Inter-and Intraindividual Variation in Emotion Recognition: Investigating the Role of the Ovulatory Cycle, Sex, and Personality

Dissertation

for the award of the degree

"Doctor rerum naturalium" (Dr. rer. nat.)

of the Georg-August-Universität Göttingen

within the doctoral program Behavior and Cognition (BeCog)

of the Georg-August University School of Science (GAUSS)

Submitted by

Yasaman Rafiee

Born in Tehran, Iran

Göttingen, 2023

Thesis Committee

Prof. Dr. Anne Schacht
Department for Cognition, Emotion and Behavior,
University of Göttingen

Prof. Dr. Julia Ostner
Department of Behavioral Ecology,
University of Göttingen

Prof. Dr. Lars Penke
Department of Biological Personality Psychology,
University of Göttingen

Members of the Examination Board

Reviewer: Prof. Dr. Anne Schacht
Department of Cognition, Emotion and Behavior,
University of Göttingen

Second Reviewer: Prof. Dr. Julia Ostner
Department of Behavioral Ecology,
University of Göttingen

Further members of the Examination Board:

Prof. Dr. Lars Penke
Department of Biological Personality Psychology,
University of Göttingen

Prof. Dr. Nivedita Mani
Department of Psychology of Language
University of Göttingen

Dr. Igor Kagan
Cognitive Neuroscience Laboratory,
German Primate Center (DPZ)

Dr. Tanya Behne
Department of Cognitive Developmental Psychology
University of Göttingen

Date of Oral Examination: 19.06.23

Acknowledgment

First, I would like to sincerely thank my supervisor, Anne Schacht, for her invaluable support throughout my PhD journey. Her guidance, patience, and encouragement helped me to pursue my research interests despite the challenges I faced. Anne, I am so grateful for all the amazing opportunities you have given me and for believing in me!

I would like to express my gratitude to Julia Ostner, and Lars Penke for their supervision, scientific insight to my dissertation, and helping me to develop my research interest. Also, a big thank to Roger Mundry and Peter Zezula for their statistical support and always inspiring me.

I would like to thank all the members of the ANaP lab (past and present) for creating such an inspiring and welcoming atmosphere. A special thanks to my wonderful colleagues and friends, Ronja, Annika, Francesco, Anna, Adi, Lauriane, Sriranjani, Ricarda, Martina, Raji, Fatih, Weicheng, Vanessa, Jin, Esther, Behnaz, Bahar, Farnaz, Yasna, and Soheil. Your friendship and encouragement have been a constant source of motivation for me. I would like to thank all the students who helped me with data collection: Marlene Meyer, Maria Bathow, Luiza Beltrame, and Jana Thiel. I also thank my BeCog and RTG colleagues for their constructive feedback and companionship on this journey. A big thank you to Rebecca Jürgens, Ute Dorra, Valerie Liebs, and Christian Schlögl for their patience and support. Furthermore, I would like to express my gratitude for all the opportunities DAAD, RTG, and LSC (Primate Cognition) have provided me.

On a personal note, I would like to thank André for his incredible support, enthusiasm for my research, and engaging discussions about human cognition. Finally, I am deeply thankful to my family, especially my parents, for believing in me and supporting my dreams. I am forever grateful for the invaluable role they have played in every stage of my life.

"*All men have stars, but they are not the same things for different people. For some, who are travelers, the stars are guides. For others they are no more than little lights in the sky. For others, who are scholars, they are problems... But all these stars are silent. You-You alone will have stars as no one else has them.*"

— Antoine de Saint-Exupéry, The Little Prince

Table of Contents

CHAPTER 4: Study 3

CHAPTER 5: General Discussion

Preface

This dissertation is not a cumulative, publication-based dissertation, but follows it in form. It includes three manuscripts, two published and one under review.

Rafiee, Y., Stern, J., Ostner, J., Penke, L., Schacht, A. (2023). Does emotion recognition change across phases of the ovulatory cycle? *Psychoneuroendocrinology*, 148 (105977), https://doi.org/10.1016/j.psyneuen.2022.105977

Rafiee, Y., & Schacht, A. (2023). Sex differences in emotion recognition: investigating the moderating effects of stimulus features. *Cognition & emotion*, 37(5), 863–873. https://doi.org/10.1080/02699931.2023.2222579

Rafiee, Y., Heine, C., Schacht, A. (2023). Does the interplay of emotion-related personality traits and reproductive hormones predict individual variation in emotion recognition? [*Manuscript submitted for publication*] Preprint retrieved from: https://psyarxiv.com/rpybw

Declaration of own work

In the present thesis, I integrated three separate manuscripts (Study 1-3). I am first author in all manuscripts. The table below represents my responsibilities and contributions to each manuscript. Please note that my co-authors were involved in each process as well, meaning that every step is not my contribution alone but the result of a shared project.

	Study 1	Study 2	Study 3
Study Design	✓	✓	✓
Data Collection	✓	✓	✓
Data Analysis	✓	✓	✓
Manuscript Writing	✓	✓	✓

Note: Data collection in Study 2 was partially performed by the doctoral student.

Doctoral Student's Declaration

I hereby declare that all parts of this dissertation were written by myself, that assistance of third parties was only accepted if scientifically justifiable and acceptable in regards to the examination regulations, and that all sources have been quoted.

Göttingen, 2023

_____

Yasaman Rafiee

Abstract

English Abstract

Previous research has shown that emotion recognition varies within and between individuals, influenced by physiological and psychological factors such as hormones, biological sex, and dispositional traits. However, inconsistent results and methodological shortcomings have hindered our understanding of this variation and its contributing factors. Therefore, using two large datasets, this dissertation investigated whether variation in emotion recognition accuracy is associated with person-related sources in methodologically rigorous research. Study 1 examined whether women's emotion recognition accuracy varied as a function of the ovulatory cycle and ovarian hormone levels (estradiol and progesterone) in a within-subject, pre-registered study. However, no significant variation was found, questioning the overemphasis placed on the ovulatory cycle shift in emotion recognition accuracy in women. Study 2 examined sex differences in emotion recognition accuracy and found small but significant differences favoring women, particularly for recognizing negative emotions and across all modalities presented. Finally, Study 3 used an exploratory approach to examine the potential interactive effect of person-related sources, including dispositional traits and ovulatory cycle, in predicting variation in emotion recognition accuracy. Results showed that when personality traits such as openness and neuroticism were considered, emotion recognition accuracy varied significantly across the ovulatory cycle. Women with higher neuroticism scores had impaired emotion recognition during the mid-luteal phase compared to the late follicular phase and when progesterone levels were elevated. In addition, women who scored higher on openness had improved emotion recognition in the late follicular phase compared to the mid-luteal phase. The significant effects of sex and the interaction between personality traits and biological markers, such as the ovulatory cycle and ovarian hormones, highlight the importance of considering biology-environment interactions in understanding

individual variability in emotion recognition. Therefore, to gain a better understanding of individual variability in emotion recognition accuracy and the factors that contribute to it, future research needs to take these interactions into account.

Keywords: Emotion recognition, individual variation, ovulatory cycle, ovarian hormones, sex differences, personality

deutsche Zusammenfassung

Frühere Forschungen haben gezeigt, dass die Erkennung von Emotionen innerhalb und zwischen Individuen variiert und von physiologischen und psychologischen Faktoren wie Hormonen, biologischem Geschlecht und dispositionellen Merkmalen beeinflusst wird. Inkonsistente Ergebnisse und methodische Mängel haben jedoch unser Verständnis dieser Variation und der dazu beitragenden Faktoren behindert. Daher wurde in dieser Dissertation anhand von zwei großen Datensätzen methodisch streng untersucht, ob die Unterschiede in der Genauigkeit der Emotionserkennung mit personengebundenen Quellen in Verbindung stehen. In Studie 1 wurde untersucht, ob die Genauigkeit der Emotionserkennung von Frauen in Abhängigkeit von dem Eisprungzyklus und den Hormonspiegeln in den Eierstöcken (Östradiol und Progesteron) variiert. Es wurden keine signifikanten Unterschiede festgestellt, was die Überbetonung der Verschiebung des Eisprungzyklus bei der Genauigkeit der Emotionserkennung bei Frauen in Frage stellt. Studie 2 untersuchte geschlechtsspezifische Unterschiede in der Genauigkeit der Emotionserkennung und fand kleine, dafür aber signifikante Unterschiede zugunsten der Frauen, insbesondere bei der Erkennung negativer Emotionen und bei allen dargebotenen Modalitäten. Schließlich untersuchte Studie 3 in einem explorativen Ansatz den potenziellen interaktiven Effekt personenbezogener Quellen, einschließlich dispositioneller Merkmale und des Eisprungzyklus, bei der Vorhersage von Unterschieden in der Genauigkeit der Emotionserkennung. Die Ergebnisse zeigten, dass bei Berücksichtigung von Persönlichkeitsmerkmalen wie Offenheit und Neurotizismus die Erkennungsgenauigkeit von Emotionen während des Eisprungzyklus signifikant variierte. Bei Frauen mit höheren Neurotizismus-Werten war die Emotionserkennung in der mittleren Lutealphase schlechter als in der späten Follikelphase und auch bei erhöhtem Progesteronspiegel. Darüber hinaus hatten Frauen, die bei der Offenheit besser abschnitten, in der späten Follikelphase eine bessere Emotionserkennung als in der mittleren Lutealphase. Die signifikanten Auswirkungen

des Geschlechts und die Wechselwirkung zwischen Persönlichkeitsmerkmalen und biologischen Markern, wie dem Ovulationszyklus und den Eierstockhormonen, unterstreichen die Bedeutung der Berücksichtigung von Wechselwirkungen zwischen Biologie und Umwelt für das Verständnis der individuellen Variabilität der Emotionserkennung. Um ein besseres Verständnis der individuellen Variabilität in der Emotionserkennungsgenauigkeit und der Faktoren, die dazu beitragen, zu erlangen, muss die zukünftige Forschung diese Wechselwirkungen berücksichtigen.

Schlüsselwörter: Emotionserkennung, individuelle Variation, Ovulationszyklus, Ovarialhormone, Geschlechtsunterschiede, Persönlichkeit

CHAPTER 1

General Introduction

1 Emotions and Their Functions

Emotions are an essential part of our humanity, connecting us to each other and shaping our daily experiences. They play a vital role in our sense of self and our relationship with the world around us, affecting our perceptions, actions, and social interactions (Solomon, 2008; Van Kleef & Côté, 2021). Understanding emotions and their functions is crucial to comprehending human cognition and behavior, but their complex and multifaceted nature makes studying them a challenge (Lench et al., 2015; Lench & Carpenter, 2018). Although there is still no consensus on the definition of emotion, and the definition might vary depending on the discipline or approach (for an overview, see Sander, 2013), major theories identify several components of emotions, including cognitive appraisal (e.g., evaluating situation as dangerous), expression (e.g., upper eyelid raised), autonomic response (e.g., increased heartrate), action tendency (e.g., tendency to flee), and feeling (e.g., feeling scared) (Sander, 2013; Scarantino & de Sousa, 2021). Furthermore, researchers generally agree that emotions serve as an interface between individuals and their environment and continually influence the individual's reactions and experiences, while responding to the changes in social contexts and events (Mulligan & Scherer, 2012).

From an evolutionary perspective, emotions have been hard-wired as responses to certain prototypical situations, with the aim of survival and well-being (Levenson, 1999), and they help the organism to react to relevant stimuli with a complex behavioral pattern manifesting in different modalities like faces, vocal cues, or bodily gestures (Keltner et al., 2019). Adaptive functions of emotions help individuals respond to environmental challenges as well as opportunities in a quick and efficient way, and therefore they coordinate various physiological and psychological responses to different environmental stimuli to help the organism to navigate the environment and survive (see Tooby & Cosmides, 2008). In summary, the adaptive role of emotions enables the organism to successfully cope with life's challenges, such as mating, finding resources, identifying dangers, and parenthood (Ekman, 1992), as well as with the challenges posed by the social environment (Fischer and Manstead, 2009).

1.1 Expression and Recognition of Emotions

Individuals convey their emotional states through emotional expressions, which serve critical communicative functions in social interactions and interpersonal coordination (see Scherer, 2005; Van Kleef, 2009; 2016). Furthermore, emotional expressions convey important information about the expresser, such as status, power, dominance, and intention (Van Kleef & Côté, 2021). In order to understand the internal state of other people, individuals usually infer their external states presented through faces which carry important social information such as individual identity and emotional state and therefore have been extensively studied in social cognition research (Todorov et al., 2013). In the real world, however, emotions are expressed through faces and other modalities such as voices, bodily (postural) movements, or touch (see De Gelder & Vroomen, 2000; Gerdes et al., 2014).

In addition to the expression of emotions, the recognition of emotions is also essential for effective social interaction. Emotion recognition is the ability to interpret other people's emotional states from their facial, vocal, and physical expressions. This ability, therefore, plays a crucial role in facilitating effective interpersonal interactions, as it allows individuals to better understand their internal state and respond to the emotional cues of those around them (Elfenbein, 2007; Schlegel et al., 2019). Furthermore, emotion recognition plays an important adaptive role, as accurate and rapid emotion recognition may increase the likelihood of self and offspring survival (e.g., avoidance of environmental hazards or contamination), mating, and ultimately reproductive success (e.g., Shariff & Tracy, 2011).

1.2 Variability in Emotion Recognition Across Individuals

Individual variation has been observed in various cognitive abilities and behavioral manifestations (Revelle et al., 2010), as well as in emotion recognition (Bänziger, 2016). Given the crucial role of emotion recognition in social interactions, such variations may have significant effects on social functioning and mental well-being. In addition, problems with emotion recognition may contribute to the development and persistence of emotional disorders and other psychopathological symptoms. For example, difficulties in emotion recognition have been observed in various domains of disorders such as major depressive disorder (for review, see Dalili et al., 2015), borderline personality disorder (for review, see Domes et al., 2009), individuals with subthreshold psychotic symptoms (e.g., Amminger et al., 2012), and schizophrenia (for review, see Edwards et al., 2002), resulting in impaired social functioning.

Even among healthy individuals, there is variability in the ability to recognize emotions (Laukka et al., 2021). The underlying reasons for the variability in the recognition of emotions are not yet fully understood. However, individual differences in person-related sources, ranging from

psychological to physiological factors, may contribute to the observed inter- and intra-individual variation in emotion recognition. Research suggests that these person-related sources play a modulatory role in the neural processing of emotions, suggesting their importance in shedding light on the underlying mechanisms of emotion recognition (see Hamann & Canli, 2004).

Current findings concerning the association of person-related sources and individual variation in emotion recognition is rather inconsistent, potentially stemming from limitations in methodology and small sample sizes. Robust findings in studies investigating individual variation in emotion recognition research require large sample sizes, which have been lacking in many previous studies due to practical and financial constraints. In addition, previous studies have focused primarily on facial expressions. However, their ecological validity has been debated because, in real-world settings, emotions are expressed not only through faces but also through other modalities, such as voices (e.g., Holleman et al., 2020). Thus, this research gap highlights the need for conducting further studies that examine individual variations in emotion recognition using multiple perspectives, such as physiological and psychological measures, as well as incorporating well-established research methodologies.

2 Hormones: Biomarkers of Individual Differences

Hormones play an important role in the coordination of physiological and behavioral mechanisms to enhance the fitness of individuals in the face of various challenges (Ketterson et al., 2013). They act as effective messengers by being released into the bloodstream, where they transmit information to the brain and body parts tasked with responding to potential problems (Roney, 2016). One hormone can have multiple physiological and psychological effects on the brain and body (O. C. Schultheiss & Stanton, 2009). From an adaptive perspective, neurochemicals

such as hormones contribute to the processing of emotional expressions because quick and efficient recognition of emotions can increase the chances of survival and reproductive success (Gingnell et al., 2019). Investigating a possible relationship between hormones (both exogenous and endogenous) and emotion recognition is a thriving field that has been explored in a small number of studies. The following sections provide a brief overview of the current state of the field.

2.1 Exogenous Hormones

Exogenous hormones refer to any externally administered hormone not produced by the individual's endocrine glands. Some studies have used the administration of hormones, including oxytocin, testosterone, and cortisol, to examine the possible link between the endocrine system and the processing of emotional expressions. For example, Schulze et al. (2011) provided evidence that a single dose of intranasally administered oxytocin enhanced recognition of emotional stimuli, particularly happy expressions. Similarly, Domes et al. (2007) found that oxytocin improved performance on the Mind in the Eyes Test (RMET). In another study, Di Simplicio and Harmer (2016) demonstrated that oxytocin facilitated the processing of positive emotions over negative emotions, resulting in lower misclassification rates for happiness, surprise, and neutral expressions compared to the placebo group.

Regarding exogenous testosterone, Van Honk and Schutter (2007) administered testosterone (0.5 mg) to healthy women and reported decreased accuracy in recognizing angry expressions compared to the placebo group. In the case of exogenous cortisol, Duesenberg et al. (2016) administered 10 mg of hydrocortisone to participants to determine if cortisol was associated with facial emotion recognition, but found no compelling evidence for the hypothesized association. A few studies have examined the association between oral contraceptives and emotion recognition,

suggesting impaired emotion recognition (e.g., Pahnke et al., 2019) as well as no association (e.g., Shirazi et al., 2020).

Despite the distinct functions and varied associations of these hormones, such as testosterone, cortisol, oxytocin, and hormonal contraceptives, with emotion recognition, research examining the relationship between exogenous hormones and emotion recognition has yielded inconsistent results.

2.2 Endogenous Hormones

Endogenous hormones are the category of hormones that are produced naturally in an individual's body. There have been limited studies investigating the relationship between endogenous hormones and emotion recognition. For instance, Lausen et al. (2020) conducted research on healthy young men and found no significant correlation between cortisol or testosterone levels and emotion recognition accuracy. These findings were consistent with the null results reported by Derntl et al. (2009) investigating testosterone and emotion recognition. In contrast, Rukavina et al. (2018) discovered a negative correlation between salivary testosterone levels and emotion recognition, and Vongas and Al Hajj (2017) identified a positive relationship between testosterone levels and emotion recognition.

Similar to the findings regarding the relationship between exogenous hormones and emotion recognition, existing research on the relationship between endogenous hormones and emotion recognition has yielded inconsistent results. Some studies have reported no significant correlation between hormone levels and emotion recognition accuracy, while others have found negative or positive associations.

2.2.1 Menstrual Cycle: A Hormonal Model

A recent area of research has delved into the link between female reproductive hormones i.e., estradiol and progesterone, and emotion recognition, particularly through the examination of the menstrual cycle. This approach has several advantages: first, due to the natural fluctuation of ovarian hormones during the menstrual cycle, it provides a hormonal model for steroid hormones to study the link between (within- and between-person) endogenous hormone levels and emotion recognition (Allen et al., 2016; Gurvich & Thomas, 2021; Poromaa & Gingnell, 2014). Steroid hormones cross the blood-brain barrier and affect brain functions. Their levels in the central nervous system corresponds to their levels in the rest of the body (Schultheiss et al., 2019). This quality makes these hormones serve as ideal subjects for investigating the relationship between the endocrine system and cognition. Second, because of hormonal fluctuations, it is possible to detect high and low levels of hormones in individuals, thus eliminating the need for hormone administration, which could raise ethical concerns (Puts, 2006). Third, the distribution of estradiol and progesterone receptors (ERα, ERβ, PRA, and PRB) in brain regions involved in emotional processing, such as the hypothalamic and limbic systems, provides further evidence for a possible link between menstrual cycle phases, ovarian hormones, and emotion recognition (Poromaa & Gingnell, 2014). Finally, exploring the potential impacts of the menstrual cycle can provide valuable insights into the ways in which women's reproductive hormones influence their cognition and behavior. (Allen et al., 2016; Sacher et al., 2013).

The menstrual cycle consists of two main phases: the follicular phase and the luteal phase (Bull et al., 2019; Kiesner et al., 2020). During these phases, the levels of estradiol and progesterone levels fluctuate in a cyclical pattern. In the late follicular phase, estradiol levels peak, triggering a surge of luteinizing hormone (LH) that ultimately leads to ovulation approximately 24-48 hours

later (Blake et al., 2016). The luteal phase begins with the corpus luteum's secretion of progesterone, which continues to increase throughout the mid-luteal phase. The mid-luteal phase is characterized by the highest progesterone levels and a secondary peak in estradiol (Allen et al., 2016).

Among the limited number of studies that have examined the relationship between menstrual cycle phases and emotion recognition, some have provided evidence of an existing relationship. They report better performance in the follicular phase compared to the luteal phase, which is associated with higher levels of estradiol (e.g., Derntl et al., 2013; Derntl, Kryspin-Exner, et al., 2008; Derntl, Windischberger, et al., 2008; Osório et al., 2018; R. Pearson & Lewis, 2005; Rubin, 2012), and a negative correlation between progesterone levels and emotion recognition accuracy (Derntl et al., 2013). In contrast, some studies have not found such a relationship between menstrual cycle phases, ovarian hormones, and emotion recognition (e.g., Di Tella et al., 2020; Kamboj et al., 2015; Shirazi et al., 2020; Zhang et al., 2013).

While the field is developing, inconsistent findings make it difficult to draw definitive conclusions. The discrepancy in findings may be due to the lack of standardized methods, such as different methods for estimating cycle phase and measuring ovarian hormones, the use of within-subject versus between-subject comparisons, and the application of different experimental designs (see Allen et al., 2016; Gangestad et al., 2016; Gingnell et al., 2019; Schmalenberger et al., 2021). Another factor contributing to the inconsistent findings in the literature is the low ecological validity of the experimental task; most research on the hypothesized change in emotion recognition across the menstrual cycle has focused on facial expressions, whereas in the natural environment emotions are expressed through more modalities, such as voices (e.g., De Gelder & Vroomen, 2000). Furthermore, like any other field in psychology, this field suffers from reproducibility issues

due to insufficient reliability, validity, and statistical power of some previous studies (for a review, see Schultheiss et al., 2019).

3 Biological Sex

Sex is one of the most commonly studied person-related sources of individual variation in emotion recognition (Connolly et al., 2019). The origin and underlying mechanisms of the sex effect in emotion recognition remain complex. A wide range of explanations, ranging from biological (chromosomal, hormonal, structural, and functional differences at the neural levels), environmental (culture and society), and their interaction, have tended to explain the observed sex differences (Kret & De Gelder, 2012; Maner & Miller, 2014).

Sex differences in emotion recognition have been widely explained from an adaptive perspective, interpreting these differences in the context of parental investment. In most species, males and females contribute differently to parental care, with females generally providing more care (see Liker et al., 2015). Females, due to their limited reproductive success and lower reproductive variance as compared to men, may invest more in their offspring to increase their chances of reproductive success (Kokko & Jennions, 2008; Taylor et al., 2000). Evidence suggests that in nonhuman primates, females tend to provide more care for infants than male (Hames et al., 1985). This pattern also exists in humans, although males are generally more involved in parenting than other species (Hames et al., 1985). Therefore, it is possible that men and women differ in their ability to recognize certain emotions in infants as a result of sex differences in parental investment in their offspring (Hames et al., 1985). In this line, the "*primary caretaker hypothesis*" proposes that women have an advantage in recognizing facial expressions critical to offspring survival, regardless of their prior experience in childcare (Babchuk et al., 1985; Hames et al., 1985).

Similarly, the "*fitness threat hypothesis*" suggests that women have a superior ability to recognize negative emotions (Hampson et al., 2006, 2021). This hypothesis is based on the idea that negative emotions convey more important survival-related information than other emotions. Therefore, women may have evolved to develop this skill as a means of increasing their offspring's chances of survival in challenging environments.

Although a large body of empirical evidence has shown that women are significantly better than men at recognizing emotional expressions (for a review, see Hall, 1978; McClure, 1999; Thompson & Voyer, 2014), some studies reveal the opposite (null) effect (Hoffmann et al., 2010; Lambrecht et al., 2014; Lyusin & Ovsyannikova, 2016; Rahman et al., 2004). The observed discrepancy could be attributed to factors such as limited statistical power and variations in experimental design. The reported effect size associated with sex differences in emotion recognition is relatively small (e.g., Cohen's $d = 0.19$; see Thompson & Voyer, 2014) to moderate (e.g., Cohen's $d = 0.40$; see Hall, 1978) indicating that a large sample is required to observe the sex effect. Furthermore, the existing evidence lacks providing a more comprehensive and robust view of sex differences in emotion recognition as the function of stimulus features. Only a few studies have shown that the encoder's sex, stimulus modality, and specific emotions moderate the magnitude of sex differences in emotion recognition (for a review, see Thompson & Voyer, 2014). For example, some studies have suggested a female advantage in recognizing anger expressions (Di Tella et al., 2020; Duesenberg et al., 2016; Grimshaw et al., 2004; Hall & Matsumoto, 2004), and conversely, some studies have reported a male advantage in recognizing anger (e.g., Rotter & Rotter, 1988; Wagner et al., 1986; for an overview see Kret & De Gelder, 2012). The inconsistency also exists in sex differences in recognition of emotions expressed by male and female encoders. A few studies reported better performance of females in recognizing emotions from male faces (Collignon et al., 2010; Davitz, 1964; Tagiuri, 1969; Thompson & Voyer, 2014); however, this

finding was not supported in some other studies (Hall, 1978; Lausen & Schacht, 2018; Thayer & Johnsen, 2000). Finally, the moderating effect of stimulus modality on sex differences in emotion recognition has been largely understudied, as most previous studies have focused only on facial emotion recognition. Overall, while many studies have investigated the sex effect in emotion recognition, some previous methodological limitations, such as small sample sizes and limited stimulus variety and number, may hinder a more comprehensive picture of the sex effect.

4 Interplay of Personality Traits and Biological Markers

Individual differences exist in a wide variety of species and have been proposed to have fitness consequences (Blaszczyk, 2020). These differences are assumed to broadly affect evolutionarily relevant outcomes such as survival, mating, parenthood, and, ultimately, reproductive success (Buss, 2009; Buss & Greiling, 1999). In this line, some individual differences are thought to function as individual strategies to solve social adaptive problems (Buss, 2009).

One of the main sources of individual differences in psychology is dispositional personality traits. Personality traits are distinctive characteristics that influence an individual's behavior, thoughts, and feelings in different situations and over an extended period of time (Ashton, 2013). The predictive value of personality traits in various behavioral outcomes is recognized in the extant literature (McAdams & Pals, 2006). In this research area, the five-factor model of personality, known as the Big Five, has been widely accepted and used by researchers (Costa & McCrae, 1994; McAdams & Pals, 2006) and helps to determine the structure of individual differences in personality psychology (Eisenlohr-Moul & Owens, 2016). The Big Five model proposes that nearly all traits can be grouped into five broad domains: openness, extraversion, conscientiousness, neuroticism, and agreeableness (DeYoung & Gray, 2009). Interestingly, evolutionary psychology

suggests that these five factors represent individual strategies for survival and reproductive success and are involved in social adaptation and cooperation (Buss, 2009; Nettle, 2006).

The affect-personality relationship has been explored in previous research, and the focus has tended to be on the relationship between particular personality dispositions and the experience of different emotions. For example, why some people become angry and others become afraid in response to threats (for a review see Revelle & Scherer, 2009). It is plausible that there is a connection between personality traits and the ability to recognize emotions. However, there have been limited studies that have investigated this relationship. Existing research suggests that openness and extraversion are positively associated with emotion recognition ability, whereas neuroticism is negatively associated (Matsumoto et al., 2000; Scherer & Scherer, 2011; Terracciano et al., 2003).

Previous research has predominantly examined the main effect of person-related resources in investigating individual variation in emotion recognition, neglecting the potential joint role of different person-related sources. As a result, there is a gap in the literature regarding how different aspects of human personality traits and biological markers, such as the ovulatory cycle and ovarian hormones, jointly contribute to individual variation in emotion recognition.

5 Aims and Overview

Overall, the existing literature has provided evidence, albeit mixed, for the association of person-related sources such as the menstrual cycle, ovarian hormones, sex, and personality with emotion recognition. Thus, this dissertation aimed to extend current knowledge on the role of these potentially person-related sources in predicting inter- and intra-individual variation in emotion recognition. To this end, using two large datasets, three studies examined the association of the

ovulatory cycle (as a model for tracking high vs. low levels of ovarian hormones), biological sex, and the potential joint role of personality traits and the ovulatory cycle (and thus ovarian hormones) in emotion recognition. Each study further investigated specific hypotheses within a particular framework, intending to shed light not only on the theoretical and practical implications of the relationship between person-related factors and individual variation in emotion recognition but also on the moderating role of stimulus characteristics in this relationship. In addition, this dissertation sought to advance the field by implementing methodological rigor, including high statistical power, ecologically valid experimental design, appropriate statistical analysis, and the practice of open and transparent science.

CHAPTER 2

Does Emotion Recognition Change across Phases of the Ovulatory Cycle?

Abstract

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

ovulatory cycle phases (late follicular and mid-luteal) and emotion recognition ability might have been overestimated in previous studies, and may contribute to theoretical and practical implications of socio-cognitive neuroendocrinology.

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

1 Introduction

Emotions evolved to help individuals to deal with various life tasks, including mating, resource finding, danger identification, and parenting (Al-Shawaf et al, 2015). One of their functions is to improve individuals' chances of survival and ultimately their reproductive success (Fischer & Manstead, 2009). In social contexts, emotional expressions carry important information supporting individuals to regulate their responses to environmental opportunities and risks (Keltner et al, 2019), also in the service of successful interpersonal interactions (Schlegel et al. 2019). The underlying mechanisms of emotion recognition are not completely understood; nevertheless, it is plausible to assume that this ability is influenced by neurochemicals, including hormones (Thagard, 2002). Some previous studies examined the link between endogenous or exogenous hormones and emotion recognition, including the specific role of female sex hormones fluctuating across women's menstrual cycle (Derntl, Kryspin-Exner, et al., 2008; Derntl et al., 2013; Di Simplicio & Harmer, 2016a; Duesenberg et al., 2016a; Ellenbogen et al., 2012; Lausen et al., 2020a; Lischke et al., 2012; Maner & Miller, 2014; Marečková et al., 2014; Pahnke et al., 2019; R. M. Pearson et al., 2009; Radke & Derntl, 2016; Roos et al., 2012; Schulze et al., 2011; Shirazi et al., 2020; Young et al., 2017).

The menstrual cycle, due to its periodicity, provides a natural model to study relationships between female sex hormones, cognition, and emotion (Allen et al., 2016; Poromaa & Gingnell, 2014; Sacher et al., 2013; Sundström-Poromaa, 2018), and can roughly be divided into two main phases, namely follicular and luteal, across which the levels of ovarian hormones, i.e. estradiol and progesterone fluctuate in a cyclic fashion. The fluctuation of ovarian hormones, which is highly related to the reproductive state, could be associated with the processing of emotional expressions as an important component of reproductive success (Gingnell et al., 2019). Ovarian hormones could

potentially yield alterations in women's recognition of emotions assumed to be involved in the facilitation of social interactions (Derntl, Kryspin-Exner, et al., 2008) and flagging social threats (Maner & Miller, 2014), respectively. A high ability to recognize emotions might thus increase the chance of successful social interaction and, as a result, a higher chance for reproductive success (Gingnell et al., 2019).

A recent review (Osório et al., 2018) and some studies (Derntl, Kryspin-Exner, et al., 2008; Pearson and Lewis, 2005; Rubin, 2012) suggested an improved emotion recognition accuracy (ERA) in the follicular compared to the luteal phase, presumably regulated by estradiol levels. Limited evidence indicated an overall impairment of emotion recognition by increased progesterone levels (see Osório et al., 2018). On the contrary, several studies found no evidence for a relationship between cycle phase or ovarian hormones with emotion recognition in healthy naturally cycling women (Di Tella et al., 2020; Gingnell, 2013; Kamboj et al., 2015; Shirazi et al., 2021; Zhang et al., 2013). In sum, the evidence concerning a possible association between cycle phases, ovarian hormones, and emotion recognition ability is inconsistent.

Regarding the interplay of ovarian hormones and specific emotions, previous findings are also mixed or even contradictory (for an overview see Gamsakhurdashvili et al., 2021; Osório et al., 2018). Sakaki and Mather (2012) suggested that increased levels of estradiol are related to a reduced reaction to negative stimuli, supported by reports of negative relationships between estradiol levels and accuracy in recognizing anger (Guapo et al., 2009; Kamboj et al., 2015) and disgust (Kamboj et al., 2015). Contradictory, Pearson and Lewis (2005) found a higher ability to recognize fear during the fertile phase, employing a between-subject study with a small sample (N = 50) and lacking hormone level measurement. Some studies reported higher levels of progesterone to be associated with an enhanced ability to recognize negative emotions, i.e., expressions of fear

and disgust (Conway et al., 2007), and of anger, fear, disgust, and sadness (Maner and Miller, 2014), which was explained by an assumed behavioral defense mechanism to avoid physical danger or contamination (Conway et al., 2007). Other studies failed to demonstrate an association between ovarian hormones and recognizing specific emotions (e.g, Zhang et al., 2013).

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

In this preregistered study (https://osf.io/dkpf5/), we investigated within-subject differences in cycle phase and associated ovarian hormones levels and the recognition of different emotions expressed by faces, voices, and face-voice combinations across two cycle phases namely late follicular and mid-luteal. The reason to choose these phases across the menstrual cycle is to capture the peak of estradiol levels in the late follicular phase, and progesterone levels in the mid-luteal

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

2 Methods

The study was approved by the ethics committee of the Psychological Institute of the University of Goettingen. Each participant signed a consent form following the Declaration of Helsinki (DoH) ethical principles for human subjects. Before data collection, the study aims, hypotheses, and study design were preregistered at the open science framework (https://osf.io/dkpf5/). The open data and

analysis script are available (https://osf.io/2pucf/). Participants were compensated with either course credit or monetary rewards of 25€. As an additional incentive, participants who completed the study had the opportunity to win one of four Amazon gift vouchers with a maximum value of 50€.

2.1 Participants

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

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

Included participants were German native speakers, self-reported being healthy, heterosexual, non-pregnant, naturally-cycling women with a cycle length of 25 to 35 days ($M_{\text{Cycle length}} = 28.77$, $SD_{\text{Cycle length}} = 2.03$), between 18 and 35 years old ($M_{\text{Age}} = 24.1$, $SD_{\text{Age}} = 3.5$), and having had a regular ovulatory cycle for at least three months before their first participation in the study. All participants reported normal or corrected-to-normal vision and normal hearing, without any history of psychiatric, neurological, metabolic, or hormonal disorders. In addition, they did not use any sort of hormonal medications such as contraceptives, nor did they breastfeed for at least three months before their first participation in the study. Of the 131 included participants, sixty participants reported being single, sixty-three were in a relationship, two were engaged, three were married, and three reported to be in different forms of relationships[1]. One-hundred-nineteen participants were righthanded.

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

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

2.2 Procedure

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

2.2.1 Introductory Session

First, participants were screened according to the inclusion criteria. We estimated the onset of the next menstruation and used the backward-counting method to predict the ovulation date (Puts, 2006). Moreover, to validate the fertile phase estimate, participants were asked to use highly sensitive (10mIU/ml) urine ovulation test strips from Runbio Biotech Co., Ltd., as soon as their menstruation ended and report to us whenever they saw a positive test. To standardize the influence of possible physiological factors, we asked participants to use LH strips between 10 am and 8 pm, preferably at the same time of the day. We also asked participants to send us photos of their LH tests on a voluntary basis.

2.2.2 Testing Sessions

Sessions two and three took place in the late follicular and the mid-luteal phase of each participants' ovulatory cycle. Following Jünger et al. (2018) the late follicular phase was estimated as reverse

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

2.3 Saliva Sampling

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

2.3.1 Hormone Measures

Levels of estradiol (E2) and progesterone (P4) were measured via the Chemiluminescence Immunoassays method at the Endocrinology Laboratory at the Technical University of Dresden.

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

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

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

2.3.2. Handling Hormonal Data[4]

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

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

To track the within-subject fluctuation of ovarian hormones across the ovulatory cycle, hormonal measures were subject mean-centered and scaled by being divided by a constant. The values varied from −.5 to .5, which eases the calculation in the linear mixed model (e.g. Jünger et al., 2018). Subject mean centering distinguishes the effect of within- and between-subject variation of hormones, and therefore this method is recommended to track the influence of hormonal fluctuations across the ovulatory cycle (see Schmalenberger et al., 2021). To investigate the association between ovarian hormone variation between different individuals and their emotion recognition ability, hormonal measures were averaged across two sessions for each participant. Importantly, adding between-subject effects to our analyses did not affect any of the within-subjects results. We then log-transformed (base e) the average hormonal measures representing the between-subject levels of estradiol and progesterone. To facilitate the interpretation of model outcome and model convergence, we then z-transformed previously log-transformed between-subject levels of estradiol and progesterone.

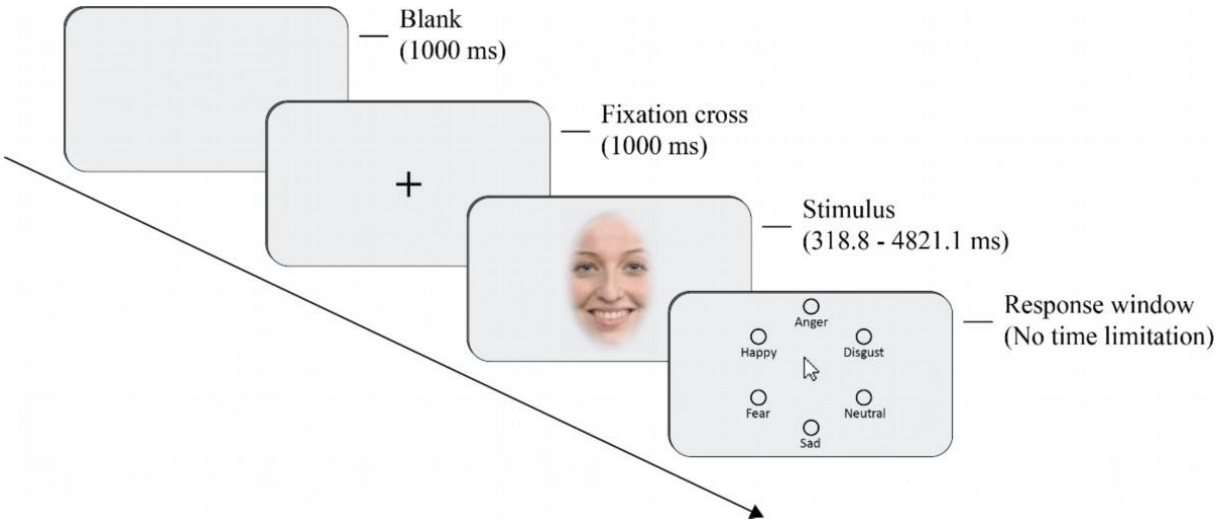
2.4 Emotion Recognition Task

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

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

Figure 1

Schematic overview of the trial scheme.



2.5 Stimuli

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

2.6 Statistical Analyses[5]

Data analyses were performed using R Software version 4.0.3 and R studio version 1.4.1106. Generalized Linear Mixed Models (GLMMs) with binomial error structure and logit link function was applied. To make inferences, standard p-value 0.05 was used as the cut-off criterion for two-tailed distributions. To preprocess the data, the following packages were used: tidyverse 1.3.1, knitr 1.33, dplyr 1.0.5. We used ggplot2 3.3.3, and sjplot 2.8.7 for data visualization, lme4 1.1.26 for computing models, and car 3.0.10 for assessing collinearity among predictors. Variance Inflation Factors (VIF) with a model lacking the interaction showed no collinearity issue in our models (maximum VIF: 1.044). To deal with the convergence issue we added "bobyqa" optimizer to fit the models (see supplementary document on OSF: https://osf.io/2pucf/).

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

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

3 Results[6]

3.1 Descriptive statistics

Across both cycle phases, women had the highest performance in recognizing expressed emotions in the audiovisual modality ($M_{\text{proportion correct responses (PCR)}} = .96$) and the lowest performance in recognizing auditory emotional expression ($M_{PCR} = .81$) (see Table 1 for emotion recognition performance in each cycle phase). The most recognized emotion was the neutral expression ($M_{PCR} = .94$) and the least recognized emotion was the disgust expression ($M_{PCR} = .81$) across all phases of the ovulatory cycle. The recognition of emotions expressed by female actors was ($M_{PCR} = .91$) and for male actors was ($M_{PCR} = .89$).

Table 1

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

Late Follicular Phase

Modality	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Audiovisual	0.962	0.037	1.275	0.348
Auditory	0.813	0.076	1.629	0.49
Visual	0.912	0.049	1.299	0.28
Emotion Category	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Anger	0.932	0.044	1.378	0.35
Disgust	0.81	0.105	1.489	0.427
Fear	0.905	0.072	1.551	0.373
Happy	0.93	0.047	1.206	0.262

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

Neutral	0.935	0.063	1.349	0.366
Sad	0.863	0.089	1.433	0.343

Sex of the actor	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Female	0.905	0.044	1.381	0.314
Male	0.886	0.05	1.421	0.316

Mid-Luteal Phase

Modality	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Audiovisual	0.963	0.038	1.233	0.29
Auditory	0.816	0.079	1.612	0.563
Visual	0.914	0.053	1.282	0.316

Emotion Category	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Anger	0.933	0.059	1.326	0.342
Disgust	0.809	0.096	1.430	0.392
Fear	0.906	0.07	1.524	0.388
Happy	0.934	0.042	1.218	0.292
Neutral	0.938	0.07	1.355	0.71
Sad	0.865	0.088	1.402	0.332

Sex of the actor	**Accuracy**		**Reaction Times**	
	Mean	SD	Mean	SD
Female	0.908	0.044	1.360	0.367
Male	0.887	0.056	1.392	0.321

3.2 Cycle Shifts and Facial Emotion Recognition (*H1*)

First, we investigated potential ovulatory cycle shifts in facial emotion recognition. We included the cycle phase as the fixed effect, session number as the control variable, and subject ID as the random intercept. The reference category for comparison was the mid-luteal phase. The included participants (n = 74) in this model observed a positive LH test during the optimal days (maximum of three days before and two days after their estimated day of ovulation) (see Blake et al., 2016). The model showed no significant differences in women's emotion recognition

performance between the late follicular and mid-luteal phase of the ovulatory cycle (β = -0.031, SE = 0.028, 95% CI = [0.92; 1.02], z = -1.122, OR= 0.97, *p* = *.262*; Fig. 2, right panel). We further compared the log-likelihood of this model with a model lacking the ovulatory cycle phases to examine the goodness of fit of our model. The result showed no significant difference between the main model and the model lacking the ovulatory cycle phase (χ^2 = 1.25, *df* = 1, *p* = 0.262). However, participants showed a better performance in the second session compared to the first session (β = 0.315, SE = 0.028, 95% CI = [1.30; 1.45], z = 11.388, OR= 1.37, *p <0.001*), see Table 1. To control for the robustness of the findings, we fitted an additional model including all participants (*N* = 131). The results were consistent with the initial model (see supplementary document).

Figure 2

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

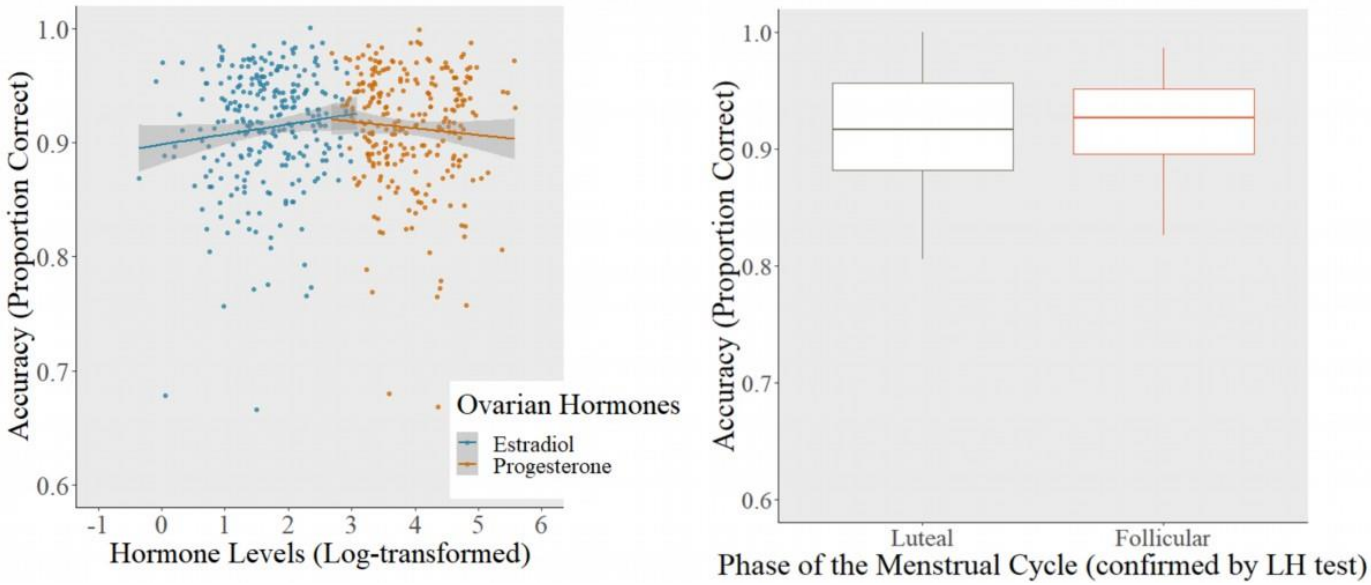


Table 2

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

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

3.3 Association between Ovarian Hormones Levels and Facial Emotion Recognition (*H2a, 2b, 2c*)

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

session number was significant and suggests that women performed better in the second session (Table 3). The model outcome also showed that the emotion category contributes significantly to predicting the outcome variable and the happy expression had the highest recognition rate compared to threat-related emotions (Table 3).

The main model was compared with three different null models including all participants (N = 131). The first comparison was between the model and the null model lacking estradiol measures (within- and between-subject). The likelihood ratio test revealed no significant difference between the two models ($\chi^2 = 1.351$, $df = 2$, $p = .509$). Further, the main model was compared with a null model lacking progesterone measures (within- and between-subject). Again, the main model did not significantly explain the outcome better than the null model ($\chi^2 = 3.35$, $df = 5$, $p = .646$). Lastly, we compared the main model with a null model lacking the interaction between progesterone measures (within-subject) and the emotion category, which showed that the main and the null model are not significantly different ($\chi^2 = 2.07$, $df = 3$, $p = .556$). These comparisons imply that ovarian hormones and the interaction term did not predict emotion recognition considerably in our model.

Table 3

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

	Estimates	*SE*	*z*	*p*	*OR*	*95% CI*
Model Hormones						
Estradiol (within-subject)	0.058	0.183	0.315	0.753	1.06	0.74 – 1.52
Progesterone (within-subject)	-1.512	1.325	-1.141	0.254	0.22	0.02 – 2.96
Estradiol (between-subject)	0.064	0.057	1.129	0.259	1.07	0.95 – 1.19
Progesterone (between-subject)	0.023	0.058	0.400	0.689	1.02	0.91 – 1.15

Progesterone (within-subject) × Anger	1.777	1.353	1.313	0.189	5.91	0.42 – 83.84
Progesterone (within-subject) × Disgust	1.425	1.356	1.051	0.293	4.16	0.29 – 59.37
Progesterone (within-subject) × Fear	1.174	1.407	0.835	0.404	3.24	0.21 – 50.97
Emotion [Anger]	-3.625	0.257	14.095	**<0.001**	0.03	0.02 – 0.04
Emotion [Disgust]	-3.861	0.257	15.016	**<0.001**	0.02	0.01 – 0.03
Emotion [Fear]	-2.473	0.261	-9.460	**<0.001**	0.08	0.05 – 0.14
Session	0.286	0.049	5.846	**<0.001**	1.33	1.21 – 1.47

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

3.4 Exploratory Analysis

As preregistered, in *N* = 131 participants, we investigated if there was a moderating effect of stimulus modality on the association of within-subject fluctuations of ovarian hormones and emotion recognition accuracy. Session number was included as a control variable, subject ID was added as a random intercept, and stimulus modality as a random slope. Our analysis did not reveal a significant interaction between ovarian hormone levels and stimulus modality. As in previous models, participants showed a better performance in the second session. The model showed the significant main effect of stimulus modality in which the audiovisual expression has the highest recognition rate compared to the other two modalities (see supplementary document).

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

4 Discussion

The current study aimed at understanding the within-subject differences between ovulatory cycle phases (late follicular and mid-luteal), associated ovarian hormone levels and the recognition of emotions expressed in visual, auditory, and audiovisual modalities within a large-scale sample of healthy, naturally cycling females. We expected a higher accuracy of facial emotion recognition in the late follicular phase as compared to the mid-luteal phase, a positive relationship between levels of estradiol and facial emotion recognition accuracy, and a negative relationship between levels of progesterone and facial emotion recognition accuracy. We also predicted a positive association between levels of progesterone and the recognition of threat-related emotions (anger, disgust, and fear) presented in faces, known as negativity bias. In an exploratory manner, we investigated whether within-subject differences in ovarian hormones fluctuation across the late follicular and the mid-luteal phase of the cycle and emotion recognition differs among visual, auditory, and audiovisual modalities. Furthermore, we examined the interplay of within-subject ovarian hormone fluctuation and stimulus sex in emotion recognition.

4.1 No Compelling Evidence that Women's Emotion Recognition Ability Shifts between the Late Follicular and Mid-luteal Phase

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

The lack of differences in emotion recognition between cycle phases, contrary to previous studies (e.g., Derntl, Kryspin-Exner, et al., 2008, Rubin et al., 2012), might be explained by the methodological specificities of the studies. For instance, the current study employed a larger sample size as compared to most of the previous studies, used a within-subject design, and a large number of trials (144 trials per modality), which together resulted in higher statistical power. Moreover, the cycle phase estimation was confirmed via LH surge tests, and levels of ovarian hormones were directly measured in saliva. It should also be noted that different experimental setups could contribute to the diversity of findings in the field. As (Gamsakhurdashvili et al., 2021) suggested, applying a standardized emotion recognition task could address this issue in the future. Another possible explanation for the heterogeneity of results in the existing literature could be publication bias which refers to publishing excessive significant results while non-significant results remain underreported (Francis, 2012). Our results, however, are in accordance with a recent large-scale study (*N = 192*) by Shirazi and colleagues (2020), examining the association of ovarian hormone levels in the fertile and mid-luteal phase, and recognizing complex emotions using the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 2001). In line with our results,

the authors report no compelling evidence for a relationship between fertile and mid-luteal phase, ovarian hormones levels, and emotion recognition ability. Together, the findings of the current study and the study conducted by Shirazi and colleagues (2020) highlight the importance of employing a study design including high statistical power, within-subjects design, and direct hormonal measurements to study the association between ovulatory cycle phases and emotion recognition.

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

experiences (e.g., the interacting effect of facial and bodily expression along with the attractiveness, intelligence, personality, and familiarity of the portrayer) and require enriched sensory stimulation.

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

4.2 Limitations

We only collected data in the estimated late follicular and mid-luteal cycle phases and assessed hormone levels therein; however, as recommended in a recent study by Stern and colleagues (2021) and a recent review by Gamsakhurdashvili and colleagues (2021), including more than two testing sessions (e.g. including the early follicular phase, or the premenstrual phase) might create a better contrast to show the possible effects of hormonal variation across the ovulatory cycle. Given that there is a second estradiol peak in the mid-luteal phase, a third session scheduled to collect data in a cycle phase characterized by low estradiol levels (e.g. early follicular or late luteal) might have provided a better insight into differential effects of estradiol levels. Furthermore, to provide a more reliable within-subjects measure for the random effect in the multilevel model, (Schmalenberger et al., 2021) suggested to include at least three observations per cycle. In this study, however, due to practical concerns, we were only able to observe each participant twice per cycle. In addition, 43% of the participants (*n = 57*) did not observe positive LH tests during the ideal days and were therefore omitted from the main model that investigated the link between ovulatory cycle phase and (facial) emotion recognition (results remained virtually identical when including all

participants in these analyses). Although the rate of observed cycles with negative LH tests only seems high, it is in a range of reported values from previous studies. Nevertheless, to ensure the detection of ovulation, future studies should rather employ more than 10 LH tests per participant to ensure captioning delayed ovulation and let participants provide pictures of the LH tests to the study team to avoid misinterpretation of positive results. It is also recommended to measure hormones on a daily basis rather than just tracking the fertile phase by LH test to identify false negative LH test results (Marcinkowska, 2020). A strong limitation to the current and previous studies is the common approach to measure salivary estradiol and progesterone with immunoassays. Although the immunoassays approach is an easy and accessible method to measure gonadal steroids in saliva, liquid chromatography-mass spectrometry (LC-MS/MS) provides more sensitivity, validity, and accuracy in measuring steroid hormone levels (Arslan et al., 2022; Schultheiss, et al., 2019). Thus, the results of our hormone models should be interpreted in light of this limitation and are in need for replication with a more valid analysis method. Moreover, to achieve reliable inter-and intra-assay CVs in hormonal samples, it is recommended to analyze hormone samples in duplicates (Stern et al., 2021). Further, we did not control for potentially confounding physiological factors associated with the menstruation such as headache, cramps, or other premenstrual symptoms which could be a threat to the internal validity of previous studies (Kiesner et al., 2020).

One task-related limitation in this study could be not implementing different intensities in emotional expressions that also explained the presence of the ceiling effect in our data which potentially explains very wide confidence intervals regarding some interaction effects (proportion correct responses = *.90*). Another limitation associated with the task might be the unbalanced number of positive and negative stimuli. One of our hypotheses particularly aimed at investigating the link between negativity bias (improved recognition of threat-related emotions) and the within-

subject fluctuation of progesterone levels across the ovulatory cycle. To detect the negativity bias it is recommended to include a balanced number of positive and negative stimuli (Norris, 2019). Since we studied only a few emotions, the number of positive and negative emotions was not balanced in our design as happy expressions were the only positive emotion. The main reason for using basic emotions in this study was due to previous studies on basic emotions that would allow us to compare our findings with the existing limited research literature. Secondly, validated auditory databases are mostly restricted to basic emotion expressions, and therefore, to create balanced modalities in the emotion recognition task we were limited to basic emotions. Nevertheless, this issue should be improved in future studies by including different emotions ranging from basic to complex expressions to provide a balanced set of stimuli in terms of valence (Gamsakhurdashvili et al., 2021). Moreover, the use of emotional prosody with still faces in the audiovisual condition might decrease the ecological validity of the study, as in the real environment we experience moving faces along with emotional prosody (Collignon et al., 2010).

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

It is also worth noting that the current study counts as quasi-experimental, which means that the females' natural hormonal fluctuation was used (see Gingnell et al., 2019), and therefore drawing causal interpretation is not feasible from such a study. Studies employing hormonal

administration may contribute more to our causal understanding of behavioral and cognitive changes moderated by hormones (Gingnell et al., 2019).

4.3 Implications

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

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

Third, studies like the current one encourages the culture of publishing null findings which contributes to reducing the replication crises and publication bias. To be able to clearly define whether results are in favor of a null hypothesis or not, we recommend future studies to conduct

Bayesian analyses with a priori defined regions of practical equivalence or smallest effect sizes of interest.

5 Conclusion

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

Funding

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

Data Availability

The data and supplementary analysis for this study can be accessed on the Open Science Framework through the following link: https://osf.io/2pucf/

CHAPTER 3

Sex Differences in Emotion Recognition:

Investigating the Moderating Effects of Stimulus Features

Abstract

Emotion recognition – a prerequisite for social interactions – varies among individuals. Sex differences have been proposed as a central source of individual differences, although the existing evidence is rather heterogeneous. In the current study (N = 426), we investigated the potential moderating effects of stimulus features, including modality, emotion specificity, and the sex of the encoder (referring to the sex of the actor) on the magnitude of sex differences in emotion recognition. Our findings replicated women's overall better emotion recognition, particularly evident for negative expressions (fear and anger) compared to men. This outperformance was observed across all modalities, with the largest differences for audiovisually expressed emotions, while the sex of the encoder had no impact. Given our findings, future studies should consider these and other potential moderator variables to better estimate sex differences.

Keywords: sex differences, emotion recognition, facial expression, vocal expression, audiovisual.

1 Introduction

Emotion recognition plays a substantial role in shaping and maintaining interpersonal communication and social interactions. The underlying mechanisms of this ability are highly complex and yet not completely understood (Bänziger et al., 2016). In addition to contextual factors, several person-related factors have been proposed to determine emotion recognition, resulting in observable individual variations in this ability (see Lewis et al., 2016).

One of the most studied sources of individual differences in emotion recognition is the biological sex of the decoders (Kret and De Gelder, 2012; Connolly et al., 2019). The existing literature suggests that males and females differ in their skills in recognizing emotional expressions (see Kret and De Gelder, 2012), with women usually outperforming men (for an overview see Hall, 1978; Hall, 1984; McClure, 2000; Thompson and Voyer, 2014). This finding seems to be consistent across several studies applying different experimental tasks (e.g., dynamic stimuli, morphs, high vs. low-intensity stimuli) (Hall, 1978, 1984; Thompson and Voyer, 2014) or even across different geographical regions (e.g., Hall et al., 2010). In this line, an extensive review by Hall (1978) and a meta-analysis by Thompson and Voyer (2014) revealed persistent findings concerning the role of sex as a source of individual differences in emotion recognition.

In contrast to these findings, some studies reported a lack of assumed sex differences in emotion recognition (e.g., Buck et al., 1974; Grimshaw et al., 2004; Hoffmann et al., 2010; Lambrecht et al., 2014, Lyusin and Ovsyannikova, 2016; Rahman et al., 2004; Wagner et al., 1986). This discrepancy may be due to methodological limitations, such as the lack of statistical power which is highly dependent on the sample size and the number of trials. Previous studies reported the effect size for observing sex differences in global emotion recognition (as overall performance across all modalities) as small to moderate (see Hall, 1978; 1984; Thompson and Voyer, 2014); therefore,

high statistical power is required to detect the hypothesized sex differences. Moreover, some previous studies only concentrated on sex differences in global emotion recognition. Hence, there are yet unanswered questions concerning the potential moderating effects of different stimulus features that might also have led to some mixed evidence or inaccurate estimation of sex differences in emotion recognition. Among the most considered features were stimulus modality, emotion specificity, and sex of the encoder (referring to the sex of the actor) (see Dores et al., 2020; Kret and De Gelder, 2012; Thompson and Voyer, 2014). Ignoring these moderating effects risks underestimating presumably existing sex differences in emotion recognition (Lausen and Schacht, 2018), and in fact, the magnitude of sex differences in emotion recognition has been demonstrated to depend on such variables (Hall, 1978, 1984; Thompson and Voyer, 2014).

Due to the multifaceted and highly relevant information that faces usually convey about individuals and their environment (Wilhelm et al., 2014), previous research mainly focused on sex differences in the recognition of *facial* expressions of emotion. However, the ecological validity of previous studies in this field is questionable, as emotions are usually expressed through more than one modality in natural interactions (de Gelder and Vroomen, 2000). Only a few studies attempted to investigate sex differences in recognizing emotions expressed through other modalities, such as by means of auditory and audiovisual stimuli. Significantly better performance of women compared to men has been indicated for auditory stimuli such as emotional prosody (e.g., Lambrecht et al., 2014) and vocal expressions of emotion (e.g., Paulmann and Uskul, 2014). This pattern was replicated in a recent large-scale study ($N > 300$) by Lausen and Schacht (2018), although the magnitude of the sex differences was rather small. In another large-scale and cross-cultural study, Scherer et al. (2001) also found a small but significant main effect of sex, reflecting women's better performance in vocal emotion recognition compared to men. Similar to emotion

recognition in the visual and auditory modalities, there is evidence for a female advantage in the recognition of audiovisual expressions of emotion (e.g., Hall 1978; 1984).

Another conceivable moderator of sex differences in emotion recognition is emotion specificity (for an overview see Hall 1978, 1984; Thompson and Voyer, 2014). Several studies reported an advantage for women over men in recognizing negative emotions; however, this pattern was not observed in positive emotions (e.g., happy expressions) as men and women performed equally well in recognizing them (Rotter and Rotter, 1988; Hoffman et al., 2010; Connolly et al., 2019; Montagne et al., 2005; Hampson et al., 2006, 2021; for an overview see Thompson and Voyer, 2014). Negative emotions are in general more difficult to recognize, and therefore women's outperformance is more evident due to higher variability in recognizing negative emotions (Thompson and Voyer, 2014). From an adaptive view, women's superiority in recognizing negative emotions could play an important functional role in the survival of self and offspring. Babchuk et al. (1985) proposed the "primary caretaker" hypothesis, which claims that a women's advantage in emotion recognition arises from within the scope of their responsibility in childcare. In this line, Taylor et al. (2000) suggested that women, due to different parental investments, tend to have different stress responses to environmental threats, which are assumed to have been selectively evolved during evolution to increase the survival of self and offspring. Related to this is the "fitness threat hypothesis" which emphasizes the relevance of recognition of negative emotions in reproductive success (Hampson et al., 2006). The "fitness threat hypothesis" gained preliminary support in a study by Hampson et al. (2006), showing women's quicker responses in recognizing negative emotions compared to men. Interestingly, the previous childcare was not a predictive factor, which suggests that the women's advantage in recognizing negative emotions is an evolved mechanism and is not learned through childcare experience. In addition, this finding

was replicated in a recent study by Hampson et al. (2021), where they showed women's higher accuracy in recognizing negative emotions in infants, toddlers, and adult faces compared to men.

Sex differences were studied not only in the ability to recognize but also to express emotions. Some studies suggested women's higher expressiveness and expression accuracy compared to men (e.g., Wagner et al., 1986) implying the potential moderating effect of the sex of the encoder on sex differences in emotion recognition (see Hall, 1984; Hall et al., 2010). Evidence is however mixed concerning the interaction of the sex of the decoder (referring to the sex of the participant) and the sex of the encoder, with some studies showing better women's recognition of male expressions (Collignon et al., 2010; Davitz, 1964; Tagiuri, 1969; Thompson and Voyer, 2014) and others showing no moderating effects of the encoder's sex (Hall, 1978).

In sum, the partially inconsistent findings on sex differences in emotion recognition might not only be due to the disregard of the potential moderator variables described above. In addition, several earlier studies had potential methodological limitations. Remarkably, only a few studies investigated the hypothesized sex differences a) within a large sample size and b) by implementing enough experimental trials, both necessary for gaining appropriate statistical power to detect sex differences. Further, many previous studies were constrained to only one stimulus modality as well as to very few distinct emotions. Thus, their findings have restricted generalizability to fundamental variations in emotion recognition abilities (e.g., Collignon et al., 2010).

In light of of these limitations, the present study aimed, first, to replicate indicated sex differences in emotion recognition, by studying a large-scale healthy population (N = 426) in a well-established study design, including a considerable number of trials (432/session). Second, we aimed to investigate to what extent the assumed sex differences might be moderated by different stimulus features, i.e., modality, emotion specificity, and sex of the encoder. We predicted:

(*H1*) a global advantage in emotion recognition for women over men;

(*H2*) an advantage in emotion recognition specifically for expressions of negative emotions (fear, anger, disgust, sadness) for women over men, indicated by an expected sex of decoder × emotion specificity interaction;

We attempted to explore the following hypotheses due to inconsistencies and lack of evidence:

(*H3*) the possible interaction effect between encoder's and decoder's sex in emotion (sex of encoder × sex of decoder);

(*H4*) if there are sex differences in recognizing emotions from visual, auditory, and audiovisual modalities (sex of decoder × stimulus modality).

Finally, we examined the possible confusion in the inferred and intended category of emotions, known as the confusion matrix, in men and women.

2 Methods

The data was obtained by merging two studies that were using the same set of stimuli and experimental procedure in two samples of healthy men (Lausen et al., 2020) and women (Rafiee et al., 2023). Both studies have been approved by the local ethics committee of the Institute of Psychology at Goettingen University. Each participant signed a written consent form according to the Declaration of Helsinki (DoH) ethical principles for human subjects before their participation. After completing the certain experiment, participants were reimbursed with a reasonable amount of money or course credits. The data and script can be found on Open Science Framework (OSF) under this link: https://osf.io/u2khx/

2.1 Participants

Included participants were healthy, native German speakers with no history of psychiatric or neurological diseases, and had normal or corrected to normal vision and hearing, according to self-report. In total $N = 426$ ($M_{Age} = 24.2$, $SD_{Age} = 3.63$, Range = 18-36; 145 females) individuals were included in the analysis. Participants' characteristics are summarized in **Table 1**.

Table 1

Demographics of the participants. Please note that the numbers in parentheses within the categories "Relationship Status" and "Hand Dominance" indicate the exact number of participants.

	Female (N = 145)	Male (N = 281)
Age in years, mean (SD)	24.0 (3.39)	24.3 (3.75)
Relationship Status		
Single	47.6% (69)	54.1% (152)
In a relationship	46.2% (67)	40.2% (113)
Engaged	1.4% (2)	1.1% (3)
Married	2.1% (3)	1.4% (4)
Others	2.1% (3)	3.2% (9)
Hand		
Right	91% (132)	90% (252)
Left	9% (13)	10% (29)

2.2 Procedure

The dataset was merged from two separate studies applying the same experimental design and task requests. The study collecting data from women was a longitudinal study (cf. Rafiee et al., 2023); however, we only included women's performance in their first experimental session to eliminate any effects of learning and practicing, respectively. Upon arrival, female participants filled out a computer-based screening questionnaire related to their health and hormonal assessment (adapted from Jünger et al., 2018), followed by a German version of the Positive and Negative

Affect Schedule (PANAS) questionnaire (Breyer and Blumke, 2016). Next, saliva samples were collected prior the beginning of the emotion recognition task described below.

Men were recruited for one session only. Participants filled out a computer-based screening questionnaire related to their health and hormonal assessment (adapted from Schultheiss and Stanson, 2009), Multi-Motive Grid Questionnaire (MMG) (Sokolowski et al., 2000), and PANAS questionnaire (Breyer and Blumke, 2016) at the beginning of the session. Next, their saliva samples were collected and then participants performed the emotion recognition task. Neither the obtained questionnaire data nor the hormonal measures extracted from saliva samples were in the scope of the current study.

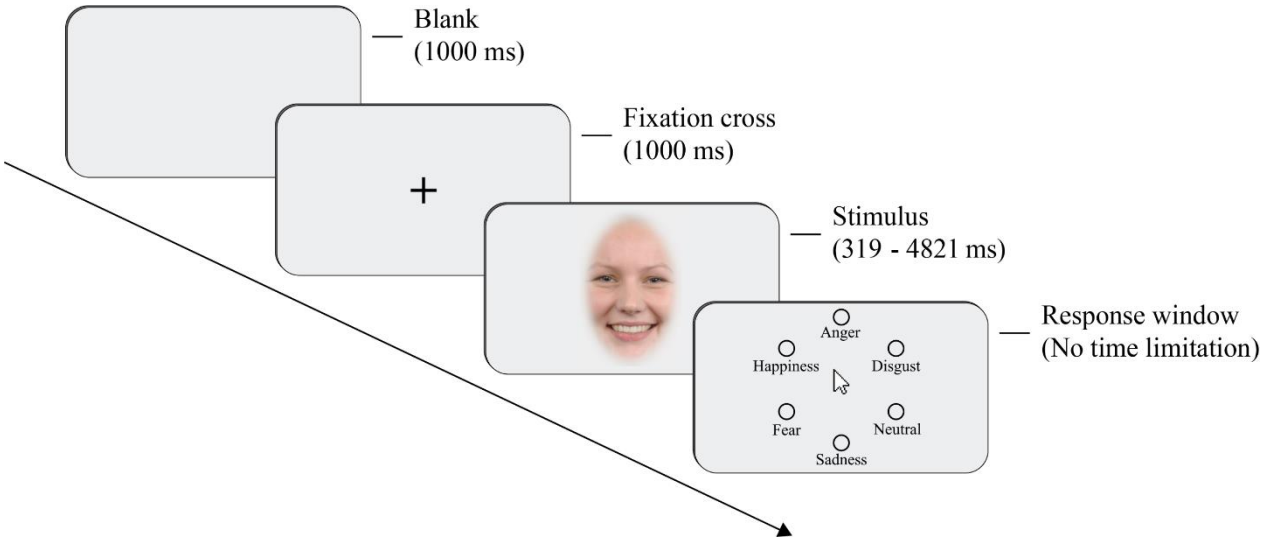
2.3 Emotion Recognition Task

All participants performed a computer-based emotion recognition task, including visual, auditory, and audiovisual stimuli. The visual modality was extracted from Radboud face database (RaFD, Langner et al., 2010), and the auditory modality included affect burst, pseudo-words, and pseudo-sentences extracted from Montreal Affective Voices (Belin et al., 2008), Magdeburg Prosody Corpus (Wendt and Scheich, 2002), and Paulmann Prosodic Stimuli (Paulmann et al., 2008) respectively. In each experimental session, a total of 432 stimuli were presented in three different modalities (144 each) and containing of five emotional expressions (anger, fear, sadness, happiness, and disgust), and neutral expressions (24 per emotion × modality). The order of stimuli and each modality block was randomized across participants (see Lausen et al., 2020).

Trials started with a blank screen for 1000 ms followed by a fixation cross in the center of the screen for 1000 ms. Following the presentation of the target stimulus (varying in duration between 319 and 4821 ms), a circular answer display appeared, containing all six (emotion) categories of

interest, together with the selection cursor, in the center of the display. The arrangement of the emotion labels was counterbalanced but remained constant throughout the task. Participants had to select an emotion category by using the mouse to move the cursor. They were required to select the corresponding emotion label as accurately and quickly as possible, although there was no time limitation for responding to stimuli. Emotion recognition accuracy and response time (relative to the onset of the answer display) were measured. Responses were considered hits when they matched the expected label with the name of the target expression. Before the main experiment, three practice trials were presented to familiarize the participants with the experimental procedure.

Figure 1: *Schematic overview of an exemplary trial displaying happiness in visual modality.*



2.4 Statistical Analysis

Using G*Power, the post-hoc power analysis revealed 99% statistical power for a small effect size (f2 = 0.15) in the population (r = .15; α = .05; 1 – β = .99) (Faul et al., 2007; 2009). We further conducted a comparison between our sample size and an estimated sample size obtained from a prior power analysis. The analysis indicated that a sample size of *N* = 211 participants was required for the same effect size (f2 = 0.15), with α = .05 and 1 – β = .99. Consequently, our sample size

exceeded the required number of participants, providing evidence that our study possesses statistical power greater than $1 - \beta > .99$.

To analyze the data, R statistical software (version 4.0.5) and R studio (version 1.4.1106) were applied. A p-value of 0.05 was set as the cut-off criterion for two-tailed distributions. Data was preprocessed using the following packages, tidyverse (version 1.3.1), knitr (version 1.39), and dplyr (version 1.0.9). We used ggplot2 (version 3.3.6) and sjPlot (version 2.8.10) for data visualization.

To test the first hypothesis - sex differences in global emotion recognition - we used the Mann-Whitney U test by comparing the relative frequency of hits between the two sexes. To examine the moderation effects of stimulus modality, sex of encoder, and emotion specificity with the sex of decoder (H2, exploratory analysis: H3, H4), we fitted one Generalized Linear Model with quasibinomial error structure and logit link function (McCullagh and Nelder, 1989), with the outcome variable as emotion recognition accuracy (correct vs. false), and stimulus features as predictors including stimulus modality, sex of encoder, and emotion specificity. Stimulus duration (log-transformed) was entered into the regression model as the control variable.

To examine the absence of collinearity among predictors, Variance Inflation Factors (VIF; Field et al., 2012) was applied, using the function vif of the package car (version 3.0.10; Fox and Weisberg 2019). This showed no collinearity issue among predictors (maximum VIF: 1.059). In addition, model stability was assessed to find influential observations, which were not present in our model. The goodness of fit of the fitted model was assessed by likelihood ratio test by comparing the deviances of the fitted model with the minimal model lacking the predictor of interest.

3 Results

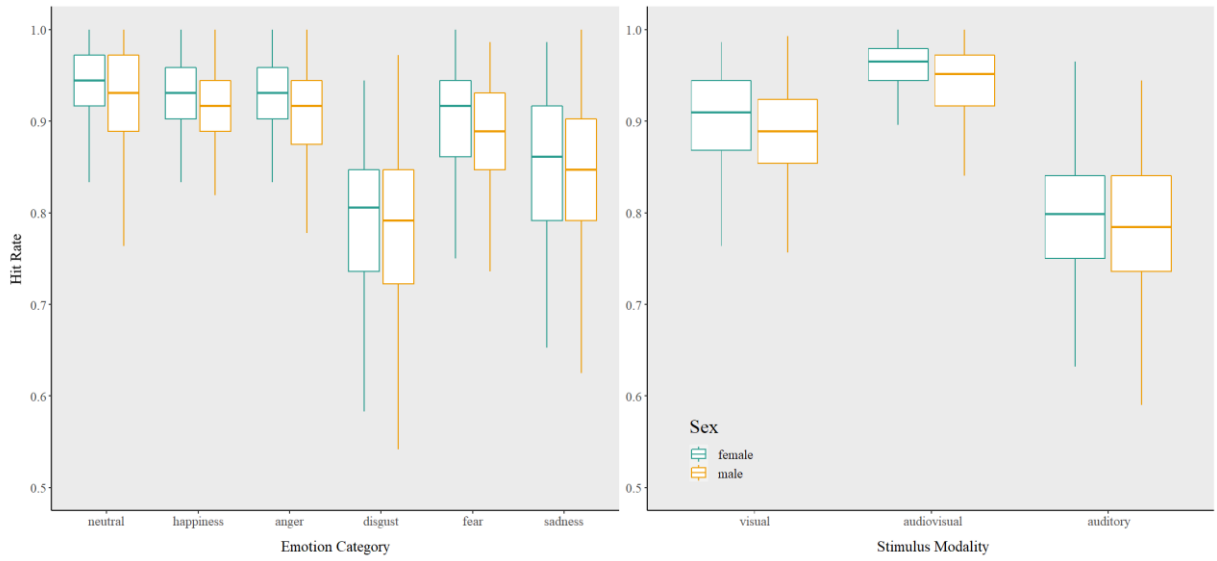
The non-parametric Mann-Whitney U indicated that the global emotion recognition accuracy was higher for women than for men (W = 16116, $P < 0.001$). To further investigate the potential moderating effect of stimulus modality, sex of encoder, and emotion specificity on the magnitude of sex differences in emotion recognition, we fitted a generalized linear model (GLM), with the mentioned predictors and accuracy as the outcome variable. In addition, the relationship between emotion specificity and sex of decoder (emotion specificity × sex of decoder), stimulus modality and sex of decoder (stimulus modality × sex of decoder), and sex of encoder and sex of decoder (sex of encoder × sex of decoder) were entered as interaction terms. To control for the effect of stimulus duration on responses, we included duration (log-transformed – base e) as a control variable. The reference categories were visual for stimulus modality, neutral expression for emotion specificity, and male for both sex of the encoder and sex of the decoder.

The GLM revealed a significant moderating effect of emotion specificity on sex differences in emotion recognition. Women showed a better performance than men in recognizing anger and fear but not sadness and disgust, partially supporting H2 (*female decoder* × *anger*: β = 0.302 , SE = 0.09, 95% CI = [1.19 – 1.54], t = 4.675 , OR= 1.35 , $p < 0.001$; *female decoder* × *fear*: β = 0.193, SE =, 95% CI = [1.08 – 1.37], t = 3.170, OR= 1.21 , $p = 0.002$; Figure 2, left panel). The model outcome did not show a significant interaction of either encoder's sex on sex differences in recognizing emotions (Table 2). We also investigated the moderating effect of stimulus modality and sex differences on emotion recognition. The model showed a significant moderating effect of stimulus modality × sex of decoder in a way that women outperformed men in recognizing emotions expressed in all modalities (*female decoder* × *auditory modality:* β = -0.173, SE = 0.03, 95% CI = [0.78 – 0.90], t = -4.839, OR= 0.84, $p < 0.001$*; female decoder* × *audiovisual modality:*

β = 0.123, SE = 0.06, 95% CI = [1.03 – 1.24], t = 2.519, OR= 1.13, *p* = 0.012; Table 2 contains the GLM results). The most pronounced differences between the sexes were found in the ability to recognize emotions in the audiovisual modality, whereas the least pronounced differences were found in the recognition of emotions expressed in the auditory modality (note that the intercept corresponds to the performance of males in recognizing neutral expressions conveyed by male encoders in the visual modality. These findings are based on dummy coding and compared to the intercept).

Subsequently, we compared the main model with a model lacking sex of decoder and its interactions with each of the stimulus features. The likelihood ratio test revealed that the main model, including sex and its interaction with stimulus features, predicts the outcome variable significantly better (*Deviance* = 236.91, df = *9, p* <0.001). Furthermore, the main model was compared to a null model omitting only the interaction terms. Again, the main model explains the outcome variable significantly better than the null model (*Deviance* = 109.736, df = *8, p* <0.001). Overall, these comparisons indicate that not only sex of decoder but also features of the stimulus are significant predictors of emotion recognition.

Figure 2: *The boxplots depict the distribution of the mean hit rate (accuracy) with medians per emotion category (left panel) and modality (right panel), for female and male participants.*



The results of the confusion matrix indicated that the least recognized expression of emotion was disgust in the auditory modality and the most recognized emotion expression was happiness in the visual modality in men and women (see Table 3a and 3b in the supplementary document).

Table 2

Results of Generalized Linear Model investigating the association of stimulus features and sex differences in emotion recognition, N = 426 (females = 145). Note, fem = female.

	Estimates	*SE*	*t*	*p*	*OR*	*95% CI*
Decoder's sex [fem]	0.158	0.06	2.915	**0.004**	1.17	1.05 – 1.30
Emotion [Anger]	-0.283	0.03	-8.035	**<0.001**	0.75	0.70 – 0.81
Emotion [Disgust]	-1.264	0.01	-39.986	**<0.001**	0.28	0.27 – 0.30
Emotion [Fear]	-0.563	0.02	-16.651	**<0.001**	0.57	0.53 – 0.61
Emotion [Happiness]	-0.086	0.03	-2.347	**0.019**	0.92	0.85 – 0.99
Emotion [Sadness]	-0.895	0.01	-27.215	**<0.001**	0.41	0.38 – 0.44
Encoder's sex [fem]	0.22	0.02	12.557	**<0.001**	1.25	1.20 – 1.29
Modality [Audiovisual]	0.722	0.05	27.892	**<0.001**	2.06	1.96 – 2.17
Modality [Auditory]	-0.736	0.01	-37.108	**<0.001**	0.48	0.46 – 0.50
Stimulus duration (log- transformed)	-0.002	0.01	-0.175	0.861	1.00	0.98 – 1.02
Decoder's sex [fem] × Emotion [Anger]	0.302	0.09	4.675	**<0.001**	1.35	1.19 – 1.54
Decoder's sex [fem] × Emotion [Disgust]	-0.066	0.05	-1.195	0.232	0.94	0.84 – 1.04
Decoder's sex [fem] × Emotion [Fear]	0.193	0.07	3.170	**0.002**	1.21	1.08 – 1.37
Decoder's sex [fem] × Emotion [Happiness]	0.084	0.07	1.296	0.195	1.09	0.96 – 1.24
Decoder's sex [fem] × Emotion [Sadness]	0.033	0.06	0.573	0.567	1.03	0.92 – 1.16
Decoder's sex [fem] × Encoder's sex [fem]	0.044	0.03	1.382	0.167	1.04	0.98 – 1.11
Decoder's sex [fem] × Modality [Audiovisual]	0.123	0.06	2.519	**0.012**	1.13	1.03 – 1.24
Decoder's sex [fem] × Modality [Auditory]	-0.173	0.03	-4.839	**<0.001**	0.84	0.78 – 0.90

4 Discussion

Emotion recognition, as an essential ability for social beings like humans, varies among individuals. One of the most frequently discussed sources of individual differences in emotion recognition has been the sex of the decoder, with a trend for women performing better than men. In contrast, there were contradictory findings indicating no sex differences (e.g., Buck et al., 1974; Hoffmann et al., 2010; Lambrecht et al., 2014; Wagner et al., 1986), likely due to low statistical power such as insufficient sample size and trial number. The effect size reported in previous studies tended to be small to moderate, indicating that high statistical power is required to detect the effect of interest. In addition, other potential moderating effects of stimulus features on the magnitude of sex differences in emotion recognition are understudied, which potentially lead to inaccurate estimation of sex differences. The present study aimed to fill this gap by investigating the moderating effects of stimulus modality, emotion specificity, and sex of the encoder on sex differences in emotion recognition, in a large-scale healthy population (N = 426). Going beyond previous studies focusing on emotion recognition in faces, this study also used vocal and bimodal expressions with the aim of increasing ecological validity.

4.1 Sex Differences in Emotion Recognition

In line with previous studies, our findings indicate that women are more accurate in recognizing emotions. In the current study, the obtained effect size concerning sex differences in global emotion recognition was rather small (d = .17), which also corresponds with the reported effect size (d = .19) in an extensive meta-analysis including 215 studies (see Thompson and Voyer, 2014). Our findings also indicated that stimulus features including emotion specificity and stimulus modality

significantly contribute to variance in effect size and thus supports the proposed necessity for considering these factors as predictors of sex differences in emotion recognition.

In our second hypothesis, we predicted women's advantage specifically in recognizing negative emotions (anger, disgust, fear, and sadness). As hypothesized, women were more accurate in recognizing anger and fear than men; however, this pattern was not observed in the recognition of disgust and sadness. We suggest two potential explanations for not observing sex differences in the recognition of disgust and sadness expressions.

First, our results may have been influenced by the study design (e.g., exposure time, response time limitation) and the stimulus material used. The recognition rate of disgust and sadness expressions was found to be the lowest among the emotional expressions respectively (see Table 2), with poor performance of both men (hit rates: disgust = 0.78; sadness = 0.83) and women (hit rates: disgust = 0.78; sadness = 0.84). This likely led to no significant differences in recognition performance between the sexes. However, there were significant differences in the recognition rates for anger and fear expressions between female (hit rates: anger = 0.93; fear = 0.89) and male decoders (hit rates: anger = 0.89; fear = 0.87; see also supplementary material, Tables 3a, 3b). Even, using the same stimuli can also result in varying outcomes probably due to different study designs. For example, a study using the same visual stimuli (RaFD) in a large sample of 1249 participants (15.4% male) with a 10-second exposure time and 1-second response time indicated sex differences only in the recognition of anger (and contempt) expressions (Dores et al., 2020).

Second, the higher survival priority of fear and anger recognition relative to other emotions may be another explanation for the discrepancy in results. Recognizing fear and anger conveys information about potential threats and dangers, making their recognition highly relevant to survival (e.g., Skuse, 2003). Although disgust expressions also carry (biologically) important

information, its level of indicated danger might be perceived as less immediate and severe compared to fear and anger in the context of laboratory studies. Related to this point, many previous studies have shown sex differences mostly in the recognition of anger expressions, likely due to its high saliency, while sex differences in recognizing other negative emotions have been inconsistently reported. A meta-analysis by Thompson and Voyer (2014) showed that only anger had a significant impact on the variability of sex differences in emotion recognition.

Like emotion specificity and in line with previous studies (Hall, 1978; 1984, Thompson and Voyer, 2014), stimulus modality had a significant moderating effect on the magnitude of sex differences in the current study. Women performed better across all modalities compared to men and, remarkably, the largest difference in hit rates was observed when comparing women's performance in audiovisual modality to men. These findings support Collignon et al.'s (2010) study, where women were not only better at recognizing unimodal emotional expressions (visual and auditory), but also at integrating visual and auditory information, leading to their advantage in recognizing audiovisually expressed emotions.

Our study did not reveal that the interaction between sex of the encoder and sex of the decoder influences emotion recognition, in line with some previous reports (e.g., Hall, 1978; Harris et al., 2016; Lausen and Schacht, 2018; Thayer and Johnsen, 2000), but contrary to others (e.g., Collignon et al., 2010; Davitz, 1964; Tagiuri, 1969; Thompson and Voyer, 2014). The heterogenous findings might be due to utilizing a different set of stimuli and or the authenticity of the presented stimuli. Although we used an established and validated stimulus set in our study, it could be argued that the moderating influence of encoder's sex on emotion recognition depends to some extent on other contextual factors. Emotions expressed by female encoders in our study were recognized with greater accuracy, suggesting a potential advantage in their expressive abilities compared to male

encoders; however, null findings regarding the interaction effect of encoder and decoder's sex in emotion recognition call into question the own-sex bias in emotion recognition.

Overall, our study demonstrated that emotion specificity and stimulus modality, but not the sex of the encoder have a moderating effect on sex differences in emotion recognition. Therefore, we, recommend including both variables in future studies of emotion recognition in order to obtain a more precise assessment of sex differences in this domain.

4.2 Implication

Sex differences in emotion recognition have been consistently replicated in numerous studies, including the current study. The methodological strengths of this study support the reliability of the effect, although it is small in magnitude. The small effect may not have significant implications in everyday life. However, it may have a greater implication in clinical settings and in the diagnosis of gender-specific psychiatric disorders. For example, considering sex differences in emotion recognition is important when designing diagnostic tools for disorders with different prevalence in women and men. Evidence suggests that autism is often diagnosed at a later age in females than males autism due to their higher functioning in social interaction (Wood-Downie et al., 2021), highlighting the need to consider sex differences in the implementation of diagnostic instruments.

4.3 Limitations and Outlook

One of the potential limitations of the current study may be the combination of two datasets collected in two large studies (see Lausen et al., 2020; Rafiee et al., 2023). Consequently, the proportion of men and women in the current study was imbalanced (N = 426, 145 women), due to the strict inclusion criteria for recruiting women in the study by Rafiee et al. (2023) that focused

on ovulatory cycle shifts in emotion recognition. A gender-balanced sample should be preferred in future studies to control for the main and interaction effects of encoder and decoder sex in emotion recognition in future studies (Lausen and Schacht, 2018; Dores et al., 2020). Another limitation relates to the stimulus material: In the audiovisual condition, we presented static faces combined with congruent emotional prosody. A combination of dynamic facial expressions with emotional prosody could further increase ecological validity (Collignon et al., 2010). In addition, implementing a different range of intensities of the emotional expressions could have sex differences become more evident. For instance, Hoffmann et al. (2010) did not find sex differences in the recognition of facial expressions with 100% intensity, i.e. when the expressions could be easily recognized; however, they did find sex differences with women outperforming in the recognition of only subtle facial expressions.

Some studies have suggested that the multiple-choice format has low ecological validity as real-life situations often do not provide a list of potential responses (Georgopoulos et al., 2022). This might lead to draw inferences about the displayed emotion not elicited by the stimulus itself (e.g., Betz et al. 2019; Georgopoulos et al., 2022; Russell, 1994). However, the multiple-choice format has several advantages, including reduced missing data due to the requirement to choose a label, decreased response variability, and easier data coding (Dores et al. 2020).

To achieve a coherent understanding of the moderating effects of stimulus features in sex differences in emotion recognition, we encourage future studies to include more variation in stimulus intensities and to implement a wider range of emotional expressions (e.g., complex emotions), more lifelike stimuli (e.g., authentic expressions), as well as different stimulus modalities (e.g., bodily expression). As a prerequisite, the generation and validation of new sets of stimuli would be recommended.

5 Conclusion

The present study replicated significant differences in emotion recognition between men and women, indicating that sex is a crucial factor in explaining individual variations in this ability. Remarkably, our study indicates that the characteristics of the stimuli, such as emotion specificity and stimulus modality, play a role in the magnitude of these sex differences. Exploring possible moderating factors that influence sex differences in emotion recognition could provide a deeper understanding of both variation and commonalities across individuals.

Funding

This work was supported by funding from the German Academic Exchange Service (DAAD), the Research Training Group (RTG) 2070 “Understanding Social Relationships” under the grant (254142454 / GRK 2070), and the Leibniz ScienceCampus “Primate Cognition” under the grant (LSC-SF2018–10).

Data availability statement

The data that support the findings of this study are openly available on Open Science Framework at https://osf.io/u2khx/

CHAPTER 4

Does the Interplay of Emotion-related Personality Traits and Reproductive Hormones Predict Individual Variation in Emotion Recognition?

Abstract

Person-related variation has been identified in many socio-cognitive domains, and there is evidence for links between certain personality traits and individual emotion recognition. Some studies, using the menstrual cycle as a hormonal model, attempted to demonstrate that hormonal fluctuations could predict variations in emotion recognition, but with merely inconsistent findings. Remarkably, a potential interplay of hormone fluctuations and other potentially influential person-related factors in emotion recognition is yet understudied. In the current study, we examined if the interactions of emotion-related personality traits, namely openness, extraversion, and neuroticism, and the ovulatory cycle predict individual variation in facial emotion recognition in healthy naturally cycling women. We collected salivary ovarian hormones measures from $N = 131$ ($n = 74$ validated via LH test) women across their late follicular and mid-luteal phases of the ovulatory cycle. Individuals with higher scores in openness performed better in emotion recognition in the late follicular phase, confirmed via LH test, as compared to the mid-luteal phase. We also found that

higher neuroticism scores were negatively associated with emotion recognition in the mid-luteal phase and elevated levels of progesterone. Overall, the current study emphasized the significant role of person-related factors' interactions in predicting individual variation in emotion recognition.

Keywords: facial emotion recognition, individual variation, personality, menstrual cycle, ovarian hormones

1 Introduction

1.1 Individual Differences in Cognitive Capacities

Individual variations have been observed in many cognitive domains, including emotion recognition (Bänziger, 2016), which is the ability to decode the emotional states of others presented through facial, vocal, and body expressions, and therefore contributes significantly to interpersonal communications (Schlegel et al. 2019). While some person-related factors have been identified to associate with individual variation in emotion recognition (e.g., Laukka et al., 2021), the specific determinants of individual variation and their potential interactions in this domain are not completely understood, revealing incomplete and inconsistent evidence.

1.1.1 Person-related Factors Relevant to Emotion Recognition

A proposed factor contributing to individual variation in emotion recognition is personality, which is "a dynamic organization, inside the person, of psychophysical systems that create the person's characteristic patterns of behavior, thoughts, and feelings" (Carver and Scheier, 2004, p. 5). One of the most recognized models to assess personality is the *Big Five* representing traits including extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience (Costa & McCrae, 1980; Sundin et al., 2020). The big five model asserts that almost all traits and their descriptions can be classified in five broad domains (DeYoung & Gray, 2009).

The relationship between personality traits and emotion recognition has been investigated in a limited number of studies with inconsistent results. There is evidence for positive associations between emotion recognition and openness (Matsumoto et al., 2000; Terracciano et al., 2003) and extraversion (Matsumoto et al., 2000; Scherer and Scherer, 2011), whereas neuroticism has been linked to both poorer (Matsumoto et al., 2000) and better (Cunningham, 1977) recognition of facial

emotional expressions. Individuals with high openness scores have been supposed to be more receptive to environmental stimuli, including emotional states and expressions of others (Matsumoto et al., 2000). Similarly, the personality-relationship transaction framework (Asendorpf & Wilpers, 1998) suggests that extraverted individuals tend to be more social and have broader social contacts, making it likely that those with high openness and extraversion are better at recognizing emotions. Individuals with higher neuroticism scores might have impaired emotion recognition due to their tendency to experience negative emotions that might induce avoidance of recognizing emotional states of others (Matsumoto et al., 2000).

1.1.2 Hormones and Emotion Recognition

Another person-related factor related to individual variations in emotion recognition is levels of neurochemicals such as hormones. Hormones are recognized as biomarkers of individual differences in behavior and cognition (Sundin et al., 2020). Studies have investigated the link between hormone levels (endogenous and exogenous) and emotion recognition ability, mainly with inconsistent results. Different hormones such as cortisol (Duesenberg et al., 2016), oxytocin (Di Simplicio & Harmer, 2016; Lischke et al., 2012; Schulze et al., 2011), the interplay of cortisol and testosterone (Lausen et al., 2020), and ovarian hormones including oral contraceptives (Derntl, Kryspin-Exner, et al., 2008; Derntl et al., 2013; Maner & Miller, 2014; Pahnke et al., 2019; Radke & Derntl, 2016; Marecková et al., 2014) have been investigated in relation to emotion recognition.

Overall, research suggested that not only inter-individual but also intra-individual differences in hormone levels could explain individual variation in emotion recognition (see Gamsakhurdashvili et al., 2021; Schmalenberger et al., 2021). One hormonal model, that represents both inter- and intraindividual hormonal levels, is the menstrual cycle. The considerable fluctuation of ovarian hormones, i.e. of estradiol and progesterone, across the menstrual cycle, has gained

noticeable research interest to study the link between the menstrual cycle and socio-cognitive functions in the last years (Poromaa & Gingnell, 2014). The menstrual cycle consists of follicular and luteal phases (Bull et al., 2019), with highest estradiol levels in the late follicular phase around the ovulation and highest progesterone levels in the mid-luteal phase (Kiesner et al., 2020).

Studying conceivable inter- and intra-individual variations in emotion recognition associated with the menstrual cycle is a recent and under-researched field; however, limited and inconsistent findings are prevalent due to a lack of standardization in research methodologies. Some studies indicate emotion recognition to be improved in the late-follicular phase (high estradiol) and impaired in the mid-luteal phase (high progesterone; for an overview see Osório et al., 2018; Gamsakhurdashvili et al., 2021). In contrast, other studies did not provide evidence for alteration in emotion recognition across the menstrual cycle (e.g., Di Tella et al., 2020; Kamboj et al., 2015; Pahnke et al., 2019; Rafiee et al., 2023; Shirazi et al., 2020; Zhang et al., 2013). Moreover, the lack of consideration for the joint role that different sources of individual differences might play in emotion recognition, contributes to inconsistent results in the literature.

1.2 Study Rationale

Women's unique personality dispositions might play a role in determining their emotional responding during a menstrual cycle (Wu et al., 2014a, b). Previous studies have focused on either investigating personality traits or menstrual cycle phases (associated with ovarian hormones' fluctuation) and their links to emotion recognition, while their potential interplay has been neglected so far. In the current study, we attempted to explore whether the interaction of emotion-related personality traits, i.e., openness, extraversion, neuroticism, and menstrual cycle phases and associated ovarian hormones as a model for inter- and intra-individual hormonal variation, play a role in predicting facial emotion recognition in women.

Considering the largest likelihood of conception in the late follicular phase, successful social interaction, requiring fast and accurate emotion recognition, might increase the chance of mating and maximizing reproductive success (see Gingnell et al., 2019). Individuals with extraversion or openness traits have been assumed to have higher mating opportunities (see Nettle, 2006), due to their tendencies to seek novelty in the environment and having more social contacts. Therefore, from an adaptive point of view it seems plausible that emotion recognition is pronounced in the late follicular phase vs. mid-luteal phase in these individuals which could also be associated with high levels of estradiol in the late follicular phase.

With regard to openness and extraversion, we hypothesized:

H1a) Individuals with higher score in openness have better emotion recognition performance in the late-follicular phase compared to the mid-luteal phase (openness × phase).

H1b) Higher scores in openness and higher levels of estradiol (within- and between-subject) is related to improved emotion recognition (openness × estradiol).

H2a) Individuals with higher score in extraversion have better emotion recognition performance in the late-follicular phase compared to the mid-luteal phase (extraversion × phase).

H2b) Higher scores in extraversion and higher levels of estradiol (within- and between-subject) is related to improved emotion recognition (extraversion × estradiol).

The mid-luteal phase is related to the highest levels of progesterone, which is correlated with negative mood (see Sundström-Poromaa, 2018) and claimed to increase responses to negative emotions (e.g., Conway et al., 2007; Guapo et al., 2009; Derntl, Kryspin-Exner, et al., 2008). Based on these assumptions, the 'window of vulnerability model' (Andreano et al., 2018) would indicate that levels of ovarian hormones in the luteal phase increases stress-related physiological and neural

responses that are mostly seen in affective disorders. This association between high progesterone level and beneficial recognition performance in negative expressions of emotion has not been fully replicated (e.g., Rafiee et al., 2023), although it should be considered that a higher responsiveness to emotional stimuli does not need be associated with an improved recognition performance. However, also the 'window of vulnerability model' was not supported by a recent well-established study indicating no increased affective symptoms during the mid-luteal phase (Guevarra et al., 2023).

Neuroticism is associated with stress and vigilance (Nettle, 2006) and has been demonstrated to be negatively related to emotion recognition (Matsumoto et al., 2000). The lack of consideration for the moderating effect of neuroticism on the relationship between menstrual cycle phase and emotion recognition has potentially contributed to the inconsistent results in previous studies. While ongoing debate and mixed evidence exists about the relationship between subjective emotional experience and the ability to decode others' emotions (e.g., Goldman and Sripada, 2005; Holland et al., 2021; Oberman et al., 2007; Wearne et al., 2018), it should be noted that the current evidence linking neuroticism and menstrual cycle phases (and hence ovarian hormone levels) to emotion processing is primarily limited to studies that investigate subjective emotional experiences such as emotional arousal, emotion regulation, and emotional eating disorders. Evidence from psychophysiological studies suggests that women with high levels of neuroticism show increased neural activity towards emotional stimuli during the luteal phase, compared to the follicular phase (Zhang et al., 2015). Additionally, these women showed heightened arousal levels, as indicated by galvanic skin response and heart rate, during different phases of their menstrual cycle (Wu et al., 2014b). These findings raised the question whether the assumed association between luteal phase (high levels of progesterone), and emotion recognition tended to be dependent on the personality traits and particularly neuroticism. Thus, there is a need for further research to directly examine the

link between the moderating effect of neuroticism on the association between cycle phase and emotion recognition.

To bridge the existing gap in the literature, we tested the following hypotheses:

H3a) Individuals with higher score in neuroticism have impaired emotion recognition performance in the mid-luteal phase compared to the late-follicular phase (phase × neuroticism).

H3b) Higher scores in neuroticism and higher levels of progesterone (within- and between-subject) are related to impaired emotion recognition (progesterone × neuroticism).

Due to the lack of prior evidence regarding other personality traits and emotion recognition, we did not hypothesize about consciousness and agreeableness.

2 Methods

The study was part of a broader project preregistered here (https://osf.io/dkpf5/) and approved by the local ethics committee of the Institute of Psychology at the University of Göttingen prior to data collection. Following the Declaration of Helsinki (DoH) for human experimentation, all participants acknowledged to take part in the experiment by signing written consent forms. Each participant received either course credit or monetary rewards (25 €) as compensation. The data was collected at the Affective Neuroscience and Psychophysiology laboratory, University of Göttingen. The data is available on the open science framework under this link: https://osf.io/fts2n/

2.1 Participants

One-hundred-eighty women were recruited in the initial data collection; however, the final sample consisted of N = 131 (M_{Age}= 24.1 in years, SD_{Age} = 3.47, $Range_{Age}$ = 18 – 35) women, who completed the experiment. Included participants were healthy, heterosexual, and German

native speakers, who had a natural and regular menstrual cycle (*Cycle length* = 25 to 35 days, $M_{Cycle\ length}$ = 28.77, $SD_{Cycle\ length}$ = 2.03) for at least three months prior to their participation in the study. Forty-nine participants were excluded from the study, primarily because they voluntarily withdrew due to personal reasons like lack of time or interest (44%). Other exclusion factors were an irregular menstrual cycle (32%) or the use of contraceptives (10%). Further details can be found in Rafiee et al. (2023). In the current study, n = 60 participants reported as single, n = 63 were in a relationship, two participants were engaged, three were married, and three reported to be in different forms of relationships. Note that two participants changed their relationship status from the first testing session to the second testing session, from "single" into "in an open relationship" (one participant), and from "other" into "in an open relationship" (one participant). One-hundred-nineteen participants were reported as righthanded.

2.1.1 Study Design and Procedure

The study was longitudinal and designed to include three sessions. In the first (introductory) session, the eligibility of participants was assessed according to the inclusion criteria (for more details see Rafiee et al., 2023). Next, participants filled out personality (BFI-2, Danner et al., 2016), empathy (MET, Dziobek et al., 2008), and motives (MMG, Sokolowski, et al., 2002) questionnaires (MET and MMG questionnaires were not in the scope of this study). Moreover, participants were instructed to use highly sensitive (10mIU/ml) urine ovulation test strips from Runbio Biotech Co., Ltd. in order to validate the next fertile phase of their menstrual cycle. They were asked to use the LH strips as soon as their menstruation ended until they saw positive results and send us the pictures of LH tests voluntary. LH tests were asked to use between 10 am to 8 pm to control for the possible diurnal factors.

The emotion recognition task took place in two separate sessions that were scheduled on the estimated dates of the late follicular and mid-luteal phases of the menstrual cycle, based on backward-counting methods (See Jünger et al., 2018; Puts et al., 2006). The late follicular phase was estimated as reverse cycle days 16-18, with reverse cycle day 16 as the most ideal date. The mid-luteal phase was considered reverse cycle days 4 to 11, with ideal days as days 6 to 8 (Jünger et al., 2018). In each session, upon arrival participants first completed a computer-based screening questionnaire concerning their health status and saliva sampling (adapted from Schultheiss & Stanton, 2009; Jünger et al., 2018), and the PANAS mood questionnaire (Breyer & Bluemke, 2016) (The mood questionnaire was out of the scope of the current study). As the next step, their saliva samples were collected via passive drooling into salicaps and were immediately stored in a fridge in -80 C. To control for the carry-over effect between sessions two and three, we randomized the menstrual cycle phases between two session. Consequently, n = 63 women started session two in their late-follicular phase and n = 68 started their second session in their mid-luteal phase. The intervals between the two testing sessions were 19.55 days on average (SD = 14.03). Due to possible diurnal fluctuations hormones (Bao et al., 2003; Liening et al., 2010) we scheduled session two and three between 12.00 to 4.00 pm and most of the participants were tested at the same time of the day in both session (see Rafiee et al., 2023).

2.2 Study Power

To achieve 80% statistical power to detect medium size effect (d = .5) in a within-subject study including two sessions across the menstrual cycle (fertile vs. non-fertile) and validation of fertile phase with LH test, a sample size of N = 48 participant was required (Gangestad et al., 2016). Our sample size (N = 131, n = 74 validated fertile phases via LH tests) therefore has sufficient statistical analysis to detect small to medium-sized effects (see Rafiee et al., 2023).

2.3 Measures

2.3.1 Assessment of the Menstrual Cycle and Ovarian Hormones

In the introductory session, we estimated the next day of menses based on the three last dates of participants' menstruation. The estimated date of ovulation was determined according to backward-counting method. According to these estimations, each participant took part in the experiments in their late-follicular and mid-luteal phases of the menstrual cycle. The late-follicular phase was validated via LH urine test. Due to the peak of estradiol in the late-follicular and progesterone's peak in the mid-luteal phases, we decided to study these two phases (see Blake et al., 2016; Gangestad et al., 2016).

The levels of estradiol (E2) and progesterone (P4) were measured in saliva during each phase of the menstrual cycle using Chemiluminescence Immunoassays at the Endocrinology Laboratory of the Technical University of Dresden. The samples were analyzed as single determination which is less accurate comparing to duplicate determination, and the reported coefficient of variation was below 10%. Furthermore, as an external validation to hormonal measures we performed a regression model between estradiol to progesterone ratio (E/P) and cycle phases, which showed a significant association between levels of hormones and cycle phases ($\beta = 0.116$, SE = 0.000, 95% CI = [0.11; 0.12], t = 133.1, $p < .001$).

Moreover, to investigate if levels of ovarian hormones fluctuate significantly across the late follicular and mid-luteal phases of the menstrual cycle, we fitted two linear mixed models for each of the hormones. In each model, one of the hormones (log-transformed) was included as the outcome variable, phase of the cycle as the predictor, and participant ID as the random effect. The model with estradiol as the outcome, revealed a significant drop of estradiol levels in the mid-luteal phase compared to the late follicular phase ($\beta = -0.09$, SE = 0.002, 95% CI = [-0.09; -0.10], p

<.001), and the model with progesterone as the outcome showed a significant rise of the progesterone levels in the mid-luteal phase compared to the late follicular phase (β = 0.84, SE = 0.002, 95% CI = 0.84; 0.85], *p* <.001).

We excluded hormone outliers ± 3 SDs from the sample mean (see Jones et al., 2018). Five hormonal measures including three estradiol and two progesterone measures were excluded from the data.

2.3.2 Assessment of Personality Measures

Personality traits were measured via the German version of the Big-Five Inventory (BFI-2), consisting of 60 sentences (Danner et al., 2016). The BFI-2 evaluates five main personality domains: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Extraverts are described as being outgoing, optimistic and enjoying social contacts, whereas neuroticism is related to stress and tendency to be worried (Canli et al., 2001; Costa & McCrae, 1980). Openness is often described as being aesthetically sensitive, intellectually curious, attentive to inner feeling, and imaginative (McCrae & Costa, 1985; Terracciano et al., 2003). Conscientiousness is associated with being reliable, hard-working and goal-directed (Matsumoto et al., 2000; McCrae & Costa, 1985). Individuals with high agreeableness are identified as being empathetic toward others, and typically are compassionate and polite (Eisenlohr-Moul & Owens, 2016).

Participants identified their answers to each sentence on a likert scale from one to five (1 = strongly disagree, 2 = somewhat disagree, 3 = neither agree nor disagree, 4 = somewhat agree, 5 = strongly agree). Each domain (personality trait) consisted of three subcategories called facets, represented by four sentences each. Personality trait scores were calculated by averaging the values selected for each personality trait and the associated facets (Danner et al., 2016).

2.3.3 Emotion Recognition Task

The emotion recognition task was adopted from Lausen et al. (2020) and consisted of three separate blocks of stimuli from the visual, auditory, or audiovisual modality, respectively. In the current study we only focused on facial expression, because the relevant available literature is limited to this domain of expression. Stimuli consisted of 24 face portrays (half females) per emotion (anger, disgust, sadness, fear, happiness, neutral), chosen from the Radboud face database (Langner et al., 2010), and matched in their luminance (see Lausen et al., 2020). The emotion recognition task measured emotion recognition accuracy and reaction times; however, there was no time limitation for the responses to the trials.

To familiarize participants with the experimental procedure, the task began with three practice trials. All trials started with a blank screen for 1000 ms, followed by a fixation cross in the center of the screen for 1000 ms. Afterwards, the target stimulus, varying in duration between 319 and 4821 ms, appeared in the center of the screen. Subsequently, a circular multiple-choice answer display with labelled emotion categories was presented in the center of the screen, along with a selection cursor. The arrangement of emotion labels within this circle was counterbalanced but remained constant for each participant. Participants were instructed to select the corresponding emotion label as accurately and quickly as possible using the mouse to move the cursor to the certain label. Correct responses were considered as hits (see Lausen et al., 2020, Rafiee et al., 2023).

2.4 Statistical Analysis

To test our hypothesis, the data was analyzed using Generalized Linear Mixed Models (GLMMs; Baayen, 2008) with binomial error structure and logit link function implemented in R

Software version 4.0.3 and R studio version 1.4.1106. In each model, emotion recognition accuracy (correct vs. false) was the outcome variable, session number served as the control variable, and subject ID as the random intercept. In models investigating the interaction of ovulatory cycle and personality traits, the reference category for the ovulatory cycle was determined as the mid-luteal phase. Standard p-value 0.05 was determined as the cut-off criterion for two-tailed distributions.

Before including variables in the model, the distribution of hormone measure was inspected by visual check and Shapiro-Wilk test. As expected, the distribution of both hormones was departed significantly from normality (estradiol: $W = 0.90$, $p < 0.01$, progesterone: $W = 0.83$, $p < 0.01$). The distribution of personality traits (openness, extraversion, and neuroticism) was inspected as normal.

To track the within-subject fluctuation of hormonal measures across the late follicular and mid-luteal phases (following Jones et al., 2018), hormonal measures were subject mean-centered and the scaled by being divided by a constant (see Rafiee et al., 2023) resulting in variation of values from −.5 to .5, which facilitates the calculation in the linear mixed model. Subject mean centering was used to isolate the effect of within- and between-subject variation of hormones (van de Pol & Wright, 2009). To create the between-subject hormone variable we averaged the measures of each hormone across two phases, then we log-transformed between-subject levels of estradiol and progesterone to achieve normal distribution. To facilitate the interpretation of model outcome, log-transformed between-subject levels of ovarian hormones and personality scores were z-transformed (see Schielzeth, 2010).

For each model, we assessed Variance Inflation Factors (VIF; Field et al., 2012) with a model lacking the interaction term. No collinearity issue was found for each model (maximum VIF: 1.03). To evaluate the goodness of fit of the fitted model, the log-likelihood function of the fitted model

was compared with the log-likelihood function of the minimal (reduced) model lacking the predictor or the interaction of interest (see Dobson, 2002). Model stability was good for all models.

The following packages were used: ggplot2 3.3.3 (Wickham, 2016), and sjplot 2.8.7 (Luedecke, 2021) for data visualization, lme4 1.1.26 (Bates et al., 2015) for computing models, and car 3.0.10 (Fox & Weisberg, 2019) for assessing collinearity within predictors.

3 Results

3.1 Descriptive Analysis

3.1.1 Emotion Recognition

The average proportion of hits in facial emotion recognition was $M_{accuracy}$ = 0.912 (SD = 0.049) in the late follicular and $M_{accuracy}$ = 0.914 (SD = 0.053) in the mid-luteal phase. The average reactions times for facial emotion recognition was $M_{reaction\ times}$ = 1.29 s (SD = 0.28) in the late follicular phase and $M_{reaction\ time}$ = 1.28 s (SD = 0.316) in the mid-luteal phase.

3.1.2 Personality

The average score for personality traits was as following: $M_{agreeableness}$ = 3.96 (SD = 0.65), $M_{consciousness}$ = 3.45 (SD = 0.75), $M_{extraversion}$ = 3.45 (SD = 0.75), $M_{openness}$ = 3.75 (SD = 0.75), $M_{neuroticism}$ = 2.81 (SD = 0.67).

3.1.3 Hormone Measures

The average levels of estradiol excluding outliers (±3 SDs) in the late follicular phase was $M_{Estradiol}$ = 6.76 (SD = 4.47, range = 1.05 to 21.76) pg/mL and the average progesterone levels was $M_{Progesterone}$ = 39.4 (SD = 25.2, range = 14.60 to 218.16) pg/mL. In the mid-luteal phase, the average

of estradiol levels excluding outliers (±3 SDs) was $M_{Estradiol}$ = 6.12 (*SD = 3.82*, range = 0.70-19.62) pg/mL and the average of progesterone levels was $M_{Progesterone}$ = 83.8 (SD = 45.8, range = 20.18-266.16) pg/mL.

3.2 Cycle Phase × Openness Model (H1a)

In the first hypothesis, we assumed an interacting effect of the menstrual cycle (late-follicular vs. mid-luteal) and openness on emotion recognition. A GLMM was fitted, with cycle phase (confirmed by LH test, n = 74) and openness scores (scaled) as fixed effects, and their interaction. The model revealed a significant interaction of cycle phase and openness (β = 0.105, SE = 0.05, 95% CI = [1.02 – 1.21], z = 2.457, OR= 1.11, p = .014) (Fig. 1). Comparison between the log-likelihood of the main model with the log-likelihood of the model lacking the interaction effect (reduced model) revealed that the main model predicts the outcome variable significantly better (χ^2 = 6, df = 1, p = .014; see Table 1). In line with our hypothesis, individuals with higher scores in openness outperformed in the emotion recognition task in the late-follicular phase compared to the mid-luteal phase. Women performed better in the second session compared to the first session. We ran an additional model, including menstrual cycle phase – regardless of LH results – (N = 131) to control for the robustness of our findings. The additional model showed no significant interaction of phase and openness on emotion recognition (β = 0.049, SE = 0.04, 95% CI = [0.98 – 1.13], z = 1.338, OR= 1.05, p = .181) (See supplementary – Table S1).

Figure 1:

Individuals with higher scores in openness performed better in emotion recognition in the late-follicular phase as compared to mid-luteal phases (n = 74). Each point represents the relationship between the personality trait of openness and emotion recognition accuracy (hit rate). The regression line shows the interacting effect of openness scores and the menstrual cycle phases in emotion recognition accuracy.

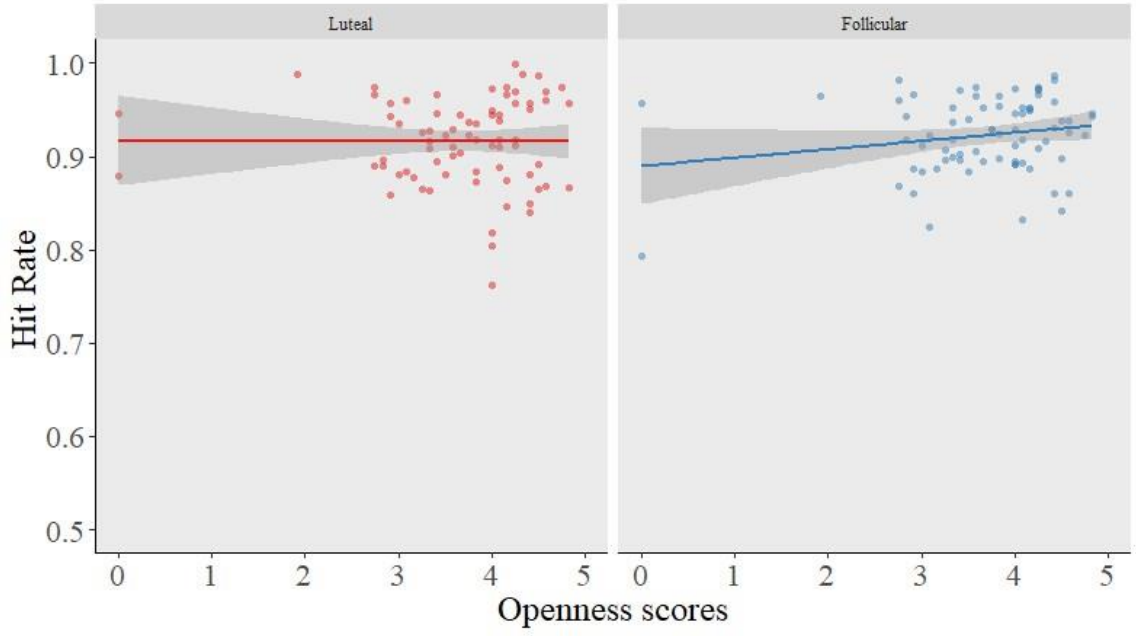


Table 1:

Results of Generalized Linear Mixed Model (GLMM) testing the interaction of menstrual cycle phase (late-follicular vs. mid-luteal) and personality traits on facial emotion recognition (n= 74).

	Estimates	*SE*	*z*	*p*	*OR*	*95% CI*
Openness						
Phase [mid-luteal]	0.058	0.052	1.131	0.258	1.06	0.96 – 1.17
Openness	-0.013	0.059	-0.218	0.827	0.99	0.88 – 1.11
Phase × Openness	0.105	0.043	2.457	**0.014**	1.11	1.02 – 1.21
Session	0.261	0.052	5.052	**<0.001**	1.30	1.17 – 1.44
Extraversion						
Phase [mid-luteal]	0.048	0.052	0.932	0.351	1.05	0.95 – 1.16
Extraversion	0.008	0.059	0.142	0.887	1.01	0.90 – 1.13
Phase × Extraversion	0.031	0.04	0.737	0.461	1.03	0.95 – 1.12

Session	0.252	0.051	4.890	**<0.001**	1.29	1.16 – 1.42
Neuroticism						
Phase [mid-luteal]	0.051	0.051	0.986	0.324	1.05	0.95 – 1.16
Neuroticism	-0.135	0.062	-2.176	**0.030**	0.87	0.77 – 0.99
Phase × Neuroticism	0.123	0.045	2.745	**0.006**	1.13	1.04 – 1.24
Session	0.257	0.052	4.994	**<0.001**	1.29	1.17 – 1.43

3.2.1 Estradiol × Openness Model (H1b)

We fitted a GLMM, including levels of estradiol (within- and between subject), scores of openness, and their interaction. The model indicated no significant interacting effect between estradiol levels (neither within- nor between-subject) and openness scores (see Table 2). Women performed better in the second than in the first session, independent of the hormone levels. Comparing the log-likelihood of the reduced model (lacking the interaction term) with the log-likelihood of main model indicated no significant differences in predicting the outcome variable (χ^2 = 0.32, df = 2, p = .851).

3.3 Cycle Phase × Extraversion (H2a)

The GLMM, including cycle phase, extraversion scores, and their interaction indicated no significant interaction of cycle phase and extraversion (β = 0.031, SE = 0.04, 95% CI = [0.95 – 1.12], z = 0.737, OR= 1.03, p = .461; Table 1). Comparison between the log-likelihood of the main model with the log-likelihood of the reduced model (lacking the interaction term) revealed no significant difference in predicting the outcome variable (χ^2 = 0.538, df = 1, p = .463).

To test for robustness of the results, we fitted an extra model including cycle phase, regardless of LH confirmation (N = 131), extraversion, their interaction, session number and Subject ID. The

model revealed consistent non-significant results (β = 0.030, SE = 0.04, 95% CI = [0.95 – 1.11], z = 0.847, OR= 1.03, *p* = .397; see supplementary – Table S1).

Table 2

Results of the Generalized Linear Mixed Model (GLMM), testing interaction of ovarian hormones and personality traits on emotion recognition (N=131).

	Estimates	*SE*	*z*	*P*	*OR*	*95% CI*
Estradiol × Openness						
Estradiol (within-subject)	-0.036	0.140	-0.257	0.797	0.96	0.73 – 1.27
Estradiol (between-subject)	0.057	0.050	1.142	0.253	1.06	0.96 – 1.17
Openness	0.014	0.056	0.253	0.800	1.01	0.91 – 1.13
Estradiol (within-subject) × Openness	0.003	0.153	0.017	0.986	1.00	0.74 – 1.35
Estradiol (between-subject) × Openness	-0.036	0.063	-0.568	0.570	0.96	0.85 – 1.09
Session	0.253	0.037	6.781	**<0.001**	1.29	1.20 – 1.39
Estradiol × Extraversion						
Estradiol (within-subject)	-0.030	0.141	-0.212	0.832	0.97	0.74 – 1.28
Estradiol (between-subject)	0.056	0.050	1.111	0.267	1.06	0.96 – 1.17
Extraversion	-0.020	0.053	-0.375	0.707	0.98	0.88 – 1.09
Estradiol (within-subject) × Extraversion	-0.096	0.153	-0.627	0.531	0.91	0.67 – 1.23
Estradiol (between-subject) × Extraversion	0.024	0.057	0.422	0.673	1.02	0.92 – 1.15
Session	0.253	0.037	6.767	**<0.001**	1.29	1.20 – 1.39
Progesterone × Neuroticism						
Progesterone (within-subject)	-0.125	0.148	-0.844	0.399	0.88	0.66 – 1.18
Progesterone (between-subject)	-0.023	0.048	-0.481	0.630	0.98	0.89 – 1.07
Neuroticism	-0.105	0.049	-2.140	**0.032**	0.90	0.82 – 0.99
Progesterone (within-subject) × Neuroticism	-0.299	0.150	-1.991	**0.046**	0.74	0.55 – 1.00
Progesterone (between-subject) × Neuroticism	-0.128	0.043	-2.966	**0.003**	0.88	0.81 – 0.96
Session	0.253	0.037	6.760	**<0.001**	1.29	1.20 – 1.39

3.3.1 Estradiol × Extraversion Model (H2b)

The GLMM including estradiol (within- and between-subject), extraversion, and their interaction revealed no significant interaction between estradiol levels (neither within- nor between-subject) and extraversion scores (see Table 2). Moreover, comparing the log-likelihood of the main model with the log-likelihood of the reduced model (lacking the interaction term) revealed no significant differences in predicting the outcome variable (χ^2 = 0.571, df = 2, *p* = 0.752).

3.4 Cycle Phase × Neuroticism Model (H3a)

A GLMM model was fitted to assess the interaction between menstrual cycle phases (late-follicular vs mid-luteal, confirmed via LH) and neuroticism on emotion recognition. The results showed a significant interaction effect of cycle phase and neuroticism score (β = 0.123, SE = 0.05, 95% CI = [1.04 – 1.24], z = 2.745, OR= 1.13, p = .006). Women performed better in the second session than in the first session. The log-likelihood ratio test revealed significant differences between the main and reduced models indicating that the main model predicts the outcome variable better (χ^2 = 7.502, df = 1, p = .006). Emotion recognition in the mid-luteal phase was impaired as compared to the late-follicular phase (see Table 1, Fig 2 – top panel) for women scoring high in neuroticism.

We fitted an additional model, including the menstrual cycle phase regardless of LH results (N = 131) as robustness check. The model confirmed the significant interacting effect of phase and neuroticism on emotion recognition accuracy (β = 0.099, SE = 0.04, 95% CI = [1.03 – 1.19], z = 2.733, OR= 1.10, p = .006; see supplementary, Table S1)

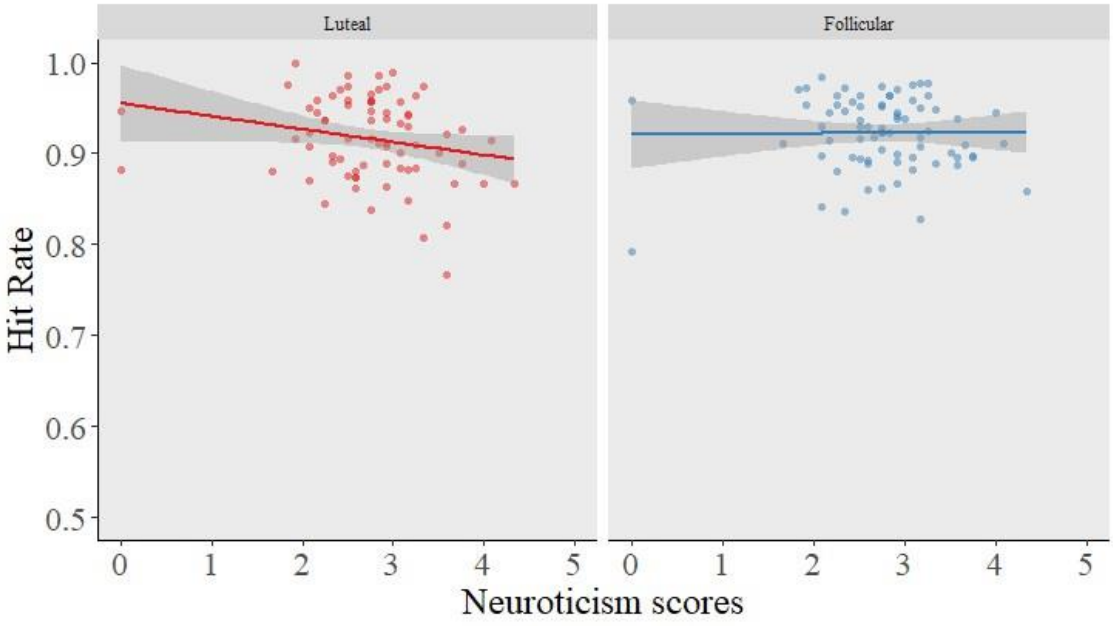
3.4.1 Progesterone × Neuroticism Model (H3b)

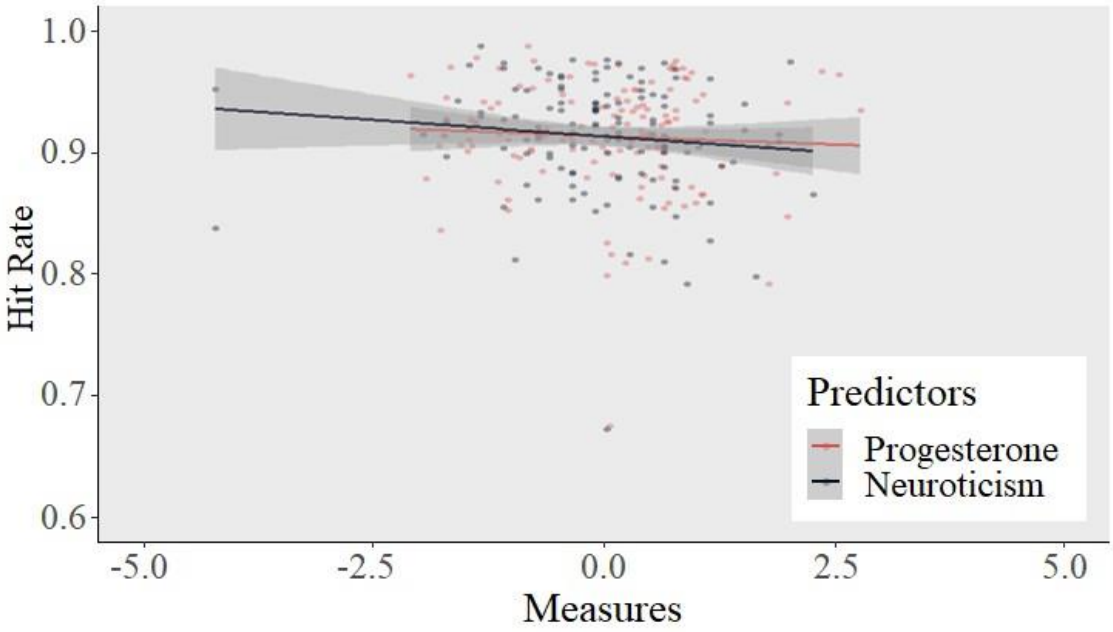
Next, we tested the potential interaction of progesterone levels and neuroticism by fitting a GLMM, including progesterone levels (within- and between-subject), neuroticism, and their interaction. The model revealed a significant interaction of progesterone (within- and between-subject) × neuroticism (see Table 2). Aligned with our prediction, increasing levels of progesterone in individuals with higher neuroticism scores is associated with significantly impaired emotion recognition (Figure 2). The likelihood ratio test showed that the main model predicted the outcome significantly better than the reduced model (χ^2 = 12.44, df = 2, p = .002).

Figure 2:

Women with higher neuroticism score had lower emotion recognition performance in the mid-luteal phase compared to the late follicular phase (n = 74). Each point represents the relationship between the personality trait of neuroticism and emotion recognition accuracy (hits). The regression line shows the interacting effect of neuroticism scores and the menstrual cycle phases in emotion recognition accuracy (**upper panel**). Elevated progesterone levels in individuals with higher neuroticism score decreases emotion recognition performance. Orange line shows the changes in emotion recognition accuracy (hits) as the function of neuroticism scores. Blue line shows the changes in emotion recognition accuracy (hits) as the function of progesterone levels (between-subject). Note: Both progesterone and neuroticism measures were scaled for plotting purposes. Between-subject progesterone measures were averaged, then log-transformed and then scaled (**lower panel**).

[The figure can be found on the following page.]


Luteal
Follicular
Hit Rate
1.0
0.9
0.8
0.7
0.6
0.5
0
1
2
3
4
5
Neuroticism scores


Hit Rate
1.0
0.9
0.8
0.7
0.6
-5.0
-2.5
0.0
2.5
5.0
Measures
Predictors
Progesterone
Neuroticism

3.5 Exploratory Analysis

We conducted an exploratory analysis to examine the three-way interaction between ovulatory cycle, personality traits, and emotion category in predicting facial emotion recognition. We also repeated the analysis using estradiol and progesterone instead of the ovulatory cycle phases. The script and results can be found at the following link: https://osf.io/fts2n/

4 Discussion

The current study provides initial evidence on a potential moderating role of emotion-related personality traits in the assumed relationship between the ovulatory cycle and facial emotion recognition. Our findings indicate that 1) women with higher openness scores performed better in emotion recognition during the late follicular phase than during the mid-luteal phase (n = 74, confirmed by LH tests). 2) Higher neuroticism scores were associated with lower emotion recognition in the mid-luteal phase compared to late follicular phase. Neuroticism was also associated with impaired emotion recognition when progesterone levels were high (within- and between-subjects). 3) Contrary to our predictions, no significant shifts in emotion recognition revealed when investigating the interactions between openness and ovarian hormone levels, between extraversion and cycle phase, and between extraversion x ovarian hormone levels.

Previous research on the assumed association between ovulatory cycle, and thus the level of ovarian hormones, and emotion recognition reported mixed evidence. Some studies suggested that recognizing emotions is improved in the follicular phase associated with higher levels of estradiol (e.g., Derntl, Kryspin-Exner, et al., 2008; Pearson and Lewis, 2005; Rubin, 2012; for an overview see Osório et al., 2018). On the contrary, other studies did not find this association (e.g., Shirazi et al., 2020; Di Tella et al., 2020; Kamboj et al., 2015; Pahnke et al., 2019; Zhang et al., 2013). It has

been proposed that these heterogeneities might result from different methodologies (e.g., hormonal assessment, estimation of cycle phase, various experimental paradigms) and the lack of methodological rigor (e.g., low statistical power, cross-sectional comparisons) (see Gamsakhurdashvili et al., 2021).

In a recent study based on the dataset analyzed here (N = 131), Rafiee et al. (2023) investigated possible alterations in emotion recognition across ovulatory cycle phases (late follicular vs. mid-luteal) and reported null findings. Remarkably, the current study indicates significant changes in emotion recognition across the ovulatory cycle, after including moderating effects of openness and neuroticism. This result might explain the equivocal findings in the existing literature and ultimately highlights the relevance of moderating variables on the relationship between the psycho-endocrine system and emotion recognition. Such approach is further supported by the complex and interactive effects of reproductive steroid hormones on the body and brain that impact molecular, organ, tissue, and behavioral levels (Kiesner, 2017). These interplays – studied here in conjunction with personality traits – seem to manifest also in individual differences, and their comprehensive understanding requires a multi-level approach, going beyond their potential main effects (Kiesner, 2017).

Aware of their novelty and the need for replication, we would like to provide some cautious interpretations of the present study's findings: Improved emotion recognition in the late follicular (fertile) phase for individuals with higher openness scores might be due to increased activation of reward-related brain regions (e.g., striatum) and dopamine system in the follicular phase (see Sacher et al., 2013). According to previous evidence, women might be more responsive to reward stimuli in the follicular phase than luteal phase (Dreher et al., 2007; Sakaki and Mather, 2012). Dopamine might not only have impacts restricted to reward-related behavior; as the proposed

biological substrate of openness, it also modulates exploratory behavior (DeYoung & Gray, 2009; DeYoung, 2013; Käckenmester et al., 2019). Considering the crucial role of emotion recognition in exploring the social environment, it seems plausible that openness improves this socio-cognitive capacity particularly in the fertile phase.

From an evolutionary perspective, it has been proposed that openness facilitates mating (see e.g., Nettle, 2006). Personality traits are thought to express differences in individuals' approaches to tackling (social) adaptive challenges such as acquiring and maintaining dominance and resources (see e.g., Buss, 2009). Individuals with a higher openness score tend to seek new experiences, which likely leads to broader social connections and as a result higher chances for reproductive success. As the probability of conception during the late follicular phase is highest, menstrual cycle phase and openness might jointly facilitate mating through increased emotion recognition. The connection between reproductive state and steroid hormones suggests that changes in steroid hormones may impact emotional processing crucial for social interaction, such as the perception of emotional stimuli, recognition of facial expressions, and regulation of one's own behavior (Gingnell et al., 2019).

Similar to openness, extraversion has been also linked to social behavior, but it was not found to significantly influence emotion recognition with variations across the menstrual cycle in our sample. This null result requires further validation through replication and studies that distinguish the unique contributions of openness and extraversion to emotion recognition. For instance, research has shown that extraversion is more strongly associated with improved ability in emotional encoding rather than decoding (Riggio and Riggio, 2002).

Despite the potential moderating effect of neuroticism on emotion recognition alteration during the mid-luteal phase and when progesterone levels are elevated, this topic has been largely

understudied. A limited number of studies have attempted to investigate the joint role of neuroticism and the mid-luteal phase in emotional processing across the menstrual cycle, including (subjective) emotional experiences, emotion regulation, and eating disorders, yet these studies have produced mixed results (e.g., Wu et al., 2014a, b; Racine et al., 2013, Zhang et al., 2015).

Further evidence for the potential joint role of neuroticism and mid-luteal phase in emotion recognition may be provided by studies focusing on premenstrual syndrome (PMS) symptoms during the luteal phase. Studies suggest that higher levels of neuroticism can intensify PMS symptoms during the luteal phase (e.g., Treloar et al., 2002; del Mar Fernández et al., 2019). PMS affects approximately 80% of women, and it has been suggested that its high frequency may indicate some evolutionary advantage, possibly related to preventing sexual or relationship breakdowns, but more research is needed to understand its significance (Gillings, 2014).

Neuroticism, which is linked to heightened vigilance to danger (Nettle, 2006), may also serve an adaptive purpose during the luteal phase. PMS has been associated with different individual responses to the fluctuation of ovarian hormones, highlighting the relevance of individual differences in reproductive steroid hormones (Kiesner, 2017). The mid-luteal phase of the menstrual cycle, with elevated progesterone levels, is believed to increase vulnerability to negative affective symptoms, according to the "window of vulnerability" model (Andreano et al., 2018). However, a recent study by Guevarra et al. (2023) found that perceived stress has a greater impact on affective symptoms than the menstrual cycle phase and did not find compelling evidence for an increase in affective symptoms during the mid-luteal phase. The primary driver of perceived stress is neuroticism, according to the stress perception hypothesis (Conard and Matthews, 2008). It's worth noting that there is still ongoing debate, with mixed findings, about the relationship between emotional experience and emotion recognition, as highlighted in recent studies such as Holland et

al. (2021) and Wearne et al. (2018). Altogether, it is possible that the combined effect of neuroticism and the luteal phase may further exacerbate the detrimental impact on emotion recognition in women.

4.1 Limitations

There are some limitations to the current study that needs to be acknowledged. One of the limitations of this study is the method used to measure salivary estradiol and progesterone, which was an immunoassay approach. Arslan et al. (2022) recommend using the Liquid chromatography–mass spectrometry (LC-MS/MS) method, considered the gold standard for ovarian hormones measurement, for more accurate results in hormonal studies. Our findings thus should be interpreted with this limitation in mind (see Rafiee et al., 2023). This constraint could be a possible explanation for the absence of association between openness and estradiol's interaction on emotion recognition.

Additionally, the study only included women in the late follicular and mid-luteal phases of their menstrual cycle, thus to have a more valid comparison a wider examination of emotion recognition including pre and post menstrual points where hormone levels are at their lowest is required.

Another limitation could be the context-dependence of the relationship between hormones and behavior. As suggested by Sundin et al (2020), hormone levels (e.g., cortisol and testosterone) measured in neutral settings may not accurately reflect their role in behavior and cognition. Thus, future studies should consider context and situational factors when exploring the link between hormones and behavior (Sundin et al., 2020).

We limited our sample to healthy young women of reproductive age, primarily consisting of students. Future research should examine the moderating effect of personality traits on the

relationship between emotion recognition and menstrual cycle in a more diverse and representative population. It would also be interesting for future research to explore the same research questions in women who use contraceptives. Additionally, it would be valuable to compare these findings in a clinical setting, particularly investigating the relationship between neuroticism and emotion recognition in women with Premenstrual Dysphoric Disorder (PMDD) during the luteal phase.

5 Conclusion

This study provides valuable insight into the joint role of dispositional traits and biological markers (menstrual cycle and hormones) on predicting the individual differences in emotion recognition. Understanding individual differences in emotion recognition can help shedding light on the complex interplay of various factors in this cognitive ability. Recognizing and appreciating variations in cognitive abilities and the factors that contribute to them helps to develop new approaches to understanding cognitive development and performance (Boogert et al., 2018).

Funding

This project was funded by the Leibniz ScienceCampus “Primate Cognition” under project number LSC-SF2018-10, with additional support provided by the German Academic Exchange Service (DAAD) and the Research Training Group (RTG) 2070 “Understanding Social Relationships.”

Data Availability

The data and supplementary analysis for this study can be accessed on the Open Science Framework through the following link: https://osf.io/fts2n/

CHAPTER 5

General Discussion

The overarching aim of this dissertation was to investigate the role of potential person-related sources of individual variation in emotion recognition. Emotion recognition is one of the key components of social cognition. Therefore, understanding how person characteristics are associated with variation in this ability helps to advance current knowledge of the underlying mechanisms involved in the processing of emotional expression.

Previous research has suggested that emotion recognition varies as a function of the menstrual cycle phases, ovarian hormones, biological sex, and personality traits. However, the field has produced ambivalent results due to methodological shortcomings and a lack of ecological validity. Therefore, by adding methodological rigor and using an ecologically valid task including visual, auditory, and audiovisual emotional expressions, three large-scale studies examined inter- and intra-individual variation in emotion recognition.

In Study 1, a preregistered study investigated the association between the ovulatory cycle (and ovarian hormone levels, i.e., estradiol and progesterone) with inter- and intra-individual variability in facial emotion recognition in healthy, naturally cycling women. Study 2 aimed to replicate previously observed sex differences in emotion recognition and to examine the moderating role of

stimulus features, including emotion specificity, stimulus modality, and encoder's sex, on the magnitude of sex differences in emotion recognition.

Given that the majority of previous studies have mainly focused on the main effect of the person-related source, the potential joint role of person-related sources, particularly biological markers, and personality, has been understudied. Study 3, in an exploratory approach, aimed to fill this gap by investigating whether women's personality traits are related to their emotion recognition across the ovulatory cycle and with ovarian hormone levels.

In the following sections, I will discuss key findings and potential implications within each study's broad and particular aims. Furthermore, I will highlight the limitations and outlook for future studies.

1 It Is Not the Cycle

In recent years, investigating the possible alteration in emotion recognition across the menstrual cycle has gained increasing interest in research (for an overview, see Gamsakhurdashvili et al., 2021; Gingnell et al., 2019; Osório et al., 2018b). There are potential reasons to study the assumed association. First, the menstrual cycle provides an excellent hormonal model to track high vs. low levels of ovarian hormones (Poromaa & Gingnell, 2014). Second, from an adaptationist perspective, emotion recognition might vary in different phases of the menstrual cycle in women to maximize the reproductive success. It is proposed that a better ability to recognize emotions in the fertile phase might increase the likelihood of social interaction and consequently the chance of mating (Derntl, Kryspin-Exner, et al., 2008; Derntl, Windischberger, et al., 2008; Gingnell et al., 2019; Kamboj et al., 2015). Finally, there is evidence for the high expression of estradiol and progesterone receptors in brain areas involved in emotion processing, such as the hypothalamus

and limbic system (see Poromaa & Gingnell, 2014), suggesting a possible link between ovarian hormones and emotion recognition.

Since this line of research is relatively new, the number of studies examining the hypothesized association is still limited, and there are inconsistencies in the findings, likely due to methodological shortcomings (for an overview, see Gamsakhurdashvili et al., 2021; Osório et al., 2018). Following the existing literature, in Study 1, we predicted inter and intra-individual variation in emotion recognition as the function of the ovulatory cycle and ovarian hormone levels. Adding methodological rigor such as high statistical power, within-subject design, relatively more accurate estimation of the ovulatory cycle phases (e.g., validation of the fertile phase via LH test) and ovarian hormones (direct measurement of estradiol and progesterone via saliva), and conducting an ecologically valid experiment that included visual, auditory, and audiovisual modalities, we found no compelling evidence for a change in emotion recognition as a function of either the ovulatory cycle or ovarian hormone levels.

There are several possible explanations for the lack of significant results. Considering that our environment has changed significantly compared to that in which our ancestors lived, we are not subject to the same selection pressures as they were. As a result, there may be a "matching problem" between our current modern environment and the hypothesized association, which could explain why we did not observe the expected results (e.g., Smith, 2020). Furthermore, while we utilized the best available methods and experimental settings, our stimuli may not have been successful or relevant in eliciting the assumed association on the behavioral level. Overall, the main finding of Study 1 suggests that relying solely on biological markers, such as the ovulatory cycle and ovarian hormone levels, may not be sufficient to predict inter- and intra-individual variation in emotion recognition in women. Thus, it is imperative that further research be conducted not only to replicate

the findings, but also to examine other person-related sources and their potential interactions that may contribute to individual variation in emotion recognition. To this end, we conducted Study 3 as a follow-up using secondary data analysis.

2 Exploring the Missing Link: Personality

Expanding on the results of Study 1, we explored whether women's personality traits and their position in the ovulatory cycle (late follicular vs. mid-luteal) could jointly predict individual variation in emotion recognition. Interestingly, and in line with our predictions, the results showed that when women's personality traits were taken into account, emotion recognition varied significantly across the ovulatory cycle and in relation to ovarian hormone levels. In the previous section, I highlighted that methodological limitations could be a plausible explanation for the observed inconsistencies in findings on emotion recognition alteration across the ovulatory cycle and in relation to ovarian hormones. However, in light of the findings of Study 3, it is worth considering that not incorporating personality traits as a predictor variable in prior research may also have played a role in contributing to the inconsistent results.

One of the key findings worth highlighting is the lower rates of emotion recognition accuracy among individuals with higher neuroticism scores in the mid-luteal phase compared to the late follicular phase. The interaction of neuroticism and cycle phase in predicting emotion recognition accuracy was consistent in the robustness analysis (including the subsample with confirmed fertile phase by LH tests and the full sample), as well as in the model, including the interaction between neuroticism and progesterone levels. Interestingly, empirical evidence suggests that women with a diagnosis of premenstrual dysphoric disorder (PMDD) tend to have higher levels of neuroticism-related personality traits, with the effect being more pronounced in those with severe luteal phase

symptoms than in a control group (Gingnell et al., 2010). Furthermore, women with higher levels of neuroticism may experience more severe premenstrual syndrome (PMS) symptoms during the luteal phase of their menstrual cycle (del Mar Fernández et al., 2019). The existing evidence linking neuroticism with more severe symptoms of PMS and PMDD suggests that there may be potential alterations in women's cognition, modulated by the interaction between cycle phase (luteal phase and progesterone levels) and neuroticism. The results of Study 3, therefore, provide new avenues for further investigation in healthy women, shedding light on how neuroticism and the ovulatory cycle (and ovarian hormones) may be linked to possible changes in cognitive functions. In particular, considering the role that neuroticism plays in psychological distress in healthy individuals (see Ploubidis & Frangou, 2011), these findings may have important implications for intervention strategies aimed at promoting mental health.

This study must be seen as a first attempt to explore the joint role that personality (as a more stable trait) and the ovulatory cycle and ovarian hormones (as a more variable physiological state) play in predicting individual variation in emotion recognition. Replication studies are crucial and encouraged to further elucidate this joint role in women's emotion recognition. Ovarian hormones play a crucial role in regulating a number of physiological and psychological processes throughout a woman's life (Farage et al., 2008). Thus, investigating their association with cognitive domains such as emotion recognition is not only of theoretical interest, but also has practical implications for promoting women's well-being (Del Río et al., 2018; Farage et al., 2008). In addition, conducting these types of studies can contribute to increasing menstrual literacy, dispelling the "menstrual cycle myth", and ultimately improving the quality of life for women of all ages (Bobel, 2019).

3 Rethinking Sex Differences

Sex differences are one of the main known individual differences in emotion recognition (Proverbio, 2021). Consistent with the majority of previous studies, Study 2, with high statistical power (*.99*), found that women performed better than men in overall emotion recognition (Hall, 1978; McClure, 1999; Thompson & Voyer, 2014). The overall significant differences in emotion recognition performance between men and women were rather small (d = .17), as reported in previous findings (d = .19) (for an overview see Thompson and Voyer, 2014). In recent years, there has been an increasing recognition concerning the importance of reporting effect sizes – which are independent of significance testing – as a statistically appropriate and sophisticated measure for comparing men's and women's performance. Effect sizes are particularly useful in assessing the stability or variability of psychological sex differences across studies and provide a meaningful estimate of their average size (see Eagly & Wood, 2013). Interestingly, 78% of the observed psychological sex differences showed a size effect of .35 or smaller across 46 reviewed meta-analyses (Hyde, 2005). In this line, the "*gender similarity*" hypothesis proposes that there are generally small differences in psychological characteristics between women and men, suggesting that they are more alike than different in most cognitive abilities (Hyde, 2005, 2014). The adaptive perspective, however, posits that sex differences in cognitive abilities (and potentially emotion recognition) are due to different selective pressures regarding reproductive success in women and men, who have consequently developed different psychological mechanisms over the course of evolution (Buss & Schmitt, 2011). Empirical findings on sex differences in emotion recognition have been largely explained by the adaptive view, which emphasizes the evolution of distinct adaptive mechanisms related to parental investment between the two sexes (e.g., the primary caretaker hypothesis and the fitness-threat hypothesis)(Babchuk et al., 1985; Hames et al., 1985;

Hampson et al., 2006, 2021). However, the opponents argue that this perspective overlooks the influence of social and cultural environments on women and men (Eagly & Wood, 1999). Nevertheless, the evolutionary psychology perspective does not refute the gender similarity hypothesis, as it acknowledges that men and women are largely similar in many aspects while also providing potential explanations for domains of difference (for an overview, see Buss & Schmitt, 2011). Empirical evidence has shown that there are many similarities between women and men at the neural level, but there are also some differences (Cahill, 2006). Overall, a biopsychosocial perspective that recognizes the intertwined relationship between biological and environmental factors is needed to investigate factors that contribute to cognitive gender differences (Miller & Halpern, 2014). Such a perspective can help us to better understand and maximize the cognitive potential of individuals, which in turn has implications for policy-making and addressing societal issues arising from gender inequities and stereotypes (Choleris et al., 2018; Miller & Halpern, 2014). To understand the potential effect of biology and environment and culture on sex differences in emotion recognition, further research is needed to follow these findings across different developmental stages of individuals and across diverse societies (e.g., patriarchy vs. matriarchy) (see Henrich et al., 2010; McClure, 1999).

It is worth noting that sex differences in emotion recognition have been consistently replicated in numerous studies, including the current dissertation project. The methodological strengths of this study support the reliability of the effect, albeit small in magnitude. The small effect may not have significant implications in everyday life; however, it can have a greater impact in clinical areas and diagnostics concerning sex-specific psychiatric disorders. For instance, accounting for sex differences in emotion recognition comes to importance while designing diagnostics tools for disorders with different prevalence among women and men. Evidence has shown that women with autism are often diagnosed later than autistic men due to their higher functioning in social

interaction (Wood-Downie et al., 2021), highlighting the need to account for sex differences in the implementation of diagnostic instruments. Finally, basic and clinical research on sex differences can have a significant impact on policy-making (Choleris et al., 2018; Miller & Halpern, 2014). For example, in countries such as the United States, Canada, and Europe, there are policies that require researchers to include sex as a biological variable in biomedical research. This ensures that both sexes are taken into consideration and provides a more comprehensive understanding of biomedical data (Choleris et al., 2018).

4 Limitations

4.1 Stimulus-related Limitations

Although methodological rigor was applied to the studies included in this dissertation, there are still limitations that need to be addressed. Since all three studies were conducted on the basis of data collected in two large-scale studies with the same experimental design, the limitation regarding the stimulus material applies to all of them. First, the ceiling effect in the hit rate was observed in all three studies, which is explained by the use of full-blown emotional expressions. As a consequence, the ceiling effect reduces the variance, which is a potential source of model misfit (see Schweizer et al., 2019). This issue can be addressed by implementing different intensities of emotion. Second, the use of static facial expressions in the audiovisual modality may limit the external and ecological validity of the stimuli. To increase the validity of future studies, it is recommended to implement dynamic facial expressions instead. Third, the ratio of negative to positive stimuli was unbalanced. This was due to the inclusion of basic emotional expressions in the experimental design. Future research would benefit from using balanced emotional categories in terms of valence. In addition, future studies should include a variety of emotional expressions to

improve ecological validity, as we are exposed to a wide range of emotional expressions, from basic to complex, in our daily lives. Fourth, the use of a multiple-choice response format in all three studies may reduce their ecological validity, as real-life situations rarely provide a list of predefined emotional categories (Georgopoulos et al., 2022).

4.2 Sample-related Limitations

Another limitation of the present dissertation is the lack of diversity in the sample studied. The majority of participants were from WEIRD populations (Western, Educated, Industrialized, Rich, Democratic countries), which limits the generalizability of the findings to other populations with different cultural backgrounds and socioeconomic status (see Henrich et al., 2010; Rad et al., 2018). Study 1, due to a specific hypothesis – which was predicting women's better recognition of emotions expressed by male encoders in the late follicular vs. mid-luteal phase – only examined heterosexual women. To further expand our understanding of possible emotion recognition alteration across the ovulatory cycle, future research should include women with diverse sexual orientations. Moreover, most of the existing studies on emotion recognition across the menstrual cycle have predominantly included young women, leaving a gap in our current knowledge of this phenomenon in other important reproductive stages such as pregnancy, postpartum, puberty, and menopause (Osório et al., 2018).

A potential limitation of Study 2 is the lack of explicit differentiation between biological sex and gender. To address this, it may be beneficial to include additional questions in the research, such as those related to biologically assigned sex, preferred pronouns, and gender identity spectrum (e.g., Reproductive Statues Questionnaire) (Hopkins & Richardson, 2020; Schmalenberger et al., 2021). In addition, there are other confounding factors such as personality (sex-role-related

personality differences), physical characteristics, and socialization that interact with sex that need to be considered in future studies (Hall, 1978).

5 Implication

On a broader scale, this dissertation has the potential to contribute significantly to the theoretical and practical implications of the study of individual variations, which is critical in understanding different cognitive domains. By studying individual variations, we can gain insight into the genetic and environmental influences that shape human behavior and cognition (Boogert et al., 2018). In particular, advancing our understanding of individual variations in emotion recognition can provide a better insight of the underlying mechanisms and neural basis of emotions (see Hamann & Canli, 2004; Kanai & Rees, 2011). As neuroimaging studies can be costly, these findings may serve as a cost-effective starting point for further investigations. Furthermore, variations in an individual's affect, cognition, and desire can provide insight into their behavior and serve as predictors for future research (Revelle et al., 2010). Finally, recognizing and valuing individual differences among people allows us to appreciate how these differences contribute to the diversity and richness of our society and culture.

6 Conclusion

The primary objective of this dissertation was to examine the association between person-related sources, namely biological markers (ovulatory cycle, ovarian hormones), sex, and the joint role of biological markers and dispositional traits with inter- and intra-individual variations in emotion recognition. In addition, this study aimed to address methodological limitations of previous research that have contributed to inconsistencies in the literature, using the best available methods to improve upon these limitations. The results of this dissertation indicate that while biological

markers such as the ovulatory cycle and ovarian hormones do not significantly predict within- and between-individual variation in emotion recognition, other person-related sources, including sex and the joint role of personality traits and ovulatory cycle, along with progesterone levels, play a significant role in predicting variation in emotion recognition. These findings highlight the importance of considering person-related sources and their potential joint role in contributing to variation in emotion recognition. Therefore, a more interactive approach that considers both biological and environmental factors is needed to better understand individual variation in emotion recognition.

References

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

Al-Shawaf, L., Conroy-Beam, D., Asao, K., & Buss, D. (2015). Human Emotions: An Evolutionary Psychological Perspective. *Emotion Review*. 10.1177/1754073914565518.

Amminger, G. P., Schäfer, M. R., Papageorgiou, K., Klier, C. M., Schlögelhofer, M., Mossaheb, N., Werneck-Rohrer, S., Nelson, B., & McGorry, P. D. (2012). Emotion recognition in individuals at clinical high-risk for schizophrenia. Schizophrenia bulletin, 38(5), 1030–1039. https://doi.org/10.1093/schbul/sbr015

Andreano, J. M., Touroutoglou, A., Dickerson, B., & Barrett, L. F. (2018). Hormonal Cycles, Brain Network Connectivity, and Windows of Vulnerability to Affective Disorder. *Trends in neurosciences*, 41(10), 660–676. https://doi.org/10.1016/j.tins.2018.08.007

Arslan, R.C., Schilling, K.M., Gerlach, T. M., & Penke, L. (2021). Using 26 thousand diary entries to show ovulatory changes in sexual desire and behavior. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000208

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology*, 74(6), 1531–1544. https://doi.org/10.1037/0022-3514.74.6.1531

Ashton, M., 2013. Individual differences and personality. Amsterdam: Academic Press.

Baayen, R. H. (2008). Analyzing Linguistic Data. Cambridge University Press, Cambridge.

Babchuk, W. A., Hames, R. B., & Thompson, R. A. (1985). Sex differences in the recognition of infant facial expressions of emotion: The primary caretaker hypothesis. *Ethology and Sociobiology*, *6*(2), 89–101. https://doi.org/10.1016/0162-3095(85)90002-0

Bänziger, T. (2016). Accuracy of judging emotions. In *The Social Psychology of Perceiving Others Accurately* (pp. 23–51). Cambridge University Press. https://doi.org/10.1017/cbo9781316181959.002

Bao, A. M., Liu, R. Y., van Someren, E. J. W., Hofman, M. A., Cao, Y. X., & Zhou, J. N. (2003). Diurnal rhythm of free estradiol during the menstrual cycle. *European Journal of Endocrinology*, *148*(2), 227–232. https://doi.org/10.1530/eje.0.1480227

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531–539. https://doi.org/10.3758/BRM.40.2.531

Betz, N., Hoemann, K., & Barrett, L. F. (2019). Words are a contextfor mental inference. *Emotion*, 19(8), 1463–1477.https://doi.org/10.1037/emo0000510

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

Blaszczyk, M. B. (2020). Primates got personality, too: Toward an integrative primatology of consistent individual differences in behavior. *Evolutionary Anthropology*, *29*(2), 56–67. https://doi.org/10.1002/evan.21808

Bobel, C. (2019). Beyond the Managed Body: Putting Menstrual Literacy at the Center. In C. Bobel, *The Managed Body* (pp. 281–321). Springer International Publishing. https://doi.org/10.1007/978-3-319-89414-0_8

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170280. https://doi.org/10.1098/rstb.2017.0280

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

Buck, R. W., Miller, R. E., & Caul, W. F. (1974). Sex, personality, and physiological variables in the communication of emotion via facial expression. Journal of Personality and Social Psychology, 30, 587-596

Bull, J. R., Rowland, S. P., Scherwitzl, E. B., Scherwitzl, R., Danielsson, K. G., & Harper, J. (2019). Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *Npj Digital Medicine*, *2*(1), 1–8. https://doi.org/10.1038/s41746-019-0152-7

Buss, D. M. (2009). How Can Evolutionary Psychology Successfully Explain Personality and Individual Differences? *Perspectives on Psychological Science*, *4*(4), 359–366. https://doi.org/10.1111/j.1745-6924.2009.01138.x

Buss, D. M., & Greiling, H. (1999). Adaptive individual differences. Journal of Personality, 67, 209-243. http://dx.doi.org/10.1111/1467-6494.00053

Buss, D. M., & Schmitt, D. P. (2011). Evolutionary Psychology and Feminism. *Sex Roles*, *64*(9), 768–787. https://doi.org/10.1007/s11199-011-9987-3

Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience*, *7*(6), 477–484. https://doi.org/10.1038/nrn1909

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. E. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience*, 115(1), 33–42. https://doi.org/10.1037/0735-7044.115.1.33

Carver, C.S., Scheier, M.F. (2004). Perspectives on Personality, 5th ed. Pearson Education, Boston.

Choleris, E., Galea, L. A. M., Sohrabji, F., & Frick, K. M. (2018). Sex differences in the brain: Implications for behavioral and biomedical research. Neuroscience and biobehavioral reviews, 85, 126–145. https://doi.org/10.1016/j.neubiorev.2017.07.005

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia*, *48*(1), 220–225. https://doi.org/10.1016/j.neuropsychologia.2009.09.007

Conard, M. A., & Matthews, R. A. (2008). Modeling the stress process: personality eclipses

dysfunctional cognitions and workload in predicting stress. Personality and Individual Differences, 44(1), 171–181.

Connolly, H. L., Lefevre, C. E., Young, A. W., & Lewis, G. J. (2019). Sex differences in emotion recognition: Evidence for a small overall female superiority on facial disgust. Emotion, 19(3), 455–464. https://doi.org/10.1037/emo0000446

Conway, C., Jones, B., DeBruine, L., Welling, L., Law Smith, M., Perrett, D., Sharp, M., and Al-Dujaili, E. (2007). Salience of emotional displays of danger and contagion in faces is enhanced when progesterone levels are raised. *Hormones and behavior*. 51. 202-6. 10.1016/j.yhbeh.2006.10.002.

Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, 38(4), 668–678. https://doi.org/10.1037/0022-3514.38.4.668

Costa, P. T., Jr., & McCrae, R. R. (1994). Stability and change in personality from adolescence through adulthood. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 139–150). Lawrence Erlbaum Associates, Inc.

Cunningham M. R. (1977). Personality and the structure of the nonverbal communication of emotion. *Journal of personality*, 45(4), 564–584. https://doi.org/10.1111/j.1467-6494.1977.tb00172.x

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

Davitz, J. R. (1964). A review of research concerned with facial and vocal expressions of emotion. In J. R.Davitz (Ed.), The communication of emotional meaning. New York: McGraw-Hill.

De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311. https://doi.org/10.1080/026999300378824

del Mar Fernández, M., Regueira-Méndez, C., & Takkouche, B. (2019). Psychological factors and premenstrual syndrome: A Spanish case-control study. *PLOS ONE*, *14*(3), e0212557. https://doi.org/10.1371/journal.pone.0212557

Del Río, J. P., Alliende, M. I., Molina, N., Serrano, F. G., Molina, S., & Vigil, P. (2018). Steroid Hormones and Their Action in Women's Brains: The Importance of Hormonal Balance. *Frontiers in Public Health*, *6*(May), 1–15. https://doi.org/10.3389/fpubh.2018.00141

Derntl, B., Hack, R. L., Kryspin-Exner, I., & Habel, U. (2013). Association of menstrual cycle phase with the core components of empathy. *Hormones and Behavior*, *63*(1), 97–104. https://doi.org/10.1016/j.yhbeh.2012.10.009

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., Kryspin-Exner, I., Gur, R. C., Moser, E., & Habel, U. (2009). Amygdala activity to fear and anger in healthy young males is associated with testosterone. *Psychoneuroendocrinology*, 34(5), 687–693. https://doi.org/10.1016/j.psyneuen.2008.11.007

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

DeYoung, C. G. (2013). The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. *Frontiers in Human Neuroscience, 7*, Article 762. https://doi.org/10.3389/fnhum.2013.00762

DeYoung, C. G., & Gray, J. R. (2009). Personality Neuroscience. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (p. 850). Cambridge University Press.

Di Simplicio, M., & Harmer, C. J. (2016). Oxytocin and emotion processing. *Journal of Psychopharmacology*, *30*(11), 1156–1159. https://doi.org/10.1177/0269881116641872

Di Tella, M., Miti, F., Ardito, R. B., & Adenzato, M. (2020). Social cognition and sex: Are men and women really different? *Personality and Individual Differences*, *162*(April). https://doi.org/10.1016/j.paid.2020.110045

Dobson, A. (2002) An Introduction to Generalized Linear Models. 2nd Edition, Chapman & Hall/CRC, USA.

Domes, G., Heinrichs, M., Gläscher, J., Büchel, C., Braus, D. F., & Herpertz, S. C. (2007). Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biological psychiatry*, *62*(10), 1187–1190. https://doi.org/10.1016/j.biopsych.2007.03.025

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

Dores, A. R., Barbosa, F., Queirós, C., Carvalho, I. P., & Griffiths, M. D. (2020). Recognizing

Emotions through Facial Expressions: A Largescale Experimental Study. *International journal of environmental research and public health*, 17(20), 7420. https://doi.org/10.3390/ijerph17207420

Dreher, J. C., Schmidt, P. J., Kohn, P., Furman, D., Rubinow, D., & Berman, K. F. (2007). Menstrual cycle phase modulates reward-related neural function in women. *Proceedings of the National Academy of Sciences of the United States of America*, 104(7), 2465–2470. https://doi.org/10.1073/pnas.0605569104

Duesenberg, M., Weber, J., Schulze, L., Schaeuffele, C., Roepke, S., Hellmann-Regen, J., Otte, C., & Wingenfeld, K. (2016a). Does cortisol modulate emotion recognition and empathy? *Psychoneuroendocrinology, 66*, 221–227. https://doi.org/10.1016/j.psyneuen.2016.01.011

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). Journal of Autism and Developmental Disorders, 38(3), 464–473. https://doi.org/10.1007/s10803-007- 0486-x

Eagly, A. H., & Wood, W. (1999). *The Origins of Sex Differences in Human Behavior Evolved Dispositions Versus Social Roles*.

Eagly, A. H., & Wood, W. (2013). The Nature-Nurture Debates: 25 Years of Challenges in Understanding the Psychology of Gender. *Perspectives on Psychological Science, 8*(3), 340–357. https://doi.org/10.1177/1745691613484767

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: a methodological review. *Clinical psychology review*, 22(6), 789–832. https://doi.org/10.1016/s0272-7358(02)00130-7

Eisenlohr-Moul, T. A., & Owens, S. A. (2016). Hormones and Personality. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1–23). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-28099-8_762-1

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Elfenbein, H. A. (2007). Emotion in organizations: A review and theoretical integration. The Academy of Management Annals, 1(1), 315–386. https://doi.org/10.1080/078559812

Ellenbogen, M. A., Linnen, A. M., Grumet, R., Cardoso, C., & Joober, R. (2012). The acute effects of intranasal oxytocin on automatic and effortful attentional shifting to emotional faces. *Psychophysiology*, *49*(1), 128–137. https://doi.org/10.1111/j.1469-8986.2011.01278.x

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

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Field, A., Miles, J., & Field, Z. (2012). Discovering statistics using R. SAGE Publications.

Fischer, A.H., Manstead, A.S.R. (2009). Social functions of emotions. In M. Lewis, J.M.

Haviland-Jones, & L. Feldman Barrett (Eds.): Handbook of Emotions. New York/ London: The Guilford Press

Fox J, Weisberg S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

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

Gamsakhurdashvili, D., Antov, M. I., & Stockhorst, U. (2021). Facial Emotion Recognition and Emotional Memory From the Ovarian-Hormone Perspective: A Systematic Review. *Frontiers in Psychology*, *12*(641250), 1–14. https://doi.org/10.3389/fpsyg.2021.641250

Gangestad, S. W., Haselton, M. G., Welling, L. L. M., Gildersleeve, K., Pillsworth, E. G., Burriss, R. P., Larson, C. M., & Puts, D. A. (2016). How valid are assessments of conception probability in ovulatory cycle research? Evaluations, recommendations, and theoretical implications. *Evolution and Human Behavior*, *37*(2), 85–96. https://doi.org/10.1016/j.evolhumbehav.2015.09.001

Georgopoulos, M. A., Brewer, N., Lucas, C. A., & Young, R. L. (2022). Speed and accuracy of emotion recognition in autistic adults: The role of stimulus type, response format, and emotion. *Autism Research*, *15*(9), 1686–1697. https://doi.org/10.1002/aur.2713

Gerdes, A. B. M., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology*, *5*(DEC). https://doi.org/10.3389/fpsyg.2014.01351

Gillings M. R. (2014). Were there evolutionary advantages to premenstrual syndrome?.

Evolutionary applications, 7(8), 897–904. https://doi.org/10.1111/eva.12190

Gingnell, M., Comasco, E., Oreland, L., Fredrikson, M., & Sundström-Poromaa, I. (2010). Neuroticism-related personality traits are related to symptom severity in patients with premenstrual dysphoric disorder and to the serotonin transporter gene-linked polymorphism 5-HTTPLPR. *Archives of Women's Mental Health*, *13*(5), 417–423. https://doi.org/10.1007/s00737-010-0164-4

Gingnell, M., Hornung, J., & Derntl, B. (2019). Emotional processing and sex hormones. *Routledge International Handbook of Social Neuroendocrinology*, *30443*, 403–419. https://doi.org/10.4324/9781315200439-24

Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, 94(3), 193–213. https://doi.org/10.1016/j.cognition.2004.01.005

Grimshaw, G. M., Bulman-Fleming, M. B., & Ngo, C. (2004). A signal-detection analysis of sex differences in the perception of emotional faces. Brain and Cognition,54(3), 248-250.

Guapo, V., Graeff, F., Zani, A., Labate, C., Dos Reis, R., and Del-Ben, C. (2009). Effects of sex hormonal levels and phases of the menstrual cycle in the processing of emotional faces. *Psychoneuroendocrinology*. 34. 1087-94. 10.1016/j.psyneuen.2009.02.007.

Guevarra, D. A., Louis, C. C., Gloe, L. M., Block, S. R., Kashy, D. A., Klump, K. L., & Moser, J. S. (2023). Examining a window of vulnerability for affective symptoms in the mid-luteal phase of the menstrual cycle. *Psychoneuroendocrinology*, *147*, 105958. https://doi.org/10.1016/j.psyneuen.2022.105958

Gurvich, C., & Thomas, N. (2021). Hormones and Cognition. Brain sciences, 11(3), 318. https://doi.org/10.3390/brainsci11030318

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*(4), 845–857. https://doi.org/10.1037/0033-2909.85.4.845

Hall, J. A. (1984). Nonverbal sex differences: Communication accuracy and expressive style. Baltimore: Johns Hopkins University Press

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*(2), 201–206. https://doi.org/10.1037/1528-3542.4.2.201

Hall, J. A., Carter, J. D., & Horgan, T. G. (2010). Gender differences in nonverbal communication of emotion. In A. H. Fischer (Ed.), Gender and emotion: Social psychological perspectives (pp. 97–117). Cambridge University Press. https://doi.org/10.1017/CBO9780511628191.006

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology*, *14*(2), 233–238. https://doi.org/10.1016/j.conb.2004.03.010

Hames, R. B., Babchuk, W. A., & Hames, L. (1985). *Sex Differences in the Recognition of Infant Facial Expressions of Emotion: The Primary Caretaker Hypothesis* (p. 159). http://digitalcommons.unl.edu/anthropologyfacpubhttp://digitalcommons.unl.edu/anthropologyfacpub/159

Hampson, E., Istasy, P., Owais, S., Chow, J. A., Howidi, B., & Ouellette, S. J. (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

Harris, D. A., Hayes-Skelton, S. A., & Ciaramitaro, V. M. (2016). What's in a Face? How Face Gender and Current Affect Influence Perceived Emotion. Frontiers in psychology, 7, 1468. https://doi.org/10.3389/fpsyg.2016.01468

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica*, *135*(3), 278–283. https://doi.org/10.1016/j.actpsy.2010.07.012

Holland, A. C., O'Connell, G., & Dziobek, I. (2021). Facial mimicry, empathy, and emotion recognition: a meta-analysis of correlations. *Cognition & emotion*, 35(1), 150–168. https://doi.org/10.1080/02699931.2020.1815655

Holleman GA, Hooge ITC, Kemner C and Hessels RS (2020) The 'Real-World Approach' and Its Problems: A Critique of the Term Ecological Validity. *Front. Psychol*. 11:721. doi: 10.3389/fpsyg.2020.00721

Hopkins, S., & Richardson, L. (2020). *Gender Identity: From Biological Essentialism Binaries to a Non-binary Gender Spectrum* (pp. 1–10). https://doi.org/10.1007/978-3-319-70060-1_87-1

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65 , 373–398.

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

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

Käckenmester, W., Bott, A., & Wacker, J. (2019). Openness to experience predicts dopamine effects on divergent thinking. Personality neuroscience, 2, e3. https://doi.org/10.1017/pen.2019.3

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, 12(4), 231-242. https://doi.org/10.1038/nrn3000

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

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional Expression: Advances in Basic Emotion Theory. *Journal of Nonverbal Behavior*, *43*(2), 133–160. https://doi.org/10.1007/s10919-019-00293-3

Ketterson, E., Atwell, J., & McGlothlin, J. (2013). VII.2. Evolution of Hormones and Behavior. 10.1515/9781400848065-086.

Kiesner J. (2017). The Menstrual Cycle-Response and Developmental Affective-Risk Model: A

multilevel and integrative model of influence. *Psychological review*, 124(2), 215–244. https://doi.org/10.1037/rev0000058

Kiesner, J., Eisenlohr-Moul, T., & Mendle, J. (2020). Evolution, the Menstrual Cycle, and Theoretical Overreach. *Perspectives on Psychological Science*, *15*(4), 1113–1130. https://doi.org/10.1177/1745691620906440

Kokko, H., & Jennions, M. D. (2008). Parental investment, sexual selection and sex ratios. *Journal of Evolutionary Biology*, *21*(4), 919–948. https://doi.org/10.1111/j.1420-9101.2008.01540.x

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

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition and Emotion*, *28*(3), 452–469. https://doi.org/10.1080/02699931.2013.837378

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

Laukka, P., Bänziger, T., Israelsson, A., Cortes, D. S., Tornberg, C., Scherer, K. R., & Fischer, H. (2021). Investigating individual differences in emotion recognition ability using the ERAM test. *Acta psychologica*, *220*, 103422. https://doi.org/10.1016/j.actpsy.2021.103422

Lausen, A., Hammerschmidt, K. (2020). Emotion recognition and confidence ratings predicted by vocal stimulus type and prosodic parameters. *Humanities and Social Sciences*

Communications. 7, 2. https://doi.org/10.1057/s41599-020-0499-z

Lausen, A., & Schacht, A. (2018). Gender differences in the recognition of vocal emotions. *Frontiers in Psychology*, 9(JUN), 1–22. https://doi.org/10.3389/fpsyg.2018.00882

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

Lench, H. C., & Carpenter, Z. K. (2018). What do emotions do for us? In *The Function of Emotions: When and Why Emotions Help Us* (pp. 1–7). Springer International Publishing. https://doi.org/10.1007/978-3-319-77619-4_1

Lench, H. C., Bench, S. W., Darbor, K. E., & Moore, M. (2015). A functionalist manifesto: Goal-related emotions from an evolutionary perspective. *Emotion Review*, *7*(1), 90–98. https://doi.org/10.1177/1754073914553001

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition and Emotion*, *13*(5), 481–504. https://doi.org/10.1080/026999399379159

Lewis, G.J., Lefevre, C.E., Young, A.W. (2016). Functional architecture of visual emotion recognition ability: A latent variable approach. *J Exp Psychol Gen.* 145(5):589-602. doi: 10.1037/xge0000160. Epub 2016 Mar 17. PMID: 26986040.

Liening, S. H., Stanton, S. J., Saini, E. K., & Schultheiss, O. C. (2010). Salivary testosterone, cortisol, and progesterone: Two-week stability, interhormone correlations, and effects of time of day, menstrual cycle, and oral contraceptive use on steroid hormone levels. *Physiology & Behavior*, *99*(1), 8–16. https://doi.org/10.1016/j.physbeh.2009.10.001

Liker, A., Freckleton, R. P., Remeš, V., & Székely, T. (2015). Sex differences in parental care:

Gametic investment, sexual selection, and social environment. *Evolution*, *69*(11), 2862–2875. https://doi.org/10.1111/evo.12786

Lischke, A., Berger, C., Prehn, K., Heinrichs, M., Herpertz, S. C., & Domes, G. (2012). Intranasal oxytocin enhances emotion recognition from dynamic facial expressions and leaves eye-gaze unaffected. *Psychoneuroendocrinology*, *37*(4), 475–481. https://doi.org/10.1016/j.psyneuen.2011.07.015

Luedecke, D. (2021). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.10, https://CRAN.R-project.org/package=sjPlot.

Lyusin, D., & Ovsyannikova, V. (2016). Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity. Learning and Individual Differences, 52 , 129–136.

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

Marcinkowska, U. M. (2020). Importance of Daily Sex Hormone Measurements Within the Menstrual Cycle for Fertility Estimates in Cyclical Shifts Studies. *Evolutionary Psychology*, *18*(1), 1–8. https://doi.org/10.1177/1474704919897913

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

Marečková, K., Perrin, J. S., Khan, I. N., Lawrence, C., Dickie, E., McQuiggan, D. A., & Paus, T. (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

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., Yrizarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L., & Goh, A. (2000). A New Test to Measure Emotion Recognition Ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior*, *24*(3), 179–209. https://doi.org/10.1023/A:1006668120583

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*(3), 204–217. https://doi.org/10.1037/0003-066X.61.3.204

McClure. (1999). A Meta-Analytic Review of Sex Differences in Facial Expression Processing and Their Development in Infants, Children, and Adolescents. *Journal of Pediatrics*, *135*(6), 657–661. https://doi.org/10.1016/S0022-3476(99)70079-X

McCullagh, P., Nelder, J. (1989). Generalized Linear Models, Second Edition. *Chapman & Hall.* ISBN: 9780412317606

Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences*, *18*(1), 37–45. https://doi.org/10.1016/j.tics.2013.10.011

Montagne, B., Kessels, R., Frigerio, E., de Haan, E., & Perrett, D. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive processing*. 6. 136-41. 10.1007/s10339-005-0050-6.

Mulligan, K., & Scherer, K. R. (2012). Toward a working definition of emotion. *Emotion Review*, *4*(4), 345–357. https://doi.org/10.1177/1754073912445818

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622–631. https://doi.org/10.1037/0003-066X.61.6.622

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

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: blocking facial mimicry can selectively impair recognition of emotional expressions. *Social neuroscience*, 2(3-4), 167–178. https://doi.org/10.1080/17470910701391943

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

Pahnke, R., Mau-Moeller, A., Junge, M., Wendt, J., Weymar, M., Hamm, A. O., & Lischke, A. (2019). Oral Contraceptives Impair Complex Emotion Recognition in Healthy Women. *Frontiers in Neuroscience*, *12*, 1041. https://doi.org/10.3389/fnins.2018.01041

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

Paulmann, S., & Uskul, A. K. (2014). Cross-cultural emotional prosody recognition: evidence from Chinese and British listeners. *Cognition & emotion*, 28(2), 230–244. https://doi.org/10.1080/02699931.2013.812033

Pearson, R. M., Lightman, S. L., & Evans, J. (2009). Emotional sensitivity for motherhood: Late pregnancy is associated with enhanced accuracy to encode emotional faces. *Hormones and Behavior*, *56*(5), 557–563. https://doi.org/10.1016/j.yhbeh.2009.09.013

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

Ploubidis, G., & Frangou, S. (2011). Neuroticism and psychological distress: To what extent is their association due to person-environment correlation? *European Psychiatry*, 26(1), 1-5. doi:10.1016/j.eurpsy.2009.11.003

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

Proverbio, A. M. (2021). Sex differences in the social brain and in social cognition. *Journal of Neuroscience Research*. https://doi.org/10.1002/jnr.24787

Puts, D. A. (2006). Cyclic variation in women's preferences for masculine traits: Potential hormonal causes. *Human Nature*, *17*(1), 114–127. https://doi.org/10.1007/s12110-006-1023-x

Racine, S. E., Keel, P. K., Burt, S. A., Sisk, C. L., Neale, M., Boker, S., & Klump, K. L. (2013). Individual differences in the relationship between ovarian hormones and emotional eating across the menstrual cycle: a role for personality?. *Eating behaviors*, 14(2), 161–166. https://doi.org/10.1016/j.eatbeh.2013.02.007

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11401–11405. https://doi.org/10.1073/pnas.1721165115

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

Rafiee, Y., Stern, J., Ostner, J., Penke, L., Schacht, A. (2023). Does emotion recognition change across phases of the ovulatory cycle? *Psychoneuroendocrinology*, 148 (105977), https://doi.org/10.1016/j.psyneuen.2022.105977

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition*, *54*(3), 179–185. https://doi.org/10.1016/j.bandc.2004.01.002

Revelle, W. R., & Scherer, K. (2009). Personality and emotion. In D. Sander, & K. Scherer (Eds.), Oxford Companion to Emotion and the Affective Sciences (pp. 304-305). Oxford University Press.

Revelle, W., Wilt, J., & Rosenthal, A. (2010). *Individual Differences in Cognition: New Methods for Examining the Personality-Cognition Link* (pp. 27–49). https://doi.org/10.1007/978-1-4419-1210-7_2

Riggio, H. R., & Riggio, R. E. (2002). Emotional expressiveness, extraversion, and neuroticism: A meta-analysis. Journal of Nonverbal Behavior, 26(4), 195–218. https://doi.org/10.1023/A:1022117500440

Roney, J. R. (2016). Evolutionary Psychology and Endocrinology. *The SAGE Encyclopedia of Economics and Society*, 1067–1083. https://doi.org/10.4135/9781452206905.n259

Roos, A., Lochner, C., Kidd, M., van Honk, J., Vythilingum, B., & Stein, D. J. (2012). Selective attention to fearful faces during pregnancy. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *37*(1), 76–80. https://doi.org/10.1016/j.pnpbp.2011.11.012

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. (2012). Sex-Specific Associations Between Peripheral Oxytocin, Symptoms, and Emotion Perception in Schizophrenia. *Schizophrenia Research*, *136*(February), S70. https://doi.org/10.1016/s0920-9964(12)70261-x

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 (07), 1814–1827.

Russell, J.A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*. 115, 102–141

Sacher, J., Okon-Singer, H., & Villringer, A. (2013). Evidence from neuroimaging for the role of the menstrual cycle in the interplay of emotion and cognition. *Frontiers in Human Neuroscience*, *7*(JUN), 1–7. https://doi.org/10.3389/fnhum.2013.00374

Sakaki, M., & Mather, M. (2012). How Reward and Emotional Stimuli Induce Different Reactions Across the Menstrual Cycle. Social and personality psychology compass. 6. 1-17. 10.1111/j.1751-9004.2011.00415.x.

Sander, D. (2013). Models of emotion: The affective neuroscience approach. In J. Armony & P. Vuilleumier (Eds.), The Cambridge Handbook of Human Affective Neuroscience (pp. 5-53). Cambridge University Press.

Scarantino, A., & de Sousa, R. (2021). Emotion. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Summer 2021 Edition). Stanford University. Retrieved from https://plato.stanford.edu/archives/sum2021/entries/emotion/

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Scherer, K. R., & Scherer, U. (2011). Assessing the Ability to Recognize Facial and Vocal Expressions of Emotion: Construction and Validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior*, *35*(4), 305. https://doi.org/10.1007/s10919-011-0115-4

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology*, *32*(1), 76–92. https://doi.org/10.1177/0022022101032001009

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Schlegel, K., Fontaine, J. R. J., & Scherer, K. R. (2019). The nomological network of emotion recognition ability: Evidence from the Geneva Emotion Recognition Test. European Journal of Psychological Assessment, 35(3), 352–363. https://doi.org/10.1027/1015-5759/a000396

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

Schultheiss, O. C., & Stanton, S. J. (2009). Assessment of salivary hormones. In *Methods in Social Neuroscience*. Guilford Press.

Schultheiss, O. C., Dlugash, G., & Mehta, P. H. (2019). *Hormone measurement in social neuroendocrinology: A comparison of immunoassay and mass spectrometry methods*. *1*, 1–

35.

Schulze, L., Lischke, A., Greif, J., Herpertz, S. C., Heinrichs, M., & Domes, G. (2011). Oxytocin increases recognition of masked emotional faces. *Psychoneuroendocrinology*, *36*(9), 1378–1382. https://doi.org/10.1016/j.psyneuen.2011.03.011

Schweizer, K., Ren, X., & Zeller, F. (n.d.). On modeling the ceiling effect observed in cognitive data in the framework of confirmatory factor analysis. In *Psychological Test and Assessment Modeling* (Vol. 61, Issue 3, pp. 333–353).

Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395–399. https://doi.org/10.1177/0963721411424739

Shirazi, T. N., Levenberg, K., Cunningham, H., Self, H., Dawood, K., Cárdenas, R., Ortiz, T. L., Carré, J. M., Breedlove, S. M., & Puts, D. A. (2021). Relationships between ovarian hormone concentrations and mental rotations performance in naturally-cycling women. *Hormones and Behavior*, *127*(August 2020). https://doi.org/10.1016/j.yhbeh.2020.104886

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*(November 2019). https://doi.org/10.1016/j.yhbeh.2019.104647

Skuse D. (2003). Fear Recognition and the Neural Basis of Social Cognition. *Child and adolescent mental health*, 8(2), 50–60. https://doi.org/10.1111/1475-3588.00047

Smith, S. E. (2020). Is Evolutionary Psychology Possible? *Biological Theory*, *15*(1), 39–49. https://doi.org/10.1007/s13752-019-00336-4

Sokolowski, K., Schmalt, H.-D., Langens, T. A., & Puca, R. M. (2000). Assessing achieve- ment,

affiliation, and power motives all at once: The multi-motive grid (mmg). *Journal of Personality Assessment*, 74 (1), 126–145.

Solomon, R. C. (2008). The philosophy of emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), Handbook of emotions (pp. 3–16). The Guilford Press.

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

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

Sundin, Z., Chopik, W. J., Welker, K. M., Ascigil, E., Brandes, C. M., Chin, K. A., Ketay, S., Knight, E., Kordsmeyer, T. L., McLarney-Vesotski, A., Prasad, S., Reese, Z. A., Roy, A. K., Sim, L. W. X., Stern, J., Carré, J., Edelstein, R. S., Mehta, P., Penke, L., … Tackett, J. L. (2020). *Estimating the Associations between Big Five Personality Traits, Testosterone, and Cortisol*. https://doi.org/10.31234/osf.io/xfcjv

Sundström-Poromaa, I. (2018). The Menstrual Cycle Influences Emotion but Has Limited Effect on Cognitive Function. *Vitamins and Hormones*, *107*, 349–376. https://doi.org/10.1016/bs.vh.2018.01.016

Tagiuri, R. (1969). Person perception. In G. Lindzey & E.Aronson (Eds.), Handbook of social psychology (2nd ed., Vol. 3). Reading, Mass.: Addison-Wesley

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-

flight. *Psychological Review*, *107*(3), 411–429. https://doi.org/10.1037/0033-295X.107.3.411

Terracciano, A., Merritt, M., Zonderman, A. B., & Evans, M. K. (2003). Personality Traits and Sex Differences in Emotions Recognition Among African Americans and Caucasians. *Annals of the New York Academy of Sciences*, *1000*, 309–312. https://doi.org/10.1196/annals.1280.032

Thagard, P. (2002). How Molecules Matter to Mental Computation. *Philosophy of Science*, *69*(3), 429–446. https://doi.org/10.1086/342454

Thayer, J., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology*, *41*(3), 243–246. https://doi.org/10.1111/1467-9450.00193

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition and Emotion*, *28*(7), 1164–1195. https://doi.org/10.1080/02699931.2013.875889

Todorov, A., Mende-Siedlecki, P., & Dotsch, R. (2013). Social judgments from faces. *Current Opinion in Neurobiology*, *23*(3), 373–380. https://doi.org/10.1016/j.conb.2012.12.010

Tooby, J., & Cosmides, L. (2008). *The Evolutionary Psychology of the Emotions and Their Relationship to Internal Regulatory Variables* (pp. 114–137).

Treloar, S. A, Heath, A. C, Martin, N. G. (2002) Genetic and environmental influences on premenstrual symptoms in an Australian twin sample. *Psychological Medicine*.32:25–38.

Van de Pol, M., & Wright, J. (2009). A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behaviour*, *77*(3), 753–758. https://doi.org/10.1016/j.anbehav.2008.11.006

Van Honk, J., & Schutter, D. J. (2007). Testosterone reduces conscious detection of signals

serving social correction: implications for antisocial behavior. Psychological science, 18(8), 663–667. https://doi.org/10.1111/j.1467-9280.2007.01955.x

Van Kleef, G. A. (2009). How Emotions Regulate Social Life: The Emotions as Social Information (EASI) Model. Current Directions in Psychological Science, 18(3), 184–188. https://doi.org/10.1111/j.1467-8721.2009.01633.x

Van Kleef, G. A. (2016). The Interpersonal Dynamics of Emotion: Toward an Integrative Theory of Emotions as Social Information. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107261396

Van Kleef, G. A., & Côté, S. (2021). The Social Effects of Emotions. *Annual review of psychology*, 73, 629–658. https://doi.org/10.1146/annurev-psych-020821-010855

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

Vongas, J. G., & Al Hajj, R. (2017). The effects of competition and implicit power motive on men's testosterone, emotion recognition, and aggression. *Hormones and Behavior*, 92 , 57–71

Wagner, H. L., Macdonald, C. J., & Manstead, A. S. R. (1986). Communication of Individual Emotions by Spontaneous Facial Expressions. In *Journal of Personality and Social Psychology* (Vol. 50, Issue 4, pp. 737–743).

Wearne, T., Osborne-Crowley, K., Rosenberg, H., Dethier, M., & McDonald, S. (2018). Emotion recognition depends on subjective emotional experience and not on facial expressivity: evidence from traumatic brain injury. *Brain* injury, 1–11. Advance online publication.

https://doi.org/10.1080/02699052.2018.1531300

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

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York.* ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wilhelm, O., Hildebrandt, A., Manske, K., Schacht, A., & Sommer, W. (2014). Test battery for measuring the perception and recognition of facial expressions of emotion. *Frontiers in psychology*, 5, 404. https://doi.org/10.3389/fpsyg.2014.00404

Wood-Downie, H., Wong, B., Kovshoff, H., Cortese, S., & Hadwin, J. A. (2021). Research Review: A systematic review and meta-analysis of sex/gender differences in social interaction and communication in autistic and nonautistic children and adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *62*(8), 922–936. https://doi.org/10.1111/jcpp.13337

Wu, M., Zhou, R., & Huang, Y. (2014a). Effects of menstrual cycle and neuroticism on females' emotion regulation. International journal of psychophysiology: *official journal of the International Organization of Psychophysiology*, 94(3), 351–357. https://doi.org/10.1016/j.ijpsycho.2014.10.003

Wu, M., Zhou, R., HUANG, Y., WANG, Q., ZHAO, Y., LIU, . (2014b). Effects of Menstrual Cycle and Neuroticism on Emotional Responses of Healthy Women. *Acta Psychologica Sinica*. 46. 58-68. 10.3724/SP.J.1041.2014.00058.

Young, K. S., Parsons, C. E., LeBeau, R. T., Tabak, B. A., Sewart, A. R., Stein, A., Kringelbach, M. L., & Craske, M. G. (2017). Sensing emotion in voices: Negativity bias and gender

differences in a validation study of the oxford vocal ('OxVoc') sounds database. *Psychological Assessment*, *29*(8), 967–977. https://doi.org/10.1037/pas0000382

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

Zhang, W., Zhou, R., Wang, Q., Zhao, Y., & Liu, Y. (2015). Progesterone mediates the late positive potentials evoked by affective pictures in high neuroticism females. *Psychoneuroendocrinology*, 59, 49–58. https://doi.org/10.1016/j.psyneuen.2015.04.023