

13. What is your height? _____ (feet & inches) or _____ (cm)
14. What is your weight? _____(pounds) or _____ (kg)

Health History

15. Are you currently diagnosed with or being treated for depression? Yes; No; Maybe
16. Have you ever been diagnosed or treated for depression? Yes; No; Unsure.
17. Are you currently diagnosed with or being treated for any psychological/mental disorder (anxiety, bipolar disorder, schizophrenia, etc.)? Yes, please specify; No; Unsure.
18. Are you currently diagnosed with or being treated for attentional problems (ADHD, etc.)? Yes, please specify; No
19. Are you currently diagnosed with or being treated for any hormonal based disorders?
    Polycystic ovary syndrome. Yes; No
    Thyroid disorder. Yes; No
    Premenstrual dysphoric disorder. Yes; No
    Other: Yes, please specify: _____; No
20. Are you currently taking any hormonal medication NOT INCLUDING contraceptives (e.g., hormonal therapy for transitioning, hormone replacement therapy for menopause, progestin-only for endometrial cancer, etc.) Yes; No
21. Are you currently pregnant or lactating? Yes; No
22. Do you have 20/20 vision or vision corrected to 20/20 with corrective lenses? Yes; No
23. If you need glasses/contact lenses, are you currently wearing them? Yes; No; Not applicable

24. Have you ever had COVID-19? Yes, most recently in the past 3 months; Yes, most recently between 3-6 months ago; Yes, most recently between 6-12 months ago; Yes, most recently over 12 months ago; No.

Substance Use

When answering the following questions, keep in mind that ONE unit of alcohol is equal to a 1.5 oz distilled alcohol [e.g., vodka ,rum, whiskey], 5oz glass of wine, or a 12oz bottle/can of beer.

25. On average how many units of alcohol did you consume in the past week? <1; 1-4; 5-10; 11-15; 15-20; 20+
26. How many units of alcohol have you consumed in the past 5 hours? ___
27. How many units of alcohol have you consumed in the past 10 hours? ___
28. If you have consumed alcohol in the past 24 hours, how many hours ago was your last drink? ___
29. What is your average number of drinks per drinking occasion over the past 6 months? 0; 1-3; 4-7; 8-12; 12+.
30. What is your typical frequency of alcohol consumption over the past 6 months? Never; Once or twice a month or less; once or twice a week; three to four times a week; almost every day.
31. On average how many times do you use THC-dominant cannabis products (aka marijuana/weed) products a month (e.g., smoking or vaping them, eating edibles,)? <1; 1-4; 5-10; 11-20; 20-30; 30+
32. Have you used THC-dominant cannabis products in the past 10 hours? Yes; No; Maybe
33. Have you used any recreational drugs other than alcohol in the past 10 hours? Yes; No; Maybe
34. Have you used any prescription drugs (e.g., opiates, amphetamines) that affect your thinking or perception in the past 5 hours? Yes; No; Maybe
35. Are you currently experiencing symptoms of caffeine withdrawal? Yes; No
36. Are you currently experiencing symptoms of nicotine withdrawal? Yes; No

Sleep and Lifestyle

37. One average how many hours of sleep do you get per night? (number drop down menu)
38. How many hours of sleep did you get last night? (number drop down menu)
39. What best describes your sexual activity and likelihood of pregnancy?
    [ ] I have never had sexual intercourse.
    [ ] I have had sexual intercourse but am not presently sexually active.
    [ ] I am presently sexually active and trying to get pregnant.
    [ ] I am presently sexually active but it is not possible for pregnancy to result (e.g. same sex partner(s), infertile).
    [ ] I am presently sexually active, but I do not have the intent of pregnancy although it is physiologically possible (e.g. opposite sex partner(s)).

(Women Only – items 36 to 59)

40. Have you ever been diagnosed with Premenstrual Dysphoric Disorder? Yes, presently; Yes, previously; No

41. In the week prior to the beginning of your period, do you experience premenstrual symptoms or premenstrual syndrome? Never; Yes, mild; Yes, moderate; Yes, severe; Not sure

Oral Contraceptives

42. Are you CURRENTLY using any of the following hormonal contraceptives?
    Oral contraceptives (the pill)? Yes; No
    Hormonal patch? Yes; No
    Hormonal implant? Yes; No
    Hormonal intrauterine device (IUD)? Yes; No
    Other:
43. Have you EVER used any of the following hormonal contraceptives?
    Oral contraceptives (the pill)? Yes; No
    Hormonal patch? Yes; No
    Hormonal implant? Yes; No
    Hormonal intrauterine device (IUD)? ? Yes; No
    Other:

44. What best describes your history with using oral contraceptives? Current user; Previous user; Never user

45. If you previously used oral contraceptives, how long has it been since you last took oral contraceptives? _____ years and _____ months

If you are currently taking an oral contraceptive:

46. Please check the type of oral contraceptives you are currently taking?

- o Alesse
- o Alysena
- o Apri
- o Aviane
- o Brevicon 0.5/35
- o Brevicon 1/35
- o Cyclen
- o Demulen 30
- o Demulen 50
- o Indayo
- o Linessa
- o Loestrin 1.5/30
- o Lolo
- o Lutera
- o Marvelon
- o Micronor
- o Min Ovral
- o MinEstrin 1/20
- o Mirvala
- o Movisse
- o Mya
- o Norinyl
- o Norlestin 1/50
- o Ortho 0.5/35
- o Ortho 1/35
- o Ortho 10/11
- o Ortho 7/7/7
- o Ortho Micronor
- o Ortho-Cept
- o Ortho-Novum 1/50
- o Ovima
- o Ovral
- o Portia
- o Reclipsen
- o Seasonale
- o Seasonique
- o Select 1/35
- o Synphasic
- o Synphasic
- o Tri-Cyclen
- o Tri-Cyclen Lo
- o Tricira Lo
- o Triphasil
- o Triquilar
- o Triquilar
- o Yasmin
- o Yaz
- o Yaz plus
- o Zamine
- o Zarah
- o Other: _____

47. Why did you start taking oral contraceptives? (Check all that apply)
    [ ] Birth Control [ ] Treat acne [ ] For cycle regularity
    [ ] Due to a hormonal medical condition (Specify): _____
    [ ] I was taking another medication that could have produced birth defects
    [ ] Other: _____
48. For how long have taken oral contraceptives in total (i.e., the total amount of time you have taken on any/all brands of OCs)? _____ years and _____ months
49. For how long have taken the current oral contraceptives you are taking in total? _____ years and _____ months
50. What week of your current pill pack are you in? Week 1 active pills; Week 2 active pills; Week 3 active pills; Week 4 active pills (if applicable); Pill-free/inactive or sugar pill week (if applicable)
51. Are you taking an oral contraceptive that you take continuously so that you don't have a period once a month? Yes; No
52. How many pills have you missed (e.g., forgot to take at the usual time) in the past week of your pill pack? ___
53. Have oral contraceptives ever made you feel "not like yourself"? Yes; No; Maybe, please describe ___
54. Have oral contraceptives ever made you feel like your emotions are dulled? Yes; No; Maybe, please describe ___
55. Please describe if oral contraceptives have had any other negative effects on your life. _____

Menstrual Cycle

56. Which statement best describes your menstrual cycle right now?
    - I have not had my period in the past three months (but not due to menopause).
    - Some months I get my period and some months I don't.
    - I usually get my period every month, but it is irregular and I cannot predict when it will start.
    - I usually get my period within two to three days of when I expect it.
    - My period is like clockwork and the same number of days elapse between periods each month.
    - I am currently going through menopause, have gone through menopause, or have had a hysterectomy
    - I am currently pregnant
    - I am currently lactating
57. Are you currently menstruating today? Yes; No
58. If you are currently menstruating, for how many days have you had your period (including today)? ___
59. If you are not currently menstruating, what day are you at in your cycle (day 1 = the first day of your last period)? ___
60. Using the calendars below, please indicate the first day of your last menstrual period. If you are not completely sure, please estimate the day that you believe you started menstruating on. Also, please indicate the day that you believe your next period will start.

61. How confident are you that the day indicated above was the first day of your last period? 0%; 25%; 50%; 75%; 100%
62. How confident are you that the day indicated above is the day that you will next get your period? 0%; 25%; 50%; 75%; 100%
63. Which week of your menstrual cycle are you in?
    [ ] Week 1: My last period started within the past 7 days.
    [ ] Week 2: My last period started within the past 8-14 days.
    [ ] Week 3: My last period started more than 2 weeks ago, and I don't expect my next period to start for over a week.
    [ ] Week 4: I expect to get my period within the next 4 - 7 days.
    [ ] Week 4: I expect to get my period within the next 3 days.
64. How confident are you that the week/days indicated above is where you are at in your menstrual cycle? 0%; 25%; 50%; 75%; 100%

Self-Report Quick Inventory of Depressive Symptomatology (QIDS-SR$_{16}$)
Please circle the one response to each item that best describes you for the **past seven days.**

1. Falling Asleep:
   0 I never take longer than 30 minutes to fall asleep.
   1 I take at least 30 minutes to fall asleep, less than half the time.
   2 I take at least 30 minutes to fall asleep, more than half the time.
   3 I take more than 60 minutes to fall asleep, more than half the time.

2. Sleep During the Night:
   0 I do not wake up at night.
   1 I have a restless, light sleep with a few brief awakenings each night.
   2 I wake up at least once a night, but I go back to sleep easily.
   3 I awaken more than once a night and stay awake for 20 minutes or more, more than half the time.

3. Waking Up Too Early:
   0 Most of the time, I awaken no more than 30 minutes before I need to get up.
   1 More than half the time, I awaken more than 30 minutes before I need to get up.
   2 I almost always awaken at least one hour or so before I need to, but I go back to sleep eventually.
   3 I awaken at least one hour before I need to,and can't go back to sleep.

4. Sleeping Too Much:
   0 Most of the time, I sleep no longer than 7-8 hours/night, without napping during the day.
   1 Most of the time, I sleep no longer than 10 hours in a 24-hour period including naps.
   2 Most of the time, I sleep no longer than 12 hours in a 24-hour period including naps.
   3 Most of the time, I sleep longer than 12 hours in a 24-hour period including naps.

5. Feeling Sad:
   0 I do not feel sad

1 I feel sad less than half the time.
2 I feel sad more than half the time.
3 I feel sad nearly all of the time.

6. Decreased Appetite:
   0 There is no change in my usual appetite.
   1 I eat somewhat less often or lesser amounts of food than usual.
   2 I eat much less than usual and only with personal effort.
   3 I rarely eat within a 24-hour period, and only with extreme personal effort or when others persuade me to eat.

7. Increased Appetite:
   0 There is no change from my usual appetite.
   1 I feel a need to eat more frequently than usual.
   2 I regularly eat more often and/or greater amounts of food than usual.
   3 I feel driven to overeat both at mealtime and between meals.

8. Decreased Weight (Within the Last Two Weeks):
   0 I have not had a change in my weight.
   1 I feel as if I've had a slight weight loss.
   2 I have lost 2 pounds or more.
   3 I have lost 5 pounds or more.

9. Increased Weight (Within the Last Two Weeks):
   0 I have not had a change in my weight.
   1 I feel as if I've had a slight weight gain.
   2 I have gained 2 pounds or more.
   3 I have gained 5 pounds or more.

10. Concentration/Decision Making:
   0 There is no change in my usual capacity to concentrate or make decisions.
   1 I occasionally feel indecisive or find that my attention wanders.
   2 Most of the time, I struggle to focus my attention or to make decisions.
   3 I cannot concentrate well enough to read or cannot make even minor decisions.

11. View of Myself:
   0 I see myself as equally worthwhile and deserving as other people.
   1 I am more self-blaming than usual.
   2 I largely believe that I cause problems for others.
   3 I think almost constantly about major and minor defects in myself.

12. Feelings of Hopelessness
   0 I do not feel any more discouraged about the future than most people.
   1 I feel discouraged about the future.
   2 I feel a little hopeless about the future and have nothing to look forward to.
   3 I feel hopeless about the future and that things are unlikely improve.

13. General Interest:
    0 There is no change from usual in how interested I am in other people or activities.
    1 I notice that I am less interested in people or activities.
    2 I find I have interest in only one or two of my formerly pursued activities
    3 I have virtually no interest in formerly pursued activities.

14. Energy Level:
    0 There is no change in my usual level of energy.
    1 I get tired more easily than usual.
    2 I have to make a big effort to start or finish my usual daily activities (for example, shopping, homework, cooking or going to work).
    3 I really cannot carry out most of my usual daily activities because I just don't have the energy.

15. Feeling slowed down:
    0 I think, speak, and move at my usual rate of speed.
    1 I find that my thinking is slowed down or my voice sounds dull or flat.
    2 It takes me several seconds to respond to most questions and I'm sure my thinking is slowed.
    3 I am often unable to respond to questions without extreme effort.

16. Feeling restless:
    0 I do not feel restless.
    1 I'm often fidgety, wringing my hands, or need to shift how I am sitting.
    2 I have impulses to move about and am quite restless.
    3 At times, I am unable to stay seated and need to pace around.

Positive and Negative Affect Schedule (PANAS) – NA Scale

This scale consists of a number of words to describe different feelings and emotions. Read each item and indicate to **what extent you have felt this way over the past 2 weeks**.

	1	2	3	4	5
	Very slightly or not at all	A little	Moderately	Quite a bit	Extremely
Distressed					
Upset					
Guilty					
Scared					

Hostile					
Irritable					
Ashamed					
Nervous					
Jittery					
Afraid					

DSM-5-Based Screening for Premenstrual Symptoms (DSPMS)

Some women experience changes in mood and physical functioning during the week prior to their menstrual period.

As best as you can, please indicate the severity, level of impairment, and frequency of the following 11 symptoms during your PREMENSTRUAL PHASE over the past year.

Please rate the degree of severity of these symptoms.

Not at all, Mild, Moderate, Severe, Extremely severe.

To what extent do these symptoms cause impairment in work, school, usual social activities, or relationships with others?

Not at all, Mild, Moderately, Quite a bit, Extremely.

Number of menstrual cycles in which the symptom(s) have been experienced over the past 12 months.

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

I. Marked affective lability (e.g., mood swings, feeling suddenly sad or tearful, or increased sensitivity to rejection).

2. Marked irritability or anger or increased interpersonal conflicts.

3. Marked depressed mood, feelings of hopelessness or self- deprecating thoughts.

4. Marked anxiety, tension, and/or feelings of being keyed up or on edge.

5. Decreased interest in usual activities (e.g., work, school, friends, hobbies).

6. Subjective difficulty in concentration.

7. Lethargy, easy fatigability, or marked lack of energy.

8. Marked change in appetite; overeating; or specific food cravings.

9. Sleeping too much or too little.

10. A sense of being overwhelmed or out of control.

11. Physical symptoms such as breast tenderness or swelling, joint or muscle pain, a sensation of "bloating" or weight gain.

Positive and Negative Affect Schedule (PANAS)

This scale consists of a number of words to describe different feelings and emotions. Read each item and indicate to **what extent you feel this way right now**, that is, at the present moment.

	1	2	3	4	5
	Very slightly or not at all	A little	Moderately	Quite a bit	Extremely
Interested					
Distressed					
Excited					
Upset					
Strong					
Guilty					
Scared					
Hostile					
Enthusiastic					
Proud					
Irritable					
Alert					
Ashamed					
Inspired					
Nervous					
Determined					
Attentive					
Jittery					
Active					
Afraid					

Final Questionnaire

Behavioural Inhibition System/Behavioural Activation System (BIS/BAS)

24 items were used, retrieved from Carver and White (1994).

OC Side Effects Scale

Below is a list of symptoms that can be positive or negative. Please indicate the extent to which you have **experienced an increase in each symptom *after* starting to take oral contraceptives that you believe may be due to oral contraceptives.** Please also indicate whether the experience of the symptom has affected you in a positive (good) or negative (bad) way. If you have taken more than one type of oral contraceptive, please indicate if you experienced these symptoms during use of **any** of the types/brands.

Emotional	Not at all 0	Mild 1	Moderate 2	Strong 3	Extreme 4	Check if you are currently experiencing this symptom due to oral contraceptives	Was this symptom change positive or negative?
1. Slept more than usual						[ ]	+ / -
2. Slept less than usual						[ ]	+ / -
3. Disrupted sleep						[ ]	+ / -
4. Depression[a]						[ ]	+ / -
5. Sadness						[ ]	+ / -
6. More content/happy						[ ]	+ / -
7. Feelings of inferiority						[ ]	+ / -
8. More pessimistic						[ ]	+ / -
9. More irritable						[ ]	+ / -
10. More jealous						[ ]	+ / -
11. More sensitive to criticism						[ ]	+ / -
12. Less trust in partner (fidelity)						[ ]	+ / -
13. More moody[b]						[ ]	+ / -
14. Less moody						[ ]	+ / -
15. Lower self-esteem						[ ]	+ / -
16. Cried more than usual						[ ]	+ / -
17. More self-critical						[ ]	+ / -
18. More aggressive[a]						[ ]	+ / -
19. Less aggressive						[ ]	+ / -
20. Nervousness[a]						[ ]	+ / -

	OC Side Effects Scale

Physical	Not at all 0	Mild 1	Moderate 2	Strong 3	Extreme 4	Check if you are currently experiencing this symptom due to oral contraceptives	Was this symptom change positive or negative?
1. Nausea/vomiting						[ ]	+ / -
2. Weight gain						[ ]	+ / -
3. Increased appetite[a]						[ ]	+ / -
4. Decreased appetite						[ ]	+ / -
5. Weight loss						[ ]	+ / -
6. Headaches						[ ]	+ / -
7. Tiredness/fatigue[a]						[ ]	+ / -
8. Heavier periods[b]						[ ]	+ / -
9. Lighter periods[a]						[ ]	+ / -
10. Increased sex drive/arousal						[ ]	+ / -
11. Decreased sex drive/arousal						[ ]	+ / -
12. Positive mood change						[ ]	+ / -
13. Negative mood change						[ ]	+ / -
14. Swelling of breast/abdomen						[ ]	+ / -
15. Breast size increase[b]						[ ]	+ / -
16. Fewer menstrual cramps						[ ]	+ / -
17. More menstrual cramps						[ ]	+ / -
18. Dizziness/faintness[b]						[ ]	+ / -
19. Painful or tender breasts						[ ]	+ / -
20. Clearer complexion						[ ]	+ / -
21. Complexion problems (e.g., acne)[a]						[ ]	+ / -
22. Leg cramps[b]						[ ]	+ / -
23. Hot flashes[a]						[ ]	+ / -

Note. [a] representative of androgenic profile. [b] representative of estrogenic profile.

i. Please indicate the oral contraceptive(s) you were taking while experiencing any of the above symptoms – check all that apply.
[ ] Alesse [ ] Brevicon 0.5/35 [ ] Brevicon 1/35 [ ] Cyclen [ ] Demulen 30 [ ] Loestrin 1.5/30 [ ] Marvelon [ ] MinEstrin 1/20
[ ] Ortho-Cept [ ] Ortho 7/7/7 [ ] Ortho 10/11 [ ] Synphasic [ ] Tri-Cyclen [ ] Triphasil [ ] Triquilar [ ] Demulen 50
[ ] Min-Ovral [ ] Norlestin 1/50 [ ] Norinyl [ ] Ovral [ ] Ortho 1/35 [ ] Ortho-Novum 1/50
[ ] Ortho 0.5/35 [ ] Other: _____

ii. I believe that oral contraceptives affected my mood and emotions:

1	2	3	4	5
Very Negatively	Slightly Negatively	In No Way At All	Slightly Positively	Very Positively

iii. I believe that oral contraceptives affected my physical health:

1	2	3	4	5
Very Negatively	Slightly Negatively	In No Way At All	Slightly Positively	Very Positively

iv. I believe that oral contraceptives have affected my sexual functioning (i.e., arousal, sex drive)

1	2	3	4	5
Very Negatively	Slightly Negatively	In No Way At All	Slightly Positively	Very Positively

v. If you have experienced any symptoms of oral contraceptives that were not noted above, please note them here:

Toronto Alexithymia Scale (TAS-20)

20 items were used. For further information see Bagby et al. (1994).

Adverse Childhood Experiences Questionnaire (ACE-Q)

10 items were used, retrieved from Felitti et al. (1998).

Personality Assessment Inventory Validity Scales

26 items were used, retrieved from Morey (2007).

Email Use Question

1. Are you are interested in receiving emails about future studies conducted in the Health, Hormones, and Behaviour lab? Please note there would be no obligation to participate in any future studies and your response will not affect your submission to the $50 Visa card draw. Additionally, your email will not be connected to any data you have just provided in the current study.

    - Yes , I want to be contacted about future studies
    - No, I do not want to be contacted

Department of Psychology
Lakehead University
Email: bboboc@lakeheadu.ca

Appendix E

Letter of Information & Consent for Potential Participants

HORMONES AND FACIAL EMOTION DETECTION STUDY

Dear Potential Participant,

You are invited to participate in the *Hormones and Facial Emotion Detection* Study, a research study being conducted to investigate the links between mood, hormonal health, and facial emotion detection.
Taking part in this study is voluntary. Before you decide whether or not you would like to take part, please read this letter carefully to understand what is involved. After you have read the letter, please feel free to email us to ask any questions you may have.

PURPOSE
This study is being conducted by Bianca Boboc and Dr. Kirsten Oinonen of the Health, Hormones, and Behaviour Lab, Department of Psychology, Lakehead University. The purpose of the study is to better understand the relationship between social processes, such as facial emotion detection, and aspects of health, such as mood and hormones.

WHAT INFORMATION WILL BE COLLECTED?
Information regarding your general demographics, medical history, substance use, and sleep will be asked about. Additional questions will be asked to assess your mood, motivation, emotion detection abilities, and everyday functioning. Finally, your performance on a cognitive task assessing facial emotion detection will be measured. For female participants only, information regarding your contraceptive use and menstrual cycle will also be collected. All of this information is integral to investigating the purpose of this study.

WHAT IS REQUESTED OF ME AS A PARTICIPANT?
Within this study you will be requested to answer several questions and participate in a cognitive task over an online platform. You will first complete an initial demographics questionnaire. Following this you will complete a facial emotions detection task. Within this task you will see several pictures of neutral expression faces that will gradually morph to an emotion. You will be asked to indicate when you perceive the emotion and what emotion it is. Following the facial emotions task, you will be asked to complete another short questionnaire including various questions about your everyday functioning. In total this study should take between 45 minutes and an hour, but no longer than 1.5 hours.

WHAT ARE MY RIGHTS AS A PARTICIPANT?
Your participation is entirely voluntary. You may withdraw from the study or refuse to participate in any part of the study at any time without explanation or penalty. You may decline to answer any question. You also have the right to request a copy of research results once the study has been completed.

Department of Psychology
Lakehead University
Email: bboboc@lakeheadu.ca

WHAT ARE THE RISKS AND BENEFITS?

There are no known or anticipated risks from participating in this study. Some participants may feel uncomfortable answering certain personal questions, or may have new positive or negative thoughts about oneself after answering the questions. At the end of the study all participants will receive a debriefing form that will include mental health resources should participants wish to seek out support.
Potential benefits include developing a better understanding about psychological research, learning more about hormones, contraceptives, and emotions, and potentially gaining self-insight. You will also have your name entered into a draw for a chance to win a $50 pre-paid Visa gift card. Finally, Lakehead University students will receive 1.5 bonus points towards a participating Lakehead University psychology course.

HOW WILL MY CONFIDENTIALITY BE MAINTAINED?
All information collected in this study will be anonymous and confidential. A unique code number will be used to link the data from Survey Monkey to data from the Facial Emotion Detection Task and this code will be deleted once the data is linked. Researchers will not be able to identify your responses. Additionally, you will not be asked any identifying information, and you may choose to skip answering any questions. Survey instruments will not be labeled in any way that will make identifying you possible. Additionally, individual participants will not be able to be identified in published results (data will be published in aggregate form). Please note that the online survey tool used in the study, (Survey Monkey), is hosted by a server located in the USA. The US Patriot Act permits U.S. law enforcement officials, for the purpose of anti-terrorism investigation, to seek a court order that allows access to the personal records of any person without the person's knowledge. In view of this we cannot absolutely guarantee the full confidentiality and anonymity of your data. With your consent to participate in this study, you acknowledge this.

WHAT WILL MY DATA BE USED FOR:
Your data will be used to inform the main research purpose of this study, including the Masters thesis of Bianca Boboc. It may also be used to examine additional related exploratory research questions in the laboratory. There is no intention to commercialize the research findings. Your data will also be kept in the Health, Hormones, and Behaviour Lab for a minimum of 5 years, and may be used to inform future projects within the lab. There is also the intention to publish findings from this study in peer-reviewed journals.

WHERE WILL MY DATA BE STORED?
Please remember that data collected during this study is anonymous. Once data is linked the participant codes will be removed. All data will be stored within Lakehead University for a minimum of 5 years following completion of the project. Upon completion of this study, anonymous data may also be deposited in an online public repository/database to support the Tri-Agency policy on open data. Potential identifying information will not appear in any database, report, publication, or presentation resulting from the study.

Department of Psychology
Lakehead University
Email: bboboc@lakeheadu.ca

HOW CAN I RECEIVE A COPY OF THE RESEARCH RESULTS?
If you are interested in learning the outcomes of the study, please feel free to email any of the researchers involved in the study at any time. Eventually, the conclusions of this study will be

shared with the research community through seminars, conferences, presentations, and journal articles. Results are estimated to be available by the end of August 2023. Subsequent published papers will also be listed on the primary investigator's Lakehead University website, allowing participants to read peer-reviewed descriptions of the findings.

WHAT IF I WANT TO WITHDRAW FROM THE STUDY?
You may withdraw from the study at any time without loss of remuneration or any effect on your academic status; you may simply close your browser or stop responding.

RESEARCHER CONTACT INFORMATION:

Graduate Student Researcher:
Bianca Boboc, H.B.Sc.
M.A. Student
Health, Hormones and Behaviour Lab
Lakehead University
955 Oliver Road
Thunder Bay, Ontario P7B 5E1
email: bboboc@lakeheadu.ca

Faculty Researcher
Dr. Kirsten Oinonen Ph.D., C. Psych.
Professor
Health, Hormones and Behavour Lab
Lakehead University
955 Oliver Road
Thunder Bay, Ontario P7B 5E1
email: koinonen@lakeheadu.ca

RESEARCH ETHICS BOARD REVIEW AND APPROVAL:
This research study has been reviewed and approved by the Lakehead University Research Ethics Board. If you have any questions related to the ethics of the research and would like to speak to someone outside of the research team, please contact Sue Wright at the Research Ethics Board at 807-343-8283 or research@lakeheadu.ca.

Department of Psychology
Lakehead University
Email: bboboc@lakeheadu.ca

Consent Form for Potential Participants

MY CONSENT:
I agree to the following:

- ✓ I have read and understand the information contained in the Information Letter
- ✓ I agree to participate
- ✓ I understand the potential risks and benefits to the study
- ✓ That I am a volunteer and can withdraw from the study, and may choose not to answer any question
- ✓ That the data will be securely stored at Lakehead University for a minimum period of 5 years following completion of the research project
- ✓ I understand that the research findings will be made available to me upon request
- ✓ I will remain anonymous
- ✓ All of my questions have been answered

By consenting to participate, I have not waived any rights to legal recourse in the event of research-related harm.

I have read and agree to the above information and consent to proceed to the online survey [INSERT LINK]

Department of Psychology
Lakehead University
Email: bboboc@lakeheadu.ca

Debriefing Form

HORMONES AND FACIAL EMOTION DETECTION STUDY

We appreciate your participation in our study, and thank you for spending your time to help us with our research. The purpose of this study was to investigate how mood and hormones influence facial emotion detection. In particular, we wanted to see how depressive symptoms, oral contraceptive use, and premenstrual symptoms (PMS) affect facial emotion detection. Previous research has suggested that facial emotion detection is influenced by a variety of factors including sex, presence or absence of certain diagnoses (e.g., depression, hormonal disorders), and hormone levels. For the present study, we have formed hypotheses based on this past literature, such as that women may be better at detecting some emotions. Please see the references below if you are interested in reading more about this. Also, in order to learn more from this experience, you may want to consider the following question: Why do you think we asked about your sleep and drug/alcohol use? How might the researchers use this information in this study?

In case you have any concerns about your mood and would like to see a mental health professional, we have provided you with a list of such resources on the attached sheet.

Please keep the study details confidential and do not discuss your experiences in the study with other students until the end of the school year. If participants have prior knowledge of the task it could influence the results, and the data we collect would not be useable. Please also keep the details of this feedback form confidential.

Should you have further questions, do not hesitate to contact Bianca Boboc or Dr. Kirsten Oinonen, using the information listed below. This study was approved by the Lakehead University Research Ethics Board and they can also be contacted about any concerns (807-343-8283 or research@lakeheadu.ca).

We hope that you have enjoyed participating in our study and thank you very much for your assistance. As noted on the consent form, you will receive a summary of the results of the study at its completion if you have indicated an interest.

Principal Investigators:

Graduate Student Researcher:
Bianca Boboc, H.B.Sc.
M.A. Student
Health, Hormones and Behaviour Lab
Lakehead University
955 Oliver Road
Thunder Bay, Ontario P7B 5E1
email: bboboc@lakeheadu.ca

Faculty Researcher
Dr. Kirsten Oinonen Ph.D., C. Psych.
Professor
Health, Hormones and Behavour Lab
Lakehead University
955 Oliver Road
Thunder Bay, Ontario P7B 5E1
email: koinonen@lakeheadu.ca

Mental Health Resource Sheet

Sometimes people can feel upset when thinking about their mood. Thus, it is possible that something occurred during your participation in the study that may have upset you. If you feel as though you would like to talk to a mental health practitioner for any reason please consider the resources listed below:

- Lakehead University Health and Counseling Centre: (807) 343-8361
- Family Services Thunder Bay: (807) 343-6100
- Emergency services are available at the Thunder Bay Health Sciences Centre
- Thunder Bay Crisis Response phone line (24 hours): (807) 346-8282.

The following are some references in case you are interested in reading more about research that is related to the study that you just participated in:

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: A review. *Australian & New Zealand Journal of Psychiatry*, *44*(8), 681-696. https://pubmed.ncbi.nlm.nih.gov/20636189/

Hamstra, D. A., De Rover, M., De Rijk, R. H., & Van der Does, W. (2014). Oral contraceptives may alter the detection of emotions in facial expressions. *European Neuropsychopharmacology*, *24*(11), 1855-1859. https://pubmed.ncbi.nlm.nih.gov/25224104/

Rubinow, D. R., Smith, M. J., Schenkel, L. A., Schmidt, P. J., & Dancer, K. (2007). Facial emotion discrimination across the menstrual cycle in women with premenstrual dysphoric disorder (PMDD) and controls. *Journal of Affective Disorders*, *104*(1-3), 37-44. https://pubmed.ncbi.nlm.nih.gov/17367867/

Appendix F

Descriptive Supplementary Data

Depression Groups

Table F1

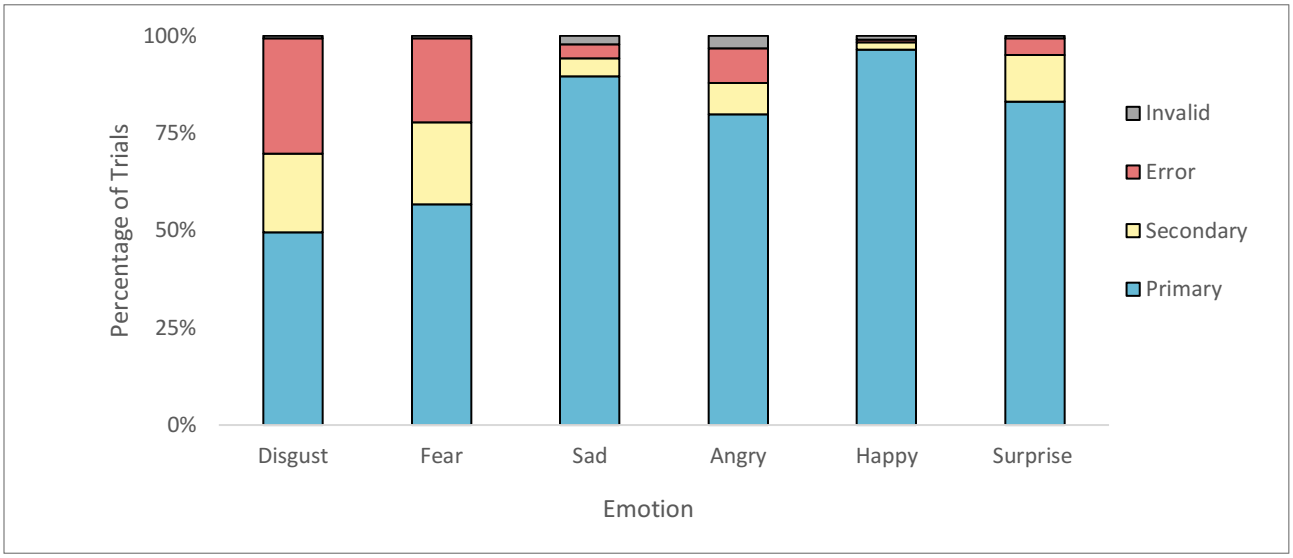
Unadjusted Means (SDs) of Average Image Number at Detection per Emotion for Low and High Female Depression Score Groups

Emotion	Mean (*SD*) of Image Number at Detection (Unadjusted)	
	Low depression (*n* = 49)	High depression (*n* = 52)
Fear	9.62 (1.76)	9.06 (1.62)
Sad	9.14 (1.2)	8.92 (1.91)
Angry	10.07 (1.77)	9.23 (1.72)
Disgust	8.45 (1.38)	8.56 (1.8)
Happy	7.79 (1.79)	6.55 (1.97)
Surprise	7.12 (1.36)	6.45 (1.74)

Note. Mean calculations are unadjusted for missing data and contain only image number at detection for primary and secondary trials (i.e., error trials are excluded). Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection on trials in which the final emotion was detected.

Figure F1

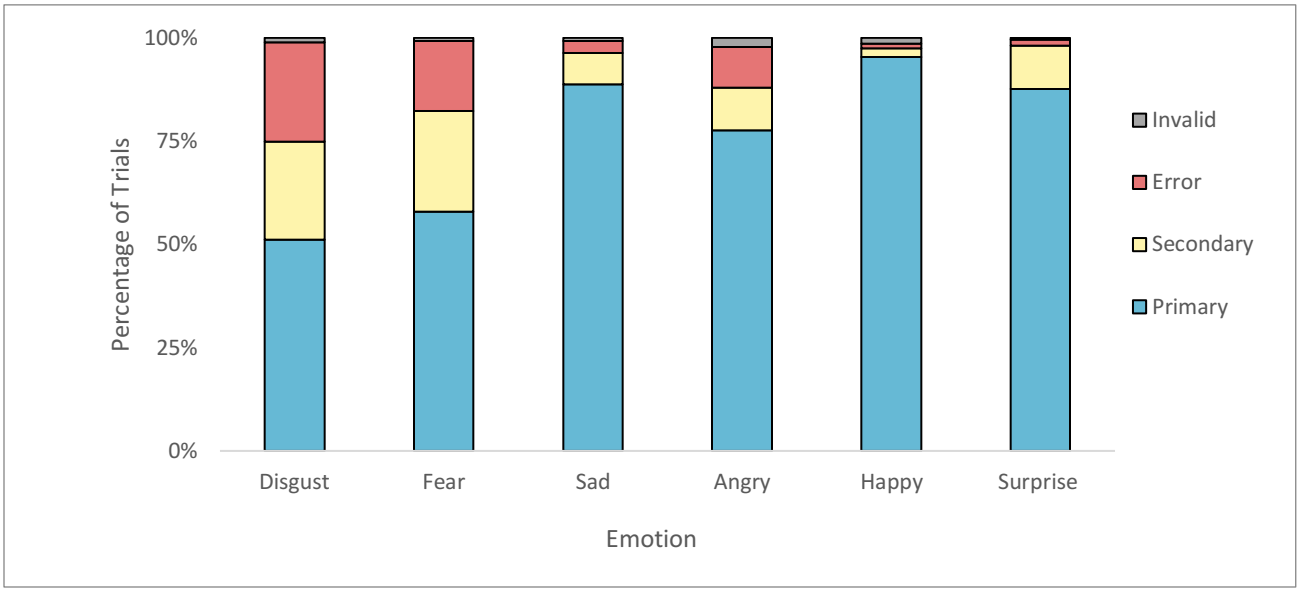
Trial Types per Emotion for Low Depression Score Groups



Note. The percentage of trial type based on emotion for the low depression score group. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

Figure F2

Trial Types per Emotion for High Depression Score Groups



Note. The percentage of trial type based on emotion for the high depression score group. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

OC Users, FC Women, and Men

Table F2

Unadjusted Means (SDs) of Average Image Number at Detection per Emotion for OC Users, FC Women, and Men

Emotion	Mean (*SD*) of Image Number at Detection (Unadjusted)		
	OC users (*n* = 37)	FC women (*n* = 72)	Men (*n* = 35)
Fear	9.22 (1.61)	8.91 (1.62)	9.69 (1.58)
Sad[t]	9.21 (1.52)	8.53 (1.71)	8.81 (1.35)
Angry	9.65 (1.79)	9.23 (1.84)	9.34 (1.71)
Disgust	8.98 (1.46)	9.16 (1.54)	8.74 (1.35)
Happy	7.56 (1.88)	6.33 (1.76)	6.55 (1.9)
Surprise	6.86 (1.63)	6.63 (1.67)	7.03 (1.19)

Note. Mean calculations are unadjusted for missing data and contain only image number at detection for primary and secondary trials (i.e., error trials are excluded). Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection on trials in which the final emotion was detected.

Figure F3

Trial Types per Emotion for OC Users

Note. The percentage of trial type based on emotion for OC using women. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

Figure F4

Trial Types per Emotion for FC Women

Note. The percentage of trial type based on emotion for FC women. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

Figure F5

Trial Types per Emotion for Men

Note. The percentage of trial type based on emotion for the high depression score group. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

PMDD Groups

Table F3

Unadjusted Means and SD of Average Image Number at Detections per Emotion for No/Minimal PMDD, Mild PMDD, and Moderate-Severe PMDD Groups

Emotion	Mean (*SD*) of Image Number at Detection (Unadjusted)		
	No/Minimal PMDD (*n* = 30)	Mild PMDD (*n* = 34)	Moderate-Severe PMDD (*n* = 8)
Fear	9.32 (1.93)	8.5 (1.3)	9.07 (1.3)
Sad	8.83 (1.95)	8.31 (1.58)	8.31 (1.28)
Angry	9.33 (2.02)	9.08 (1.84)	9.5 (1.09)
Disgust	9.03 (1.19)	9.28 (1.85)	9.15 (1.35)
Happy	6.17 (1.74)	6.36 (1.87)	6.78 (1.4)
Surprise	6.76 (1.67)	6.42 (1.77)	7.03 (1.24)

Note. Mean calculations are unadjusted for missing data and contain only image number at detection for primary and secondary trials (i.e., error trials are excluded). Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection on trials in which the final emotion was detected.

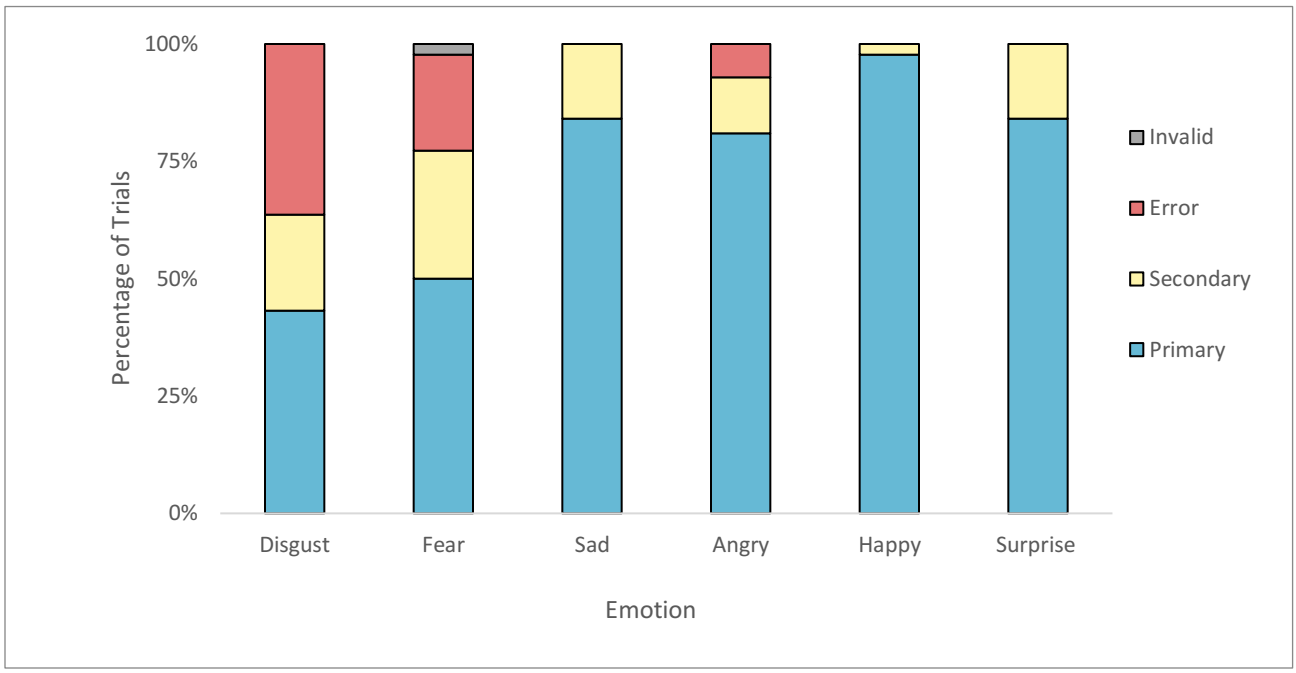
Figure F6

Trial Types per Emotion for No/Minimal PMDD Group

Note. The percentage of trial type based on emotion for the no pmdd group. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

Figure F7

Trial Types per Emotion for Mild PMDD Group



Note. The percentage of trial type based on emotion for the moderate PMDD group. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

Figure F8

Trials Type per Emotion Moderate-Severe PMDD

Note. The percentage of trial type based on emotion for the high MDD group. Primary = detected the correct emotion directly; Secondary = detected an incorrect emotion, then ultimately detected the correct emotion; Error = detected an incorrect emotion, and never detected the correct emotion; Invalid = Trial not analyzable.

Appendix G

Supplemental Analyses

Depression Diagnosis IV MANCOVA Tables

Table G1

Unadjusted Means (SDs) of Image Number at Detection per Emotion for Female Depression (No vs. Yes) Groups

Emotion	Mean (*SD*) of Image Number at Detection	
	No Depression (*n* = 85)	Yes Depression (*n* = 32)
Fear	10.65 (2.39)	10.07 (2.59)
Sad	9.05 (1.91)	8.77 (1.44)
Angry	10 (2.09)	9.94 (2.02)
Disgust	11.06 (2.24)	10.74 (2.35)
Happy	7.02 (2.00)	6.66 (1.82)
Surprise	7.12 (1.87)	6.53 (1.54)

Note. Lower scores indicate earlier (i.e., requiring lower intensity of emotion) detection. Variables used in the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The means are unadjusted for covariates, but all analyses controlled for hours of sleep last night.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table G2

ANCOVA Results for Image Number at Detection per Emotion for Female Depression (No vs. Yes) Groups

Emotion	*df*	*F*	*p*	η^2
Fear	2, 98	1.635	.204	.014
Sad	2, 98	0.751	.388	.007
Angry	2, 98	0.07	.792	.001
Disgust	2, 98	0.394	.532	.003
Happy	2, 98	1.096	.297	.010
Surprise	2, 98	3.353	.070[t]	.029

Note. Results of the follow-up univariate ANCOVAs testing group differences in Image Number at Detection for individual emotions. All analyses controlled for hours of sleep last night.
[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table G3

Unadjusted Means (SDs) for Percentage of Incorrect Responses per Emotion for Female Depression (No vs. Yes) Groups

Emotion	Mean (*SD*) of Percentage of Incorrect Responses	
	Low depression (*n* = 49)	High depression (*n* = 52)
Fear	15.45 (11.99)	12.99 (12.39)
Sad	3.01 (5.88)	2.5 (4.48)
Angry	6.31 (9.39)	7.66 (11.15)
Disgust	22.41 (15)	21.55 (16.87)
Happy	0.95 (3.68)	1.41 (4.19)
Surprise	4.9 (7.12)	4.32 (6.56)

Note. Lower scores indicate a lower percentage of errors when detecting trials with the identified emotion. Variables pertaining to the negative emotion MANCOVA are shown above the dotted line. The full table shows the dependent variables used within the overall emotion MANCOVA. Significance values reflect the results of follow-up ANCOVAs. The data here are unadjusted for covariates, but all analyses controlled for hours of sleep last night.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table G4

ANCOVA Results for Percentage of Incorrect Responses per Emotion for Female Depression (No vs. Yes) Groups

Emotion	*df*	*F*	*p*	η^2
Fear	2, 98	0.917	.340	.008
Sad	2, 98	0.148	.701	.001
Angry	2, 98	0.611	.436	.005
Disgust	2, 98	0.003	.958	.000
Happy	2, 98	0.428	.514	.004
Surprise	2, 98	0.225	.636	.002

Note. Results of the follow-up univariate ANCOVAs testing group differences in Image Number at Detection for individual emotions. All analyses controlled for hours of sleep last night.
$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Percentage of Correct Responses Depression Groups

Differences in Percentage of Correct Responses was examined as an additional variable encompassing intensity and accuracy. That is, a higher Percentage of Correct Responses is representative of both earlier detection (i.e., lower intensity) and less incorrect responses (i.e., higher accuracy), and was included as an overall measure of performance on the FEDT.

A two-group (low vs. high depression) MANCOVA (with follow-up ANCVOAS) was conducted with a DV of Percentage of Correct Responses for all (i.e., fear, sad, angry, disgust, happy, and surprise) emotion trials. Hours of sleep the night before testing was included as a covariate.

Table G5 contains the unadjusted means and *SD*s of scores for the MANCOVA. Visual examination of the Percentage of Correct Responses means revealed that they were all in the same direction such that the low depression group had a lower Percentage of Correct Responses than the high depression group. Despite this, the MANCOVA testing Percentage of Correct Responses for overall emotions was non-significant, $F(6, 101) = 1.536$, $p = .175$, $\eta^2 = .090$. This suggests that low depression and high depression groups do not significantly differ in the number of correct responses they provide.

Table G6 contains the follow-up ANCOVA results. The ANCOVAs revealed that the low depression group had a significantly higher Percentage of Correct Responses than the high depression group on surprise emotions ($p = .009$, $\eta^2 = .068$). This suggests that the low depression group are providing more correct responses on surprise emotions.

Table G5

Unadjusted Means and SD of Percentage of Correct Responses per Emotion for Low and High Female Depression Score Groups

Emotion	Mean (*SD*) of Percentage of Correct Responses	
	Low depression (*n* = 49)	High depression (*n* = 52)
Fear	34.76 (14.01)	38.4 (14.8)
Sad	45.4 (11.85)	47.07 (11.63)
Angry	38.5 (13.29)	40.44 (13.39)
Disgust	27.7 (10.67)	30.39 (12.04)
Happy	60.75 (13.4)	61.6 (12.41)
Surprise**	56.59 (11.04)	62.41 (12)

Note. Lower scores indicate less correct responses when detecting trials with the identified emotion. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use.
$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table G6

ANCOVA Results for Percentage of Correct Responses per Emotion for Low and High Female Depression Score Groups

Percentage of Correct Responses				
Emotion	*df*	*F*	*p*	η^2
Fear	1, 101	2.272	.135	.023
Sad	1, 101	0.747	.390	.008
Angry	1, 101	0.77	.382	.008
Disgust	1, 101	1.498	.224	.015
Happy	1, 101	0.222	.639	.002
Surprise	1, 101	7.193	.009**	.068

Note. Results of the follow-up univariate ANCOVAs testing group differences in Percentage of Correct Responses for individual emotions. All analyses controlled for hours of sleep last night.
$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$

Percentage of Correct Responses OC users, FC women, and Men

A two-group (OC users, FC women) and a three-group (OC users, FC women, men) MANCOVA (with follow-up ANCVOAS) was conducted with a DV of Percentage of Correct Responses for all (i.e., fear, sad, angry, disgust, happy, and surprise) emotion trials. Hours of sleep the night before testing was included as a covariate for all MANCOVAs.

Table G7 contains the unadjusted means and *SD*s of scores for the MANCOVA. Visual examination of the Percentage of Correct Responses means revealed that FC women had a higher Percentage of Correct Responses than OC users and men across all emotions. The two group (OC users, FC women) MANCOVA testing Percentage of Correct Responses for overall emotions was significant, $F(6, 108) = 2.564$, $p = .024$, $\eta^2 = .133$. However, for the three-group (OC users, FC women, men) MANCOVA testing Percentage of Correct Responses, the multivariate effect was not significant, $F(12, 140) = 1.473$, $p = .134$, $\eta^2 = .063$. These findings suggest that OC users, FC women differ in their percentage of their overall responses to all emotions that were correct, however OC users, FC women, and men do not significantly differ across all emotions.

Table G8 contains the follow-up ANCOVA results. ANCOVAs revealed that OC users had a significantly lower Percentage of Correct Responses than FC women on disgust emotions, *(*$p = .017$, $\eta^2 = .04$), and happy emotions, *(*$p = .004$, $\eta^2 = .078$). The three group ANCOVAs revealed that OC users, FC women, and men differ in their Percentage of Correct Responses for disgust, *(*$p = .042$, $\eta^2 = .046$) and happy emotions *(*$p = .012$, $\eta^2 = .063$). Pairwise comparisons revealed that OC users provided less correct responses than FC women on disgust ($p = 0.49$) and happy ($p = 0.009$). emotions. This suggests that FC women provide more correct responses than OC users on disgust and happy emotions.

Table G7

Unadjusted Means and SD of Percentage of Correct Responses per Emotion for Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men

Emotion	Mean (*SD*) of Percentage of Correct Responses		
	OC users (*n* = 37)	FC women (*n* = 72)	Men (*n* = 35)
Fear	36.76 (14.82)	39.32 (14.75)	34.01 (15.06)
Sad	44.65 (10.36)	48.58 (11.68)	45.07 (9.78)
Angry	38.06 (13.62)	40.89 (13.32)	37.81 (13.14)
Disgust *	26.03 (12.43)[y]	31.87 (10.18)[y]	27.97 (14.04)
Happy **	55.84 (13.81)[y]	64.45 (11.99)[y]	61.32 (12.87)
Surprise	59.88 (12.31)	61.37 (11.01)	58.72 (10.36)

Note. Lower scores indicate less correct responses when detecting trials with the identified emotion. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table G8

ANCOVA Results for Percentage of Correct Responses per Emotion for Oral Contraceptive (OC) Users, Free-cycling (FC) Women, and Men

Percentage of Incorrect Responses								
	2 group (OC users vs. FC women)				3 group (OC users vs. FC women vs. Men)			
Emotion	*df*	*F*	*p*	η^2	*df*	*F*	*p*	η^2
Fear	1, 108	0.248	.619	.002	2, 140	0.961	.385	.014
Sad	1, 108	1.545	.217	.015	2, 140	1.299	.276	.019
Angry	1, 108	0.466	.496	.004	2, 140	0.406	.667	.006
Disgust	1, 108	5.938	.017*	.054	2, 140	3.255	.042*	.046
Happy	1, 108	8.924	.004**	.078	2, 140	4.557	.012*	.063
Surprise	1, 108	0.048	.827	.000	2, 140	0.242	.785	.004

Note. Results of the follow-up univariate ANCOVAs testing group differences in Percentage of Correct Responses for individual emotions. The ANCOVA results for the 2 group (OC users vs. FC women) analyses are presented on the left, and for the 3 group analyses are presented on the right.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Percentage of Correct Responses PMDD Groups

A three-group (no/minimal PMDD, mild PMDD, moderate-severe PMDD) MANCOVA (with follow-up ANCVOAS) was conducted with a DV of Percentage of Correct Responses for all (i.e., fear, sad, angry, disgust, happy, and surprise) emotion trials. Hours of sleep the night before testing and typical alcohol intake was included as a covariate for all MANCOVAs.

Table G9 contains the unadjusted means and *SD*s of scores for the MANCOVA. The MANCOVA examining Percentage of Correct Responses across all emotions was significant, $F(12, 124) = 1.963$, $p = .033$, $\eta^2 = .160$. This finding suggests that PMDD groups differ in their percentage of their overall responses to all emotions that were correct.

Table G10 contains the follow-up ANCOVA results. ANCOVAs revealed that PMDD groups differed in their Percentage of Correct Responses for disgust emotions ($p = .025$, $\eta^2 = .105$). A non-significant trend also emerged in their Percentage of Correct Responses for sad emotions ($p = .079$, $\eta^2 = .074$). Pairwise comparisons revealed a trend toward the moderate-severe PMDD group providing more correct response on disgust trials than the mild PMDD group ($p = .058$). This suggests that those with more severe PMDD symptoms provide more correct responses on disgust emotions.

Table G9

Unadjusted Means and SD of Percentage of Correct Responses per Emotion for No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups

	Mean (*SD*) of Percentage of Correct Responses		
Emotion	No/Minimal PMDD (*n* = 30)	Mild PMDD (*n* = 34)	Moderate-Severe PMDD (*n* = 8)
Fear	36.78 (14.6)	42.07 (15.16)	37.5 (13.21)
Sad	46.13 (13.31)	50.2 (10.51)	51.04 (8.91)
Angry	39.7 (12.75)	43 (14.32)	36.67 (10.76)
Disgust*	34.44 (7.76)	28.32 (11.14)[y]	36.88 (10.17)[y]
Happy	64.81 (12.83)	64.85 (11.98)	61.46 (9.36)
Surprise	59.3 (10.52)	63.64 (11.84)	59.79 (8.23)

Note. Lower scores indicate less correct responses when detecting trials with the identified emotion. Significance values reflect the results of follow-up ANCOVAs. The data is unadjusted for covariates, but all analyses controlled for hours of sleep last night and typical alcohol use. [x] Group differences between the indicated group and the other two groups. [y] Group differences between the two indicated groups.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table G10

ANCOVA Results for Percentage of Correct Responses per Emotion between No/Minimal, Mild, and Moderate-Severe Provisional Premenstrual Dysphoric Disorder (PMDD) Groups

Percentage of Correct Responses				
Emotion	*df*	*F*	*p*	η^2
Fear	2, 71	0.883	.418	.026
Sad	2, 71	2.645	.079[t]	.074
Angry	2, 71	0.816	.447	.024
Disgust	2, 71	3.882	.025*	.105
Happy	2, 71	0.171	.843	.005
Surprise	2, 71	0.891	.415	.026

Note. Results of the follow-up univariate ANCOVAs testing group differences in Percentage of Correct Responses for individual emotions. The ANCOVA results for all participant (across the menstrual cycle) analyses are presented on the left, and for only the participants in the premenstrual phase are presented on the right.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Appendix H

Error Biases

Error biases were examined at the trial type level by looking at the kind of incorrect responses participants were making. That is, for the six emotion trial types (disgust, fear, sad, angry, happy, surprise), the percentage of each type of incorrect response was calculated out of the total possible valid responses (i.e., excluding invalid trials). Repeated-measures MANOVAS were run for each emotion trial type to examine interaction, and multivariate effects.

Depression Group Biases

A two-group (low depression vs. high depression) repeated-measures MANOVA was run to compare group differences in the type of error responses (the percentage of fear, sad, angry, happy, and surprise responses) on disgust trials (i.e., the percentage of non-disgust responses). Five identical repeated-measures MANOVAS were run for the other trial types (fear, sad, angry, happy, and surprise trials).

The unadjusted means and *SD*s of scores for Percentage of Incorrect Responses per trial type are presented in Table H1. The results of all the repeated-measures MANOVAs are shown in Table H2. The multivariate effects were not significant for any of the emotion trial types.

The distribution of the percentage of incorrect responses per trial type are shown in Figure H1. The univariate analyses testing whether the type of incorrect emotion per trial type differed based on group were all non-significant. This suggests that low and high depression groups do not differ in the type of errors they are making (interaction effect), or the rate of each unique type of error (univariate), across any of the trial types.

Table H1

Unadjusted Means and SD of Percentage of Incorrect Responses Per Trial Type for Low and High Female Depression Score Groups

Trial Emotion	Response Emotion	Mean (*SD*) of Percentage of Responses	
		Low depression (*n* = 49)	High depression (*n* = 52)
Disgust	Disgust	16.55 (6.46)	18.18 (7.29)
	Fear	0.41 (1.19)	0.75 (2.12)
	Sad	2.04 (4.55)	2.46 (3.97)
	Angry	11.11 (8.11)	9.67 (8.18)
	Happy	0.12 (0.73)	0.12 (0.51)
	Surprise	0.06 (0.32)	0 (0)
Fear	Disgust	0.24 (1.03)	0.81 (2.28)
	Fear	20.69 (8.52)	22.98 (8.96)
	Sad	3.12 (4.66)	3.77 (4.24)
	Angry	0.02 (0.14)	0.25 (1.2)
	Happy	0.18 (0.7)	0.37 (1.37)
	Surprise	5.27 (5.16)	4 (4.99)
Sad	Disgust	0.45 (1.47)	0.5 (1.57)
	Fear	0.08 (0.4)	0.48 (1.63)
	Sad	26.61 (7.5)	28.02 (7.2)
	Angry	0.76 (2.48)	0.46 (1.34)
	Happy	0.12 (0.73)	0.21 (0.91)
	Surprise	0 (0)	0.02 (0.14)

Trial Emotion	Response Emotion	Mean (*SD*) of Percentage of Responses	
		Low depression (*n* = 49)	High depression (*n* = 52)
Angry	Disgust	1.66 (3.29)	2.46 (4.79)
	Fear	0.23 (1.2)	0.61 (2.35)
	Sad	0.67 (2.12)	1.21 (3.35)
	Angry	22.65 (8.37)	23.93 (8.11)
	Happy	0.16 (0.72)	0 (0)
	Surprise	0.33 (1.26)	0.08 (0.44)
Happy	Disgust	0.06 (0.24)	0.04 (0.28)
	Fear	0.02 (0.14)	0 (0)
	Sad	0 (0)	0.15 (0.87)
	Angry	0.08 (0.57)	0 (0)
	Happy	35.98 (7.64)	36.37 (7.86)
	Surprise	0.4 (1.59)	0.33 (1.71)
Surprise	Disgust	0.16 (0.75)	0.04 (0.28)
	Fear	0.69 (2.73)	0.25 (1.27)
	Sad	0 (0)	0 (0)
	Angry	0 (0)	0 (0)
	Happy	2.82 (4.2)	1.96 (3.38)
	Surprise	33.83 (6.85)	37.17 (6.68)

Note. Percentages are calculated based on total valid responses.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table H2

Repeated-Measures MANOVA Results for Type of Incorrect Responses per Trial Type for Low and High Female Depression Score Groups

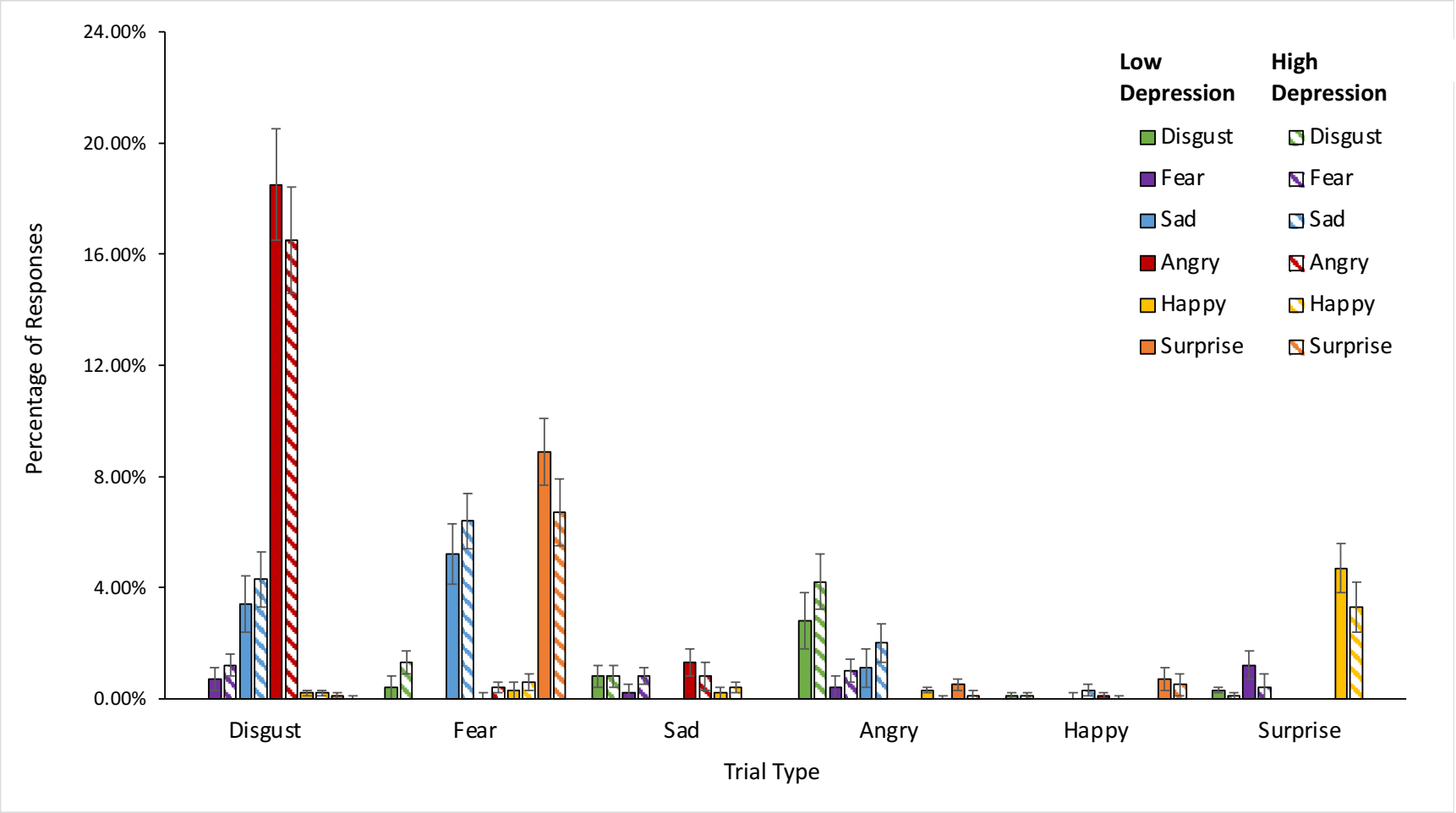
	Multivariate Effect (Depression Group * Type of Incorrect Response)			
Trial Type (Correct Emotion)	*df*	*F*	*p*	η^2
Disgust	4, 96	0.508	.730	.021
Fear	4, 96	0.895	.470	.036
Sad	4, 96	0.783	.539	.032
Angry	4, 96	1.092	.365	.044
Happy	4, 96	0.581	.677	.024
Surprise	4, 96	1.342	.265	.040

Note. Results of the repeated-measures MANOVAs testing group differences in Type of Incorrect Responses per Trial Type.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure H1

Percentage of Incorrect Responses Per Trial Type for Low and High Female Depression Score Groups



Note. Error bars represent standard error. Only incorrect responses are shown. Solid bars represent the low depression and striped bars represent the high depression group. There were no significant effects. The most common errors per trial type are as follows: Disgust trial – anger error; Fear trial – surprise error; Sad trial – anger error; Angry trial – disgust error; Happy trial – surprise error; Surprise trial – happy error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

OC Users and FC Women Biases

A two-group (OC users vs. FC women) repeated-measures MANOVA was run to compare group differences in the type of incorrect emotion responses (the percentage of fear, sad, angry, happy, and surprise responses) on disgust trials (i.e., the percentage of non-disgust responses). Five identical repeated-measures MANOVAS were run for the other trial types (fear, sad, angry, happy, and surprise trials).

The unadjusted means and *SD*s of scores for Percentage of Incorrect Responses per trial type are presented in Table H3. The results of all the repeated-measures MANOVAs are shown in Table H4. The multivariate effects were not significant for any of the emotion trial types. See Table H3 for the unadjusted means and *SD*s of scores for the error biases MANOVAs.

The distribution of the percentage of incorrect responses per trial type are shown in Figure H2. The univariate analyses testing whether the type of incorrect emotion per trial type differed based on group were all non-significant. However, three non-significant trends suggest that OC users and FC women may differ in the rate of angry responses to disgust trials, $F(1, 107) = 3.615$, $p = .060$, $\eta^2 = .033$, sad responses to fear trials, $F(1, 107) = 3.513$, $p = .064$, $\eta^2 = .032$, and sad responses to happy trials, $F(1, 107) = 2.925$, $p = .090$, $\eta^2 = .027$, as is indicated in Figure H2. This suggests that OC users and FC women do not significantly differ in their rate of overall errors (between-group effect), the type of errors they are making (multivariate effect), or the rate of each unique type of error (univariate), across any of the trial types. However, OC users show trends towards providing more angry responses on disgust trials, less sad responses on fear trials, and more sad responses on happy trials.

Table H3

Unadjusted Means and SD of Percentage of Incorrect Responses Per Trial Type for OC Users and FC Women

Trial Emotion	Response Emotion	Mean (*SD*) of Percentage of Responses	
		OC Users (*n* = 37)	FC Women (*n* = 72)
Disgust	Disgust	15.53 (7.51)	18.99 (6.18)
	Fear	0.7 (1.97)	0.69 (1.9)
	Sad	1.7 (2.89)	2.78 (4.53)
	Angry [t]	11.59 (8.71)	8.86 (6.74)
	Happy	0.05 (0.33)	0.18 (0.68)
	Surprise	0.03 (0.16)	0 (0)
Fear	Disgust	0.7 (1.84)	0.49 (1.75)
	Fear	21.84 (9.07)	23.42 (8.78)
	Sad [t]	2.49 (3.75)	4.26 (4.86)
	Angry	0.08 (0.49)	0.14 (0.97)
	Happy	0.27 (1.33)	0.25 (0.88)
	Surprise	4.35 (4.74)	3.83 (4.95)
Sad	Disgust	0.38 (1.48)	0.54 (1.64)
	Fear	0.3 (1.31)	0.28 (1.24)
	Sad	26.32 (6.66)	28.78 (7.21)
	Angry	0.51 (1.89)	0.69 (1.83)
	Happy	0.19 (0.88)	0.14 (0.63)
	Surprise	0 (0)	0.01 (0.12)

Trial Emotion	Response Emotion	Mean (*SD*) of Percentage of Responses	
		OC Users (*n* = 37)	FC Women (*n* = 72)
Angry	Disgust	2.84 (5.47)	1.82 (3.73)
	Fear	0.31 (1.34)	0.4 (1.93)
	Sad	1.05 (2.44)	1.19 (3.13)
	Angry	22.53 (8.27)	24.15 (8.36)
	Happy	0.08 (0.49)	0.25 (1.31)
	Surprise	0.41 (1.38)	0.14 (0.68)
Happy	Disgust	0.08 (0.36)	0.03 (0.17)
	Fear	0 (0)	0.01 (0.12)
	Sad [t]	0.22 (1.03)	0 (0)
	Angry	0 (0)	0.04 (0.35)
	Happy	33.3 (8.46)	38.13 (7.37)
	Surprise	0.68 (2.73)	0.39 (1.86)
Surprise	Disgust	0.16 (0.83)	0.01 (0.12)
	Fear	0.14 (0.59)	0.42 (1.58)
	Sad	0 (0)	0 (0)
	Angry	0 (0)	0 (0)
	Happy	1.73 (3.38)	2.46 (4.23)
	Surprise	35.54 (6.56)	36.53 (6.98)

Note. Percentages are calculated based on total valid responses.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$

Table H4

Repeated-Measures MANOVA Results for Type of Incorrect Responses per Trial Type for Oral Contraceptive (OC) Users and Free-cycling (FC) Women

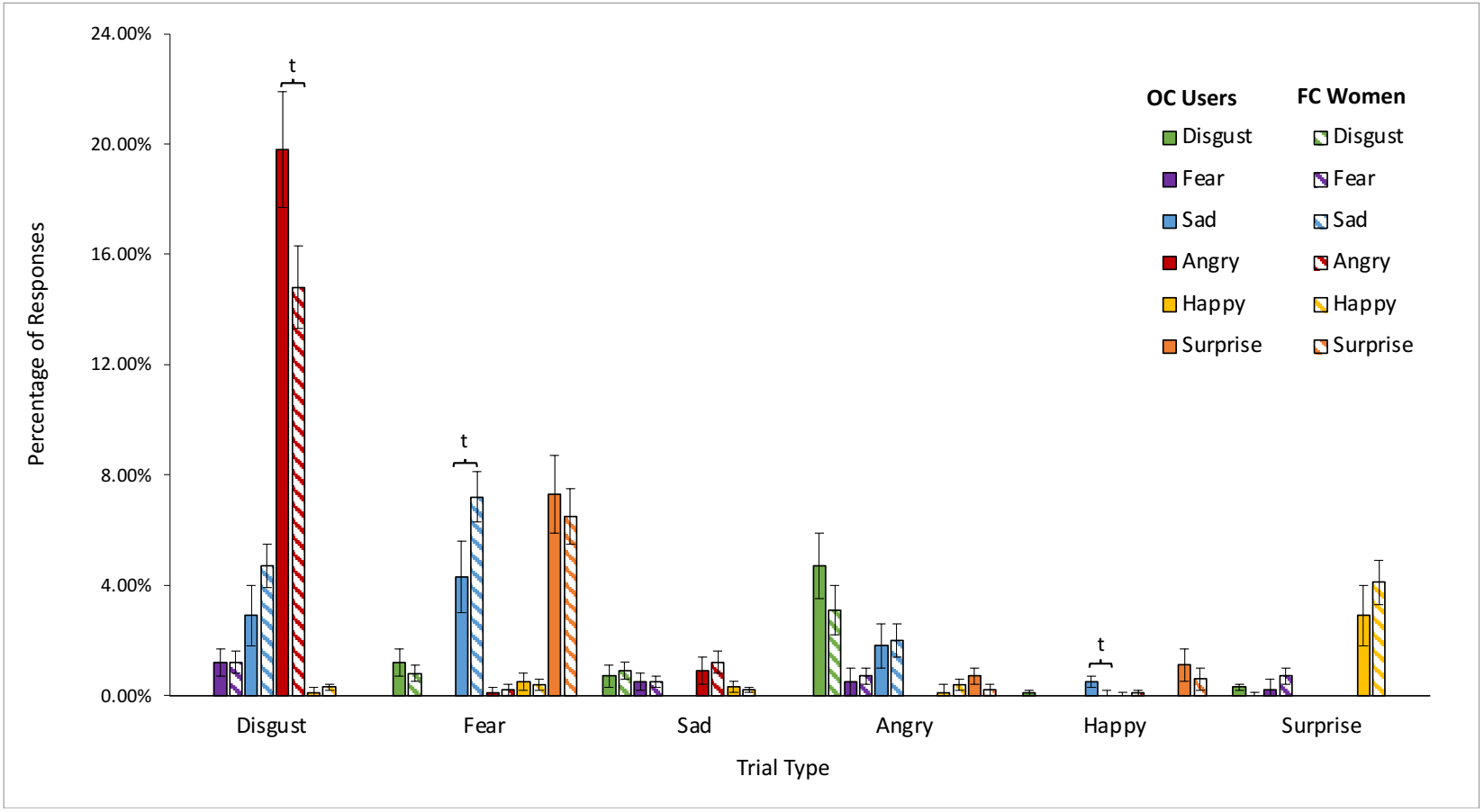
	Multivariate Effect (Group x Type of Incorrect Response)			
Trial Type (Correct Emotion)	*df*	*F*	*p*	η^2
Disgust	4, 104	1.684	.159	.061
Fear	4, 104	1.114	.354	.041
Sad	4, 104	0.143	.966	.005
Angry	4, 104	1.235	.301	.045
Happy	4, 104	1.391	.242	.051
Surprise	4, 104	1.271	.288	.035

Note. Results of the repeated-measures MANOVAs testing group differences in Type of Incorrect Responses per Trial Type.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure H2

Percentage of Incorrect Responses Per Trial Type for Oral Contraceptive (OC) Users and Free-cycling (FC) Women



Note. Error bars represent standard error. Only incorrect responses are shown. Solid bars represent OC users and striped bars represent FC women. Group difference trends emerged on angry responses to disgust trials (p = .060, η^2 = .033), sad responses to fear trials (*p* = .064, η^2 = .032), and sad responses to happy trials (*p* = 090, η^2 = .027). The most common errors per trial type are as follows: Disgust trial – anger error; Fear trial – surprise error; Sad trial – anger error; Angry trial – disgust error; Happy trial – surprise error; Surprise trial – happy error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

PMDD Groups Biases

A two-group (no PMDD, PMDD) repeated-measures MANOVA was run to compare group differences in the type of incorrect emotion responses (the percentage of fear, sad, angry, happy, and surprise responses) on disgust trials (i.e., the percentage of non-disgust responses). Five identical repeated-measures MANOVAS were run for the other trial types (fear, sad, angry, happy, and surprise trials).

The unadjusted means and *SD*s of scores for Percentage of Incorrect Responses per trial type are presented in Table H5. The results of all the repeated-measures MANOVAs are shown in Table H6. The multivariate effects were not significant for any of the emotion trial types. See Appendix F for the unadjusted means and *SD*s of scores for the error biases MANOVAs.

The distribution of the percentage of incorrect responses per trial type are shown in Figure H3. The univariate analyses testing whether the type of incorrect emotion per trial type differed based on group were all non-significant. However, three non-significant trends suggest that PMDD groups may differ in the rate of sad responses to fear trials $F(1, 70) = 3.649$, $p = .060$, $\eta^2 = .050$, sad responses to angry trials, $F(1, 70) = 2.854$, $p = .096$, $\eta^2 = .039$, and happy responses to angry trials, $F(1, 70) = 3.828$, $p = .054$, $\eta^2 = .052$, as is indicated in Figure H3. This suggests that no PMDD and PMDD groups do not significantly differ in the type of errors they are making (multivariate effect), or the rate of each unique type of error (univariate), across any of the trial types. However, the PMDD group may provide less sad responses on fear trials, more sad responses on angry trials, and less happy responses on angry trials.

Table H5

Unadjusted Means and SD of Percentage of Incorrect Responses Per Trial Type for No Provisional Premenstrual Dysphoric Disorder (PMDD) and PMDD Groups

Trial Emotion	Response Emotion	Mean (*SD*) of Percentage of Responses	
		No PMDD (*n* = 30)	PMDD (*n* = 42)
Disgust	Disgust	20.67 (4.66)	17.79 (6.88)
	Fear	0.5 (1.94)	0.83 (1.89)
	Sad	2.13 (5.13)	3.24 (4.04)
	Angry	7.6 (6.23)	9.76 (7.02)
	Happy	0.23 (0.82)	0.14 (0.57)
	Surprise	0 (0)	0 (0)
Fear	Disgust	0.3 (1.15)	0.62 (2.08)
	Fear	22.07 (8.76)	24.38 (8.76)
	Sad [t]	5.63 (5.67)	3.29 (3.97)
	Angry	0 (0)	0.24 (1.27)
	Happy	0.23 (0.82)	0.26 (0.94)
	Surprise	3.9 (4.59)	3.79 (5.25)
Sad	Disgust	0.3 (1.12)	0.71 (1.92)
	Fear	0.2 (0.81)	0.33 (1.48)
	Sad	27.2 (8.13)	29.9 (6.34)
	Angry	0.3 (1.15)	0.98 (2.16)
	Happy	0.13 (0.57)	0.14 (0.68)
	Surprise	0 (0)	0.02 (0.15)

Trial Emotion	Response Emotion	Mean (*SD*) of Percentage of Responses	
		No PMDD (n = 30)	PMDD (n = 42)
Angry	Disgust	1.77 (3)	1.86 (4.21)
	Fear	0.67 (2.75)	0.21 (1.02)
	Sad [t]	0.47 (1.48)	1.71 (3.84)
	Angry	23.43 (8.19)	24.67 (8.55)
	Happy [t]	0.6 (1.99)	0 (0)
	Surprise	0.17 (0.75)	0.12 (0.63)
Happy	Disgust	0 (0)	0.05 (0.22)
	Fear	0.03 (0.18)	0 (0)
	Sad	0 (0)	0 (0)
	Angry	0 (0)	0.07 (0.46)
	Happy	38.27 (7.4)	38.02 (7.44)
	Surprise	0.77 (2.8)	0.12 (0.5)
Surprise	Disgust	0 (0)	0.02 (0.15)
	Fear	0.27 (1.46)	0.52 (1.67)
	Sad	0 (0)	0 (0)
	Angry	0 (0)	0 (0)
	Happy	2.7 (4.64)	2.29 (3.97)
	Surprise	35.37 (6.76)	37.36 (7.09)

Note. Percentages are calculated based on total valid responses.

[t] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table H6

Repeated-Measures MANOVA Results for Type of Incorrect Responses per Trial Type for No Provisional Premenstrual Dysphoric Disorder (PMDD) and PMDD Groups

	Multivariate Effect (Group x Type of Incorrect Response)			
Trial Type (Correct Emotion)	*df*	*F*	*p*	η^2
Disgust	4, 67	1.182	.327	.066
Fear	4, 67	1.009	.409	.057
Sad	4, 67	0.894	.473	.051
Angry	4, 67	1.636	.176	.089
Happy	4, 67	1.364	.256	.075
Surprise	4, 67	0.430	.732	.019

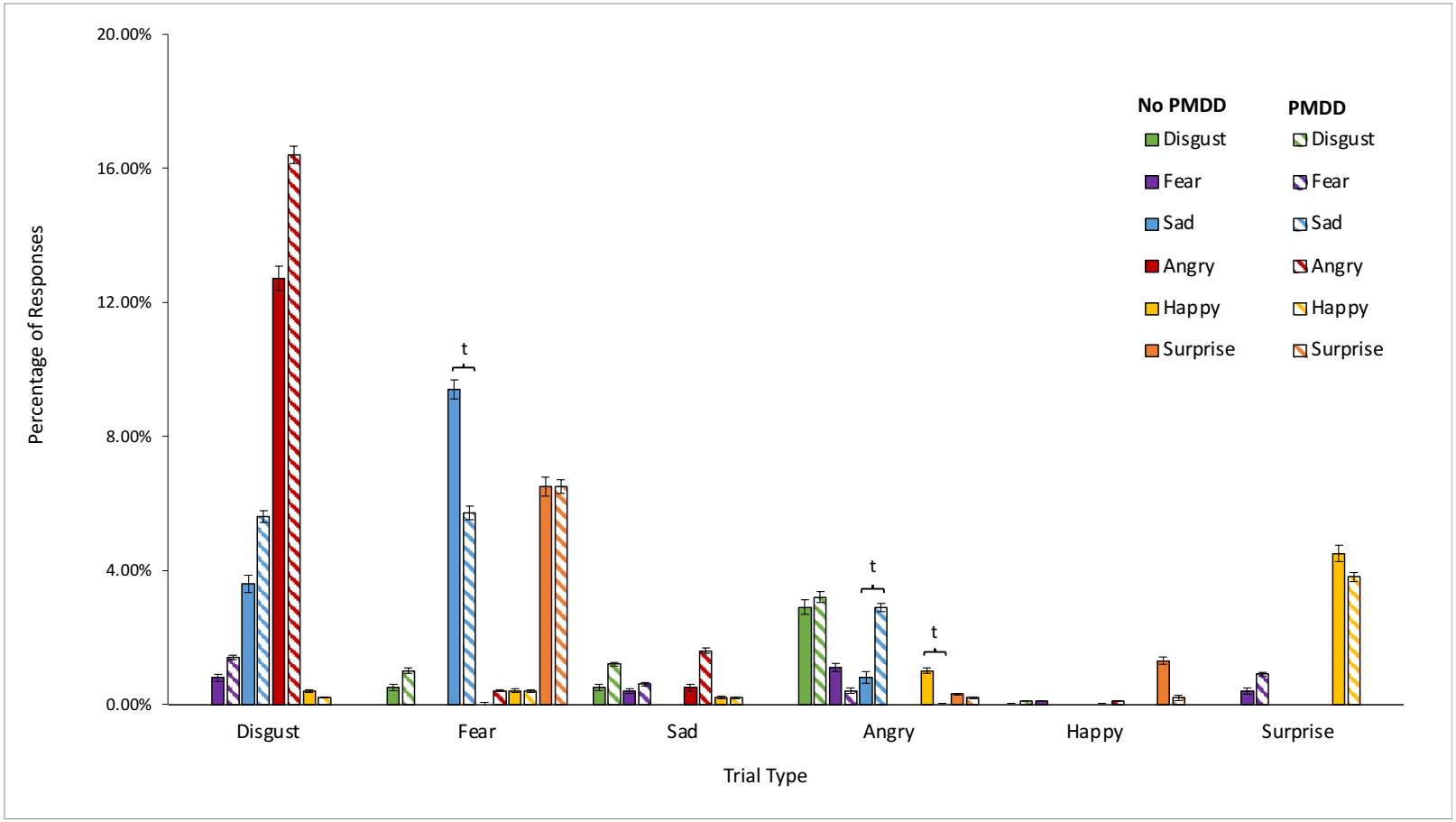
Note. Results of the repeated-measures MANOVAs testing group differences in Type of Incorrect Responses per Trial Type.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

.

Figure H3

Percentage of Incorrect Responses Per Trial Type for No PMDD and PMDD Groups



Note. Error bars represent standard error. Only incorrect responses are shown. Solid bars represent the no/minimal PMDD and striped bars represent the mild-severe PMDD group. There were no significant effects. The most common errors per trial type are as follows: Disgust trial – anger error; Fear trial – surprise error; Sad trial – anger error; Angry trial – disgust error; Happy trial – surprise error; Surprise trial – happy error.

$^{t}p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.