CYCLICAL CHANGE IN METACOGNITION ACROSS THE MENSTRUAL CYCLE

submitted in partial fulfilment of the requirements of the

Degree of Doctorate in Philosophy

NICOLE D'AMATO CARUANA

Department of Cognitive Science

University of Malta

2023

Dedicated to all the women battling silently every month

L-Università ta' Malta

FACULTY/INSTITUTE/CENTRE/SCHOOL MAKS, Cognitive Science

DECLARATION OF AUTHENTICITY FOR DOCTORAL STUDENTS

(a) Authenticity of Thesis/Dissertation

I hereby declare that I am the legitimate author of this Thesis/Dissertation and that it is my original work.

No portion of this work has been submitted in support of an application for another degree or qualification of this or any other university or institution of higher education.

I hold the University of Malta harmless against any third party claims with regard to copyright violation, breach of confidentiality, defamation and any other third party right infringement.

(b) Research Code of Practice and Ethics Review Procedure

I declare that I have abided by the University's Research Ethics Review Procedures. Research Ethics & Data Protection form code FRECMAKS_1819_029.

☑ As a Ph.D. student, as per Regulation 66 of the Doctor of Philosophy Regulations, I accept that my thesis be made publicly available on the University of Malta Institutional Repository.

□ As a Doctor of Sacred Theology student, as per Regulation 17 (3) of the Doctor of Sacred Theology Regulations, I accept that my thesis be made publicly available on the University of Malta Institutional Repository.

□ As a Doctor of Music student, as per Regulation 26 (2) of the Doctor of Music Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

□ As a Professional Doctorate student, as per Regulation 55 of the Professional Doctorate Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

Abstract

This study investigates changes in metacognition across the menstrual cycle among naturally-cycling women. Metacognition, the ability to reflect on and regulate one's own cognitive abilities, is crucial in understanding subjective cognitive impairments often reported in the absence of objective cognitive performance changes. Despite its relevance, research on metacognition during the menstrual cycle remains limited. This study aims to address this research gap by examining metacognitive change and its covariation with physical and affective symptoms. Participants (n=121) completed the Daily Online Metacognitive Evaluation (DOME), a novel questionnaire developed specifically for this study, which included measures of physical and affective symptoms. Over the course of two full menstrual cycles, a total of 7,172 questionnaires were collected. Multi-level modelling, utilizing cosine regression functions, and structural equation modelling techniques were employed to analyze within-person individual differences and the relationship between cognitive, physical, and affective changes. The results evidence a menstrual effect on metacognition and demonstrate that individual variations in cyclical change are consistently significant and hold greater importance than average levels of change. Looking at co-variations: cognitive, physical, affective, and headache symptom type changes were correlated but distinct. This study presents the first systematic documentation of metacognitive changes in the context of the menstrual cycle, shedding light on underlying causal mechanisms and informing menstrual symptom classifications. Notably, the evidence of symptom covariation supports the reclassification of metacognitive symptoms as primary symptoms alongside physical and affective symptoms, rather than as behavioural accessory symptoms as currently listed in the DSM-5. Beyond clinical implications, these findings hold significance in shaping social policies, such as those pertaining to menstrual leave.

Keywords: Metacognition, menstrual cycle, individual differences, structural equation modelling, premenstrual dysphoric disorder (PMDD), menstrual leave

Acknowledgements

I am deeply grateful to my supervisors, Professor Ian Thornton and Professor Jeff Kiesner, for their exceptional guidance and support during my PhD journey. Professor Thornton, your understanding of my working style and the positive environment you created made this experience truly valuable. Your mentorship has been invaluable, and I believe every researcher should have such an exceptional supervisor. Heartfelt thanks also go to Professor Kiesner, whose extensive expertise and warm demeanour have been an immense source of support throughout this project. Working with accomplished scholars like you has been a privilege, and your guidance has shaped my research profoundly. I am immensely grateful to Dr. Fiona Sammut for her invaluable contribution of expertise and time to this project. My heartfelt thanks go to the 121 women who generously volunteered two months of their time and who showed genuine enthusiasm for this research. I am profoundly grateful to my father for his unwavering emotional support and for granting me the invaluable opportunity to dedicate five full years to this pursuit. To my beautiful mother for her friendship, unwavering love and support; for taking over when I couldn't. I cannot thank my mother-in-law Josette enough for her help every week – this dissertation was built on Tuesdays! To my closest friends Rachel, Lynsey, Josef, and Alvin, thank you for the good times and for making life more meaningful. Heartfelt thanks to Tessa, who encouraged me to pursue this journey and the first person to call a PhD fun. I am immensely grateful to my husband, Kristian, who has been my unwavering pillar of support throughout the journey of completing this work. As my sounding board, go-to person, and essentially a third supervisor, you played an indispensable role in its success. I cannot thank you enough for being as enthusiastic and invested in this project as I am. Your insightful feedback and valuable suggestions have helped shape the direction of my research, and I truly appreciate your contribution. Lastly, my gratitude to my baby Theo for his presence and endless hugs. Thank you for always putting everything into perspective and reminding me of what truly matters. Each one of you has made a significant impact on my life, and I will forever be grateful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Chapter Outline

Since the discovery of sex hormone neuroreceptors in the 1960s (Pfaff, 1968; Eichenbaum & Otto, 1992; Bixo et al., 1995) researchers have studied the influence of sex steroids on cognitive function using the menstrual cycle as a model of hormonal flux (Toffoletto et al., 2014; Leeners et al., 2017). Despite reports of memory impairment, distractibility, confusion and indecision during certain phases of the cycle, decades of objective testing have failed to conclusively determine an observable effect (Souza et al., 2012; Sundström-Poromma & Gingnell, 2014; 2018).

Metacognition is an emerging area of research being investigated as a causal factor in multiple conditions rooted in the psychological and the physiological domains (Quattropani et al., 2014; Pahlevan et al., 2019; Ortega-Gonzales et al., 2021). Metacognition is thinking about thinking; an ability to comment on and regulate our own cognitive abilities (Winnie, 2018). Evidence for the separation of metacognition and cognition suggests that the menstrual cycle can influence metacognition, directly or via mediating factors, even if it does not have an effect on cognition (Fleming et al., 2012; Vaccaro & Fleming 2018). There is a gap in the research surrounding metacognitive changes during the menstrual cycle. Research pertaining to metacognitive symptom patterns, possible variation between individuals and co-presence with other affective and physical symptoms is virtually non-existent. This is the focus of the present thesis.

In this chapter, section 1.2 provides an overview of the workings of the menstrual cycle. Existing research relating the cycle to cognition and metacognition is outlined in sections 1.3 and 1.4 respectively. Possible theoretical pathways by which the menstrual cycle can influence metacognition are provided in section 1.5. Recom-

mendations on how to study the menstrual cycle and employ a sound methodological design are presented in section 1.6. Finally, section 1.7 provides a detailed overview of the aims and objectives of the study.

1.2 Sex hormones and the brain

1.2.1 Overview of sex hormones

Sex hormones, primarily produced in the gonads (ovaries in females) and adrenal glands, are known for their involvement in reproduction and the development of secondary sex characteristics. Their influence, however, extends to essentially every organ system, from immunity to bone development (Gennari, Merlotti, & Nuti, (2011; Moulton, 2018; Chakravarty & Sammaritano, 2019). When it comes to the brain, sex steroids are able to cross the blood-brain barrier by transmembrane diffusion because of their small size and lipophilic properties (Banks, 2012; Hampl, Bičíková & Sosvorová, 2015). Recent discoveries have shown that small amounts of gonadal steroids can also be produced directly in the brain (Azcoitia, Yagueb, & Garcia-Segurac, 2011; Biegon et al., 2020).

Once through the blood-brain barrier, sex hormones affect the central nervous system by binding with receptors in one of two ways (see Figure 1.1). The first binding option is with the nuclear membrane, intracellularly. This triggers a genomic mechanism of protein production through direct DNA interactions. Genomic pathway effects are long lasting and slow in onset. The second binding possibility is extracellular with the cell membrane, resulting in a non-genomic effect. In this situation a number of complex signalling pathways are activated to bring about an imminent transient response (overviews in Falkenstein et al., 2000; Marrocco, 2016).

Of the three classes of sex steroids (androgens, estrogens and progestogens), females mainly produce testosterone, estradiol and progesterone (Beltz & Moser, 2020). Testosterone is produced in relatively small amounts. Estradiol and progesterone con-

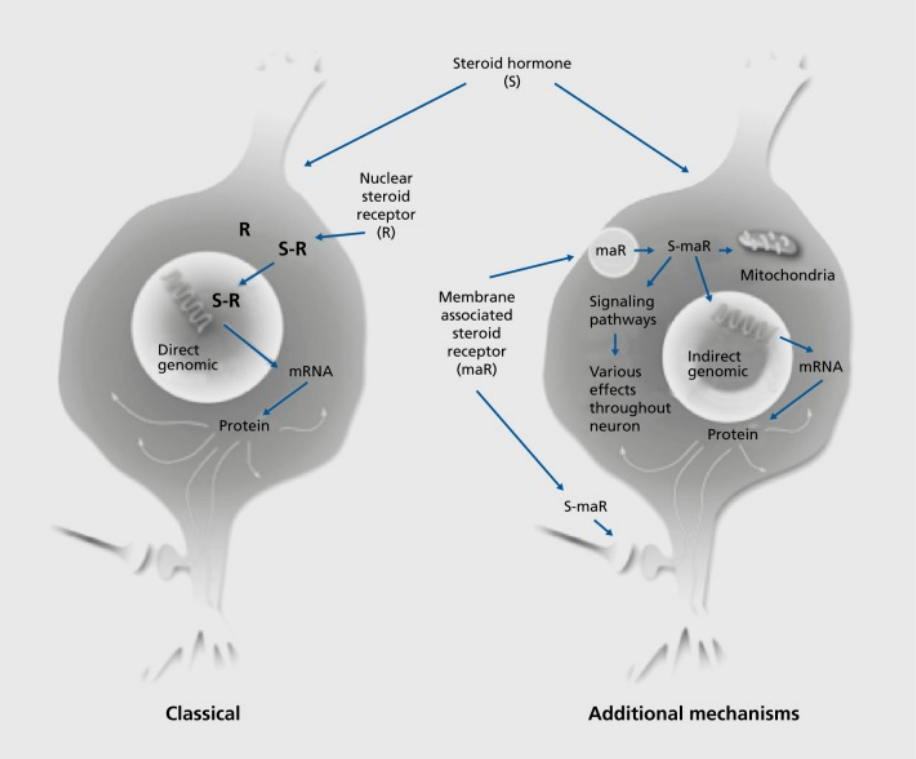


Figure 1.1: Genomic versus non-genomic mechanisms for steroid hormones. From "Sex in the brain: hormones and sex differences" by Marrocco & McEwen., 2016, pg. 374

centration levels vary significantly throughout the typical female lifespan, bringing about two types of effect.

Organizational effects occur as a result of hormonal changes during sensitive periods of development (e.g., ontogenesis and puberty). These produce lasting brain changes as a result of interactions between hormones and genes and are essential, among other functions, for sexual differentiation. Estradiol and progesterone levels also vary as a result of the cyclical ovarian and uterine changes necessary for reproduction – a process known as the menstrual cycle. When these hormones are present, activation effects – transient, reversable effects on brain activity – occur (overviews in Arnold, 2009; Berenbaum, & Beltz, 2011).

Because of these cyclical fluctuations in hormone levels, the menstrual cycle is regarded as an ideal model for the study of sex hormone effects (Le, Thomas & Gurvich, 2020; Thomas & Gurvich, 2021). Cycle influence on emotive, physical and cognitive states within and between females as well as between sexes provides significant insight into the effect of sex steroids.

1.2.2 Hormones and the menstrual cycle

The menstrual cycle is a recurring cycle of fluctuating sex hormones which trigger the production of gametes for reproduction and prepare the uterine lining for cell implantation in the case of conception. In the typical female the cycle occurs every 28 days on average, from menarche (average age of 12) until menopause around age 50 (Itriyeva, 2022).

Sex hormone fluctuations are regulated by the hypothalamus and the pituitary. The hypothalamus secretes gonadotropin-releasing hormone (GnRH), which prompts the anterior pituitary to generate and release follicle stimulating hormone (FSH) and luteinizing hormone (LH). These then stimulate estrogen and progesterone hormone secretions by the ovaries. GnRH, FSH and LH production is in turn modulated by estrogen and progesterone through a series of positive and negative feedbacks (Walker, 1997; Reed & Carr, 2015; Xu et al., 2022).

The start of the cycle is indicated by a blood show on Day 1, when low levels of estrogen and progesterone cause the shedding of the thick endometrial lining in the early follicular phase. Estrogen increases on day 6 (see Figure 1.2), staying high throughout the late follicular phase (up to day 12) as a result of developing follicles which release the hormone. It peaks just before ovulation on day 14, which occurs following a surge in LH. The ruptured follicles luteinise into the corpus luteum, which

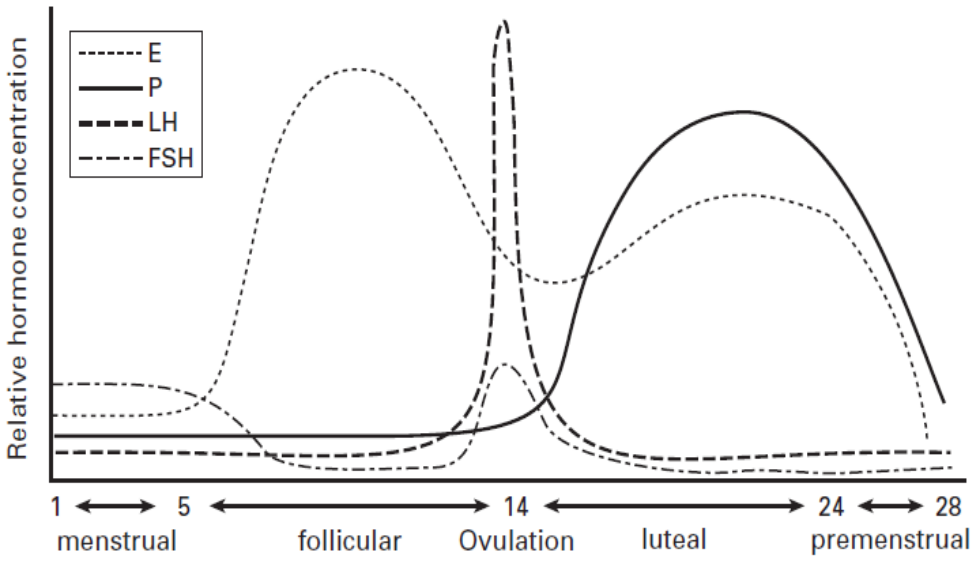


Figure 1.2: Menstrual cycle hormone fluctuations. Phases of the cycle highlighting fluctuations in pituitary and ovarian hormone levels (estrogen (E), progesterone (P), luteinizing hormone (LH) and follicle stimulating hormone (FSH)). From "The neuromodulatory properties of gonadal steroid hormones with regard to individual differences in cognition and brain organisation" by Hodgetts, 2016, p. 9.

secretes progesterone and estrogen causing estrogen to rise and progesterone to reach its maximum peak during the luteal phase (day 22). Estradiol and progesterone levels both fall rapidly during the premenstrual phase (days 24 – 28) before a new cycle begins.

1.3 The menstrual cycle and cognition

Steroid hormone receptor presence in the brain was initially established by Pfaff and Stumpf in the late 60s. Using steroid titration methods, they recorded receptor presence in the hypothalamus and amygdala, thus confirming the hypothalamus-pituitary-gonadal feedback mechanism (Pfaff, 1968; Stumpf 1968; Pfaff & Keiner, 1973). However, what was unexpected was the discovery, first demonstrated by Eichenbaum and Otto in 1992, of estrogen-activated receptors throughout the entire brain, including regions involved in cognitive processes. Receptor presence was confirmed in the hippocampus (Menard & Harlan, 1993), prefrontal cortex (Hao, Rapp & Janssen, 2007) and primary sensory-motor cortex (Chen et al., 2009), amongst others. Owing to the fluctuating estrogen levels during the menstrual cycle, there was now reason to believe that the presence or absence of gonadal hormones had an influence on functions like memory, behaviour regulation, attention and mood, in line with cognitive symptom reports (Moos, 1969).

Commonly reported cognitive symptoms can be grouped in the following categories: attention (difficulty concentrating, distractibility); memory (memory impairment, forgetfulness), decision-making (lowered judgement, lack of self-confidence when making decisions) and general mental status (confusion, decreased efficiency, motor insufficiency, feeling out of control and overwhelmed). Some women also experience positive cognitive changes, for e.g., higher creativity, inspiration and work production, but these are less frequently reported. In the case of menstrual disorders such as premenstrual syndrome (PMS) and premenstrual dysphoric disorder (PMDD), symptoms cause severe impairment and distress (reviewed in Moos, 1969; Farage et al., 2008; Souza et al., 2012).

The effect of the menstrual cycle on cognitive brain structure and function is highlighted below. In line with the scope of this thesis, focus is on the activating effects on the brain as a result of menstrual cycle changes. Brain organizational changes (as a result of puberty, menopause etc.) are not reviewed. Menstrual cycle effects on cognitive performance are discussed in section 1.3.2.

1.3.1 The menstrual cycle and the cognitive brain

1.3.1.1 Phase effect on brain structure

Studies looking at grey matter volumes during different cycle phases have consistently found changes in two areas of the brain: the hippocampus and the basal ganglia. Both areas are implicated in distinct but related memory and learning mechanisms. The hippocampus has been shown to increase in grey matter during the follicular phase when estradiol levels are high (Barth et al., 2016; Pletzer et al., 2018; Meeker et al., 2020). The opposite effect occurs with the basal ganglia, with grey matter volume decreasing during the follicular phase and increasing after ovulation in the luteal phase, when progesterone levels peak (Protopopescu, 2008; Lisofsky et al., 2015; Hidalgo-Lopez & Pletzer 2021). Structural changes have been also demonstrated in other areas: the inferior parietal lobe (Meeker et al., 2020), the prefrontal cortex (Petersen et al., 2015), the insula (Lisofsky et al., 2015), and the anterior cingulate cortex (De Bondt et al., 2016), but results have proven conflicting and difficult to replicate (reviewed in Beltz & Moser, 2020; Dubol et al., 2021).

1.3.1.2 Phase effect on brain function

Task-based fMRI literature demonstrating menstrual-cycle-phase dependence is vast and inconclusive, making it difficult to summarize. A select few studies which are of

relevance to this thesis and which have proven replicable are highlighted. Studies looking at cognitive control through go/no-go tasks found increased activity in the prefrontal cortex (PFC) during the follicular phase as compared with the premenstrual phase, indicating a relation between higher estradiol and performance monitoring PFC activity (Colzato et al., 2012; Bannbers et al., 2012). For the same task and for a reinforcement learning task, the opposite effect was observed in the anterior cingulate cortex (ACC), the area involved with decision-making, with higher activity during the luteal phase suggesting the influence of progesterone (Schoning et al., 2007; Thimm et al., 2014; Diekhof & Ratnayake, 2016). Functional response findings for the hippocampus and basal ganglia corroborated structural findings, with a higher activity response recorded in the hippocampus in the follicular phase and in the basal ganglia after ovulation for spatial navigation tasks (Toffoletto et al., 2014; Pletzer et al., 2019). Resting-state fMRI studies, which investigate functional neural connectivity without using a task, have identified neural networks which show increased connectivity both in the follicular (De Bondt et al., 2015) and luteal phases (Pletzer et al., 2016; Engman et al., 2018). Other studies, however, have found no discernible effects (Hjelmervik et al., 2014), with overall findings considered inconsistent in a systematic review by Hilago-Lopez and Pletzer (2021).

1.3.2 The menstrual cycle and cognitive performance

The majority of our understanding regarding cognitive performance and the menstrual cycle originates from research on cognitive sex differences, which continues to be a prominent area of interest in the field. Cognitive skills that are considered sexually dimorphic, favouring one sex over the other, have been extensively tested with the assumption that skills favouring men, such as visuospatial tasks, improve during the early follicular phase when estradiol and progesterone levels are lower, and decline

during the mid-follicular and luteal phases when hormone levels rise (Le et al., 2020). Conversely, it is presumed that tasks favouring women, such as verbal tasks, exhibit the opposite effect (Andreano & Cahill, 2009). Therefore, it is important to interpret study results in the context of these hypotheses.

1.3.2.1 Visuospatial ability

For male-favouring cognitive tasks, Sundström-Poromma and Gingnell (2014), the most comprehensive review to date, found no overall effect of cycle phase on visuospatial ability (reconfirmed in Sundström-Poromma and Gingnell, 2018). Studies included in the review focused mainly on mental rotation tasks (either the 3D version designed by Shepard and Metzler in 1971, or the pen and pencil variation by Vandenberg and Kuse, 1978). Out of the 12 studies conducted, only four demonstrated a phase effect, all of which consistently showed improved performance during the early follicular phase when estrogen levels were low. Le et al., (2020) reviewed the included studies and found that strong methodological flaws undermined the results. Additionally, a recent study suggests no effect of cyclical hormones on mental rotation performance (Shirazi et al., 2021). Dietrich et al., (2001) and Schöning et al., (2007), compared brain activation response for mental rotation tasks with performance in said tasks. While fMRI results showed sex differences in mental rotation there were no significant performance differences, suggesting that not all neuroactive effects are reflected at the performance level. Beltz and Mozer (2020), suggest that either neural response is more sensitive to the effect of hormonal changes than is behaviour, or that brain function compensates for hormonal variation.

1.3.2.2 Verbal tasks

Results for female-favouring tasks are similarly inconclusive with respect to menstrual cycle effects (reviewed in Sundström-Poromma & Gingnell; 2018). The majority of

tasks focus on verbal fluency and verbal recall, but tests for explicit and implicit verbal memory and verbal working memory have also been carried out (summarized in Beltz & Moser, 2020; Le et al., 2020). Of the 13 studies included, only three showed any influence of the phase on verbal tasks, and results were not entirely congruent. Maki et al., (2002) suggested improvements during the mid-luteal phase, Solis-Ortiz and Corsi-Cabrera (2008) demonstrated improvements during the late-follicular and late-luteal phases, and Rosenberg and Park (2002) found improvements during the early-follicular stage and ovulation. Of the studies included, several were subject to methodological issues (reviewed in Bernal et al., 2022).

While cycle effects on verbal task performance have been deemed unlikely (Sundström-Poromma et al., 2014; 2018; Le et al., 2020), it is interesting to note that verbal tasks (specifically verbal memory) are the only type to repeatedly and consistently show improvement following estrogen treatment administration in post-menopausal women (reviewed in Sherwin, 2012). Such findings suggest that, similarly to the visuospatial tasks, estradiol effects on verbal skills cannot be discounted, despite the influence not translating into consistent changes in gross performance across the cycle.

1.3.2.3 Non-sexually dimorphic tasks

Studies that have looked beyond the traditional sexually dimorphic tasks have mainly focused on high functioning executive tasks following evidence of menstrual cycle influence on the prefrontal cortex (Becker & Hu, 2008). Summarizing findings is challenging since a variety of incompatible tests are used to measure variable skills. Overall, though, results are once again inconclusive (Sundström-Poromma & Gingnell, 2018). Performance results for cognitive control and inhibition vary according to the task. Evidence from a stop-signal task showed performance impairment in the late follicular

phase (Colzato et al., 2010), performance improvement in late follicular in an inhibition of return task (Colzato et al., 2012), no effect in a go/no-go test (Bannbers et al., 2012) and improved attention in a Stroop Task during the follicular phase (Hatta & Nagaya, 2009). Solis-Ortiz et al., (2004; 2008) used the Wisconsin Card Sorting test to measure cognitive flexibility, the mental ability to adjust activity. Improvements in performance were reported during the early follicular and luteal phase (Solis-Ortiz et al., 2004) but no effect of menstrual cycle phase on cognitive flexibility was found in a later study by the same authors (Solis-Ortiz & Corsi-Cabrera., 2008). Studies investigating working memory suggest improvements during high estradiol phases, particularly the late follicular (Rosenberg & Park, 2002; Hampson & Morley, 2013), while other research shows the exact opposite effect, i.e., impairment in working memory during periods of high estrogen levels (Gasbarri et al., 2008).

In conclusion, despite several claims linking cognitive performance and the menstrual cycle (overviews in Farage et al., 2008; Sherwin, 2012), more recent comprehensive reviews (Sundström-Poromma & Gingnell, 2014; 2018) suggest no consistent association. Inconsistencies have been attributed to a number of factors: effects too minimal or transient (Beltz & Moser, 2020), lack of standardization in methodological designs and practices (Le et al., 2020; Schmalenberger et al., 2021), research not taking into account factors that may drastically influence results, such as individual differences and socio-cultural factors (Jacobs & D'Esposito, 2011; Colzato & Hommel, 2014; Hyde, 2016; Pletzer et al., 2019), and excessive focus on the effect of estradiol while dismissing or ignoring the influence of progesterone on cognitive performance (Le, Thomas & Gurvich, 2020). These factors are discussed in more detail in section 1.6 below.

1.4 The menstrual cycle and metacognition

1.4.1 Defining metacognition

Metacognition can be broadly defined as higher-order thinking about one's cognition, i.e., our ability to comment on and regulate our own cognitive abilities (Baker, 2010). Tracking of cognition is typically carried out automatically and covertly (Heyes et al., 2020), and is necessary for everyday behaviour, problem solving and learning (Azevedo, 2020). It also occurs explicitly in the form of self-report surveys and think-aloud protocols that demand reflection on and assessment of one's own cognitive, motivational and emotional operations (Winnie, 2018).

Metacognition theoretical frameworks originally only included **metacognitive knowledge:** "...one's stored knowledge or beliefs about oneself and others as cognitive agents, about tasks, about actions or strategies, and about how all these interact to effect the outcomes of any sort of intellectual enterprise" (Flavell, 1979, p. 906). Of the three main components (person, task and strategy), person knowledge is of special interest to the study, since it involves self-perceived competence (or self-efficacy). One's perceived cognitive competence is dependent on global beliefs (e.g., that memory retention decreases with age) and specific beliefs (e.g., that I work better after a cup of coffee; Hultsch, Hertzog, & Dixon, 1990).

The maintenance and modification of self-perceived competence occurs through metacognitive monitoring and metacognitive control processes incorporated in almost all contemporary frameworks of metacognition since Flavell (1979). **Metacognitive monitoring** involves source monitoring, error detection and the registration of emotional reactions to task difficulty (Whitfield, 2005). This emotional component is de-

tected through metacognitive experiences; conscious cognitive or affective "feelings" that occur during the task at hand allowing you to gauge how well you are doing (Schwarz, 2015). These may range from simple spontaneous incidents where you suddenly realise that you are not understanding what somebody is saying, to more complex and lengthy experiences where you wonder for some time whether you can understand what someone is doing. **Metacognitive control** is concerned with the self-regulation of behaviour following metacognitive monitoring. An example of such adaptive behaviour would be slowing down reading to better comprehend difficult text (reviewed in Panadero, 2017; Wong et al., 2019).

Metacognitive monitoring feeds metacognitive knowledge and vice-versa, making the two processes highly interdependent. For example, you may feel like you're far from achieving a goal (metacognitive monitoring). How you react to this will likely be informed by your metacognitive knowledge, and the outcome of that will likely once again re-inform and redefine or reinforce your knowledge. The ability to track performance, detect errors and control behaviour is essential in the maintenance of task specific metacognitive knowledge and self-efficacy (Whitfield, 2005; Kralik et al., 2020). Issues with metacognitive monitoring and control result in misperceptions of cognitive ability and hamper both short-term processes as well as long-term achievement efforts (Borkowski et al. 2000; Efklides 2011).

1.4.2 Metacognition vs cognition

1.4.2.1 Conceptualization and biology

The following excerpt from Azevedo (2020), highlights one of the main challenges in the field of metacognition:

> "... the complex interaction between cognition and metacognition continues to challenge researchers, as it is hard to distinguish between them since each of the constructs rely and influence each other and share processes. This represents a dilemma of having a higher-order agent overlooking and governing the cognitive system while also simultaneously being part of it" (p. 92).

In the widely accepted conceptualization by Nelson and Narens (1990), the relationship is characterized by two levels: the object level and the meta-level. The object level focuses on the processing of cognitive functions, including object discrimination, spatial representation, and semantic encoding. On the other hand, the meta-level is concerned with task representation. Bottom-up processing of information is directed upwards from the object to the meta-level (metacognitive knowledge and monitoring) and top-down regulation is imposed on object-level functions (metacognitive control). Recent discoveries have resulted in the fleshing out of this view to include 'orders of behaviour' and 'access to consciousness' to the original 'levels of representation' dimension as demonstrated in Fleming et al. (2012; Figure 1.3).

Typically, metacognition is measured through subjective report, where some form of conscious reflection or judgement on one's cognitive activity is made. However, it is well established that cognitive performance can occur without subjective experience, as evidenced by studies on blindsight patients (Weiskrantz et al., 1974). Patients with lesions to the primary visual cortex perform well above chance when responding to targets even though they report not seeing the targets and have no metacognitive insight regarding the performance (Ko & Lau, 2012; Scott et al., 2014). A similar phenomenon occurs in healthy individuals where the target is masked (Kunimoto et al., 2001; Lau & Passingham, 2006; Song et al., 2011). This supports the notion of a degree of dissociation between object-level and meta-level functioning (see Figure 1.3) and is mirrored in other cases where metacognition fails. In patients with schizophrenia, Alzheimer's

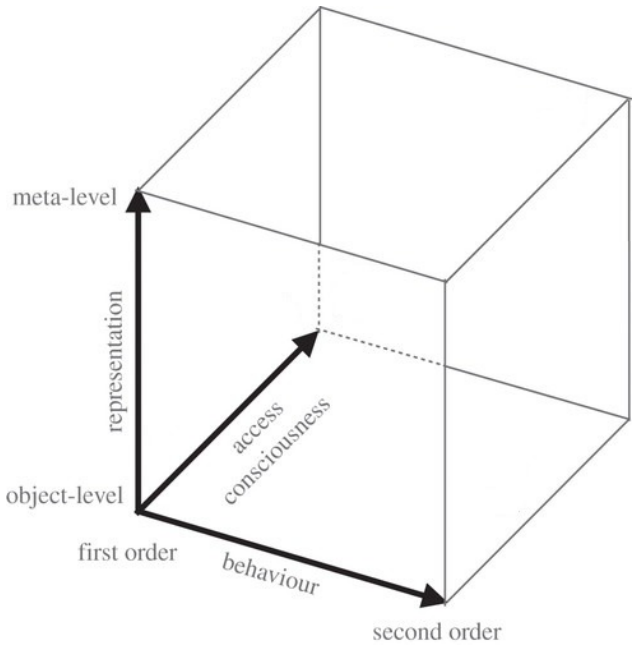


Figure 1.3: A visualization of metacognition which includes representation levels, orders of behaviour and access to consciousness (adapted from Fleming et al., 2012)

disease, and traumatic brain injury, disparities exist between how subjects perceive themselves and what others observe (Schmitz et al., 2006, David et al., 2012; Hoven et al., 2019) as well as between how well they think they perform and how they actually perform (Flashman, & McAllister, 2002; Rouault, 2018; Moses-Payne et al., 2019). In cancer patients undergoing chemotherapy, perceived cognitive impairment is reported as a main symptom yet objective cognitive performance has consistently been shown to remain intact (Diane & Tallman, 2015; Rosa et al., 2021; Oerlemans et al., 2022). Evidence has shown that perceived cognitive decline actually precedes the objective evidence demonstrated with neuroimaging (Saykin et al., 2006; Asher & Myers, 2015).

Significant effort has been made to identify the neural basis of metacognition (reviewed in Shimamura, 2000; Fleming & Dolan, 2014; Vaccaro & Fleming, 2018). Dissociation of a neural correlate of cognition from metacognition is only possible if

the two are operationalized and tested independently to each other; a challenging task considering their inter-dependencies. Typically, upon stimulus presentation a first order response (Figure 1.3) is requested so as to measure task performance. An example of a first order response would be stating whether a stimulus is present/absent or horizontal/vertical. Second order responses take on the form of **metacognitive judgements** where participants are asked to rate their confidence in their first order response (Nelson, 2001; Fleming & Dolan, 2012). **Metacognitive sensitivity** is the extent to which metacognitive judgements accurately predict task performance (Siedlecka et al., 2016).

Neuroimaging studies employing metacognitive judgements reveal distinct neural substrates between first-order and second-order responses providing support for the theoretical proposal by Nelson and Narens (1990) that cognition operates, at least to some degree, separately to metacognition (reviewed in Shimamura, 2000; Fleming et al., 2012; Vaccaro & Fleming, 2018). In all second-order judgements, irrespective of task, the role of the prefrontal cortex (PFC) is highlighted (Hayes & Kasnizak, 2005, Fleming & Dolan, 2014; meta-analysis in Vaccaro & Fleming, 2018). In their study, Qiu et al. (2018) found activation of the anterior PFC, specifically the dorsal anterior cingulate cortex and lateral frontopolar cortex, during second-order judgments. However, during first-order judgments, the inferior frontal junction was activated. In Bellon et al., (2020) the inferior frontal gyrus which is a part of the PFC, was activated during metacognitive monitoring in children but not during the actual mathematical task. And in Fleming et al., (2010), metacognitive accuracy in a perceptual decision-making task was correlated with grey matter volume in the right rostrolateral areas of the prefrontal cortex (PFC); a region which was not employed during the objective performance section of the test. Evidence for the value of the PFC in metacognition is also mirrored in lesion studies where patients with damage to the prefrontal cortex have impaired metacognitive judgements (reviewed in Shimamura, 2000). Similarly, disorders show-

casing metacognitive failure such as Korsakoff's syndrome and schizophrenia respectively demonstrate atrophy in the frontal lobe (Shimamura & Squire, 1986) and reduced anterior medial PFC activity during self-reflection when compared to controls (David et al., 2012).

Results are mixed as to whether metacognition is domain-general or domain-specific, dependent on the respective first-order task (reviewed in Rouault et al., 2018). Domain-general camps suggest that metacognitive insight into a task is a good indicator of results in another (Ais et al., 2016; Faivre et al., 2018; McCurdy et al., 2013; Song et al., 2011) while others propose that different types of insight are fractionated, with one dysfunction not necessarily predicting the other (Baird et al., 2013; Fitzgerald et al., 2017; Garfinkel et al., 2016; Morales et al., 2018). David et al., (2012) demonstrated that patients with schizophrenia showed aberrant judgements of agency but intact judgements of performance for the same task and Hayes et al., (2005), found preserved prospective judgements but impaired retrospective confidence judgments in patients with lateral frontal lesions. On a neural level, the same conflicting results and arguments for both views are found (Fleming et al., 2014; Morales et al., 2018; Valk et al., 2016). This has led some authors to believe in the co-existence of both domain-general and domain-specific systems (Fielder et al., 2019).

As an end note, it is crucial to point out that while second-order responses are taken as a marker for metacognition, it is possible that not all second-order behaviours are meta-representative (Figure 1.3). If judgements are being made directly using information provided by the stimulus, then second-order judgements may be happening at an object-level representation (Fleming et al., 2012; Vaccaro & Fleming, 2018; Fleur et al., 2020). While this scenario is theoretically possible, it still needs to be proven empirically. At this stage it is safe to say that since, as seen above, cognitive and

metacognitive behaviours have proven dissociable, with this dissociation also being demonstrated via two-layer neural network computational models (Timmermans et al., 2012), then at least some degree of separate representation at a neural level is highly probable (Fleming et al., 2012).

1.4.2.2 Metacognition vs cognition in the context of the menstrual cycle

Very few studies exist that directly assess and compare cognitive and metacognitive changes related to the menstrual cycle. In a study investigating the disparity between cognitive functioning and self-perceived cognitive effectiveness, participants underwent a series of neuropsychological assessments during different phases of the menstrual cycle. They were requested to evaluate their perceived efficacy before and after completing each task. The accuracy of their perceived efficacy was determined by comparing the efficacy ratings with actual performance. The findings did not indicate any significant impact of the menstrual phase on cognitive facilitation and did not uncover notable variations in metacognitive accuracy between phases or genders (Whitfield, 2005). Another study looked at the effect of the menstrual cycle on metamemory and everyday memory (n=60 healthy women) via questionnaire sampling at three intervals throughout the cycle (Zare et al., 2013). Results were mixed, suggesting phase influences for everyday memory and certain sub-scales of metamemory (perception of ability and strategies) but not for others (contentment with memory).

Meza-Moreno et al. (2021) examined the effectiveness of self-reported executive functioning compared to performance-based measures in understanding PMS and PMDD. Results by each measure were compared to an administered premenstrual symptoms screening test. Findings showed that self-reported complaints correlated strongly with severity of PMS/PMDD as measured by the screening tool, with associations being almost double during the luteal phase as compared to the follicular. Only

very weak associations were found between performance-based measures and symptom screening test results.

Unfortunately, the mentioned studies are marred by significant flaws, including a lack of within-subject design, underpowered, absence of hormonal measures to identify the menstrual phase, testing conducted at different times of the day, interval testing, and failure to identify sensitive populations, among others. Further, the existing research shows that in stark contrast to other areas where metacognitive research has exploded over the past few years, metacognition is not yet being considered as a factor in menstrual disorders (e.g., PMS/PMDD). Areas where metacognitive research has become well-established include fibromyalgia, chronic pain, anxiety and mood disorders, schizophrenia and chemo impairment. Similar to menstrual conditions, all exhibit subjective cognitive impairments and are rooted in the somatic and psychological (Quattropani et al., 2014; Sellers et al., 2018; Pahlevan et al., 2019; Ranieri et al., 2020; Batmaz et al., 2021; Ortega-Gonzales et al., 2021).

There is a gap in the research pertaining to self-reported cognitive symptomatology (a metacognitive measure), how it relates to other symptoms, and how symptoms differ between individuals. To begin with, however, there is the pressing question of whether metacognition truly varies as a result of the menstrual cycle.

1.5 Pathways underlying metacognition

Evidence for the separation of metacognition and cognition suggests that the menstrual cycle can influence metacognition directly or via mediating factors even if it does not have an effect on cognition. Reported cognitive improvements or deficits experienced throughout the cycle can occur as a result of a number of pathways including: direct gonadocorticoid influence on metacognitive monitoring, altered metacognitive beliefs based on personal and sociocultural expectations, and through other concurrent symptoms such as pain or emotional distress; all influences which can occur independently of objectively-identifiable cognitive changes. The different ways that the menstrual cycle can affect metacognition are addressed in more detail below.

1.5.1 Direct sex steroid influence

Metacognition has been shown to be sensitive to the effects of sex-specific hormones. In a study examining meta-memory, Hodgetts et al. (2015) found that individuals exhibiting schizotypal traits with higher levels of estradiol showed a significant decrease in the production of false memories compared to matched controls with lower levels of estradiol. Additionally, participants with higher estradiol levels displayed less confidence in their false memories compared to those with lower estradiol levels. In an exploratory study, Barrau-Sastre et al., (2022), used menstrual cycle length, amongst other factors, as a marker for cumulative estrogen levels and found an association between cycle length and self-certainty. Similarly, self-reported cognitive failures such as rumination (a repetitive pattern of thinking) have been shown to be estradiol dependent (Graham et al., 2018). Finally, studies looking at the effect of estrogen on cognition found that influence was only (Keenan et al., 2001) or more strongly (Jacobs

and D'Esposito, 2011; Hjelmervik et al., 2012) detectable in tasks that required a high level of executive functioning (possibly metacognition).

Kara et al., (2022), compared androgen levels in adolescents with polycystic ovary syndrome to measure the effects of hormones on metacognition. Cognitive monitoring scores were significantly higher in the non-hyperandrogenism group, with high androgen levels being also correlated with meta-worry (the persistent worry about one's own thoughts and cognitive processes). Research on the relationship between progesterone and metacognition is virtually non-existent with estrogen continuing to be the "prototypic" sex hormone of women. As echoed by others, more study on the function of progesterone is needed (see Henderson, 2018; Sundstrom-Poromaa et al., 2020).

Research exploring the effect of gonadal hormones on metacognition by studying sex differences is inconclusive, with the latest review papers attributing this to averaged results (Moritz et al., 2017; Ferrer-Quintero et al., 2022). Studies that have gone beyond mean differences have found that for example psychosis patients have differing symptom insight profiles based on sex (see Lysaker et., 2019; García-Mieres, 2020) and also respond differently to metacognitive training (Salas-Sender et al., 2019).

It is also important to highlight the role of neurotransmitters in modulating metacognitive effects. Jacobs and D'Esposito (2011) and Colzato and Hommel (2014) both showed that the effect of estradiol on cognition depends on individual variations in dopamine baseline levels. Their findings indicate that cognitive performance improves with moderate dopamine levels but deteriorates with high or low levels, following an inverted-U function. Additionally, Joensson et al. (2015) discovered that participants who received 100mg of L-dopa showed enhanced metacognitive ability compared to those who received a placebo. Moreover, dopamine administration was associated with increased activity in the medial prefrontal cortex. Similarly, Clos et al., (2019), found

that blocking dopamine receptors via administration of antagonist haloperidol, resulted in impaired metacognitive memory confidence. While the estrogen-dopaminergic system is given the most attention (see Sotomayor-Zárate, 2014) sex steroid hormones have been shown to interact with several other neurotransmitter systems including norepinephrine (Ganazzani et al., 1997), serotonin (McEwen & Alves, 1999), acetylcholine (Mitsushima, 2010) and GABA (Epperson et al., 2002). In turn, all these neurotransmitters have been shown to modulate metacognitive skills (in order: Hauser et al., 2017; Livermore et al., 2021).

1.5.2 Metacognitive beliefs

Metacognitive beliefs are defined as "stable knowledge or beliefs about one's own cognitive system, and knowledge about factors that affect the functioning of the system" (Wells 1995, p. 302). These beliefs about one's ability may be aligned or misaligned with one's actual ability.

The menstrual cycle is a biological phenomenon with social implications. It influences a woman's personal identity, behaviour and perception of ability, and in turn shapes culturally-shared beliefs and expectations about the role of women in society (Sommer, 1992; Richardson, 1992). Universally, menstruation has almost always been viewed negatively; as something "unclean" that needs to be controlled by means of various rituals and taboos (Delaney et al., 1976; Tan et al., 2017; Holst et al., 2022). The stigmatizing mark of menstruation can be seen in the effort made to conceal it. It is typically avoided in conversation, discussed only under certain circumstances such as amongst female friends or in a doctor's office (Kissling, 1996) and even then, euphemisms are commonly used (Golub, 1992; Chrisler, 2011). Advertisements for menstrual hygiene products focus on their unobtrusiveness under clothes and in bags (Merskin, 1999) further playing into the narrative of secrecy and the fear of being discovered

as menstruating (Oxley, 1998). The stereotypical menstruating or premenstrual woman is portrayed in books, movies, magazines and other "humorous" media like greeting cards, bumper stickers and internet memes, as fluctuating between tearful, weak and physically ill or mentally unstable, violent, irrational and out of control (Chrisler 2007, 2008; Marvan & Cortes-Iniestra, 2008).

In the workplace, women have long been stigmatized against because of the belief that work performance is negatively affected by the cycle as a result of cognitive deficits, even though research on objective performance is inconclusive at best (Sundström-Poromaa, 2018). Reports indicate that menstrual performance stigma is still very present (DPG, 2019). While extreme measures such as female pilots not being allowed to fly during "the dysfunction of their menstrual cycle" no longer apply in our culture (Cole, 1992), recent studies are still investigating whether women can perform their duties in aviation (Mumenthaler et al., 2001) and the military (Deuster et al., 2011; Manski et al., 2014) as a result of their cycle. Pushbacks against measures such as menstrual leave policies are chiefly justified on the basis that these measures will further exacerbate existing discrimination and negative attitudes toward menstruators (Levitt, & Barnack-Tavlaris, 2020; King, 2021).

Findings have shown that negative global attitudes to menstruation become internalized self-beliefs which influence girls' and women's self-esteem (Hill & Durante, 2009), body image (Roberts, 2004), self-presentation (Karlsson et al., 2014), assertiveness and self-efficacy (Blake et al., 2022). Women's self-perception of ability as a result of the menstrual cycle is overwhelming negative (Johnston-Robledo & Chrisler, 2020), with positive cognitive changes such as higher creativity, inspiration, and work production, being much less commonly reported (Farage et al., 2008; Souza et al., 2012).

Because of this, several studies have suggested that self-reports of menstrual

symptoms could be more due to global attitudes, stereotypes and expectations than the physiological reality of the menstrual cycle (Paige, 1971; Parlee, 1974; Kowalski & Chapple, 2000; Morgan & Rapkin, 2002, Marvan & Cortes-Iniestra, 2008; Feinberg, 2020). Studies have in fact shown that self-reports can be influenced by something other than direct experience. Englander-Golden et al., (1978) manipulated study purpose and made the menstrual cycle salient (retrospective questionnaire) or not salient (general daily questionnaire). Emotional and behavioural factors showed no variation when the menstrual cycle was not a salient part of the study. Marvan and Cortes-Iniestra (2001), demonstrated that women who overestimated the prevalence of PMS in the general population, reported higher premenstrual changes in retrospective questionnaires. This bias in retrospective recall, followed by misperceptions about PMS, showcases the influence of cultural stereotypes when reporting about one's menstrual experience.

The 1996 metacognitive model of Wells and Matthews (also known as The Self-Regulatory Executive Function (S-REF) model) is presented in Figure 1.4. This model has attracted widespread interest because it accounts for the subjective elements that may play a role in psychological and physical distress. The S-REF consists of three interacting levels: a low-level sensory and cognitive system, a regulatory middle system which receives and appraises cognitive information, and high-level self-beliefs which guide and provide coping responses through the metacognitive process. The model's two central tenets are that 1) distorted self-beliefs can serve as the onset for maladaptive cognitive skills, and that 2) maladaptive coping skills in turn feed and maintain distorted self-beliefs (Wells & Cartwright-Hatton, 2004). Repeated and unchallenged maladaptive coping and toxic thinking styles then contribute to psychological disorders/difficulties or distress (Wells & Simons, 2009).

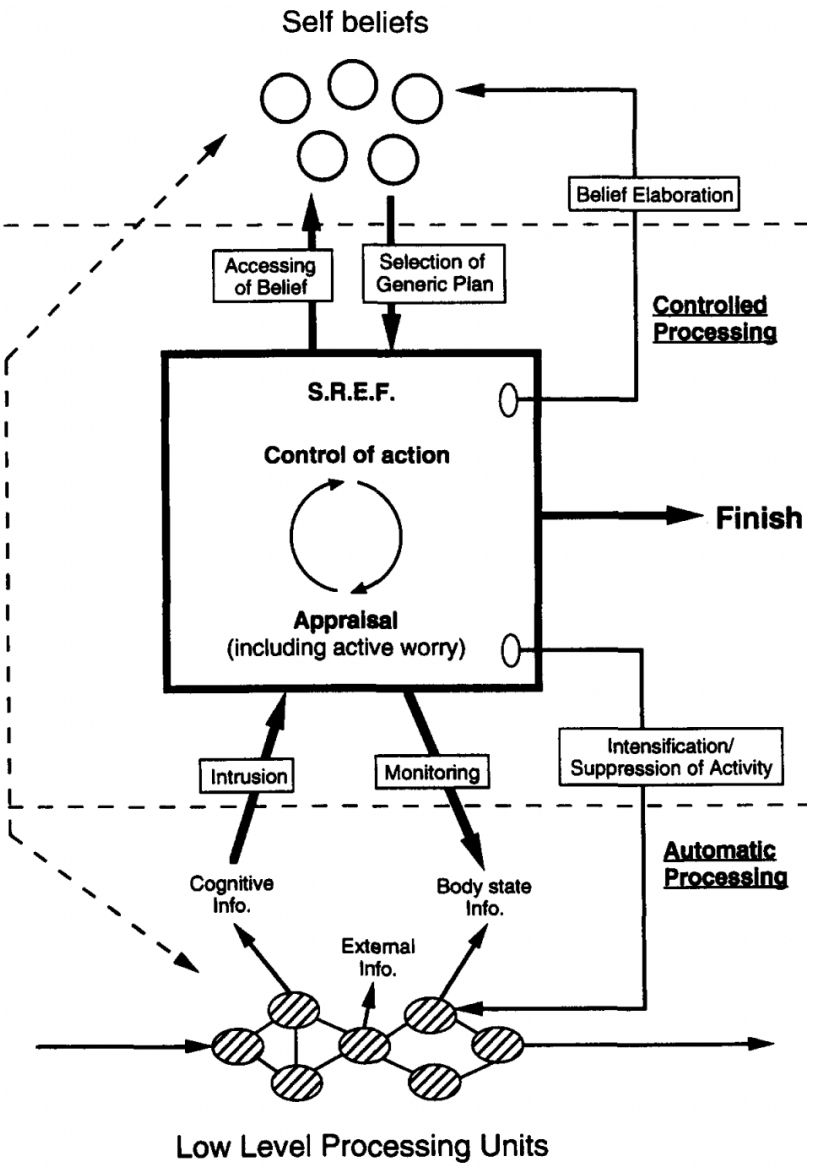


Figure 1.4: The Self-Regulatory Executive Function (S-REF) model of Wells & Matthews (1996, p. 883) consists of three interacting levels: a low level sensory and cognitive system, a regulatory middle system which receives and appraises cognitive information, and high level self-beliefs which guide and provide coping responses through the metacognitive process. Distorted or problematic self-beliefs can lead to maladaptive coping skills such as obsessive hyper-monitoring of body states.

Here I present an example demonstrating the model's value. Mary believes that her ADHD gets worse a week before her period. This leads to a series of metacognitions (e.g., beliefs about one's own worries): that she won't be able to cope at her nursing job because she'll be distracted; that she won't get anything done at home because of sensory overload; and that she'll be overwhelmed and unmotivated for a week. These disruptive metacognitions in turn lead to maladaptive coping measures such as: perseverative thinking (*constant worry, rumination, over-analysing*); inflexible self-focused attention (*monitoring for signs of bleeding, constant checking cycle days on calendar, hyper-vigilant look out for attentional deficits*); and counterproductive coping skills (*trying to distract from the problem*). As a result, Mary feels distracted and not in control; her maladaptive practices feed distress and perpetuate beliefs about symptom consequences, duration and the ability to control symptom outcome.

Metacognitive beliefs have been linked to a wide range of psychopathological conditions including anxiety disorder, clinical depression, obsessive-compulsive symptoms, schizophrenic disorders, anorexia nervosa and bipolar disorder (reviewed in Sun et al., 2017; Sellers et al., 2018; Batmaz et al., 2021). The association has also been evidenced across a number of physiological conditions including: epilepsy (Fisher et al., 2016; Parkinson's disease (Brown and Fernie, 2015), chronic fatigue (Maher-Edwards et al, 2011), diabetes (Purewal & Fisher, 2018), cancer (Cook et al., 2015a; Mutlu et al., 2018) and others (reviewed in Lenzo et al., 2020). In a number of populations metacognitive beliefs are the main predictors of anxiety, depression and trauma (Cook et al., 2015b; Quattropani et al., 2016; Ranieri et al., 2020).

With respect to the menstrual cycle, literature directly investigating the role of metacognitive beliefs on symptom outcome, metacognitive or otherwise, is non-existent. Supportive evidence comes from three streams. Firstly, evidence has shown

that metacognitive beliefs affect metacognitive symptoms. Spada et al., (2010), demonstrated a link between metacognitive beliefs about worry and a self-reported decrease in ability to shift and focus attention. Similarly, Moulds et al., (2010), showed that participants with high levels of metacognitive beliefs reported higher rumination. Secondly, strong support for the influence of metacognitive beliefs comes from demonstrated associations between menstrual attitudes and menstrual symptoms and experiences (Chang & Chen 2009; Rabiepour et al., 2017; Ghiasi et al., 2018; Larki et al., 2022). Negative or positive approaches towards menstruation are strongly linked to the perception and severity of menstrual symptoms (Yeung et al., 2005; Kulakac¸et al., 2008; Erbil et al., 2015). Finally, support for the S-REF model is also evidenced by increased maladaptive cognitive strategies as a result of the cycle. Around menstruation women reported increased self-consciousness, hypervigilance and repeated self-monitoring (Oxley, 1998). Similarly, compared to controls, women with PMDD showed premenstrual increases in self-focused attention and rumination (Craner et al., 2013; Craner, Sigmon, & Young, 2016). According to Fontana (1994), women with PMDD exhibited a tendency to perceive situations as more stressful, undesirable, and changeable during the premenstrual phase compared to the postmenstrual phase, in contrast to the control group. Notably, no significant differences in the frequency of stressful occurrences were reported between the two groups. The efficacy of cognitive therapy in treating PMS and PMDD lends further support for the effect of metacognitive beliefs and coping strategies on the menstrual symptom experience (Lustyk et al. 2009; Kleinstäuber et al., 2012; Kancheva Landolt & Ivanov, 2021).

Additionally, an interesting study by Li et al., (2020), showed that women with general anxiety disorder (GAD) reported an increase in repetitive negative thoughts from the follicular to the luteal phase. Surprisingly, this increase was not correlated with changes in estrogen or progesterone levels. It was also not associated with changes in

reported anxiety symptoms. Further, the effect was not seen in women without GAD. The possible exclusion of hormonal involvement and mediating affective symptoms could suggest the involvement of other factors such as metacognitive beliefs and maladaptive coping strategies. This is especially possible when factoring in the individuals' psychopathological disorder (GAD).

It is possible that metacognitive beliefs also mediate a wide range of menstrual symptoms. Metacognitive beliefs about worry and rumination have been associated with self-reported pain behaviour and pain catastrophizing (Yoshida et al., 2012; Spada et al. 2016; Ziadni et al., 2018; Schutze et al., 2020). Similarly, maladjusted metacognitive beliefs have been shown to negatively impact daily functionality and pain in fibromyalgia patients (Kollmann et al., 2016; Kuppens et al., 2015; Ortega-Gonzales et al., 2021) and chronic pain sufferers (De Rooij et al., 2013; Lee et al., 2017; Pahlevan et al., 2019). Zarbo et al., (2019; 2022) demonstrated that in women with endometriosis, metacognitive beliefs predicted sexual distress over and above pain severity. Coping strategies, metacognitive beliefs, and worry traits were proposed as modulators of pain experience and psychological distress. Metacognitive beliefs have also been shown to influence cognitive performance. In Kraft et al., (2017), dysfunctional metacognitive beliefs were associated with decreased executive control and reduced ability to shift between mental sets.

Similar evidence is available for affective symptoms. Metacognitive beliefs were significantly associated with perceived stress and negative emotions (Spada et al., 2008b), and were the strongest predictor for both anxiety and depression in a longitudinal study of 1,380 individuals (Spada et al., 2008a). This result was replicated by Yilmaz et al., (2011), in the context of stressful life events. Metacognitive beliefs were also shown to be better predictors of anxiety and depression than other mechanisms

such as attention, awareness and control (Solem et al., 2015; Reinholdt-Dunnel et al., 2019; Aldahadha, 2021).

To summarize, the presented S-REF model suggests that personal or global negative menstrual perceptions can result in destructive metacognitive beliefs and maladaptive cognitive coping practices. In turn these have an effect on the perception of menstrual symptom presence and severity.

1.5.3 Other menstrual cycle symptoms

The prevalence and diversity of menstrual cycle symptoms has long been recorded (Moos, 1969). Amongst the most commonly reported physical and mood disruptions are: depressed mood, irritability, tension, mood swings, bloating, breast tenderness, cramps, gastrointestinal issues, back pain, headaches and acne (Zaka & Mahmood, 2012; Ryu & Kim, 2015). Symptom correlations (Kiesner & Pastore, 2010; Eisenlohr-Moul et al., 2015) and co-variations across the cycle (Kiesner & Martin, 2013; Kiesner et al., 2016) are well-established and are a result not just of concurrent steroid effects but also of symptom influences on each other (Kiesner, 2017; Handy, 2022). For example, menstrual fatigue has an influence on social engagement, which can contribute to depression and isolation (Owens et al., 2020), and menstrual physical discomfort in the form of headaches or cramps can result in emotional changes such as irritability (Kiesner, 2009).

It is highly plausible, therefore, that even if they are not the direct cause, sex hormones may trigger a chain of causation leading to proximate symptoms that affect the metacognitive experience. Not much is known about metacognitive symptoms' relation to other cycle symptoms. Outside of the cycle, in aging populations and chemotherapy patients, research has shown that subjective memory complaints are not

consistently linked to cognitive impairment. However, these complaints consistently show a relationship with depressive symptoms, suggesting that affective symptoms may play a mediating role (Reid, 2006; Pullens et al., 2010; Jansen, 2013; Kuba et al., 2019). Similarly, in perimenopausal women, Ford et al. (2004) showed that anxiety and depressive symptoms were the primary predictors of perceived memory impairments. This finding was further supported by Weber and Mapstone (2009), where memory loss reports were not linked with performance but were related to depressive symptoms, somatic issues, and trouble sleeping. In the same population, vasomotor symptoms (hot flushes) were unrelated to objective cognitive performance but were shown to be a significant predictor of subjective forgetfulness (LeBlanc et al., 2007; Mitchell & Woods, 2011; Dragos et al., 2014). Confirmatory factor analysis (CFA) studies in PMS/PMDD women have demonstrated the same factor loading for metacognitive symptoms (e.g., the sense of being out of control; feeling overwhelmed) and physical and emotional symptoms. This supports the idea of united underlying pathways for all symptoms (Teng et al., 2004; Wang et al., 2007).

A number of possible pathways linking metacognition to other menstrual cycle symptoms are presented below. While most connections are theoretical, they pave the way for association studies which in turn lead to understanding symptom causes, consequences and treatments.

1.5.3.1 Biological pathways — CNS

Sex steroid induced stress is a possible cause of metacognitive disruptions. That stressful events have an impact on metacognitive processing is well documented (Reyes et al., 2015; Barrientos et al., 2020; Drueke et al., 2022). The stress response typically consists of a cascade of mechanisms governed by the Hypothalamic-Pituitary-Adrenocortical (HPA) axis, where released hormones, cortisol and catecholamines, are

shown to target the PFC (Arnsten, 1998; 2009). Rostral and dorsal aspects of the lateral PFC are important metacognitive accuracy regions and are stress hormone hotspots (Fleming et al., 2012; Baird et al., 2013; de Martino et al., 2013). Metacognition involves the inward orientation of attention to allow self-reflection. Acute stress is thought to shift brain activity from endogenous attention areas like the dorsolateral and medial PFC, to exogenous attention for vigilant monitoring. (Qin et al., 2010; Hermans et al., 2014; Reyes et al., 2015).

A bidirectional pathway for the influence of metacognition on affect is also possible. Metacognitive emotion regulation is the process underlying the deliberate change in thoughts and goals to ease negative emotions. Studies looking at menstrual cognitive effects in PMS/PMDD populations are notoriously inconclusive – except for when studying tasks with a cognitive-emotional basis, where a consistent, replicable decline in the luteal phase is evidenced (reviewed in Eggert et al., 2017; Slyepchenko et al., 2017; Le et al., 2020). Mahon et al., (2015) and Manikandan et al., (2016), demonstrated that women with high emotional regulation difficulty reported higher levels of menstrual severity as a result of lower perceived control (a metacognitive measure). Similarly, Graham et al., (2017), demonstrated that sex hormones interacted with emotion regulation strategies such as rumination to predict negative affect.

1.5.3.2 Biological pathways—non-CNS

The presence of steroid hormone receptors, indicating the potential effects of steroid hormones, has been demonstrated in various tissues beyond the CNS (reviewed in Farage, Neill & MacLean, 2009). One possible pathway by which gonadal steroids can influence metacognition, is through the mediation of biological mechanisms in non-CNS tissue. That estrogen and progesterone fluctuations influence inflammation is well documented (Whitcomb et al., 2014; Dragin et al., 2017; Yama et al., 2020). One effect

is increased production of circulating proinflammatory cytokines (Dias et al., 2022) which cross the blood-brain barrier (Stenfors et al., 2017) act as neuromodulators (Stuart & Baune, 2015) and have adverse effects on the CNS (Taipa et al., 2019; Janelsins et al., 2022). Proinflammatory cytokines have been shown to disrupt high level cognition and metacognition (Peters et al., 2019; Kindler et al., 2020; Bourgognon & Cavanagh, 2020; Muscat & Barrientos, 2021).

Interestingly, affective stressors have been shown to alter the immune system in the same way after activating the hypothalamic pituitary adrenal axis and sympathetic nervous system (Irwin & Cole, 2011). According to Marriman et al., (2014) genetic variations such as single nucleotide polymorphisms, underlie metacognitive symptom vulnerabilities to inflammation, highlighting the importance of identifying and targeting sub-groups when studying menstrual effects (in depth overview in Section 1.6). The fact that affective stressors are themselves sensitive to sex hormone changes and are also a cause of menstrual inflammation (reviewed in Kiesner, 2017) means that steroid hormone presence could set off a domino effect/feedback loop where metacognition is mediated both by inflammation and stress. Furthermore, physical symptoms like headaches and irritable bowel are also hypothesized to be the by-products of menstrual inflammation (Cupini et al., 2021; Thomas et al., 2020). All this serves to highlight how intricate and interwoven menstrual symptom pathways are.

Similar complex pathways are seen with other tissues like the gut. GI tissue is rich in steroid receptors, so it is susceptible to steroid hormone influences (Pfaffl et al., 2003; Garcia-Gomez et al., 2013; Rastelli et al., 2022). Issues of the gut, like altered gut microbiota (dysbiosis), are well-linked to cognitive dysfunction (reviewed in Luca, 2020). Possible routes include direct activation of the vagus nerve through infection (Bercik et al., 2011), through inflammation (Rundek et al., 2021) and through

the production of gut-derived neurotransmitters (Rea et al., 2016).

1.5.3.3 Experiential pathway

It is possible that the mere experience of a symptom influences another. According to the physical distress model, menstrual cycle physical pain or discomfort can have a direct resultant effect on psychological symptoms (Kiesner et al., 2016; van Iersel et al., 2016). So, for example, menstrual headaches can result in changes in mood. Kiesner (2017), suggests that effects can be bidirectional. Psychological symptoms, for instance, can cause physical symptoms to feel more intense (Kiesner, 2017). The same pathway can be suggested for subjective cognitive impairment, with physiological and psychological distresses bringing about possible perceived attentional or memory deficits and eliciting symptoms of 'loss of control' and 'cognitive fog'. In certain cases, symptom experience may trigger a chain of reactions. For example, Allen et al., (2016) and Hauser et al., (2017), demonstrate that the experience of pain, with its negatively-valanced and physiologically arousing qualities (e.g., increased heart rate), can affect metacognitive judgements resulting in altered (over-) confidence in perceptual decisions.

1.5.3.4 Distal stressors

So far, the pathways discussed involved endocrine changes causing rapid effects on tissues (CNS/non-CNS) through the mediation of other symptoms, physiologically (e.g., genomic/non-genomic hormone/neurotransmitter effects, gut-to-brain signaling, immune response cytokine signaling) or experientially (e.g., physical distress hypothesis; Kiesner, 2016) to result in an *immediate* plethora of physiological, affective and cognitive symptoms which at times feed-back and reinforce the effect. In Kiesner's

(2017), Menstrual Cycle-Response and Developmental Affective-Risk Model (MCR-DAR) the immediate symptoms resulting from gonadal steroids, along with underlying physiological responses and metabolic changes, are referred to as proximal stressors.

The same model proposes that proximal stressors initiate a sequence of behavioural and developmental changes that have long term consequences (see Figure 1.5). Menstrual cycle symptoms (*proximal stressors*) may affect participation and engagement in important developmental activities (*successful engagement*), such as school, social relationships, and physical activities, which in turn lead to long-term maladaptive psychological adjustments (*near-proximal adjustment*) such as negative body-image and low self-esteem, which are not directly determined by the menstrual cycle but are influenced by the experiences associated with it. Maladaptive adjustment problems can then lead to long term consequences which further propagate symptoms and pathological disorders. Cumulative consequences are the result of a snowball of life events triggered by menstrual symptoms, while contemporary consequences are the result of persistent menstrual symptoms and traits (Kiesner, 2017).

Since the MCR-DAR model is expounded in depth in Kiesner (2017), only a few examples evidencing the model will be provided here. Evidence for engagement influence on menstrual symptoms - the feedback route from engagement to proximal stressors in Figure 1.5 - can be seen from COVID-19 studies. Lifestyle changes such as social distancing, travel restrictions and working from home, forced by the pandemic, have led to changes in behaviour (different sleep, work, eating patterns) and social engagement (isolation, more time in front of screens). These changes have had a drastic influence on menstrual cycle experience. Studies have repeatedly found increases in the 30-50% range for the appearance of dysmenorrhea and the worsening of menstrual symptoms (Phelan et al., 2020; Robb et al., 2020; Lebar et al., 2022).

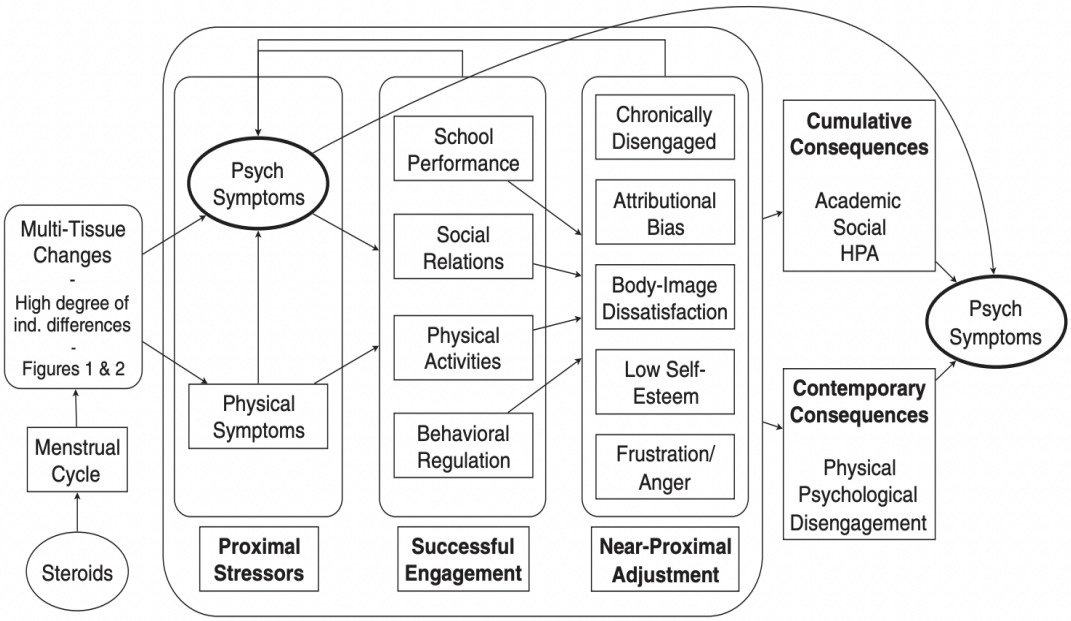


Figure 1.5: Graphic representation of Kiesner's (2017) Menstrual Cycle-Response and Developmental Affective-Risk (MCRDAR) model. Taken from Kiesner (2017), page 227.

Pandemic stress - imitative of the proximal stressors to engagement route in Figure 1.5 - was also shown to exacerbate negative behaviour, coping skills and menstrual symptoms (Muharam et al., 2022; Demir, 2021). Menstrual-symptom triggered behaviours like increased interpersonal conflicts and reduced social engagements have been linked to depression and isolation and attributed to high suicide rates in PMDD women as compared to non-PMDD controls (Owens et al., 2010; Handy, 2022). Similarly, menstrual affective and metacognitive symptoms like impulsivity and distorted judgements, have been linked to consequences like alcoholism (Carroll et al., 2015), substance use (van Iersel et al., 2016) and non-suicidal self-injury (Lockwood et al., 2017). Evidence for the role of dysfunctional psychological and cognitive adjustments as resultants and perpetrators of menstrual symptoms and as important contributors to physiological and psychological pathologies, has already been presented for maladaptive cognitive skills in the section about metacognitive beliefs (see section 1.5.2).

Summary: A series of possible pathways linking the menstrual cycle to metacognition were presented in this section. Somatic pathways include: direct biological pathways (via neurotransmitters, gut-to-brain signalling, inflammation) and symptom-mediated biological pathways. Consideration for the subjective psychological component of the menstrual cycle is given via symptom experiential routes and the psychosocial route of metacognitive beliefs and other distal effects. It is likely that effects are a result of a number of interacting bidirectional pathways.

The next section calls attention to the primary methodological pitfalls of menstrual research and how to circumvent them.

1.6 Studying the menstrual cycle

The presence of discrepancies among results is a prevalent theme in the field, highlighting the importance of understanding their underlying causes. Such understanding is crucial for critically assessing the existing literature and adopting a robust methodological design that ensures meaningful and replicable results. According to Schmalenberger et al., (2021), inconsistencies are due to three main reasons: not accounting for individual heterogeneity, lack of standardization in study designs, and low statistical power. These issues are discussed in detail below.

1.6.1 Not accounting for individual differences

Menstrual cycle differences between women have long been documented. Women have been shown to experience different symptoms (Shoep et al., 2019), for different durations (Pearlstein et al., 2005), at different points in the cycle (Kiesner, 2011), with differing symptom intensities (Lustyk et al., 2008) and symptom reactions (Kiesner, 2017).

Variations in physiological factors are a clear source for individual differences as demonstrated by different responses in effectiveness and side effects to pharmacological manipulations of reproductive steroids (Kiesner, Mendle, Eisenlohr-Moul, & Pastore, 2016). Research has shown that women have differing amounts of hormones at the same points in the cycle (Sundstrom Poromaa & Gingnell, 2014) as well as differing sensitivity to said hormones (reviewed in Kiesner, 2017). For instance, the presence of naturally varying amounts of neurotransmitters, notably dopamine, has been found to influence the extent to which estrogen affects cognition (Jacobs & D'Esposito, 2011; Colzato & Hommel, 2014; Hidalgo-Lopez & Pletzer 2017). Similarly, the effect

of estrogen on working memory has shown to be dependent on interactions with the *COMT* Val158Met genotype, further highlighting the influence of genetic variations (Jacobs & D'Esposito, 2011; Smith et al., 2014).

As previously discussed, other factors beyond physiological responses, such as sociocultural influences and personality traits, may also play a role in symptom report variation. Specific menstrual beliefs and more general values vary between individual families, cultures and sub-cultures, and affect symptom associations formed (Brooks-Gunn & Ruble, 1986; Obermeyer & Sievert, 2007; Chrisler & Gorman, 2019). For example, Clarke and Ruble (1978), revealed that a girl's experience during menarche, including the level of preparedness and the reactions of parents and friends, influenced the subsequent reporting of symptoms during later menstrual periods. These individual associations, whether positive or negative, are then integrated and reinforced by various biases in information processing related to cyclicity. In a clinical sample, Hodgetts, et al. (2015) demonstrated that the effects of estrogen on false memory rates and meta-memory differed depending on the personality of the participant (high schizotype vs low schizotype). In non-clinical samples, metacognition has been shown to be affected by self-esteem (Ehrlinger & Dunning, 2003), self-efficacy (Moores et al., 2006; Roebers et al., 2012), goal orientation (Sperling et al., 2004) and attributional style (Spencer et al., 2021).

Another factor that may vary between individuals is metacognitive ability itself. People with higher metacognitive sensitivity have a better grasp of their cognitive abilities than those with low metacognitive sensitivity (Fleming, 2014), suggesting that the two could report different symptoms even if their menstrual cycle experience was exactly the same. Interindividual differences in human metacognitive ability have been identified across a number of cognitive domains including perception, memory and

decision-making (Song et al., 2011; Fleming et al., 2012, 2010; Baird et al., 2013). Neurologically, Hu et al., (2017) suggests that grey matter volume relates to individual variations in accuracy of metacognitive judgements.

In spite of the overwhelming evidence for physiological and psychological variation in the menstrual experience, the vast majority of literature in the field manifestly fails to account for between-person variation, instead focusing on modelling effects at the group level (Kiesner et al., 2016; reviewed in Sundstrom Poromaa and Gingnell, 2014; Schmalenberger et al., 2021). Group measurements make the underlying assumption of 'sameness' (homogeneity) within groups. If this assumption is violated, measured differences get diluted and/or disappear entirely. Failing to model within-person effects will likely result in underestimated effect sizes and increase the chance of wrongfully accepting a false null hypothesis. This is especially the case if, as happens with the majority of studies, sensitive populations (e.g., subjects with PMS, PMDD etc.) are not previously identified and sample sizes are small. Moreover, unless the data is approximately normally distributed, the arithmetic mean is not only uninformative but also misleading. Schmalenberger et al., (2021), cite lack of accountability for individual differences as the single biggest contributor towards conflicting and inconclusive results in menstrual cycle research.

The gold standard in menstrual research is the modelling of between-person differences while accounting for within-person changes across the cycle (Leeners et al., 2017). Methods such as multilevel modelling (MLM) allow the estimation of both. MLM can provide estimates for average cycle effects (i.e., fixed effects) as well as individual differences in the direction and size of effects (i.e., random effects). Individual differences can be modelled top-down, using hypothesis-driven approaches, in which between-person characteristics are specified as moderators of within-person cycle ef-

fects (e.g., Eisenlohr-Moul et al., 2016; Roberts et al., 2018) or bottom-up, to identify unique groups of symptom changes across the cycle (Kiesner, 2011; Eisenlohr-Moul et al., 2019).

1.6.2 Lack of standardization

Menstrual cycle research is challenging because in order to utilise the cycle one has to operationalise it. Though the sequence of events (the hormonal flux) is fixed, the exact timing of the individual events and levels of hormones achieved at each stage varies significantly from woman to woman and cycle to cycle. For example, one cannot simply 'standardize' the cycle to a 28-day standard because this presumed 'typical' 28-day cycle is found in only 15% of women (Bull et al., 2019). Inconsistencies and imprecisions in the measurement and definition of cycle days and phases make study results questionable and study comparisons and replications difficult if not impossible (Allen et al., 2016, Hampson, 2020).

One of the key factors in ensuring proper cycle phase and reducing error variance is to screen participants for 'typical' cycles - provided of course that regular cycles are the focus of study. Schmalenberger et al., (2021) suggests that subjects to be included:

(a) menstruate monthly and do not experience missed cycles

(b) have a cycle length of 25-35 days

(c) have cycle lengths that do not vary by more than 7 days

(d) are not pregnant or breastfeeding

(e) are not on hormonal medications or devices*

(f) have experienced menarche at least three years prior; and

(g) are screened for perimenopausal symptoms**

*These are known to alter ovarian function and natural hormonal flux (Blumenthal & Edelman, 2008). **Cycle length irregularities and missed cycles increase during these stages (Fehring et al., 2006).

The start of the menstrual cycle is typically taken as the first day of menstruation and can easily be predicted through the use of smartphone menstrual cycle tracking applications (Moglia et al., 2016). Cycle phase can be defined by categorizing cycle days into phases through a series of forward (counting number of days from day 1 of first cycle prospectively) or backward (retrospectively counting backwards from day 1 of second cycle) counting mechanisms (Hampson, 2020). This method of self-reporting is cheap and effective but often makes the assumption of a fixed follicular phase at 14-days after menstruation (Levin, 2018). Using biological measurements together with self-reporting, yields more accurate identification of phases and sub-phases (reviewed in Allen et at., 2016). Options include urinary LH testing, basal body temperature, and saliva/blood testing for estrogen/progesterone. Which biological methods are chosen, if any at all, depends on available resources and cost, the participant burden of each method and the required accuracy level.

1.6.3 Low statistical power

Studies with less than 30 participants per group (e.g., Maki et al., 2002, 16ppnts; Hausmann et al., 2000; 12ppnts) and even less than 10 (Rosenberg & Park 2002, 10ppnts/group; Solis-Ortiz et al., 2004, 9ppnts; Solis-Ortiz & Corsi-Carbrera, 2008, 9ppnts) are common when studying the menstrual cycle, especially when imaging or hormonal testing is involved. Not only do these underpowered studies provide little to no insight, but they also contribute to the confusing, unclear conclusions that pollute menstrual

cycle literature. A sample should be large enough to ensure that between-person differences can be observed and modelled, particularly if sampling from the general population where no subpopulations have been identified (Button et al., 2013). For example, women who experience the worst menstrual cycle symptoms make up only a small part of the general population (Gehlert el al., 2009), and if clinically diagnosed are usually on hormonal treatments (Steiner et al., 2006), thus excluding them from the majority of studies. A large enough sample size ensures that all relevant subpopulations are represented.

For statistical robustness, three repeated measures across one cycle are considered the minimal acceptable standard when looking at intra-individual effects of the menstrual cycle (Schmidt et al., 1998). In an ideal scenario, observations are taken daily. Samples should be taken across more than one cycle to a) allow for greater confidence in reliability of between-person differences (Kiesner, 2011; Leeners et al., 2017), b) account for the changing premenstrual symptoms from one cycle to the next (Reilly & Kremer, 1999; Potter et al., 2009), and c) guarantee that observed changes in symptom cyclicity are due to the menstrual cycle and not accidental (Kiesner et al., 2016).

1.7 Aim of the present thesis

The aim of the present thesis is to investigate metacognitive fluctuations with respect to the menstrual cycle in healthy, naturally-cycling women.

For this research, metacognition will be operationalized through daily self-assessments of cognitive ability with cyclical symptom changes analysed within and across individuals, across multiple symptoms and multiple cycles. Three main research questions will be explored: Does metacognition oscillate in time with the menstrual cycle? Does metacognitive symptom change covary with other symptoms? And are there individual differences in metacognitive cyclical changes?

Research question 1: Does metacognition oscillate in time to the menstrual cycle?

The role of metacognition in the context of the menstrual cycle has been exclusively overlooked. To date, no study can make the claim that metacognition truly varies as a result of the menstrual cycle. This is because existing studies have mainly focused on the link between the menstrual cycle and cognitive performance (Section 1.3). The present work is founded on the premise of a dissociation between cognition and metacognition. Models conceptualizing metacognition distinguish between the two, with a neural degree of separation being evidenced behaviourally and computationally. Oscillating sex hormones can therefore influence metacognition even if no effect on cognitive performance is evident. DSM-V listed PMDD symptoms such as 'subjective difficulty in concentrating' and 'subjective sense of being overwhelmed or out of control' suggest that factors outside of objective cognitive performance may influence reported cognitive deficits. Possible pathways proposed in this chapter include: physiological influences on metacognitive monitoring, altered metacognitive beliefs based

on personal and sociocultural expectations, or concurrent symptoms such as pain or emotional distress. This thesis will address whether metacognitive symptoms are cyclical with respect to the menstrual cycle. Only once this question is answered can one move on to other questions: Do metacognitive symptoms relate to others? Do they vary between individuals? And what, if any, is the role of metacognition in menstruation?

Research question 2: Does metacognitive symptom change covary with other symptoms across the cycle?

Another desideratum lacking in menstrual cycle research is an understanding of symptom interplay. It is unclear to what extent and by which means the constellation of menstrual symptoms relate and affect individual pathology or well-being. This gap in the understanding of shared etiology and causal associations means that menstrual disorder diagnoses and treatments do not reflect the multifaceted nature of these conditions (Zendehdel & Elyasi, 2018). Considering that symptoms cause intense impairment and distress in at least 20% of cases (American College of Obstetricians and Gynecologists, 2022), with the typical woman experiencing menstruation monthly for half of her life, this is extremely worrying. There is currently no clear understanding of the causal mechanisms linking metacognitive abilities to other complaints. Existing literature has mainly focused on grouping symptoms in menstrual disorders based on co-presence and insights into underlying etiology is limited (Teng et al., 2004; Wang et al., 2007). This thesis will be the first to study the co-variation of cyclic *change* across metacognitive, physical and affective symptoms. Studying cyclical symptom change will ensure that symptoms are not just covaried across time, but are also synchronized with the menstrual cycle. Based on the literature (Section 1.5.3) a causal link between affective and metacognitive symptoms is predicted.

Research question 3: Are there individual differences in metacognitive cyclical changes?

There are fundamental methodological issues pertaining to menstrual cycle research (reviewed in Schmalenberger et al., 2021). This thesis draws heavily on the methodology used in Kiesner et al., (2016) which mitigates these limitations. Considerations include daily repeated measures and sampling across more than one cycle. Primarily, however, this study will tackle the long-standing issue of not accounting for individual differences. Evidence of between-person menstrual variation and the repercussions of failing to account for these heterogeneities have already been extensively conveyed in Section 1.6.1. This thesis will use multilevel modelling to model between-person differences while accounting for within-person changes. Focus will be on individual variation in *symptom change* across time. Using a similar methodology, Kiesner et al., (2016) demonstrated that significant individual differences exist in the magnitude and direction of cyclical change for a variety of different physical and emotional symptoms. This result is expected to be replicated in this thesis and extended to metacognitive symptoms.

This chapter introduced the topic of metacognition and the menstrual cycle and showed the gaps in the existing research. Chapter 2 details the process of developing and pilot testing a novel metacognitive measure for this study. Chapter 3 lays down the experiment and data analysis. Results are presented in Chapters 4 and 5. Finally, discussion of results is carried out in Chapter 6 followed by general conclusions in Chapter 7.

Chapter 2

Measure

Building DOME: an instrument for daily metacognition

2.1 Chapter Outline

This chapter details the development of the DOME (Daily Online Metacognitive Evaluation) tool, a questionnaire designed to assess daily metacognition. In the present study it will be used in the context of menstrual fluctuations, with the aim of identifying patterns in individual differences in menstrual-related metacognitive changes as well as correlations between metacognition and other menstrual symptoms. To this end, DOME will be expanded to include measures of physical and affective symptoms.

This chapter is structured follows. Section 2.2. begins by establishing the construct of interest and justifying the need of a new tool as opposed to using a pre-existing questionnaire. Section 2.3 goes over the building process with focus on questionnaire and item formats, item development, and tool length. Testing of the questionnaire with a focus on language and understandability, ease of use, tool validity, and sensitivity to metacognitive changes is covered in section 2.4. Participants' views on recruitment, questionnaire items, and overall user experience are presented in section 2.5 based on feedback from focus groups. Section 2.6 presents exploratory data analysis results, while section 2.7 provides a list of questionnaire modifications and suggested practice changes, along with the final version of DOME.

2.2 Preliminary research

2.2.1 Construct of interest

The construct to be operationalized is metacognition. Specifically, declarative (or person-centred) metacognitive knowledge which deals with one's knowledge about his or her cognitive processes, as opposed to procedural knowledge or strategic knowledge (Flavell, 1976; Paris et al., 1984). Typically, cognitive function is measured through three general approaches: self-report, administered assessment, and performance-based measures (Farias & Mungas, 2008). The last two are not suitable measures for metacognition. Since self-assessment is by definition inherently a metacognitive process, for this study, metacognition will be operationalized through daily self-assessments of cognitive ability.

Evaluations of day-to-day metacognition have many applications outside of this study. One example is with cancer patients, where cognitive function disruption — both as a result of disease as well as chemotherapy — is well established (Liou, 2019; Rodin & Ahles, 2012). Impairment is diagnosed as *perceived* cognitive impairment (PCI) since objective cognitive performance has consistently been shown to remain intact. A metacognitive evaluation tool, whose questions focus on day-to-day self-perceived cognitive competence, would allow tracking of symptom progress.

2.2.2 Pre-existing questionnaires

An existing validated and psychometrically sound questionnaire can save time and resources. However, it may not be readily available for the specific construct under study.

To meet the study's requirements, the chosen tool to measure self-assessed cognitive function should allow repeated measurement, be sensitive to changes, record improvements and deficits, and be applicable to a large sample size. An extensive literature review using relevant keywords was conducted to explore the suitability of existing questionnaires. Despite studying forty-five validated questionnaire and assessment scales in-depth, none were found suitable.

Specifically, the majority of questionnaires were not concerned with metacognition but rather focused on recording cognitive performance through observable action-based failures. Sample questions included: *"Did you forget to lock the door behind you today?"* and *"Did you jump from one task to the other without completing any?"* The tick-box or yes/no format made these measures more ideal for screening than extended monitoring. Further, these questionnaires operated with a 'deficit' perspective, focusing on capturing complaints while not allowing possible improvements to be recorded. Questionnaires included the Cognitive Failures Questionnaire (Broadbent et al., 1982), the Everyday Memory Questionnaire (Cornish, 2000; Sunderland, et al., 1986), and the Memory Complaints Questionnaire of the Cambridge Mental Disorders of the Elderly Examination (O'Connor et al., 1990).

Questionnaires concerned with metacognition tended to focus on general self-awareness of one's learning processes; e.g., *I am aware of what strategies I use when I study*, *I ask myself if I have considered all options after I solve a problem* (Metacognitive Awareness Inventory [Schraw and Dennison, 1994]; Metacognitive Awareness Scale [Song, Loyal & Long, 2021]). These types of questions tap into the construct of interest— declarative metacognitive knowledge—but are mainly concerned with behaviour regulation and cannot be used to measure *daily* change in self-assessed cognitive ability. An example of a question in the desired format would be *"Today, as*

compared to usual, how would you rate your ability to make decisions".

Another limitation was that a number of questionnaires looked only at metamemory, disregarding other reported menstrual cycle cognitive complaints such as concentration and judgement deficits. Examples include the Metamemory in Adulthood (MIA) scale (Dixon, Hultsch, & Hertzog, 1988), the Memory Functioning Questionnaire (MFQ; Gilewski & Zelinski, 1988; Gilewski, Zelinski, & Schaie, 1990), the Multifactorial Memory Questionnaire (MMQ; Troyer & Rich, 2002), and the Prospective and Retrospective Memory Questionnaire (PRMQ; Smith et al., 2000). The one questionnaire that examined more than just memory (Cognitive Functioning Self-Assessment Scale [Annunziata et al., 2011]) was still missing a number of significant cognitive domains such as language and decision-making.

With no pre-existing questionnaire found suitable for the study, the construction of a new questionnaire was deemed appropriate.

2.2.3 Development process

The process of building the questionnaire is outlined in Figure 2.1. Tsang et al.'s (2017) framework was selected from the extensive literature in this field for two primary reasons. Firstly, their guidelines are straightforward and easy to follow. Secondly, their research also focused on subjective feelings of one's current status, specifically pain measurement. The initial stage of the process involves establishing the construct of interest and ensuring no similar questionnaires already exist. The subsequent stage, development, concentrates on construct dimensions, questionnaire and item formats, item development, and questionnaire length. After a thorough review and revision, piloting is conducted to assess language, understandability, usability, sensitivity, validity, and reliability. The subsequent sections of this chapter delve into each stage in detail.

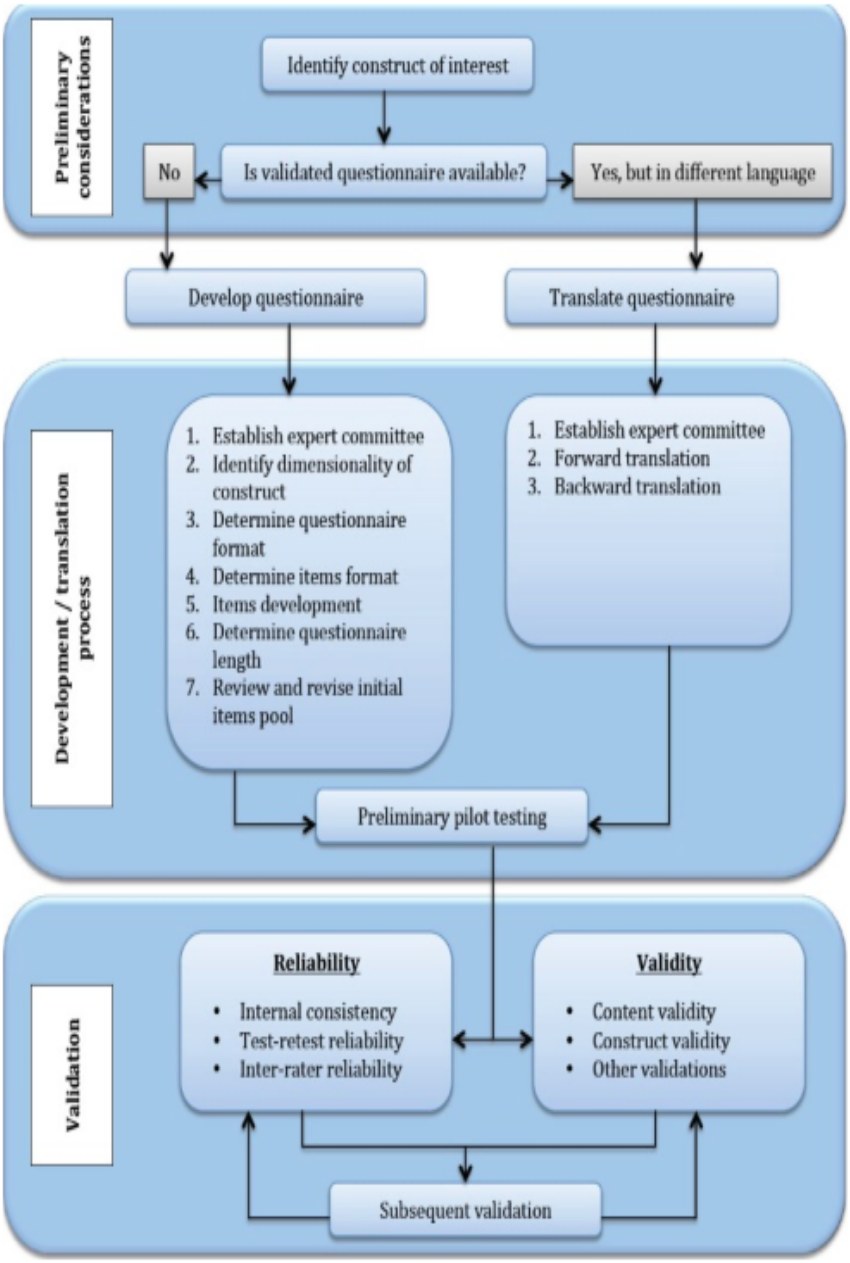


Figure 2.1: Guidelines for developing, translating, and validating a questionnaire. From "Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine" by S. Tsang, C.F. Royse and A.S. Terkawi, 2017, Journal of Saudi J Anaesth, 11(1), p. 80.

2.3 Questionnaire building

The building process was iterative and incremental and focused on the following steps: the establishment of an expert committee, identifying the key dimensions of the construct, determining the questionnaire format, determining the item format, developing the questionnaire items, determining questionnaire length, and the review and revision of questionnaire items.

2.3.1 Expert committee

A specialist in the field of cognition was involved in all aspects of the tool development process. The name DOME (Daily Online Metacognitive Evaluation) was assigned to the assessment tool.

2.3.2 Dimensionality of construct

Five key dimensions were identified: memory, attention, decision making, language and general cognitive ability (GCA). These dimensions are essentially cognitive in nature but become metacognitive phenomena once the process of self-evaluation is applied. Subscales for each dimension were created so that results from each dimension could be examined separately. Being of equal importance, all dimensions carried equal weight. The more ambiguously defined GCA sub-scale contained more question items to allow question pruning during the developmental process. Table 2.1 shows which questions fall under which construct.

Table 2.1: Metacognitive dimensions and their respective items

Dimension
GCA
1. Cognitive Competence
2. Cognitive Fatigue
3. Cognitive Control
4. Cognitive Fog
5. Ability to think creatively
Memory
10. Retention
11. Learning
13. Surprise test
19. Car park memory
Attention
5. Concentration
6. Multitasking
16. Attention to detail
Decision Making
7. Clarity and efficiency
8. Decision making
12. Clarity under pressure
20. Judging important decision
Language
14. Following Instructions
15. Writing Instructions
17. Presentation Delivery
18. Lecture Comprehension

2.3.3 Questionnaire format

The self-administered questionnaire was divided into two main sections based on two different question formats: direct questions (rate your ability today as compared to usual) and hypothetical questions (how do you think you would perform in this situation today as compared to usual). By tapping into the same construct in different ways, tool items could be more easily validated. Subjects could also test which question format they felt most comfortable using.

2.3.4 Item format

To enable easy daily administration and analysis, all question items were close-ended. A 100-point Visual Analogue Scale (VAS) was chosen as the scaling method, with the anchors 'Worse than usual' and 'Better than usual' at the minima and maxima. VAS is suited for DOME since the continuous scaling without intermittent anchors allows complex subjective phenomena to be measured more precisely than with a categorical forced choice scale (Flynn, van Schaik, & Middlesorough, 2004). Variance among respondents is also more easily detected (Funke & Reips, 2012; Klimek et al, 2017). Further, a continuous scale encourages lesser end-aversion bias (Hasson & Arnetz, 2005) and halo effects (i.e., socially desirable response behaviour), since it is harder to estimate which value is expected on the scale continuum (Klimek et al., (2017). For this reason, number markers were not displayed on the questionnaire. The slider starting point was in the middle of the scale, at the 50-point mark indicating no change from the usual condition (Figure 2.2). Users could rate by dragging the slider towards the extremes (Figure 2.3). Advancement through the questionnaire required moving the slider.

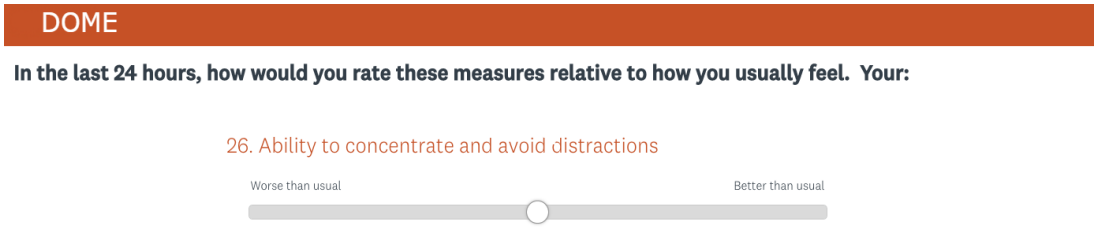


Figure 2.2: DOME scale before scoring

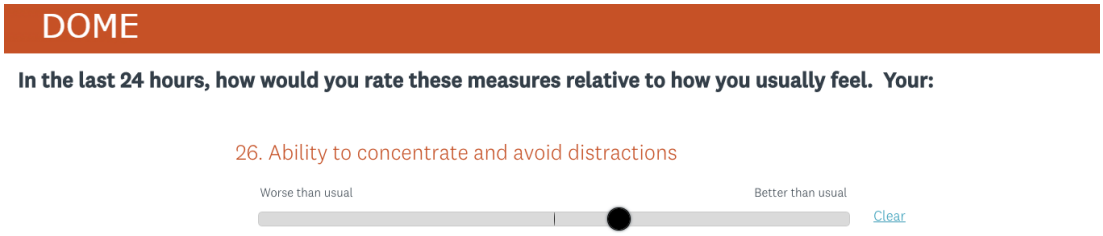


Figure 2.3: DOME scale after scoring

2.3.5 Item development

Items were developed based on questions in pre-existing questionnaires to ensure construct validity. They were tailored to suit the needs of DOME and underwent a number of modifications before a simple, direct and clear final set was decided on. A detailed explanation was provided for the more technical terms which have a specific meaning in the field (e.g., *Cognitive Competence*). Two items, *Cognitive Fatigue* and *Cognitive Fog* were reverse-scored to align rating directions.

2.3.6 Questionnaire length

A crucial element of the menstrual cycle research for which DOME was created is the repeated measure component, whereby participants fill in the questionnaire daily across

two cycles, so for approximately sixty days or more. It was therefore imperative that DOME be as concise as possible, while still tapping into the construct of interest. In Kiesner et al., (2016), forty questions pertaining to physical and affective symptoms were administered. When timed, completion took approximately five minutes. For the first version of DOME a cut-off point of twenty cognitive questions was chosen. This allowed room for an extended version with physical and affective symptoms and also permitted extra questions which could be eliminated during tool development.

2.3.7 Revision

The finalized item sets were reviewed by three experts in the field. Focus was on: question clarity (no double negatives, no double-barrelled questions), grammatical errors, elimination of jargon or words construed too technical, and the removal of leading questions or ones that could be considered too sensitive. The DOME used for testing is presented below.

Table 2.2: DOME questionnaire items pre-piloting

In the last 24 hours, how would you rate these measures relative to how you usually feel. Your:
1. Cognitive Competence (i.e., how mentally motivated and able you feel) . . .   Worse than usual — Same as usual — Better than usual |——————————|——————————|
2. Cognitive Fatigue (i.e., feeling of mental fatigue or exhaustion) . . .
3. Cognitive Control (i.e., sense of feeling in control versus being mentally overwhelmed) . . .
4. Cognitive Fog (i.e., sense of confusion or lack of mental clarity) . . .
5. Ability to concentrate and avoid distractions. . .
6. Ability to multitask. . .
7. Ability to think clearly and efficiently. . .
8. Ability to make decisions. . .
9. Ability to think creatively. . .
10. Ability to remember old information. . .
11. Ability to learn new information. . .
Compared to usual, how do you think you would perform in these hypothetical scenarios?
12. If there was an emergency today, my ability to think clearly would be   Worse than usual — Same as usual — Better than usual |——————————|——————————|
13. If I had to take a surprise test today, I think I would do. . .
14. If I had to follow complex instructions today, I would perform. . .
15. If today I had to write clear instructions for others to follow, they would be. . .
16. If I were to go for a walk today, my attention to small details around me would be. . .
17. If today I had to give a presentation, my delivery would be. . .
18. If I were to listen to a long lecture today, my level of comprehension would be. . .
19. If I had parked in a multi-storey car park today, my ability to find the car quickly would be. . .
20. If I had to make an important decision today, my judgement would probably be. . .

2.4 Pilot testing

2.4.1 Participants

Twelve participants, 6 males and 6 females, with a mean age of 23.7 ± 3.17 years were recruited to take part in the pilot testing of DOME. Eight of the participants were students at the University of Malta while the other four held full-time jobs. Participants were all Maltese citizens, residing at different locations in Malta, and all spoke English as a second language. Recruitment of participants took place via a university email call. Twenty-three people answered the call and the final twelve were chosen on a first come basis and male/female quota requirements. One participant, who dropped out because of trouble committing to the study, was replaced. In line with the research ethics and data protection guidelines of the University of Malta, participation and recording consents were collected prior to testing. All participants were given 20 Euros as compensation.

2.4.2 Procedure

2.4.2.1 Meeting set-up

To ensure a solid understanding of the study requirements, separate face-to-face meetings were held with each individual participant. Participants were told that the study was a pilot for a novel metacognitive questionnaire but were not informed of its use in the context of the menstrual cycle so as not to introduce possible gender response bias.

The aim of the pilot was to collect questionnaire responses and to get feedback on the participants' experience with the tool. Subjects were asked to evaluate the

DOME questions and to take note of anything that was unclear or hindered tool usability. They were also informed that a focus group discussion, centralized on experience and improvement suggestions, would be held at the end of the testing period. Emphasis was made on the impact their feedback would have on a larger study. Participants were guided through the first few questions and shown how to answer using the UI slider. Participants were asked to fill in DOME at the end of the day, after reflection on their overall cognition, memory, language, attention and judgment abilities on that day as compared to usual. They were told to avoid overthinking if they were unsure what to answer. Total completion time was around five minutes and participants were informed that engagement duration would be recorded and available to the researcher. This was done to try and prevent participants from rushing through the questionnaire, encouraging better reflection and data quality.

2.4.2.2 Testing period

Participants were asked to start the questionnaire on the day of the introductory session. The questionnaire had to be answered daily over a period of five days, and could not be submitted unless all questions had been answered. Access was provided through an online link and the questionnaire could be filled in using a desktop or laptop. Completion reminders were sent on a daily basis and participants were contacted if they failed to submit. In such cases, subjects were given the option of filling in the questionnaire the following day if they wanted to stay in the study. Three days in, a reminder was purposely skipped to see how many users filled in the questionnaire without prompting.

2.4.2.3 Follow-up

At the end of the testing period participants were invited to take part in a focus group made up of three people. A one-on-one option was also given for those not comfortable in a group setting. This was not an issue for any of the participants and four group

meetings were held. The focus groups were conducted in an open manner with a set of pre-prepared questions serving as prompts and with the experimenter acting as a moderator. The discussion centred around what the participants thought of the questions as well as the usability experience. The sessions lasted around thirty minutes each.

2.5 Results part 1: Focus groups

Summaries of the focus group responses were created using the notes taken during the interview session, double checking with the recordings when needed. Below, the feedback provided by participants is organized by topic: recruitment, user experience, questionnaire items, and overall experience. This feedback was used to improve the understandability and ease of use of DOME, as well as to aid in participant recruitment and reduce drop-out rates in future studies of DOME in the context of the menstrual cycle.

2.5.1 Opinions on recruitment process

Q: What motivated you to take part in the study?

Many people participated out of a sense of obligation towards the research or the researcher. Others were interested in the financial reward. The majority of people were not necessarily intrigued by the study concept but, during debriefing, were excited to learn that DOME was to be applied to menstrual cycle research. Interest was particularly prevalent among women, who found the subject matter relatable. Many started discussions about their own experiences. Participants were particularly interested in the individual differences between women and asked for the study results to be shared with them upon completion.

2.5.2 Opinions on user experience

2.5.2.1 Response rate

Q: In your opinion, what would help maintain the daily response rate?

Reminders. Reminders to fill in DOME were sent to participants daily. Some participants would have preferred to have the reminder sent to their personal email rather than the university one, since they did not check their university email in the evening. For many, this was not an issue since their university and personal email were both synced on their phone. On the day that a reminder was intentionally not sent, half of the participants forgot to fill in the questionnaire. The people who remembered were the ones who had put a personal reminder on their phone; a practice that should be encouraged in the menstrual cycle study. Although the study was only carried out over five days, some felt that answering the questionnaire had already become part of their routine and went on to fill it on extra days.

Mobile Optimisation. To help maintain a high response rate DOME should be mobile and iPad friendly. Many participants preferred to answer in the evening - a time when they did not use their laptop. Many felt that ideal responding opportunities, such as when one is idle on the bus, were missed because users could not access the questionnaire on their phone.

2.5.2.2 Response quality

Q: In your opinion, what would encourage better quality of response?

Length. The participants felt comfortable with the length of the study and did not feel the need to rush through it. Some even felt that it could have been longer. A few participants would have preferred to have had a marker indicating how many questions remained, but for the majority this was not an issue, especially after filling it in the first time.

Time of Day. Many commented on the time of day and reported this as one of

the main factors affecting their quality of response. The majority of people filled in the questionnaire within one hour of when it was sent. They felt that response quality was better when the reminder was sent earlier in the day (around 4/5 pm) rather than later (7/8 pm). Towards the end of day, they felt that their tiredness biased the results and made it difficult for them to give 'true' ratings of how they felt throughout the day.

External Events. Participants reported struggling to respond on busy days. They felt that the quality of the responses was negatively affected because they rushed to respond. For the menstrual cycle study high stress periods like examinations should be avoided. Some people had events such as accidents or vacation happen on some of the days they were meant to fill in the questionnaire. They felt that on these days their responses were strongly affected by external events, rather than internal. Having a space in the questionnaire where note of a major disturbance could be taken, could provide some perspective to the researcher analysing the data.

Randomization. Some participants reported spending less time thinking about their answers upon becoming more used to the questions. Other participants did not feel that this was an issue but recognized that it could become one over an extended period of time. Question randomization may help reduce this effect.

Missed Days. In the event of forgetting to fill in the questionnaire, participants felt that it would be better to skip the entry entirely owing to difficulty remembering one's performance the following day. Contrarily, one participant felt frustrated that she couldn't report a day she accidentally skipped the day after. She stated that on the missed day, her cognitive performance had differed significantly from its norm and that it would have been the most interesting data from the whole testing period.

Testing Day. Two participants stated mistakes in their responses on the first

day, owing to lack of familiarity with the questionnaire. They did not realize, for example, that some questions changed direction of scale. The rest of the subjects did not have this problem.

2.5.2.3 VAS

Q: What was your experience using the visual analogue scale to answer the questions?

Ten out of twelve participants complained that they were not too happy using VAS. They felt that they did not know where to place their evaluations on the scale. Many stated that they struggled to translate what they were feeling onto the scale. Participants preferred a 1–10 numerical rating scale or else a verbal rating scale (Likert).

2.5.3 Opinions on questionnaire items

2.5.3.1 Questions liked or disliked

Q: Were there any questions that you really liked or disliked?

Direct Questions. Overall, participants felt most confident answering the direct questions. They felt that the descriptors in brackets were particularly useful. Almost two thirds of the group stated that they struggled with one or more of the hypothetical questions and that they preferred the direct questions.

Relatability. A number of participants mentioned the question 'If I had parked in a multi-storey car park today, my ability to find the car quickly would be…'. They felt that they did not relate to this question since they did not own a car. Similarly, non-student participants struggled to relate to hypothetical questions directed towards

students (e.g., surprise test, long lecture). Suggestions included changing specific hypothetical questions to more general ones or else to have two versions of the tool: student and non-student.

Ceiling and Floor effects. Some participants felt that there was a floor effect on the car park question in that they would always remember where they had parked their car, irrespective of how bad their cognition might be on the day. A subject suggested changing the question to 'today compared to usual how quickly would I remember where I left my phone'. Similarly, questions like 'If today I had to give a presentation, my delivery would be…' and 'If I were to listen to a long lecture today, my level of comprehension would be…' were deemed to have ceiling effects by four participants who felt that irrespective of how cognitively able they felt they would always rate the question low because of their dislike of giving presentations and listening to long lectures.

Emergency Situations. Three participants did not like the question 'If there was an emergency today, my ability to think clearly would be…' since they felt that they would act different than usual in an emergency. It was suggested that the word emergency be replaced with 'under high stress'.

2.5.3.2 Repetitive Questions

Q: Did you think any of the questions were repetitive?

None of the participants felt that any of the questions were repetitive.

2.5.3.3 Confidence in Answers

Q: When answering, how sure were you of your answers?

Overall, all participants stated that they were confident in their ratings of improvements or deteriorations of their metacognitive ability and did not answer at random. When prompted on whether any of the questions were ambiguous, the general sentiment was that questions were easy enough to understand, but many were worried about interpreting them incorrectly. This was particularly the case with the hypothetical questions. An example is the question 'If I had to take a surprise test today, I think I would do...' A participant interpreted this as how well she could remember information and was worried that this was the wrong interpretation. Three participants felt that they struggled with the question asking them to rate their creative thinking ability. They felt that this was easy enough to answer if they had engaged in a task that required creativity during the day but struggled if not. One participant was unsure what to compare daily cognitive ability to. He would have preferred a more specific definition of what was meant by the word *usual* (e.g., defining it as the day before or the week before). The other participants did not have an issue with this, and interpreted it as their average state of mind.

2.5.4 Opinions on overall experience

Q: Is there anything the researcher could have done to improve the experience?

Participants who did not realise that some questions introduced a directional change suggested that this be explained to the participants in the future. Others who were concerned with misinterpretation, proposed going through all questions with the participants.

2.6 Results part 2: Exploratory Data Analysis

Focus group results were supported by descriptive and exploratory statistics. Questionnaire response times and scores were analysed and insights on question biases and question item correlations extracted. A key part of analysis was verifying DOME's sensitivity in measuring changes in metacognitive ability. Construct validity and internal consistency tests were run to see whether all questionnaire domains measure perceived cognitive function, whether questionnaire items in the same domain tap into the same dimension, and whether any items were redundant and could potentially be removed. Prior to the analysis, the data was processed so that the scores for *Cognitive Fog* and *Cognitive Fatigue* reflected the directional change. Analysis was carried out using SPSS V.20 and graphs were produced using Python.

2.6.1 Response times

On average it took participants 4.47 ± 2.92 minutes to answer the full questionnaire. One entry was removed since the participant took long breaks between questions. Response times are plotted in Figure 2.4.

The median response time decreases as one progresses through the questionnaire, suggesting that less time is spent reflecting on questions towards the end. Response time increases for each new section of questions (e.g., *Clarity Under Pressure*), then decreases gradually throughout the section. This finding provides support for randomizing questions in the full study.

Within individuals, the difference in mean times for direct (M=12229ms, SD =5850ms) and hypothetical questions (M=12297ms, SD=7644ms) was not significant

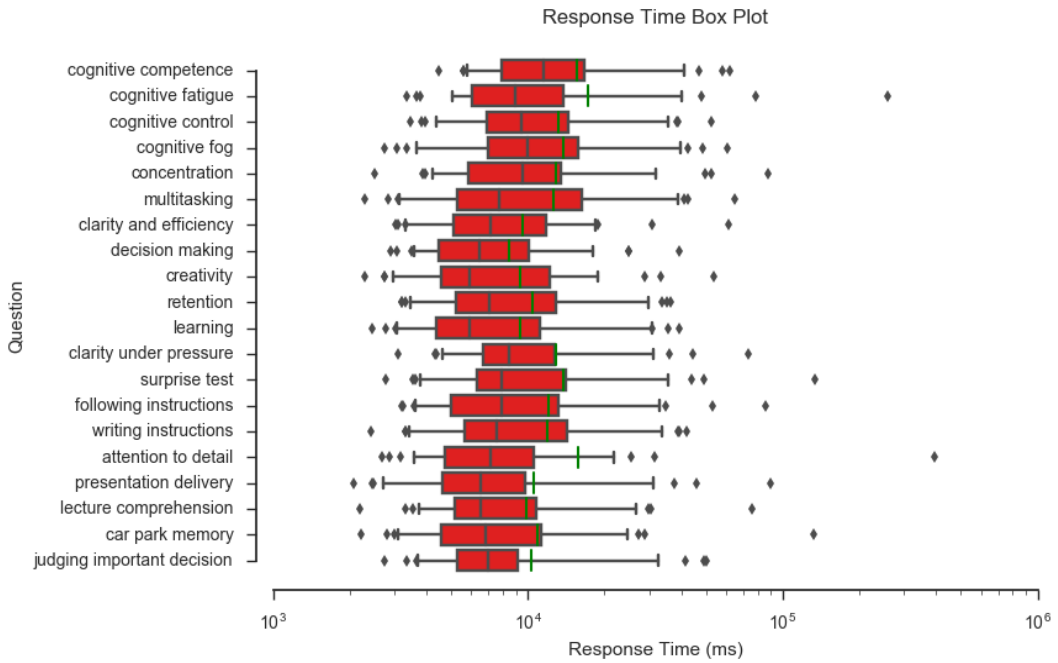


Figure 2.4: Response time box plot. Box and whisker plot displaying inter-quartile ranges (red boxes), 5^{th}–95^{th} percentile ranges (whiskers), medians (vertical black line) and means (vertical green line) for each of the 20 questions' response times. Outliers are shown as a diamond.

(t(11)= -100, p=.922). Therefore, although participants stated that they struggled with the hypothetical questions compared to the direct questions this was not reflected in response times. Similarly, there appeared to be no increase in mean times for questions which were deemed ambiguous or more open to interpretation such as 'rate your creative ability' or 'how would you do on a surprise test'. This suggests that difficulty understanding a question does not correlate with increased reflection time.

2.6.2 Bias

Tests were carried out to confirm whether any systematic bias was present in answering the questions. Since 50 is the middle of the scale, one would expect responses to have

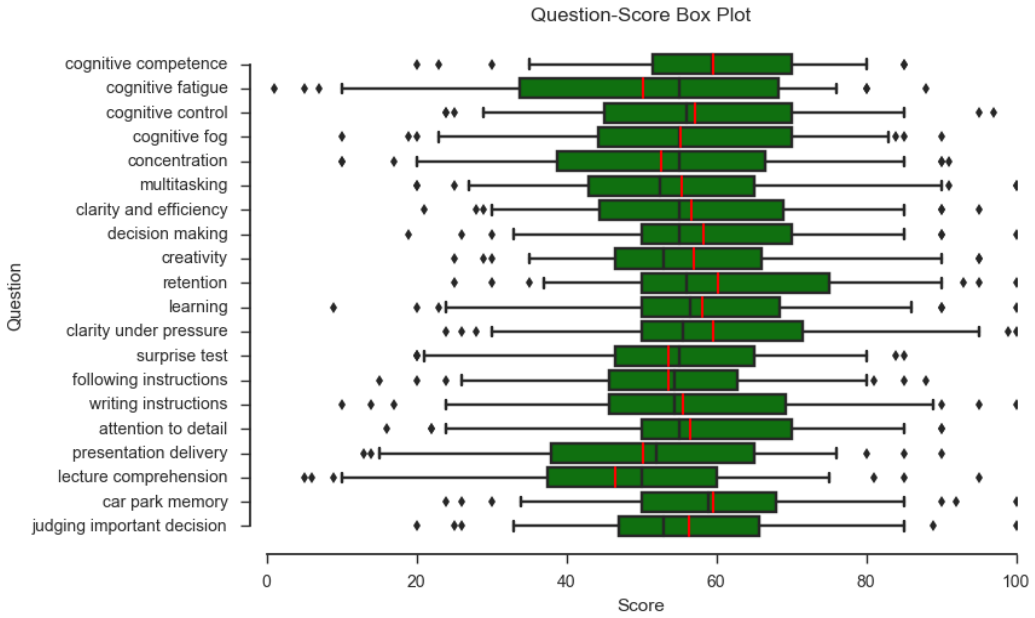


Figure 2.5: Question-score box plot. Box and whisker plot displaying inter-quartile ranges (green boxes), 5^{th}–95^{th} percentile ranges (whiskers), medians (vertical black line) and means (vertical red line) for each of the 20 questions' response times. Outliers are shown as a diamond.

temporal means around that value. If mean response values differ significantly from 50, this would confirm an offset that could in principle be down to a systematic metacognitive bias or else a questionnaire design issue. All individual responses were averaged to give 12 independent readings for each question (one value for each individual). These averages can be interpreted as mean individual responses and are mapped out in Figure 2.5.

For each question, a one-sample t-test was carried out to confirm whether the 12 individual mean responses differed significantly from the unbiased mean. Of the 20 questions, 9 had a p-value <0.05, indicating a significant difference from the expected score of 50. Interestingly, all of the significant biases were positive. Effort should be made to reduce question bias. One suggestion would be to eliminate dramatic words

such as ‘surprise’ or ‘emergency’ which may lead respondents to imagine off the norm scenarios, and therefore report abnormal behaviour. It is sensible to investigate this effect again in the larger study whose sample size will make statistical hypotheses easier to confirm. Aside from questionnaire bias, the effect could be due to a natural cognitive tendency to shift the average point towards the positive end of the scale. The absence of negative symptoms could also be predisposing respondents to overcompensate in the positive direction (Lewicka et al., 1992; Baumeister et al., 2001; Rozin & Royzman, 2001).

2.6.3 Sensitivity

To explore whether the questionnaire detected any changes over the test period, responses were analysed in two orthogonal ways. In both cases, the standard deviation of each participant's five responses was calculated as a first step. These values will be referred to as participant-question variabilities.

For question response variability, all participant-question variabilities of each question (i.e., 12 values for each question) were aggregated to create a box-and-whiskers plot. This is shown in Figure 2.6. This plot shows that questions had fairly similar mean standard deviations in the range 10-17. This is a satisfactory value that when combined with question mean responses is neither too small, nor so large as to introduce boundary effects. There were no zero participant-question variabilities. The highest participant-question variability was about 40 in the learning question.

For the participant response variability, all participant-question variabilities of each individual (i.e., 20 values for each participant) were aggregated to create a second box-and-whiskers plot shown in Figure 2.7.

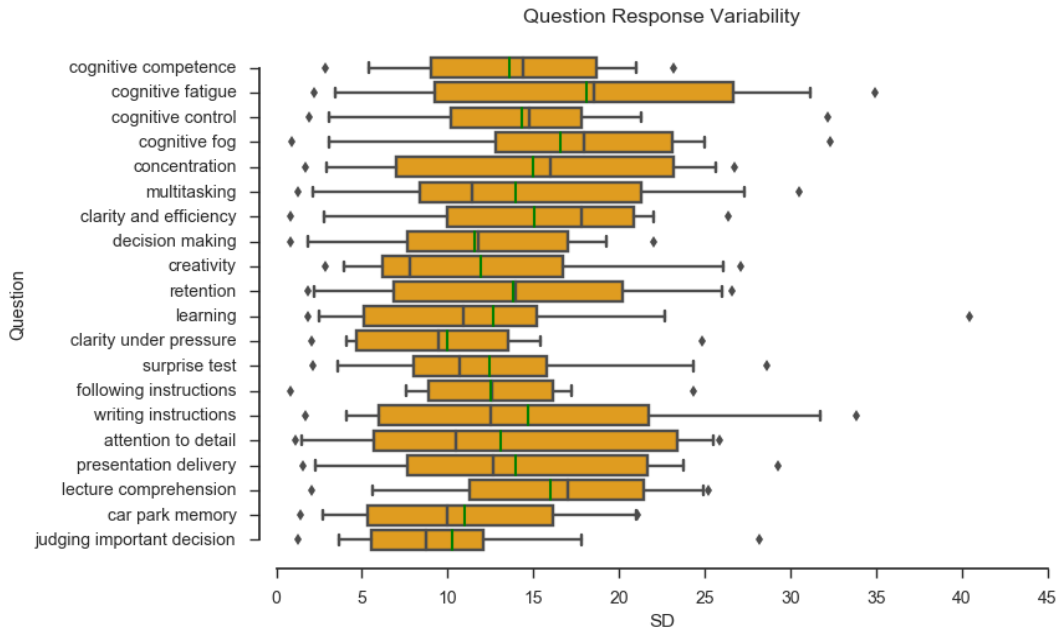


Figure 2.6: Question response variability. Box and whisker plot displaying inter-quartile ranges (orange boxes), 5^{th}–95^{th} percentile ranges (whiskers), medians (vertical black line) and means (vertical green line) for each of the 20 questions' response times. Outliers are shown as a diamond.

The Participant Response Variability plot shows greater variation between individuals. Two individuals had a mean response variability smaller than 5. This can be interpreted as either a cautiousness or conservativeness in evaluating changes, or that those individuals' changes were indeed small. The other participants had mean response variabilities ranging up to 25. Again, when combined with mean responses, the variabilities are neither too small nor too large. Overall, the data in both graphs clearly indicates that the VAS scale used was sensitive enough to detect changes in reported metacognitive abilities, but that some individuals were likely to report greater variation than others.

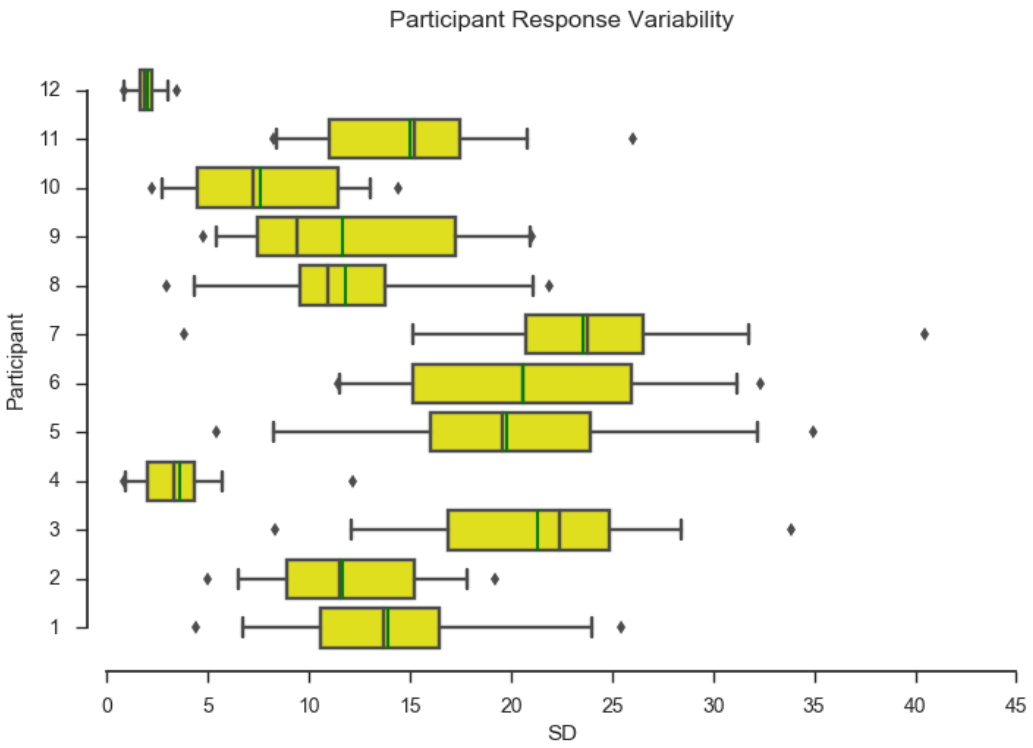


Figure 2.7: Participant response variability. Box and whisker plot displaying inter-quartile ranges (yellow boxes), 5th–95th percentile ranges (whiskers), medians (vertical black line) and means (vertical green line) for each of the 20 questions' scores. Outliers are shown as a diamond.

2.6.4 Correlations

Figure 2.8 is a heatmap showing correlations between question items. Deep red and blue indicate stronger and weaker correlations respectively. Correlations are classified as strong if they're above 0.7, moderate between 0.5 and 0.7 and weak between 0.3 and 0.5.

As expected, items in the same domain tend to be strongly correlated. Examples include *Following Instructions* and *Writing Instructions* in the Language subset, and *Multitasking* and *Concentration* for the Attention subset. Items across domains

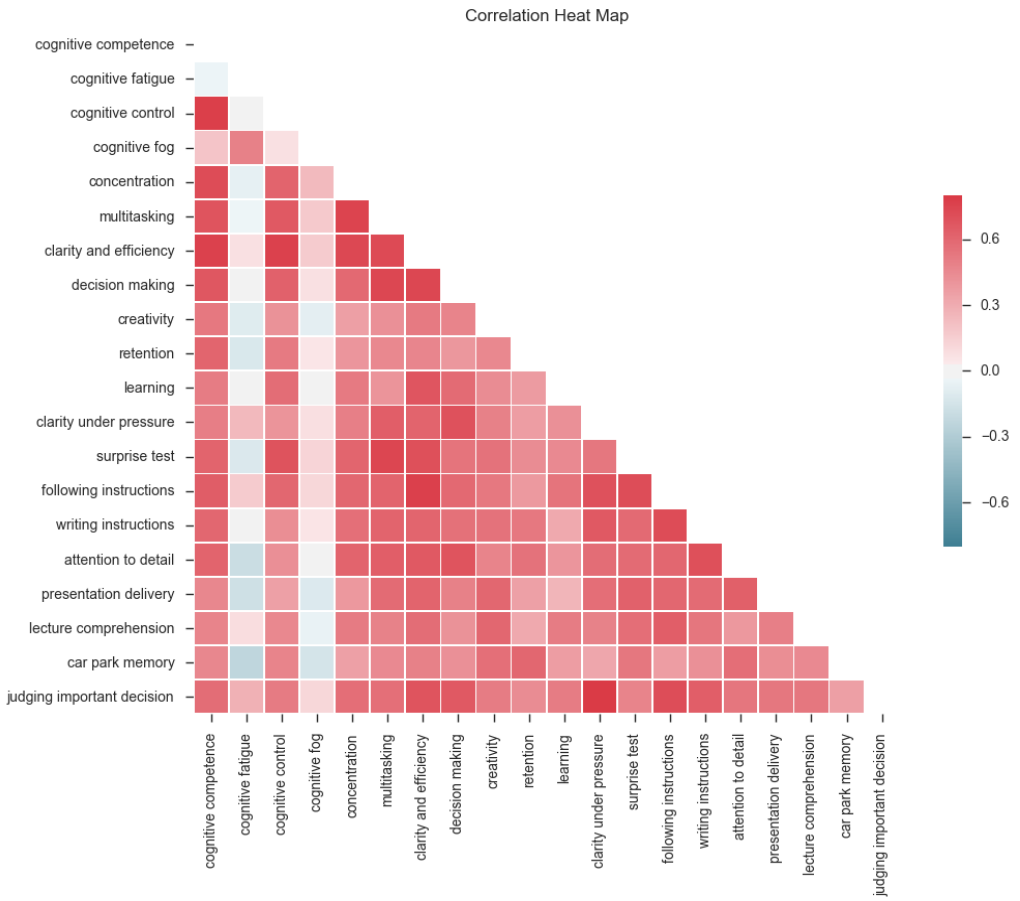


Figure 2.8: Correlation heat map for all individual DOME items.

appear to demonstrate a range of correlations. Unsurprising, there is a strong relationship between the questions asking about general cognition (*Cognitive Competence* and *Cognitive Control*) and all other items.

Interestingly, *Cognitive Fatigue* and *Cognitive Fog* are moderately correlated amongst themselves but show negligible correlation with all other items. This relationship is surprising since the direction change for the two questions has already been accounted for. Based on the feedback from the focus group, where multiple people stated that they did not immediately realise that there was a direction change, the result may suggest that negative correlations are due to scoring errors. Correlations were

re-run with the data scores for the first day of each participant removed, and improvements were visible, though a weak correlation between these two measures and the rest of the items still remained. Even though they did not report it during the focus groups, participants might not have understood what was meant by cognitive fog and fatigue. Finally, a weaker correlation can be explained by the fact that having cognitive fog or fatigue will likely have an effect on other aspects of cognition, but deficits in memory, attention, and other domains may not necessarily be caused by or be concurrent with fog and fatigue. Moving forward, participants should be given a run-through of all questions, be given an extra day to familiarise themselves with the questionnaire, and the direction change should be properly explained.

2.6.5 Construct validity

2.6.5.1 Overview

Validity was confirmed via two measures. Firstly, a Pearson's correlation test was carried out across the five DOME subscales. Values of r ranging from .50 to .75 represented a moderate to good relationship and those ranging from .75 and above an excellent relationship (Portney and Watkins, 2015). Secondly, within-domain unidimensionality was tested using exploratory factor analyses (EFA). EFAs were conducted on the five subscales, using the principal components extraction model with Varimax rotation. EFA could not be conducted on the whole scale since the requirement of a case-to-variable ratio of less than 5:1 could not be met. While a larger sample size would have been preferred, for practical reasons, this study could only be carried out with 12 participants resulting in a 3:1 (60:20) case-to variable ratio.

Sample size was not an issue when analysing subscales (minimum 12:1 case-to variable ratio). Other assumptions: KMO test being higher than .70, Bartlett's test of

sphericity being significantly different to zero and the Determinant being greater than 0.00001 were also met for each sub-scale. For the scale to have construct validity, the EFA should generate a one-factor solution for each sub-scale indicating the expected unidimensionality factor structure. According to Kaiser's criterion, factors with eigenvalues above 1 were retained. To determine the substantial loading of an item on a subscale, a criterion of .50 was established based on Munro's suggestion (2004).

2.6.5.2 DOME internal subscale structure

The DOME domains are presumed to tap into different aspects of metacognitive ability and are expected to have positive correlations. The means of the five DOME subscales were significantly correlated with each other (Table 2.3). Out of the 10 Pearson correlation coefficients, there were four moderate to good coefficients (r=.56, $p \leq .001$ to r=.69, $p \leq .001$) and six significant high correlations (r=.71, $p \leq .001$ to r=.81, $p \leq .001$). These correlations indicate that all subscales measure metacognition, providing satisfactory levels and evidence of construct validity.

The fact that the correlations range from moderate to high (.57 to .81), while not being very high (above .90) suggests that each subscale represents a distinct dimension of metacognitive ability and that for DOME to maintain its effectiveness, all five subscales should be retained.

2.6.5.3 DOME within-subscales unidimensionality structure

EFA showed unidimensionality (one-factor solution) for four of the five DOME subscales; Memory, Attention, Decisions and Language. All solutions had eigenvalues higher than 1. All subscales loaded above the established cut-off point (.50), ranging from .69 to .91 (Table 2.4). The relatively high loadings of these items provide ev-

Table 2.3: Pearson Correlations between DOME subscales

		GCA	Memory	Attention	Decisions	Language
GCA	Pearson Correlation	1	.575**	.575**	.689**	.581**
	Sig. (2-tailed)		$< .001$	$< .001$	$< .001$	$< .001$
Memory	Pearson Correlation	.575**	1	.737**	.711**	.714**
	Sig. (2-tailed)	$< .001$		$< .001$	$< .001$	$< .001$
Attention	Pearson Correlation	.575**	.737**	1	.806**	.752**
	Sig. (2-tailed)	$< .001$	$< .001$		$< .001$	$< .001$
Decisions	Pearson Correlation	.689**	.806**	.711**	1	.800**
	Sig. (2-tailed)	$< .001$	$< .001$	$< .001$		$< .001$
Language	Pearson Correlation	.581**	.714**	.752**	.800**	1
	Sig. (2-tailed)	$< .001$	$< .001$	$< .001$	$< .001$	

**Correlation is significant at the 0.01 level (2-tailed). N=60

idence of strong correlations between them and the meaningful contribution of each item within the four subscales.

The exploratory factor analysis (EFA) resulted in a two-factor solution, indicating multidimensionality for the GCA domain. Based on the correlation heatmap, this was to be expected. Loading values suggest grouping *Cognitive Fog* and *Cognitive Fatigue* in a separate group to the other items. This data suggests that the current dimension GCA does not demonstrate construct validity.

2.6.6 Internal consistency reliability

2.6.6.1 Overview

Reliability was established by computing Cronbach's alpha for the whole test and then for the items within each of the five subscales (GCA, Memory, Attention, Decisions,

Table 2.4: Factor loadings, eigenvalues, communalities and % of explained variance

Domain/Items	**Loading**	**Initial Eigenvalue**	**Communalities**	**% Variance Explained**
Memory				
10. Retention	.791	2.40	.626	59.87
11. Learning	.685		.469	
13. Surprise test	.788		.621	
19. Car park memory	.825		.680	
Attention				
5. Concentration	.896	2.34	.803	77.98
6. Multitasking	.907		.822	
16. Attention to detail	.845		.714	
Decision Making				
7. Clarity and efficiency	.859	3.11	.738	77.85
8. Decision making	.880		.774	
12. Clarity under pressure	.889		.791	
20. Judging important decision	.901		.811	
Language				
14. Following Instructions	.894	2.80	.799	70.03
15. Writing Instructions	.856		.732	
17. Presentation Delivery	.799		.648	
18. Lecture Comprehension	.795		.632	

and Language). The traditional Cronbach's alpha cut-off point of 0.7 (McGraw-Hill, 1978) was held as the indicator for a satisfactory reliability coefficient.

2.6.6.2 Entire scale reliability

Cronbach's alpha for the entire scale was .94, indicating a high level of inter-item reliability. The only item lowering the score was *Cognitive Fatigue*. By deleting this item, the alpha could be raised to .95. While it is expected that multiple indictors of the same construct should be intercorrelated, a very high reliability is not necessarily desirable, since this suggests that some questionnaire items may be redundant (Streiner, 2003). Ideally, items are related while contributing some unique information as well. In this case, a high Cronbach's alpha is not worrying for three main reasons. Firstly, DOME was designed with the help of experts in cognition who created a scale which taps into all aspects of the construct, suggesting a high content validity. Secondly, certain spe-

cific questions were intentionally designed to measure the exact same thing (e.g., *"Rate your ability to think clearly and efficiently"* and *"If there was an emergency today, my ability to think clearly would be..."*). This was done to confirm whether internal consistency would remain high irrespective of whether questions were asked directly or presented in a hypothetical format. Third and finally, no participant in the focus group felt that any of the questions were repetitive.

Since the coefficient alpha is affected by test length and is subject to inflation, it is suggested that when the alpha level is high, Cronbach's is calculated for each of the subscales rather than just for the entire test (McGraw-Hill, 1978; Tavakol & Dennick, 2011). This was done in the next section.

2.6.6.3 Within subscales reliability

Different subscales evidenced different levels of internal reliability (Table 2.5). A poor level of internal reliability was found for the subscale GCA, suggesting further that items are not measuring the same construct. This is not surprising since this subscale acts as a general housing for all measures which do not fall neatly in the other categories. Further, as indicated throughout the report, participants made clear errors when reporting on *Cognitive Fatigue* rendering data for this item unreliable. Removing this measure would, in fact, result in an alpha level increase to a .63.

Table 2.5: DOME subscales Cronbach's Alpha (N=60)

Domain	**Cronbach's Alpha**	**N. of items**	**Reliability Rating**
Decisions	.902	4	Excellent
Attention	.858	3	Good
Language	.850	4	Good
Memory	.772	4	Acceptable
GCA	.564	5	Poor

Moving forward, it is suggested that all subscales have something unique to offer and should be kept in the DOME questionnaire. Since GCA showed a lack of item homogeneity, this subscale could potentially be split into two. The very high whole scale alpha reading (.94) suggests that the removal of similar questions within subscales can be afforded as internal consistency reliability would still remain high. This would result in a more concise tool with less redundancies.

2.7 Conclusion

DOME was created to satisfy a gap in the field for a questionnaire which can measure daily self-assessed cognitive ability i.e., metacognition. The outcome of this pilot study suggests a valid and reliable tool, sensitive to metacognitive change. Recommended modifications based on verbal feedback and exploratory data analysis will be discussed for each part of the developmental process. The finalized, fine-tuned DOME is presented in section 2.7.2.

2.7.1 Suggested changes

2.7.1.1 Dimensionality of construct

All five dimensions should be retained since subscales correlated strongly, demonstrating construct validity while evidencing distinct contributions to measuring metacognitive ability.

2.7.1.2 Questionnaire format

To encourage a continuous response rate, the online questionnaire should be accessible on mobile devices. Further, users should have the option of filling it in the following day, if the entry of that day was not completed.

For higher quality data, a designated space should be available where users can note down external events which may influence responses. This could provide some perspective to the researcher analysing the data.

A case can be made for discarding the hypothetical question format and just using direct questions. Almost two thirds of the group struggled with one or more of the hypothetical questions and felt they were too open to interpretation. This was not reflected in response times, suggesting possible decrease in response quality, since participants did not spend more time reflecting on the questions.

There is strong evidence suggesting that questions should be randomized. Participants stated getting used to the questions and spending less time thinking about responses. This is reflected in the response times which decreased as one progressed through the questionnaire. If habituation occurred for a five-day pilot study it will definitely be an issue for one that will commence over a two-month period.

2.7.1.3 Item format

Results clearly show that the VAS scale was sensitive enough to detect changes in reported metacognitive abilities. The majority of users however, would have preferred a scale they were more familiar with (e.g., Likert), stating that they were unsure where to place their evaluations on the scale. This perceived weakness however, is what allows the continuous scale to have less end-aversion biases and halo effects compared to categorical measures. It is suggested that the VAS is kept with more training provided to users. Test takers need not overthink their responses as long as they are consistent throughout the questionnaire.

2.7.1.4 Item development

The question items that were flagged were mainly hypothetical questions considered too specific and difficult to relate to. Not all subjects owned a car for example, and not all were students sitting in on a lecture. Questions should be more direct and general,

particularly because questions deemed ambiguous did not demonstrate an increase in response times. This suggests that having trouble understanding a question does not translate to spending more time reflecting on it. Questions dealing with specific scenarios further create the issue of ceiling/floor effects. Participants felt they would always remember where they had parked their car, irrespective of how bad their cognition might be on the day.

Particular terminology like 'emergency', which users felt was not reflective of how they would act on a day-to-day basis should be replaced. Possible suggestions include 'under high stress'.

There is strong evidence for the question items in the GCA dimension to be reviewed. The sub-scale had poor internal reliability and did not show unidimensionality suggesting low construct validity. Scores improved when the items looking at cognitive fog and cognitive fatigue were removed. These two items, which were reverse scored, correlated amongst themselves but did not relate to any other item. One possible reason was scoring errors due to the direction change, though results improved only slightly when these errors were accounted for. It is therefore suggested these two questions be removed and that all final items be in the same direction.

2.7.1.5 Questionnaire length

On average it took participants almost five minutes to complete DOME, a time most participants reported being comfortable with. The 20 DOME items should be reduced to accommodate the extended version which will also include physical and affective symptoms. The very high reliability reading suggests that the removal of similar questions within subscales can be afforded as internal consistency reliability would still remain high.

2.7.1.6 Practice changes

A number of practice changes for the person conducting the questionnaire briefings were suggested. A summary of all suggested changes is presented in Table 2.6. To maintain high response rates participants should be encouraged to set-up daily reminders on their phone during the meeting. This is more effective than daily monitoring by the administrator since it places completion responsibility in the hands of the user and ensures that subjects that complete the study are consistent participants.

To ensure response quality, the time of the day when DOME is filled in should be specified as the end of the working day. This way there is less of a chance of tiredness biasing results and also solves the problem of people being out in the evening, particularly on weekends.

To alleviate concerns of question misinterpretation the administrator can go through a sample of questions with the users. The administrator can also explain that while the questions can be interpreted in multiple ways, the participants should not unduly worry as long as their interpretation remains consistent. It should also be highlighted that while symptoms are likely to correlate, participants can experience two seemingly opposite symptoms at the same time e.g., having cognitive fog but feeling creative.

For high quality data the results of the first day should be discarded to give the participants time to get used to the tool.

For recruitment, emphasis should be made on the relatability aspect of the menstrual cycle; an experience exclusive to females which almost all women go through. This shared phenomena, creates a bond and a sense of participation responsibility.

Table 2.6: Summary of suggested changes to DOME following testing

Questionnaire Format
• Removal of the hypothetical question format. All questions to become direct
• Accessibility of tool on mobile devices
• Randomization of question items
• A space for users to comment on external events
Item Development
• Questions to be more general to ensure relatability
• Change to terminology which suggests out of the norm situations
• Removal of items *Cognitive Fog* and *Cognitive Fatigue* in GCA dimension
• All questions to have the same direction
Questionnaire Length
• Reduction of repetitive questionnaire items to accommodate extended version
Practice Change Guidelines
• Encourage participants to put a daily reminder on their phone
• Suggest that DOME is filled in at the end of the working day
• Discard first day data to give users time to get used to the tool
• Explain that participants should not be overly concerned about question interpretation as long as it remains consistent throughout the study
• Explain that one can experience conflicting symptoms
• For recruitment emphasise the relatability aspect of the menstrual cycle

2.7.2 Final version of DOME

The final version of DOME after the implementation of suggested changes is shown below (Table 2.7).

Table 2.7: Final version of DOME after piloting

In the last 24 hours, how would you rate these measures relative to how you usually feel. Your:
1. Ability to think clearly and efficiently Worse than usual — Same as usual — Better than usual
2. Creativity level…
3. Cognitive control (i.e., sense of feeling in control versus being mentally overwhelmed) …
4. Ability to pay attention to what is happening in the present …
5. Ability to concentrate and avoid distractions…
6. Ability to multitask…
7. Ability to be decisive…
8. Ability to make good decisions…
9. Confidence in your decisions…
10. Ability to remember information you know (e.g., a phone number, fact you had learnt) …
11. Ability to learn new information…
12. Ability to remember where you put an item (e.g., phone, bag, remote) …
13. Ability to write clear instructions for others to follow…
14. Ability to comprehend a long explanation or talk…
15. Ability to give a long explanation or talk …

2.8 Chapter summary

The process behind the creation of DOME, a questionnaire that evaluates daily metacognition, was examined in this chapter. Justification for designing a new tool was provided followed by a step-by-step explanation of its development. Results from preliminary pilot testing showed its validity and reliability as a metacognitive measure. Recommended changes were implemented.

The study for which DOME was created examines daily menstrual cycle metacognitive, physical and affective symptom changes in 120 women across two cycles. Chapter 3 documents the study methodology and analysis techniques. Study results and conclusions are presented in subsequent chapters.

Chapter 3

Methodology

3.1 Chapter Outline

There is a gap in the research pertaining to menstrual metacognitive patterns, both in terms of possible variation between individuals and co-presence with other menstrual symptoms. This study looks at the influence of the menstrual cycle on metacognitive function, through the administration of DOME (Daily Online Metacognitive Evaluation) questionnaire to a large group of women over an extended period of two menstrual cycles.

In this chapter a detailed description of the study methodology is provided. Section 3.2 highlights the measure used for the assessment of metacognition. Section 3.3. goes over the data collection procedure. Special attention is given to the optimisation of recruitment and the maintenance of subject engagement over time, owing to the longitudinal nature of the study. Section 3.4 provides an explanation of the data analysis procedure. How the data was cleaned and processed is explained in detail, on account of the large data set and the fact that most of the analysis was carried out on values derived from the original scores (section 3.4.1). Advanced analysis techniques were used, and these are comprehensively explained in Section 3.4.2. To prevent dredging, analysis conditions and parameters were established before exploration. These are presented in Section 3.4.3.

3.2 Measure

3.2.1 DOME questionnaire

The Daily Online Metacognitive Evaluation (DOME) questionnaire created for this study assesses metacognition through the self-assessment of cognitive ability. Respondents rate their perceived cognition relative to how they usually feel and results are recorded on a visual analogue scale with the anchors 'Worse than usual' and 'Better than usual'. The tool assesses five domains – general cognitive ability (GCA), attention, language, memory, and decision making – that are essentially cognitive in nature but which become metacognitive phenomena once the process of self-evaluation is applied. Preliminary testing indicated correlated but distinct domains, with all three question items within the five subscales providing meaningful contribution. Scale sensitivity to the detection of changes in self-reported cognitive abilities was also evidenced. Sample questions are presented in Table 3.1.

Table 3.1: Sample questions used to evaluate self-assessed cognition

Construct	**Domain**	**Question**
Cognition	Memory	. . . ability to remember where you put an item (e.g., phone, bag, remote)
	Attention	. . . ability to concentrate and avoid distractions
	Language	. . . ability to comprehend a long explanation or talk
	Decisions	. . . confidence in your decisions
	GCA	. . . cognitive control (i.e., sense of feeling in control versus being mentally overwhelmed)

Each question was preceded with "In the last 24 hours, how would you rate these measures relative to how you usually feel. Your . . . "

3.2.2 Extended DOME questionnaire

The extended DOME questionnaire measures physical and affective symptoms in addition to self-assessed cognitive ability. The physical and affective sub-scale questions were taken from validated studies which had tested for sensitivity to menstrual-cycle related changes (Kiesner et al., 2016; Kiesner, 2011; Kiesner & Pastore, 2010). For this section of the tool, the original Likert scale response format of "not at all", "a little", "a moderate amount", "quite a bit", "very much" is used, to allow comparison with previous results. Items used are presented in Table 3.2 and all questions listed in this study refer to the past 24 hours. Each physical symptom is measured by one question while affective symptom scores are the averaged result of two questions as carried out in Kiesner (2016).

Table 3.2: Questions used to measure physical and affective symptoms

Construct	**Domain**	**Question**
Physical	Back/joint pain	... did you have back or joint pain?
	Breast	... did you experience breast swelling or pain?
	Gastrointestinal	...did you have gastrointestinal/stomach problems or pain?
	Skin	... did you experience an increase in pimples or oily skin?
	Cramps	...did you have lower abdominal cramps or pain?
	Headache	... did you have headache pain?
Affective	Mood Swings	... did you have mood swings? ... did you feel irritable?
	Anxiety	...did you feel anxious? ...did you feel tense?
	Depressed	...did you feel depressed? ...did you feel sad?

Each question was preceded with "In the last 24 hours ..."

3.3 Procedure

3.3.1 Recruitment

For this study, a target goal of 120 women was set, based on the size used in same area studies (Kiesner & Pastor, 2010; Kienser et al., 2016). The success of the research centred around keeping a large group of participants engaged for the extended duration of the study. This influenced a number of recruitment decisions taken, particularly whom to target and when to carry out the study. University students were chosen on the premise that they could potentially have more time for the study. Nursing, Psychology, Communication and Education departments were targeted because of their large, high female-to-male ratio classes, and the likelihood that these students would show interest in the study subject matter. Recruitment started in February, following mid-year exams, and had to be completed by early April so as to allow a two-month window for students to finish the study by early June. Effort was made to avoid the end-of-year examination period so that stress would not influence results. In total nine recruitment sessions were conducted: three within Psychology, and two each within Nursing, Education and Communication.

Approximately fifteen minutes were taken from either the start or end of lectures. If taken at the end of the lecture, students were informed that the lecture had not yet finished so that they would not walk out. An overview of the study was given explaining the limited conflicting research on cognition and the cycle. Volunteers needed to be female, not pregnant, and not past their menopause. The exclusions for women who were either on hormonal contraceptives, hormonal therapy or had irregular periods (defined as varying by more than one week) were also made clear. Students with psychological or medical conditions could take part and emphasis was made on the im-

portance of having women with and without menstrual difficulties participating. Other points included the personal nature of the questions and the degree of participation required i.e., daily questionnaires for two menstrual cycles. Students were told that participation was anonymous and voluntary, and that compensation would be a certificate of participation.

Forms were handed out to every female in the room and unused sheets were collected. Students who wanted to participate wrote their contact details and the expected date of their next period. Many resorted to period mobile apps to check the exact date. Students who did not want to take part, or were unable to, were told to specify the reason why from the options on the form and to leave out their personal details. All forms were then collected.

3.3.2 Individual meetings

After each class recruitment, participation forms were prioritised based on expected period due date. Students were contacted by phone and a one-to-one meeting was set up. A script was prepared so that meetings would have a consistent formula. During the meeting an explanation and demonstration of the online data collection procedure was provided and any questions were answered. Subjects were given time to read through all the DOME questions to ensure everything was clear. Participants then filled in a consent form as well as a menstrual cycle questionnaire asking for: age of menarche, period duration, cycle length and regularity, menstrual flow, menstrual cycle disorders and other psychological/physical conditions. Each person's expected next period date was established and participants were asked to set a daily reminder on their phone from the estimated date to approximately two months after. Participants were then sent an email with their assigned code and the link to the study.

3.3.3 Taking the DOME online

The link to the online questionnaire was shortened for ease of recall. Users needed to access the questionnaire on a daily basis and it was made available on both desktop and mobile for convenience. Participants were identified through their code with the time and date of completion automatically recorded and saved. Participants were required to complete the questionnaire every evening regardless of whether they were menstruating or not. The final day of their second cycle marked the end of the study. The platform automatically sent a notification if the questionnaire was not fully completed.

The full questionnaire was divided into four sections: general, self-assessed cognition, physical and emotional (Appendix A). Apart from the first page, all pages and the questions within them were randomized for each questionnaire. To start with, participants had to state whether they were starting the study on the first or second day of menstruation. If participants could not fill in the questionnaire on the day or if they accidentally missed a day, the online questionnaire allowed participants to complete one questionnaire for the previous day and one for the present day. To control for this, the second question on each questionnaire was whether it was in relation to "yesterday" or "today." Participants were then asked if they had menstrual bleeding or spotting, the number of times they changed their pad/tampon and to list any major events that happened on the day and that may influence their responses. For the Likert scale questions participants had to click radio buttons, while for the visual analogue scale questions participants had to drag a slider between the two anchors. If a slider was skipped participants could not move on to the next question.

3.3.4 Follow-up

After filling in their first entry, participants received an acknowledgement email confirming that their data was received. Follow-up emails were then sent after week one and week four. Two weeks before the end of the study participants were sent a final thank you email. Throughout data collection a daily log was taken, recording participant entries. Participants who did not complete the questionnaire were reminded through email to fill it in the next day. Those who did not fill in the questionnaire for three consecutive days were dropped from the study. Participants who had over nine missing entries across two menstrual cycles were also removed.

3.3.5 Ethics

All methods and procedures conformed to the ethics and data protection guidelines of the University of Malta. The University of Malta ethics research committee reviewed and approved this study.

3.4 Data

3.4.1 Data cleansing

Data was cleaned according to the following assumptions. If any submissions were incomplete these were removed from the study and if there were any double entries for the same day (participants sometimes forgot that they had already filled in the questionnaire) the first entry was assumed to be the correct one. Entries for the participants who had dropped out were removed from the study.

Participants also had to report whether they were filling in the questionnaire for that day — "today" or the day before — "yesterday". Since a large number of participants filled in the questionnaire late at night the 'deadline' for the day was taken to be 4am rather than 12am. For example, an entry marked "today" time-stamped 12.04.19 01.15am would be considered to be the entry for 11.04.19. Entries after 12am (but before 4am) marked "yesterday" and time stamped 12.04.19 would also be considered as entries for 11.04.19 unless the previous day (the 10^{th}) had been missed in which case participants would be contacted to verify which day the questionnaire was filled in for.

Time and date had to be rescaled to place all participants on the same metric since every person had a different cycle length. This was calculated as day within the cycle/number of days in cycle. The start of cycle one was mapped to 0, the end of cycle one to 1 and the end of cycle two to 2. Having the time variable represented as a proportion of the cycle length normalises the variable to allow easier comparison across individuals with different cycle lengths and/or across cycles of a different length within the same individual. For the purposes of visualisation, the time variable was further discretised by binning each cycle into twenty equal-width bins.

Participants had the option of starting on the second day of their period when they missed data entry on their first day. For these participants an extra blank entry was added and taken as the 1st day of their period. Some participants mistook mid-cycle bleeding as the first day of their period. This data was fixed so that the first day of the cycle represented the first day of consecutive period bleeding. All raw data prior to changes was saved.

3.4.2 Analysis techniques

Data analysis was carried out using two analysis techniques. Individual variation in symptom change differences was examined using multilevel modelling and the co-variation of cyclic changes across symptoms was studied through structural equation modelling. All parts of the analysis involved cosine regression.

3.4.2.1 Cosine regression

Cosine regression estimates the basic cyclicity parameters in non-linear data by first transforming the linear time variable into a cosine variable. This allows linear regression of the score variable against the cosine term. Performed on this transformed data, linear regression fits an intercept and slope as in the typical use case. With respect to the untransformed data however (i.e., with respect to the original linear time variable), the intercept is equal to the mean value of oscillation, and the slope corresponds to the amplitude of oscillation. Visually speaking, the process fits the nearest cosine shape to the original data, preserving the initial trough or crest.

The data follows overall trends that are well-captured by the cosine wave, even if we allow for slight variation in shape and the fact that the peaks are typically sharper than the troughs. This regularity in the collected data therefore justifies the use of cosine

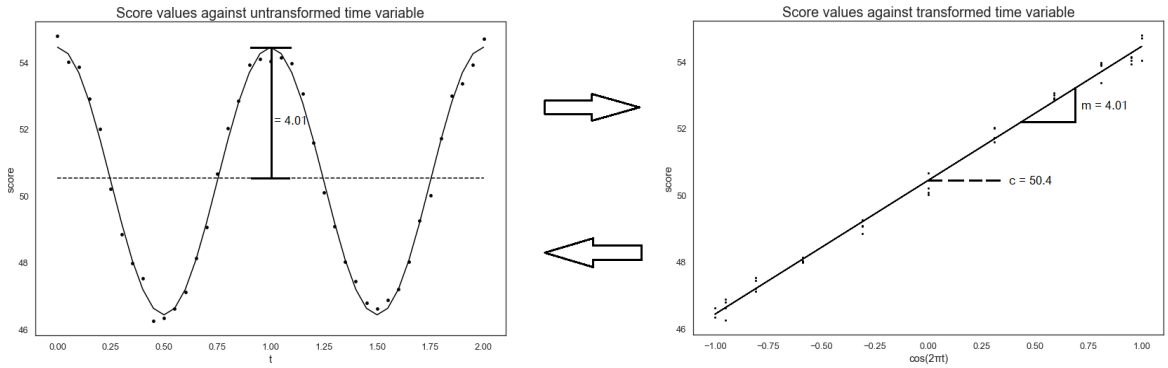


Figure 3.1: Transforming between the linear time variable and the cyclic variable cos(2πt).

regression on the score variables. It should be noted that in general, cosine regression is less successful in modelling cyclic phenomena that are time shifted with respect to cycle onset or have a complex shape. This situation, fortunately, does not arise.

3.4.2.2 Multilevel modelling

Multilevel models recognise the possibility of hierarchical relationships in data and allow residual components to be modelled at each level (Peugh, 2010). In the present case, this kind of modelling was used to confirm whether interpersonal variation is significant or not. Multilevel analyses were run on three different models, each one building on the next, to confirm whether fitting a more complex model results in a better explanation of the random variance.

The 1st level model looked at between-participants mean differences, running the analysis with the intercept as random effect.

$$Y_{ij} = B_0 + U_{0j} + r_{ij}$$

where Y_{ij} is the score of subject j at time i, B_0 is the fixed intercept (group level mean),

U_{0j} is the random intercept (individual mean difference) and r_{ij} is the within-subject residual.

The 2nd level analysis included the intercept as a random effect (random intercept) as well as the fixed effect (amplitude) of cosine.

$$Y_{ij} = (B_0 + U_{0j}) + (B_1 \times cos(2\pi t)) + r_{ij}$$

where B_1 is the fixed effect of cosine (the average amplitude across participants).

In the 3rd level the analysis was run with both the intercept and slope as random factors.

$$Y_{ij} = (B_0 + U_{0j}) + \left(B_1 + U_{1j}\right)\cos(2\pi t) + r_{ij}$$

where U_{ij} is the random effect of cosine (the individual's amplitude difference).

The proportion of variance attributed to mean differences between individuals can be captured using the ICC (Intraclass Correlation Coefficient). An ICC of 0 indicates that participants do not differ from one another in their average scores while an ICC of 1 suggests that all of the variability is between subjects (Bolger & Laurenceau, 2013). The ICC for an unconditional model is calculated by dividing the variance from the random intercepts by the total variance (random + residual) as seen in the following formula:

$$ICC = \frac{\widehat{U}_{0j}}{\widehat{U}_{0j} + \widehat{r}_{ij}}$$

where $\widehat{U}_{0j}$ is the variance of U_{0j} and $\widehat{r}_{ij}$ is the variance of r_{ij}.

The remaining ICC variance is attributed to variations between within-individual replications.

The effect size for the fixed slope model can be found by dividing the difference in residual variance between the first and second models by the residual variance of the first model. This is given by:

$$f_2 = \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2}$$

where f_2 is the effect size of the fixed slope model, and σ_1^2 and σ_2^2 are the residual variances of the first and second models respectively. The effect size quantifies the proportion of within-person variance explainable by the particular model.

The effect size for the random slope model can be found by dividing the difference in residual variance between the first and third models by the residual variance of the first model. This is given by:

$$f_3 = \frac{\sigma_1^2 - \sigma_3^2}{\sigma_1^2}$$

where f_3 is the effect size of the random slope model, and σ_1^2 and σ_3^2 are the residual variances of the first and third models, respectively. The third model will be an improvement over the second model if it can explain a significantly larger proportion of the residual variance of the first model.

Table 3.3: Fixed and random effects in each model

Type of Effects	**Model**		
	1	2	3
Random intercept (mean)	✓	✓	✓
Fixed slope (cosine amplitude)	✗	✓	✗
Random slope (cosine amplitude)	✗	✗	✓

The analysis was carried out on all cognitive abilities, and physical and affective symptoms in turn. Table 3.3 shows a summary of the fixed and random effects for each model.

3.4.2.3 Structural Equation Modelling

Co-variation across symptoms was studied through structural equation modelling (SEMs). The previously computed cosine amplitude coefficients were saved for each participant and for each variable and used in a series of nested SEMs where one model could be compared to the next. The goal was to see how *cyclical change* is best modelled and to examine the associations between symptoms at both a construct and base symptom level. These relationships are explained through the model with the best fit.

Structural equation modelling is a multivariate technique used to establish relationships among common factors; particularly latent constructs. Since these constructs are unobservable, they have to be properly defined using a measurement model prior to carrying out the SEM. A natural starting point is an EFA (exploratory factor analysis) whereby unconstrained model provides the factor structure (see chapter 2, Wang & Wang, 2019). In this case, one goes in blind, without predetermining which observed indicators load into which factors and the number of factors. With SEMs however, it is unlikely that one does not have an apriori hypothesis on what items can be linked to which constructs and on the relationship between the constructs themselves. In this case a confirmatory factor analysis (CFA) is used to test a theory-led factorial structure with specified parameters (for overview see Streiner 2006; Phakiti, 2018). CFAs fall under the SEM umbrella and differ from the latter in that they do not investigate causal links between latent constructs (Streiner, 2006). This study will stop at the measurement model and not get into structural modelling since no position will be taken on how or why the latent variables are correlated.

With SEMs it is illustrative to represent the hypothesized relations in a diagram, especially since most software accept these diagrams as input. An example of a basic SEM is given below, together with a description of the most common symbols used in CFA and SEM (See Table 3.4).

In this study, the first series of SEM runs look at physical and affective symptoms, testing the factor structure proposed in the 2016 research. The second set of SEM runs focus on the novel cognitive changes. The last set of SEMs look at cognitive changes in conjunction with physical and affective symptoms.

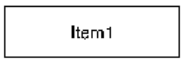
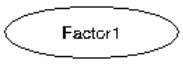
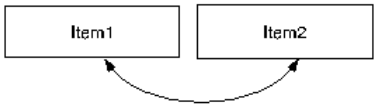
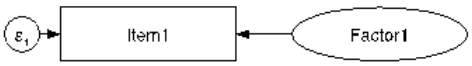
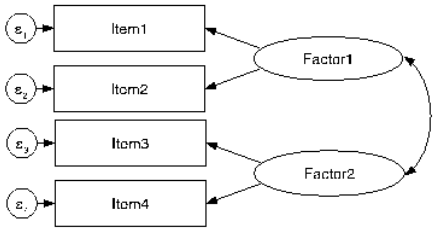
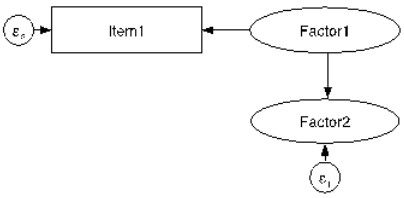
3.4.3 Assumptions, parameters and thresholds

3.4.3.1 Exploratory Factor Analysis (EFA) for SEM

This study used Factor analysis (FA) over Principal Component Analysis (PCA) since the former separates the unique variance and error of a variable from its shared variance while PCA does not discriminate between the two and assumes that the shared variance makes up all of the total variance without accounting for any unique contribution. In-depth arguments for and against, are beyond the scope of this paper but have been well covered in the following reviews, Costello & Osborne (2005), Gorsuch (1997), Howard (2015).

In cases where the data was normally distributed Maximum Likelihood was used as the factor extraction method. If the data was significantly non-normal, Principal Axis Factoring was used instead.

Table 3.4: Overview of common symbols used in CFA and SEM

Symbol	Explanation
	A rectangle indicates an observed variable
	An ellipse indicates a latent construct
	A double headed arrow between rectangles implies a non-causal relationship between observed variables. A value on a double headed arrow is a correlation coefficient
	A double headed arrow between ellipses implies a non-causal relationship between latent factors
	A unidirectional arrow implies a causal relationship from a factor to an observed variable. ε is the error associated with the measured variable
	A basic measurement model (CFA) with non-causal relationship between two correlated factors
	A basic structural model (SEM) with paths reflecting causal dependency among latent variables

The following assumptions needed to be met before the factor analysis could proceed, as reviewed in Howard (2015):

1. A case to variable ratio higher than 5:1 with above 10:1 being preferable.

2. No multicollinearity as indicated by the determinant of the correlation matrix. Calculated in SPSS a determinant greater than 0.00001 provides evidence of no multicollinearity.

3. A Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO test) higher than 0.7 with 0.5 being the absolute minimum. The KMO indicates the proportion of variance in the variables that might be caused by underlying factors.

4. A significant Bartlett's test of Sphericity. Bartlett's looks at whether variables are related enough to allow structure detection. If the correlations between variables are too small, Bartlett's is non-significant and factor analysis can't go ahead.

All variables with communalities lower than 0.2 were removed since they do not load properly onto the factor solution. The number of factors to be retained were chosen by multiple tests as recommended by Thompson (2004) and Henson and Roberts (2006). For the K-1 test, eigenvalues greater than 1 were kept. Factor retention was also tested using the scree plot where the component number indicated by the 'elbow' in the plot minus one is usually taken as the number of factors that should be retained. If the suggested number of factors varied between tests, Horn's parallel analysis was conducted to help determine results.

Varimax rotation was used as a rotation method since it produces clear results which are easy to interpret. Output results were sorted by size and coefficient values

less than .4 suppressed. In cases where coefficients were loosely correlated (majority at .6 or less) a cut-off point of .3 was used (Gaskin & Happell, 2014). Cronbach alpha was carried out on each group of variables to verify reliability. 0.7 or higher is usually cited as the acceptable cut-off point (Cronbach, 1951; Tavakol & Dennick, 2011).

3.4.3.2 Confirmatory Factor Analysis (CFA) for SEM

For CFA sample size, a basic rule of thumb is a minimum N in the range 100-150 (Anderson & Gerbing, 1984; Ding, Velicer, & Harlow, 1995; Tabachnick and Fidell, 2001) which was exceeded in this case. In other cases, sample size is often considered in light of the number of observed variables using the ratio N:q where N is the number of cases and q is the number of free parameters. These are parameters which are not fixed at a particular value and can include all elements of the SEM; observed variable variances, latent variable variances, regression paths, covariances etc. (Kenny, 2011). Whether parameters are free or fixed is usually indicated by the software being used and can simply be counted. The golden ratio is usually 10:1 case to free parameters (Jackson, 2001; Kline, 2011) with the recommended minimum being 5:1 (Bentler & Chou, 1987).

Multivariate normality was checked to decide what estimation procedure to best use for the CFAs. The Doornik-Hansen test was chosen since it is more powerful than other normality tests like the Shapiro-Wilk when dealing with multivariate distributions (Doornik & Hansen, 2008). Maximum Likelihood is usually the default estimation method but in the case of the multivariate normality assumption failing, a more robust version of ML is needed. The Satorra-Bentler rescaled chisquare statistic (Satorra and Bentler 1994), is ideal since it compensates for the nonnormality of variables (Stern, Katz-Navon, Naveh, 2008). Reliability was verified using Cronbach alpha for each group of variables loading onto a latent factor, for each model.

To assess model fit, multiple fit indices are reported and not just one. Klein (2005), suggests the use of the following tests apart from the χ^2 test: the standardized root mean square residual (SRMR; Jöreskog & Sörbom, 1981), the comparative fit index (CFI; Bentler, 1990) and the root mean-square error of approximation (RMSEA; Steiger & Lind, 1980). The fit indices chosen have different properties and measure goodness-of-fit in different ways as recommended by Brown (2006). The CFI is a comparative fit index, while the SRMR and RMSEA are indices for the model's absolute fit. The SRMR is a residual based fit index and the RMSEA is parsimony-adjusted. The cut-off points for fit indices are highly contentious and vary between researchers. The ones used here are provided by Browne & Cudeck (1993) and Hu & Bentler (1999) and are considered the most canon and most adopted (Jackson, 2009). χ^2 should fail to reject the null hypothesis with values closer to zero indicating a better fit. CFI $>.95$, RMSEA and SRMS $<.08$ is acceptable but $<.06$ provides a better fit.

Since there is the temptation of using post hoc modifications to find a fitting model, this study only used modification indices when necessary and when there was theoretical validation to do so. Following MacCallum (1995), no more than two MI recommendations were carried out on a model.

3.4.4 Software

All CFAs and SEMs were carried out using Stata Rel. 16 (Stata Corp, 2019). All parts of the EFA used SPSS V.26 (IBM Corp, 2019), except for Horn's Parallel Analysis which was done in Stata. Cronbach's alpha was run in SPSS and testing for multivariate normality in Stata. All multilevel analyses were conducted using the Stata GLLAMM (Generalized Linear Latent and Mixed Models) program with adaptive quadrature.

3.5 Chapter summary

This chapter provided an overview of the study measure and procedure used, together with an explanation of the data processing and analysis techniques. In Chapter 4, the results look into individual differences and co-variations in physical and affective symptom change. Chapter 5, on the other hand, focuses on cognitive results.

Chapter 4

Results I

Physical and Affective Symptom Changes

4.1 Chapter Outline

Kiesner et al. (2016) set out to demonstrate the extent of individual differences in women in cyclical symptom change as well as the covariation in symptom cyclicity, with focus on physical, affective symptoms and attributional style. This study replicates Kiesner et al. (2016) by using the same measures, procedures and analyses as the original work. It also extends beyond, by looking at the Maltese context, testing whether results can be generalized to other cultures. Replication study results for the physical and affective symptoms are presented in this chapter. Section 4.2 looks at participant response rates and menstrual cycle data, followed by an exploratory analysis in section 4.3. Results for multilevel modelling and structural equation modelling are presented in section 4.4. and 4.5 respectively. All reporting procedures are based on the guidelines provided by well-established works; Jackson et al. (2009) for CFA, Costello and Osborne (2005) for EFA and McDonald et al. (2002) for SEM. The following convention will be applied throughout: all symptom variables are italicised and all latent constructs are capitalised.

4.2 Demographics

4.2.1 Participants

In total 363 students were contacted to participate, 170 (47%) of whom signed up for the study. Of the 193 who declined, 56% cited hormonal drugs as the reason, 18% non-regular periods, 16% a lack of interest in the research and 10% gave other reasons; predominant ones being lack of time, lack of commitment, and not feeling comfortable doing the research.

Of the 170 individuals who agreed to participate, 121 (71%) participated for the full study, providing data for two cycles. 29 people dropped out before the meeting and 20 after starting the study. 141 one-on-one meetings were carried out. The average length of the two menstrual cycles was M=30 (SD=4.89) days for Cycle 1 and M=29.9 (SD=4.94) days for Cycle 2 with an average length of M=59.9 (SD=8.19) for the two consecutive cycles. The maximum number of days was 93 with a minimum of 41 for the two cycles. All participants reported a standard age of menarche ranging from 8-16 years of age. A heat map of the participants' bleeding days showed that 14% of participants experienced mid-cycle bleeding. This is comparable to the 13.3% (n=3941) reported in the highly cited paper by Rowland et al. (2002). Menstrual bleeding lasted for a median of 5 days (mode: 5 days; range: 3–9). One participant was excluded because of irregular extended bleeding activity in the second cycle which continued right up to ovulation.

Figure 4.1: Heat map of participants' bleeding days. Coloured cells are bleeding days. Lighter red cells are an artefact of the binning process. If two days fall within the same time bin, only one of which shows bleeding, the resulting colour is a lighter red. Grey cells are missing data. The time variable is presented as a ratio of each participant's cycle length. Participant numbers do not reflect participant code.

4.2.2 Response

Participants filled in a total of 7,172 complete questionnaires, averaging 59.7 questionnaires for each participant. Eleven participants started on their second day of the cycle. On average, participants had 1.8 (SD=2.3) missing entries across the two cycles. Out of all the complete questionnaires, 74% were completed on the day labelled "today" and 26% were labelled "yesterday."

4.3 Exploratory Data Analysis

A total of six physical and three affective variables were studied. Physical symptoms were measured by one question while affective scores are the averaged result of two questions. When analysing a variable that is expected to be cyclical, it is important to consider both the pattern of the variable over time and the distribution of the variable's values. Figures 4.2 and 4.3 show the variable scores averaged across individuals as a function of time for the physical and affective symptoms respectively. The shaded bands describe the confidence interval (to within one SD) of the score means. As described in section 3.4.1 the time variable t is the binned ratio of each participant's cycle length.

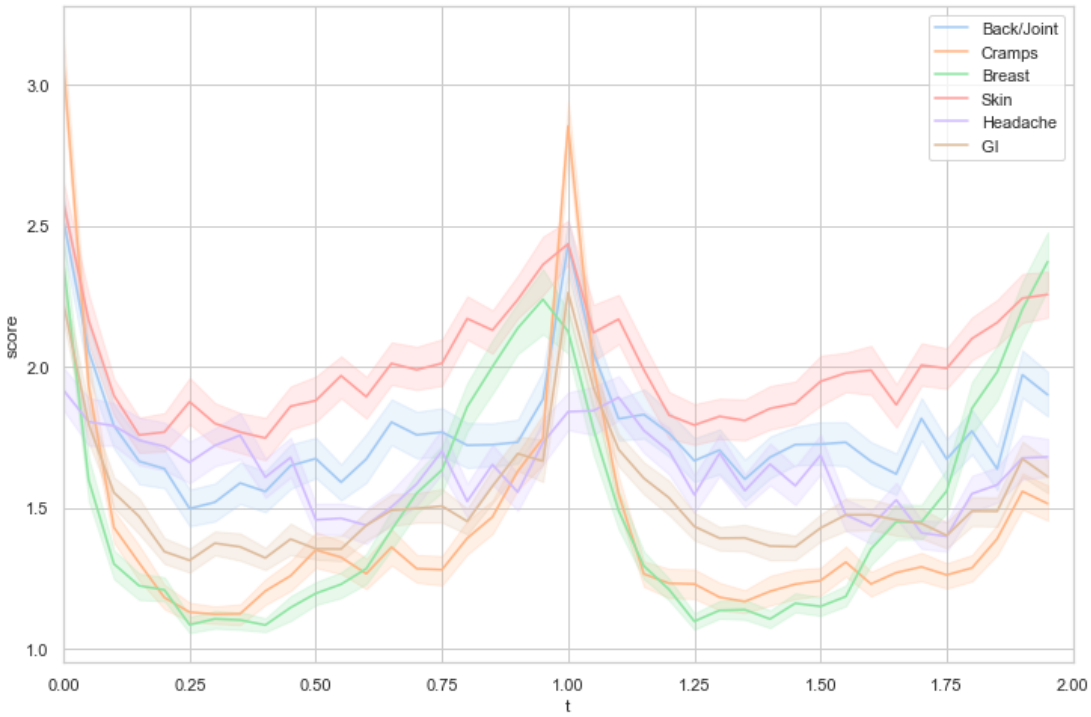


Figure 4.2: Score-time plot for physical variables

Scores for the physical variables in Figure 4.2 show clear cyclicity with relation to the cycle, encouraging further exploration. Multiple variables peak at day 1 of the cycle, with *Cramps* being the most prominent. *Breast* pain is a leading indicator, peaking just before the start of the cycle and reaching its minimum at a quarter to one third of the cycle. It is also smoother compared to other variables. *Headaches* is only mildly cyclical and relatively noisy. It is interesting to note that in the time-plot, each consecutive peak is smaller. This could suggest inflated attention to symptoms at the beginning of the study, levelling out in time.

The score-time plot for the affective symptoms is presented in Figure 4.3. As seen, cyclicity is confirmed with affective symptoms as well, even though the effect is less pronounced and noisier than for physical symptoms. This is expected since the

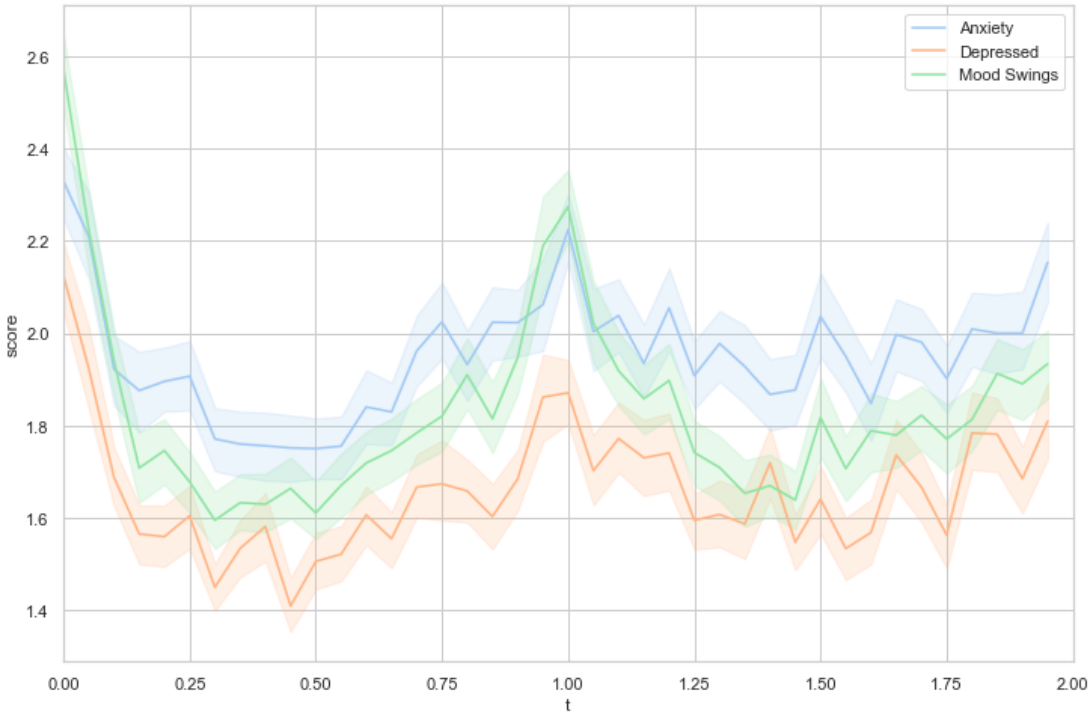


Figure 4.3: Score-time plot for affective variables

affective state is more likely to be influenced by external events. *Mood Swings* show the highest amplitude. Consecutively smaller peaks can be seen for all three symptoms as the study progresses.

Figures 4.4 and 4.5 show the box plots for the physical and affective variables respectively. The red bar indicates the mean, the grey bar the median and the green rectangles extend over the interquartile range (25th – 75th percentile) with whiskers showing the 5th and 95th percentiles. Scores in Figure 4.4 are skewed towards the lower end of the scale since most symptoms present only mildly or negligibly for most of the cycle. *Skin* shows a larger mean and spread than the other variables. In Figure 4.5 scoring is similarly skewed towards the lower end of the scale. *Depressed* shows a smaller mean and spread than *Anxiety* and *Mood Swings.*

If a variable that is expected to be cyclical shows a wide distribution of values when measured across time and subjects (box plots), but a small cyclic effect when the trajectory is averaged across subjects (time plots), it can be concluded that a significant portion of the variation is likely due to differences between subjects (or noise). This suggests that the cyclical pattern may not be as consistent across all subjects, and that individual differences may be influencing the measurements. An obvious outlier, *Skin*, has higher variance and a greater mean than the other variables (Figure 4.2). In the time plot, however, its amplitude is moderate when compared with the other variables, which could suggest high between-person variation. In contrast, *Anxiety* and *Mood Swings* have a similar distribution in the box plot, but *Mood Swings* shows more variability than *Anxiety* in the time plot, suggesting a higher portion of within-person variation in the former.

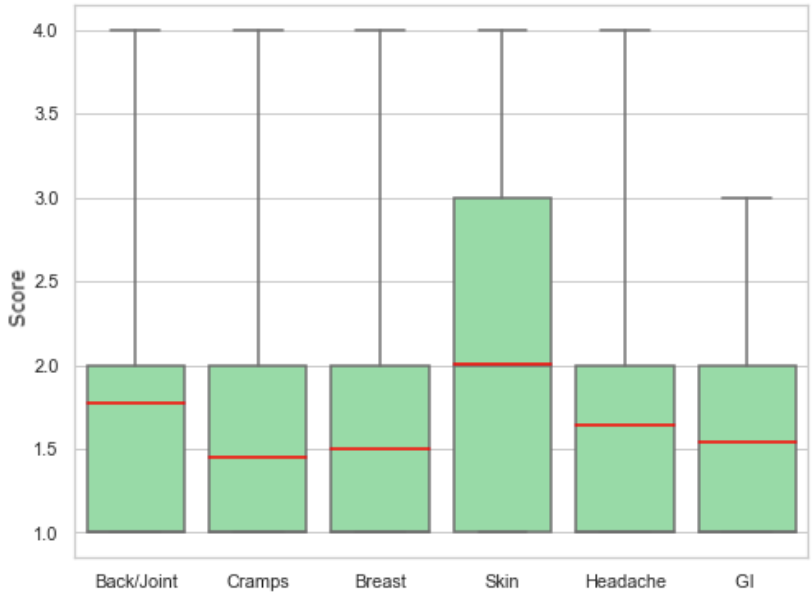


Figure 4.4: Box plot of raw scores for physical symptoms. The red bar indicates the mean, the grey bar the median, and the green rectangles extend over the interquartile range (25th – 75th percentile) with whiskers showing the 5th and 95th percentiles.

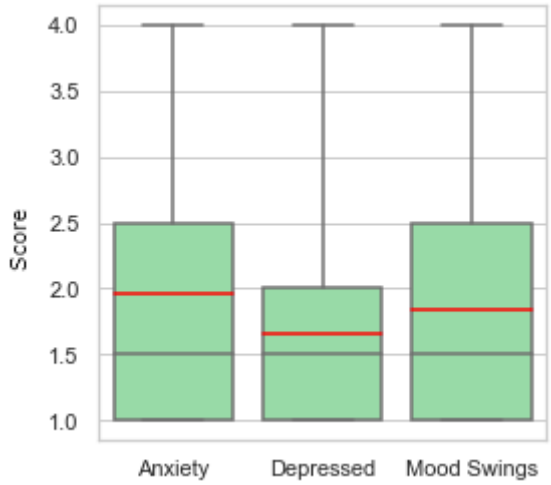


Figure 4.5: Box plot of raw scores for affective symptoms. The red bar indicates the mean, the grey bar the median, and the green rectangles extend over the interquartile range (25th – 75th percentile) with whiskers showing the 5th and 95th percentiles.

4.4 Cosine Regressions

The distribution of cosine amplitudes for physical and affective symptoms is presented in Figure 4.6.

All interquartile ranges are positive so the majority of individuals have positive cosine amplitudes (W shaped). For the symptoms *Depressed, Anxiety* and *Headache*,

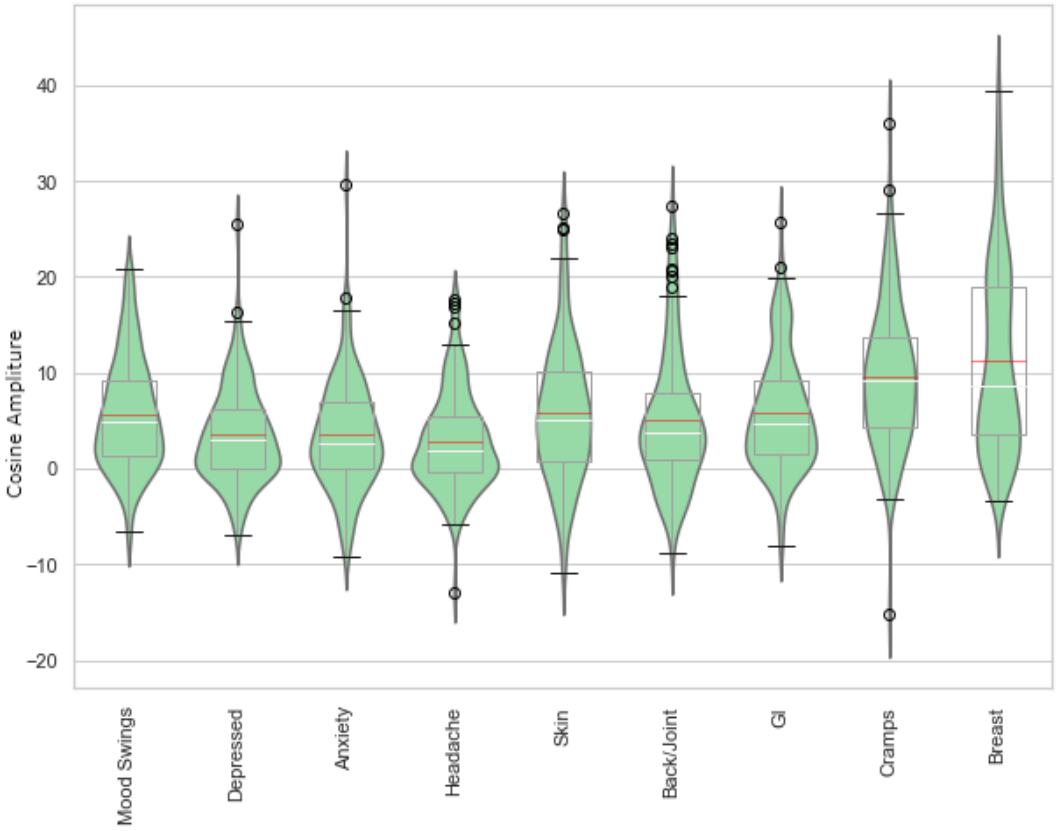


Figure 4.6: Distribution of cosine amplitudes for physical and affective symptoms. The rectangles within the violin plots show the interquartile range (25th – 75th percentile), the red line denotes the mean, and the grey line the median. Whiskers show 5th-95th percentile range and outliers are represented by circles. For the purposes of visualisation, the y-axis was rescaled to allow comparisons with the graph published in Kiesner et al., (2016). A single *Skin* reading was a distant outlier, and was therefore removed for the purposes of visualisation.

approximately 25% have a negative cosine amplitude (M shape), while other symptoms have less. *Headache* and *Depressed* have the narrowest distributions while *Breast* has the widest.

Figure 4.7 shows the distribution of cosine amplitudes in the Kiesner et al., (2016) paper. The majority of inter-quartiles ranges are positive (note that merit and blame are not being replicated) and *Depressed, Anxiety* and *Headache* have the highest percentages of negative cosine amplitudes. *Breast* has the widest distribution and *Depressed* the lowest. Absolute values fall between the same range (between -10 and 40 for 5^{th}-95^{th} percentile) for the two studies. These results show that there are strong visual similarities in cosine amplitude distributions between the Kiesner et al., (2016) and this study.

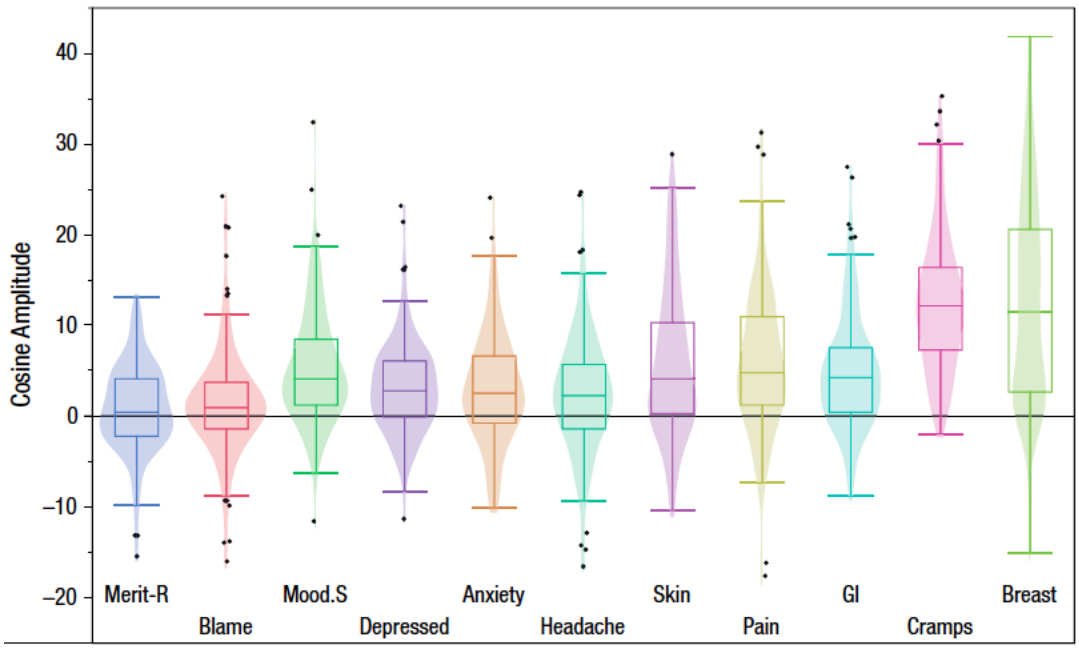


Figure 4.7: Distribution of cosine amplitudes in Kiesner et al. (2016)

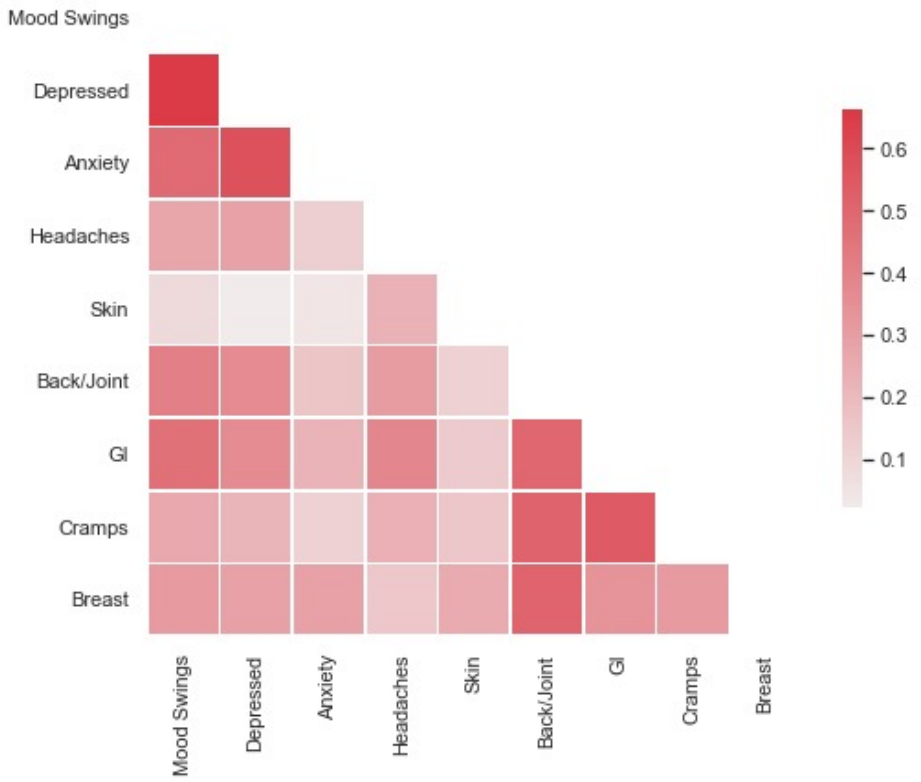


Figure 4.8: Correlation heat map of physical and affective cosine amplitudes

Figure 4.8 shows the correlations between physical and affective cosine amplitudes. The variables that share the most variance are *Depressed* and *Mood Swings* (44%) and *Depressed and Anxiety* (32%). Physical symptoms share less variance amongst themselves compared to psychological variables, with the highest being between *GI* and *Cramps* (28%) and *Back Pain* and *Cramps* (26%). *Skin* shows low shared variance across all symptoms (range of 0.04% -6%). Across different constructs *GI* and *Mood Swings* show the highest shared variance with 22%.

4.5 Multilevel Analyses

4.5.1 Outline

Multilevel analyses were run on three different models, each one building on the next, to confirm whether fitting a more complex model results in a better explanation of the random variance. All multilevel analyses were conducted using the Stata GLLAMM (Generalized Linear Latent and Mixed Models) program with adaptive quadrature.

4.5.2 Findings

The first model included the intercept as random effect. Intraclass Correlation Coefficients (ICC) are presented for each variable in Table 4.1 below, in order of increasing magnitude. The table also compares the ICC values for the Kiesner et al. (2016) study with those from the current study. The ICCs show the proportion of variance attributed to mean-level differences between participants, with higher scores indicating higher between-subject variability as opposed to within-person variability. The variables with the lower ICCs (*Cramps, Breast* and *GI)* were the same across both studies. Variables on the other end of the spectrum differed however, with *Skin* and *Mood Swings* having the higher ICCs in the Kiesner et al. (2016) study as opposed to *Anxiety* and *Back/Joint Pain* in the replication study. All nine variables had mean differences that were significantly different than zero, as suggested by the exclusion of zero from the 95% CI and confirmed by the Likelihood -ratio test ($p<0.001$).

The second model had a fixed slope and a random intercept, through the addition of the cosine function as a fixed effect. Thus, between person variability is modelled as a random effect with one separate mean level shifting up or down according to

Table 4.1: Comparison between replication and original study (Kiesner et al., 2016) of model 1 (intercept as random effect) outputs for each symptom variable

Symptom	**ICC (replication study)**	**ICC (Kiesner et al., 2016)**
Cramps	0.2	0.17
Breast	0.24	0.26
GI	0.26	0.26
Skin	0.33	0.43
Headache	0.33	0.29
Mood Swings	0.36	0.33
Depressed	0.37	0.27
Anxiety	0.42	0.29
Back/Joint	0.44	0.29

ICCs show the proportion of variance attributed to mean-level differences between participants, with higher scores indicating higher between-subject variability as opposed to within-person variability. Model 1 had a random intercept and does not account for slope. All nine variables had mean differences that were significantly different than zero ($p<0.001$).

each individual. All variables showed significant fixed effects indicating that introducing cyclicity improves the model. This finding is also seen in the Kiesner et al. (2016) study.

Fixed effect sizes are presented in Table 4.2 alongside those of the original study to allow comparison. Results show a wide variety ranging from 1% to 14.5%. The variance explainable by the second model is similar in both studies, and with the exception of *GI* (which shows a higher percentage in the replication study), symptoms follow the same increasing order in both studies.

The third model allows for a random slope. Random effect sizes are presented in Table 4.2. above. These vary widely, from 2.8% to 24%. As in Kiesner et al. (2016), random effects are approximately double the fixed effect. This shows that individual differences in cyclicity explain approximately twice the effect attributable to average effects. The near-zero significance level ($p \leq .001$ for all variables) indicates that the more complex model with a random individual-specific cosine amplitude is superior to the model that only allows for an individual-specific shift. This finding replicates Kies-

Table 4.2: Comparison between replication and original study (Kiesner et al., 2016) of fixed (model 2) and random (model 3) effects for each symptom variable.

	Replication Study		Kiesner et al. (2016)	
Symptom	**%Var Fixed**	**%Var Random**	**%Var Fixed**	**%Var Random**
Anxiety	1.6	4.1	1.0	3.8
Depressed	1.6	3.5	1.6	4.0
Mood Swings	3.9	6.7	3.4	6.6
Headache	1.0	2.8	1.0	3.1
Skin	2.9	8.7	4.3	11.3
Back/Joint	2.9	7.2	3.8	8.4
GI	6.1	11.4	2.9	6.4
Cramps	11.1	16	14.4	18.3
Breast	14.5	24.1	18.1	30.7

The effect size quantifies the proportion of within-person variance explainable by the particular model. Model 2 had a fixed slope and a random intercept. All variables showed significant fixed effects ($p<0.001$) indicating that introducing cyclicity improved the model. Model 3 had a random slope and a random intercept. All variables showed significant random effects ($p<0.001$) indicating that accounting for individual differences improved the model.

ner et al. (2016) and confirms "the presence of individual differences in the strength and direction of symptom cyclicity across individuals" (p. 888).

4.6 Structural Equation Modelling

4.6.1 Overview

To test the reliability of the best fitting factor structure in Kiesner et al. (2016), confirmatory factor analysis will be carried out following the same model configurations and nesting procedure, whereby the successive model is compared to the previous, as in the original work. CFAs are ideal as a testing method in this case since pre-existing models already exist. Since Kiesner et al. (2016) does not provide a rationale for the variables chosen to define the model factors, the current study will also carry out an EFA to test the best fitting structure suggested by data.

Three models will be replicated: a one-factor model with all variables loading onto one factor; a three-factor model with variables loading onto uncorrelated headaches, physical and affective latent factors; and finally - the Kiesner et al. (2016) model of best fit - a three-factor model with correlated headaches, physical and affective latent factors. These three models will then be compared to that generated by the EFA, provided it is not one of the already hypothesized model structures.

Before fitting the models, multiple processes will be carried out to ensure high quality data. As mentioned in the previous chapter, the saved cosine amplitudes will be used, rather than the raw scores, since we are relating symptom cyclicity and not comparing symptoms at particular points in the cycle. Working with cosine amplitudes reduces the dataset from 7000+ data points to 120 data points. The derived data set contains no missing values. Prior to the analysis multiple assumption tests are carried out to ensure the data set is suitable for factor analysis. These include multicollinearity tests, Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of spheric-

ity. All tests are listed in section 3.4.3.1 in the previous chapter. Cut-off points for CFA fit indices tests are presented in section 3.4.3.2.

4.6.2 Replication CFA

For step one of finding a model of best fit, three models were explored. These were the same as those tested in Kiesner et al. (2016): a one-factor model, an uncorrelated three-factor model and a correlated three-factor model.

4.6.2.1 Assumption tests

The case to variable ratio of the most complex model exceeded the minimum sample size threshold of 5:1 cases to free parameters, having 120 data points to twenty parameters. The Doornik-Hansen test rejected normality of the data (χ^2 (18) =119.817, $p<0.001$) and so a Maximum Likelihood estimation method with robust standard errors was employed together with the Satorra-Bentler correction.

The first CFA was a 9-item model loading onto one latent factor. A good level of reliability was demonstrated (α=.78) suggesting internal consistency without item redundancy. The three-factor model composed of three affective items (α=.82), five physical items (α=.70) and one item for headaches showed a high level of reliability.

Stata does not converge when only one observed variable loads onto a latent variable, as is the case for *Headache*. In this situation the replication value needs to be changed so that the loading is not equal to one. By fixing the reliability value to 0.8 we accounted for the realistic assumption that the observed variable might not have measured the underlying latent variable perfectly. Ideally, the constraint carried out would have been the same as in the original, however this was not specified in the Kiesner et

al., (2016) study. The reliability coefficient chosen of 0.8 is based on Nunnally's 1994 recommendations for applied research.

4.6.2.2 Replication CFA findings

Findings are presented in Table 4.3. The results show that the one-factor model which loaded all nine variables onto one factor does not fit the data well. The model failed all the required criteria imposed by the goodness of fit tests. The same is seen for the second model, a three-factor model with no set correlations among the latent variables. The third model, with the same factor structure but with correlating factors, fit the data well, with all four fit indices meeting the minimum threshold for fit.

Table 4.3: CFA results based on Kiesner et al. (2016) models

Model	χ^2	*df*	*p-value*	*CFI*	*AIC*	*RMSEA*	*SRMR*
One-Factor	71.782	27	<0.001	0.779	-4.339	0.118	0.1
Three-Factor Uncorrelated	64.85	28	<0.001	0.818	-22.968	0.105	0.199
Three-Factor Correlated	30.017	25	0.224	0.975	-65.298	0.041	0.062

CFI = comparative fit index, AIC = Akaike's information criterion, RMSEA = root mean square error of approximation, SRMR = standardized root means square residual

To allow readers to make a quick assessment of fit, a scoring system was devised whereby models meeting the criteria for two tests or less were labelled a "poor fit" and those that met the criteria for all tests marked as "good fit" (see Table 4.4). Though not present in this study, models meeting three of four tests would have been classified as "moderate fit".

The third model was re-run with an additional correlation between *Cramps* and *Mood Swings* as suggested by the modification indices in the original paper. While a good fit was produced, (χ^2 (24) = 29.673, p=.196; CFI=.972; AIC=-63.32; RMSEA=.045; SRMR=.062) fit indices were better without it, and so this additional cor-

Table 4.4: Assessing goodness of fit for one-factor and three-factor models

Model	*p-value* >.05	*CFI* >0.95	RMSEA <0.06	<0.08	*SRMR* <0.06	<0.08	*Model Fit Decision*
One-Factor	✗	✗	✗	✗	✗	✗	Poor Fit
Three-Factor Uncorrelated	✗	✗	✗	✗	✗	✗	Poor Fit
Three-Factor Correlated	✓	✓	✓	✓	✗	✓	Good Fit

CFI = comparative fit index, AIC = Akaike's information criterion, RMSEA = root mean square error of approximation, SRMR = standardized root means square residual

relation was not included in the final model. Results are in line with the Kiesner et al. (2016) paper and provide strong support for the correlated three-factor model and the variables loading onto each factor.

4.6.3 EFA

For step two of finding a model of best fit, an EFA was carried out to see what model configuration would be suggested based on the data. As in the previous section, analysis was carried out on the saved cosine amplitudes and not the raw scores.

4.6.3.1 Assumption tests

Factorability was studied for nine variables. Case to variable ratio of 120:9 was above the suggested 5:1 minimum. The Doornik-Hansen test rejected normality (χ^2 (18) = 119.871, p<0.001) and so Principal Axis Factoring was chosen as the factor extraction method instead of Maximum Likelihood.

All items were significantly correlated (p<0.05) except for *Skin* which did not show significant correlations with any variables except *Breast* (p=0.002) and *Headache* (p=0.006). The communality for *Skin* was also below the .3 cut-off at 0.07 (Table 4.5a),

Table 4.5: Communalities

	Initial	Extraction
Depressed	.576	.765
Anxiety	.454	.529
Mood Swings	.525	.602
GI	.460	.531
Back/Joint	.489	.569
Cramps	.381	.474
Breast	.348	.301
Headache	.227	.193
Skin	.125	.071

a. Before skin variable was removed

	Initial	Extraction
Depressed	.572	.749
Anxiety	.454	.543
Mood Swings	.525	.602
GI	.459	.538
Back/Joint	.486	.599
Cramps	.377	.485
Breast	.310	.281
Headache	.190	.179

b. After skin variable was removed

suggesting loose loading onto the factor solution. Therefore, *Skin* was removed from the analysis. The communality for *Headache* was also below the cut-off (0.193). The EFA was re-run without *Skin* but the communality for *Headache* remained low (Table 4.5b). *Headache* was the second variable removed from the analysis.

Kaiser-Meyer-Olkin came in at .78, above the recommended .6, and Bartlett's test was significant ($\chi 2$ (21) = 300.267, p<0.001). The determinant was greater than 0.00001 (0.075). All communalities were now well over .2. Factor analysis assumptions were considered met.

4.6.3.2 EFA findings

Using the Kaiser (K1) criterion a two-factor model was suggested with the first factor accounting for 47.7% of the variance and the second factor 18.8%; cumulative total of 66.5%. All other presented factors were discarded since they had eigenvalues below 1. Varimax rotation was used to rotate loadings and make the readings clearer to interpret. Items were sorted by size and those loading at lower than .4 suppressed. The variables loading onto the first factor were physical: *Back/Joint, Cramps, GI* and *Breast* and the variables loading onto the second variables were affective: *Depressed, Anxiety, Mood*

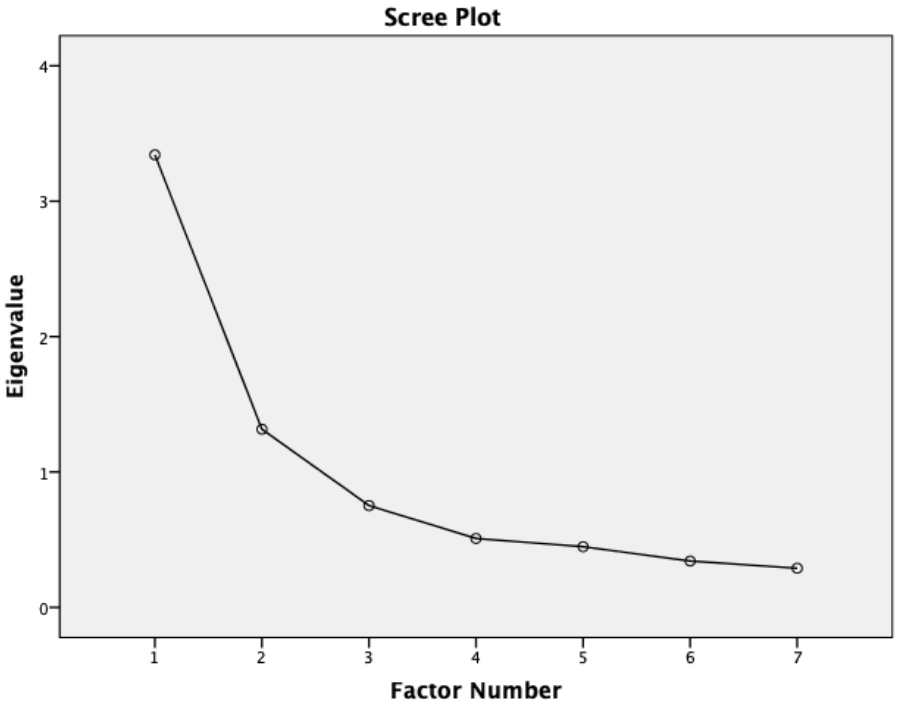


Figure 4.9: Physical and affective symptom scree plot

Swings. Factor labels were therefore kept the same as in the Kiesner et al. (2016) paper. Factor structure was then examined using a scree plot, but it was unclear at which point the eigenvalues were "levelling off" (see Figure 4.9).

Since the readings from the scree plot were not clear, factor structure was also examined using parallel analysis. PA presented seven factors, three of which were retained since the 4th factor had an eigenvalue of -0.03 (below the cut-off point of 1). Loadings were rotated using varimax with the suppression of items lower than .4. A look at the rotated matrix (Table 4.6) shows that all variables loaded onto Factor 1 and 2. None loaded on Factor 3, suggesting that Factor 3 must have contained loadings below .4 which were suppressed. Value loading suppression at .4 is justified since the rest of the variables have high loadings.

Table 4.6: Physical and affective Rotated Component Matrix

Variable	Factor 1	Factor 2	Factor 3	Uniqueness
Depressed	0.7807			0.3447
Anxiety	0.7155			0.4769
Mood Swings	0.7028			0.4026
GI		0.6366		0.4872
Back/Joint		0.6900		0.4284
Cramps		0.6704		0.5378
Breast		0.4543		0.6301

Extraction Method: Principal Axis Factoring. Rotation Method: Varimax. Loadings blank at $<.4$

The PA therefore supported the two-factor model proposed by the eigenvalues and scree plot tests. The variables which loaded onto Factor 1 and Factor 2 were the same as those in the K1 test and in the parallel analysis and were thus retained accordingly. Both physical and affective factors were tested for reliability using Cronbach's and with scores of α=.75 and α=.8 respectively, both showed high reliability.

4.6.4 CFA based on EFA modifications

For step three, CFA was once again carried out so as to test the configuration suggested by the EFA. EFA suggested that 1) the variable *Skin* correlated very weakly with the other variables, and could be a candidate for removal, the same applying to *Headache* but to a lesser extent; 2) the removal of both these variables produced a two-factor model with physical and affective latent factors.

The first model to be tested out had both *Skin* and *Headache* removed. It was therefore a two-factor model with both latent factors correlated. While this model was a good fit (see Table 4.7), it did not perform any better than the theoretically driven three-factor model presented above (see Table 4.3). The second model tested retained *Headache* but removed *Skin*, so it was a three-factor model with correlated physical, affective and headache latent variables. This model, with the just the skin variable

Table 4.7: CFA results based on EFA modification suggestions

Model	χ^2	*df*	*p-value*	*CFI*	*AIC*	*RMSEA*	*SRMR*
Two-factor without Skin & Headaches	20.350	13	0.087	0.962	-96.75	0.069	0.062
Three-Factor without Skin	23.306	18	0.179	0.975	-147.89	0.05	0.059

CFI = comparative fit index, AIC = Akaike's information criterion, RMSEA = root mean square error of approximation, SRMR = standardized root means square residual

Table 4.8: Comparing goodness of fit for theoretical and data driven models

Model	*p-value*	*CFI*	RMSEA		*SRMR*		*Model Fit*
	>.05	>0.95	<0.06	<0.08	<0.06	<0.08	*Decision*
Three-Factor with Skin	✓	✓	✓	✓	✗	✓	Good Fit
Three-Factor without Skin	✓	✓	✓	✓	✓	✓	Excellent Fit
Two-Factor without Skin & Headache	✓	✓	✓	✓	✗	✓	Good Fit

This table compares the theoretically driven model of best fit (Three-Factor with Skin) with the two models suggested by the data driven EFA (Three-Factor without Skin and the Two-Factor without Skin & Headache)

removed, performed extremely well, outperforming the original three-factor model and meeting the most conservative requirements (see Table 4.8).

4.6.5 Current study model of best fit

The model of best fit in the current study is a three-factor model with eight variables loading onto three latent factors (Figure 4.10). The latent factor Headaches has one variable loading onto it, Physical has five, and Affective has three. All latent constructs are significantly correlated and share the following variances: Headaches and Physical 22%, Physical and Affective 32%, and Headaches and Affective 13%.

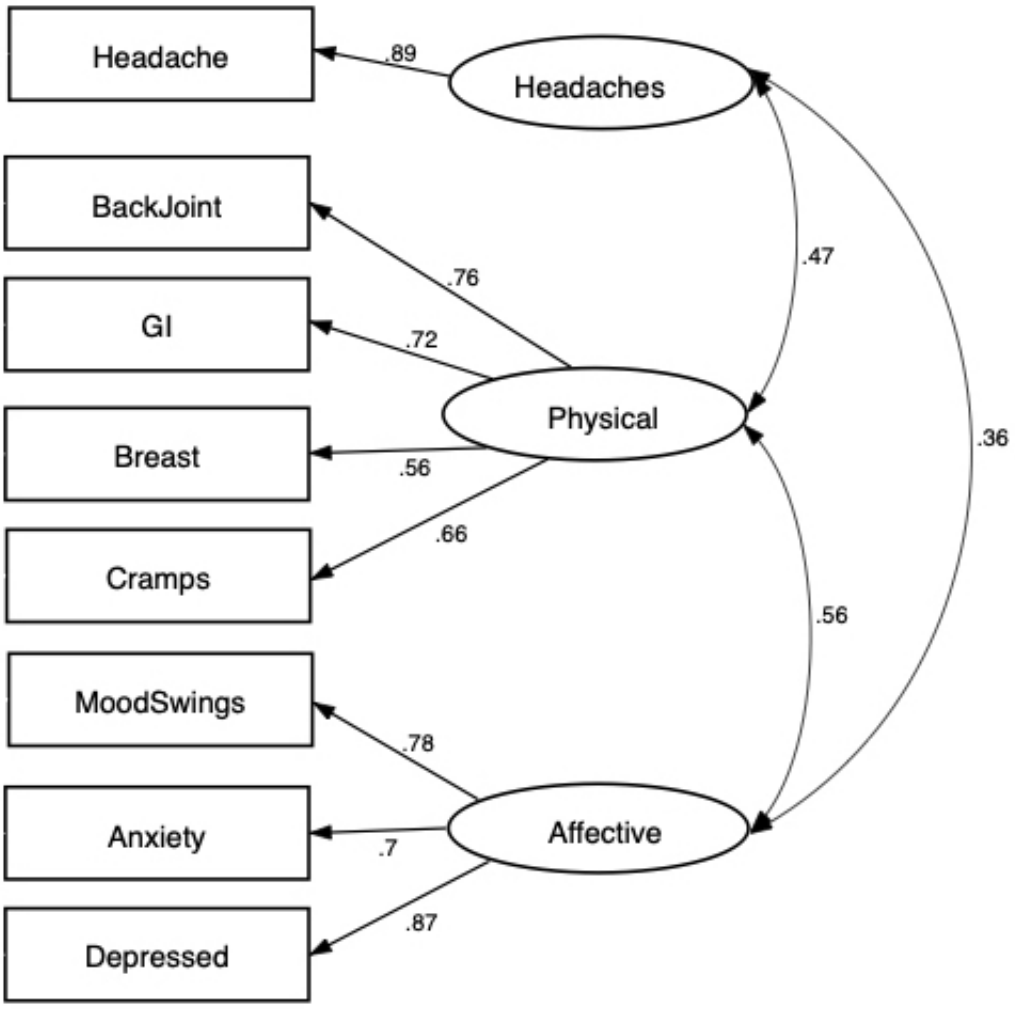


Figure 4.10: Model of best fit in replication study

4.6.6 Comparison between replication and original models

For the variables under study, both the Kiesner et al. (2016) model and replication model resulted in three separate but correlated headache, physical and affective latent factors. Path diagrams for both are presented (Figure 4.10 above & Figure 4.11 below). Table 4.9 shows the loadings and R-squared value for each model.

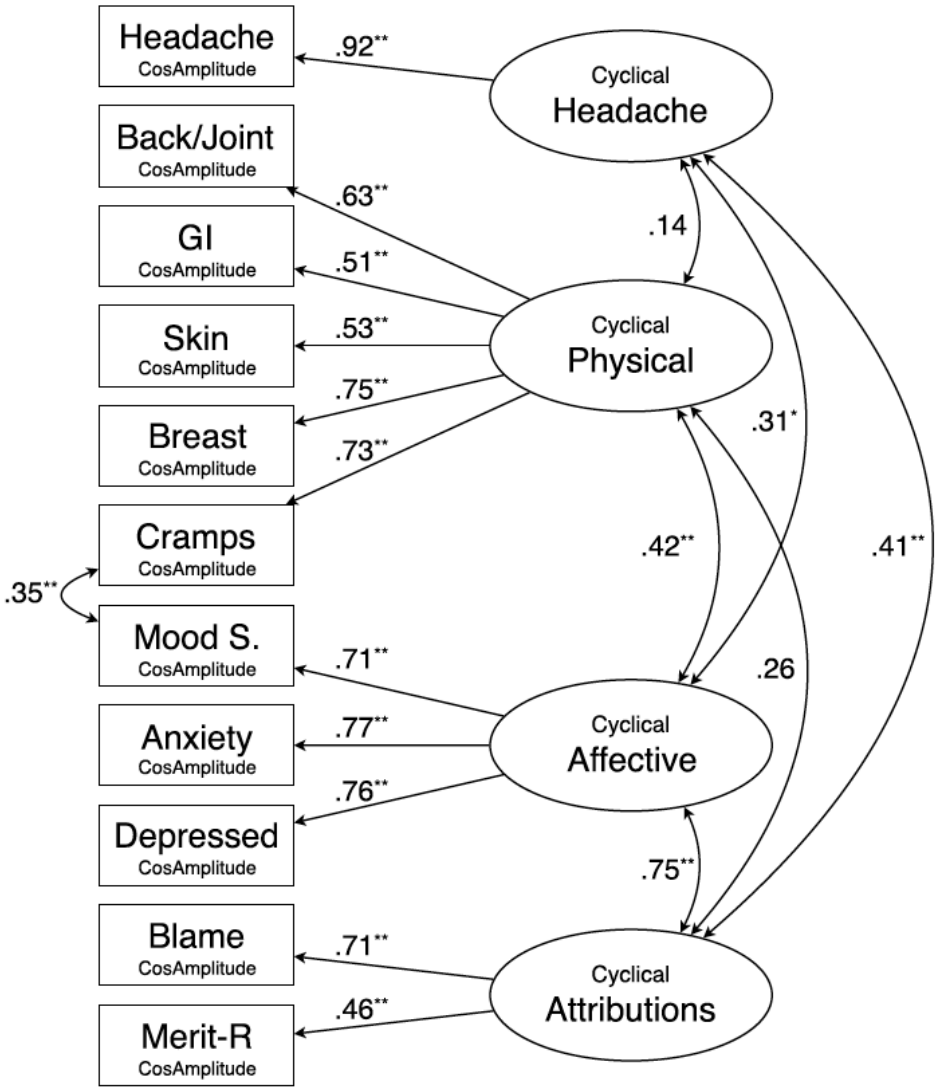


Figure 4.11: Model of best fit in Kiesner et al. (2016). The latent factor Headaches has one variable loading onto it, Physical has five, and Affective has three. The model also includes a correlated residual between cramps and mood swings, suggested by modification indices to help improve the model. A measure of attribution was not taken for the current study and so the factor and its variables were not included in the replication.

In the current study, the latent constructs Affective and Headaches are significantly associated with a shared variance of 13%. Similar associations can be seen in Kiesner et al. (2016). Physical and affective latent variables in the present study show a higher percentage of shared variance than in the original study suggesting a stronger relationship between the two than previous thought. The main difference between the

Table 4.9: Loadings and R^2 for replication and original models

	Replication Study		**Kiesner et al. (2016)**	
Latent factors	Coefficients	R^2	Coefficients	R^2
Headaches - Physical	.47	.22	.14	.02
Affective - Headaches	.36	.13	.31	.10
Physical - Affective	.56	.31	.47	.18
Variables				
Headache	.89	.79	.92	.85
Back/Joint	.76	.58	.63	.40
GI	.72	.52	.51	.26
Breast	.56	.31	.75	.56
Cramps	.66	.44	.73	.53
Mood Swings	.78	.61	.78	.61
Anxiety	.70	.49	.77	.59
Depression	.87	.76	.76	.58

two models is the relationship between headaches and physical symptoms. There is a significant shared variance in the replication model (22%), while in the original study they are not correlated. This might suggest that the two constructs are not as independent of each other as thought in the Kiesner et al. (2016) paper. The loading coefficients for affective symptoms are higher than for physical symptoms, suggesting higher internal consistency between affective symptoms than between physical ones. This result was also found in the original study.

Overall these findings support the conclusions in Kiesner et al. (2016) that symptom changes are both linked (significantly correlated constructs) and distinct (meriting different factor constructs and exhibiting different internal consistencies). Further exploration and discussion of these findings can be found in Chapter 6.

4.7 Chapter summary

This chapter explored physical and affective symptom change patterns in the context of the menstrual cycle and compared them to those in Kiesner et al. (2016). Results confirmed menstrual cyclicity for all variables as well as variation in the strength and direction of symptom cyclicity across individuals. Factor structures indicated distinct but significantly correlated headache, affective and physical constructs. Chapter 5 looks at changes in menstrual metacognitive abilities. Focus is on cyclicity patterns, variations between and within individuals and the concurrency of metacognitive changes with physical and affective symptoms.

Chapter 5

Results II

Changes in Cognitive Ability

5.1 Chapter Outline

The present study focuses on metacognitive changes in relation to menstrual hormone fluctuations. Specifically, the aim is to determine whether changes are cyclical with respect to the menstrual cycle, the extent of individual differences in ability changes, and to examine the covariance of cognitive ability changes, both among themselves and in relation to other symptoms. The findings presented in this chapter provide insight into metacognitive change patterns in terms of magnitude and direction and the importance of individual differences as compared to average levels of change. Further, findings demonstrate symptom groupings providing valuable insight into the causal mechanisms that connect them.

This chapter focuses on presenting the findings related to cognitive ability changes. The results concerning participant demographics, particularly their menstrual characteristics, recruitment, and response patterns, are discussed in the previous chapter (see section 4.2). In section 5.2, an exploratory analysis of the 15 self-assessed cognitive ability (SACA) variables and their derived variables is conducted. The results for cosine regressions and multilevel modelling are presented in sections 5.3 and 5.4, respectively. Section 5.5. demonstrates structural equation modelling (SEM) results for cognition while section 5.6 presents SEM results for cognitive, physical and affective symptom changes. All reporting procedures are based on the guidelines provided by well-established works: Jackson et al. (2009) for CFA, Costello & Osborne (2005) for EFA and McDonald et al. (2002) for SEM. The following convention will be applied throughout: variables are italicized, and latent constructs are capitalised.

5.2 Exploratory Data Analysis

A total of 15 SACAs were studied. All items were linked to one of the following constructs: GCA (General Cognitive Ability), Attention, Memory, Decisions and Language with each construct variable being the averaged result of three questions. Scoring was carried out on a 1-100 Visual Analog Scale, with participants scoring 50 if they felt the same as usual, above 50 if they felt better than usual and below if they felt worse.

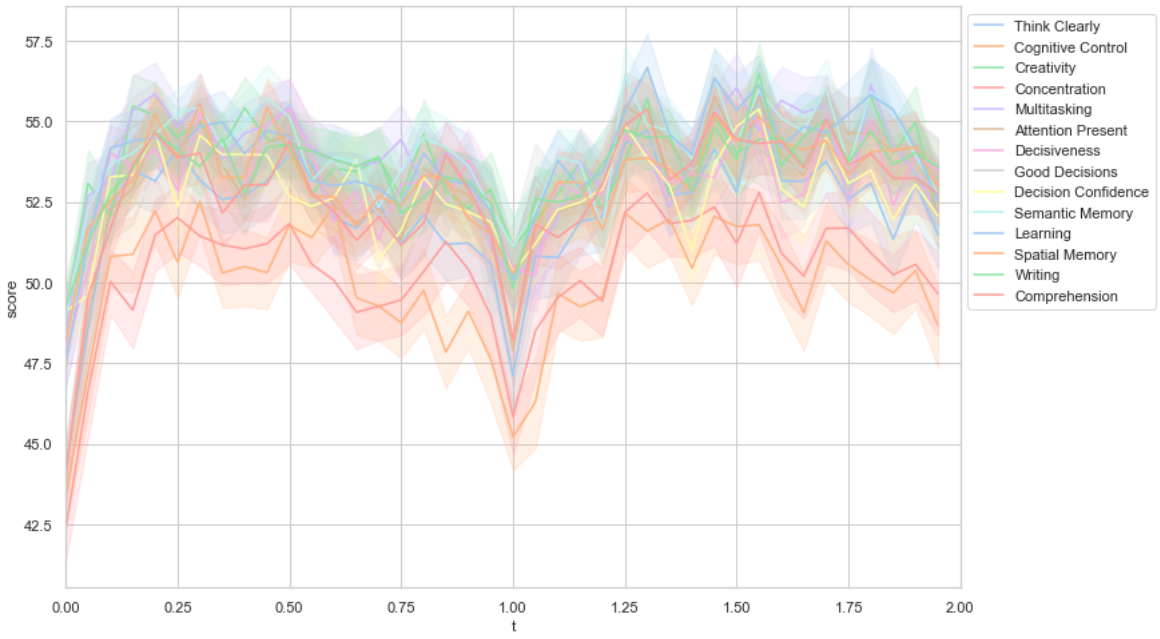


Figure 5.1: Score-time plot for all 15 SACAs. The shaded bands show the confidence intervals (to within one SD) of the score means.

Figure 5.1 shows the SACA scores averaged across individuals as a function of time. As described in section 3.4.1 the time variable t is the binned ratio of each participant's cycle length. The SACA variables show clear cyclicity with relation to the menstrual cycle. All items show an M-shaped trajectory, which in this case is indicative of a reported reduction in ability towards the beginning of the cycle and reported improvement mid-cycle. Note that although the same perimenstrual symptom

worsening was observed for physical and affective symptoms, these demonstrated a W-shaped trajectory reflective of their positive valence. Overall, all variables show a similar time profile, peaking rapidly after cycle onset, followed by a gradual decline and a sudden drop before the next cycle onset, where the pattern repeats with the lowest point reached precisely on the first day of the cycle. All time profiles are somewhat noisy.

Figure 5.2 is a collapsed version of Figure 5.1, showing derived variables instead of individual items, allowing for easy interpretation. Each derived variable is the averaged result of three questions. Among the derived variables, *Memory* and *GCA* have the highest and lowest time profiles, respectively. As with physical and affective symptoms, the amplitudes of SACAs get smaller with each cycle. At the beginning of

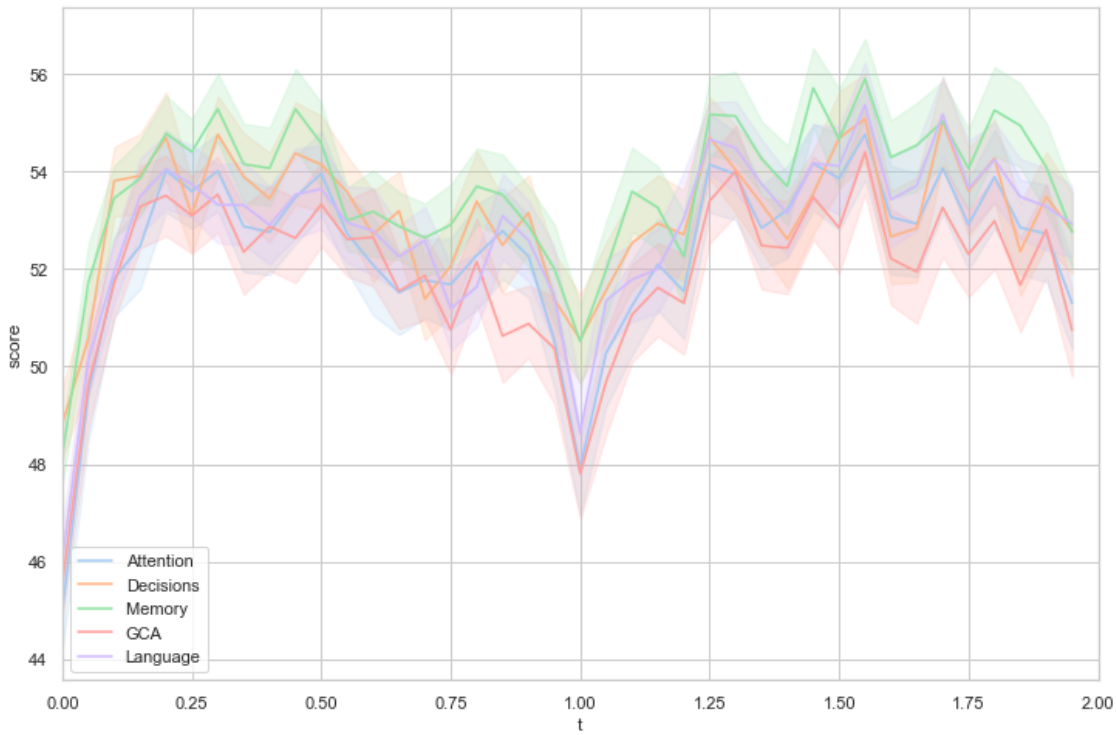


Figure 5.2: Score-time plot for the five derived variables. The shaded bands show the confidence intervals (to within one SD) of the score means.

the first cycle, all fives variables score below 50. At the end of the first cycle, only three of them score below 50, and at the end of the second cycle, none of them do. The underived items exhibit a similar phenomenon.

When analysing a variable that is expected to be cyclical, it is important to consider not only the pattern of the variable over time (Figures 5.1 & 5.2) but also the distribution of the variable's values. Figure 5.3 presents distributions for each variable. Data points are collapsed across all individuals and across both cycles. In the box plot, the red bar indicates the mean, the grey bar the median and the green rectangles extend over the interquartile range (25^{th} – 75^{th} percentile) with whiskers showing the 5^{th} and 95^{th} percentiles.

The first point of note is that item scores appear to be positively mean-biased as seen by the means and interquartile ranges which are above 50. The majority are also positively skewed since they have a higher mean than median. The second point of interest is the profile similarities between variables, with scores clustered around the 50 mark as indicated by the medians. Furthermore, the majority of variables appear to have a similar spread as seen by the interquartile ranges and whiskers. This can possibly suggest that subjects showed low discrimination between abilities. The only variables that appear to be non-biased are *Concentration* and *Cognitive Control* with the 50^{th} percentile at 50. They show symmetrical distribution since their mean and median fall in the same place.

Profile similarity is reinforced by Figure 5.4, which shows the box plot for the derived variables instead of the individual question items. All variables are positively skewed, with similar medians and spread profiles.

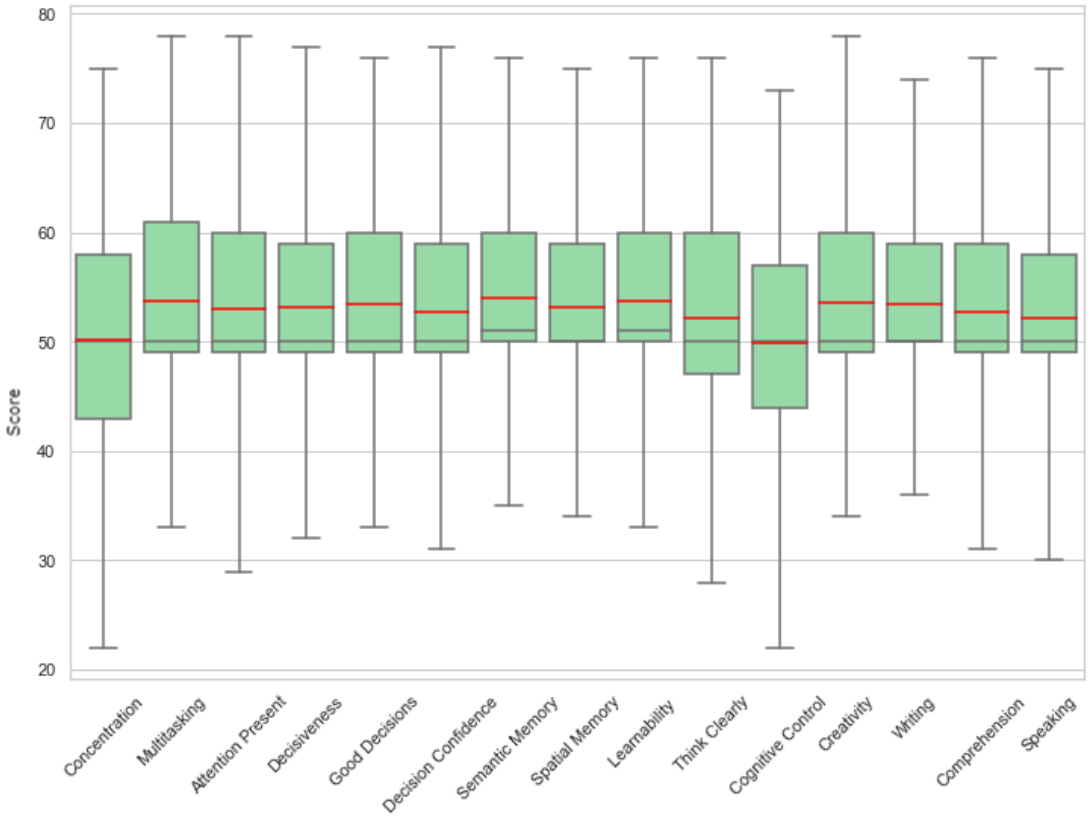


Figure 5.3: Box plot for all 15 Self-Assessed Cognitive Abilities. The red bar indicates the mean, the grey bar the median, and the green rectangles extend over the interquartile range (25th – 75th percentile) with whiskers showing the 5th and 95th percentiles.

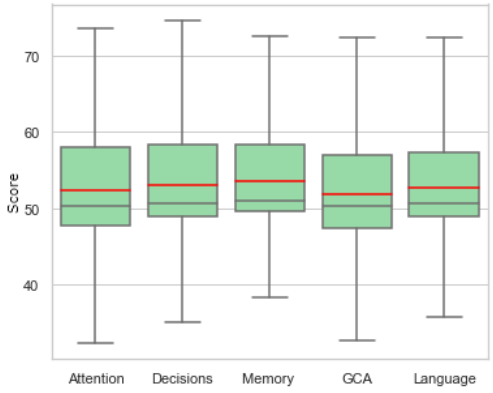


Figure 5.4: Box plot for the 5 derived Self-Assessed Cognitive Ability variables. The red bar indicates the mean, the grey bar the median, and the green rectangles extend over the interquartile range (25th – 75th percentile) with whiskers showing the 5th and 95th percentiles.

5.3 Cosine Regressions

The distribution of cosine amplitudes for SACAs is presented in Figure 5.5. The rectangles within the violin plots show the interquartile range (25th–75th percentile), the red line denotes the mean, and the grey line the median. Whiskers show the 5th–95th percentile range.

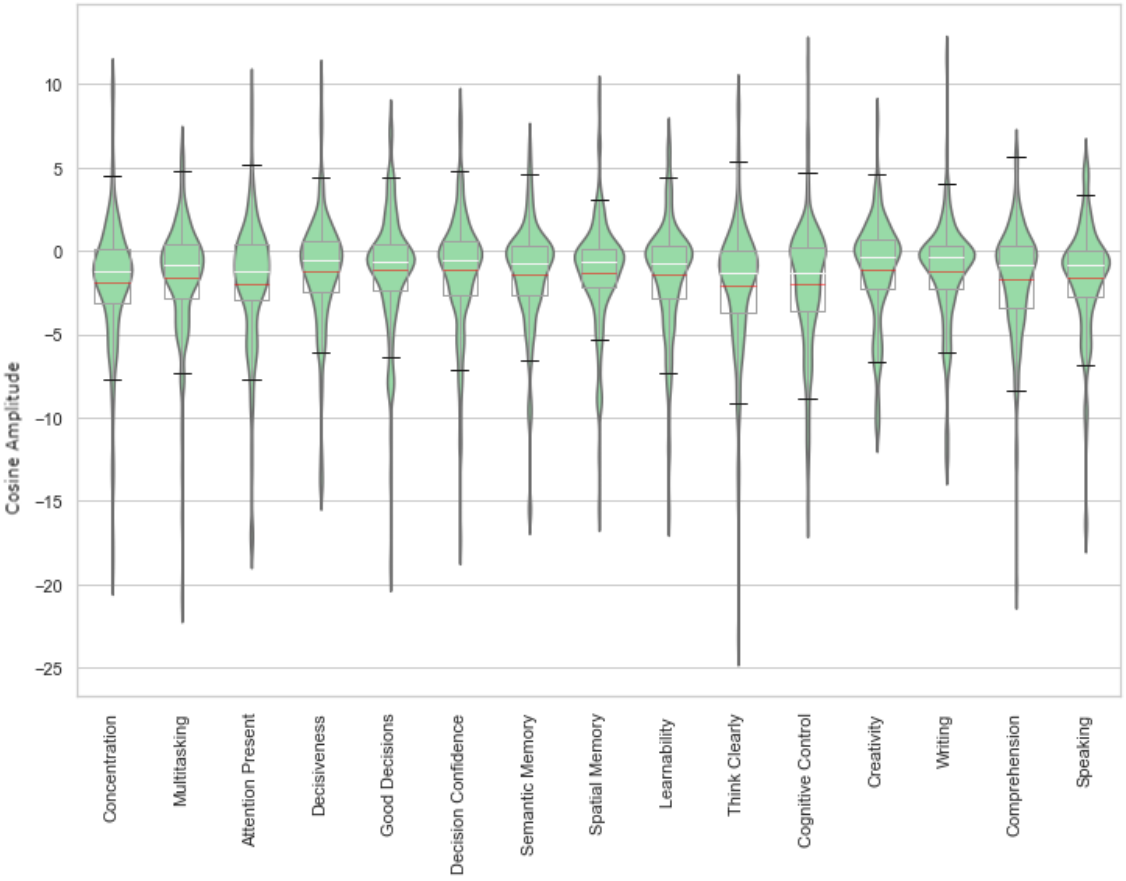


Figure 5.5: Distribution of cosine amplitudes for all 15 Self-Assessed Cognitive Abilities. The rectangles within the violin plots show the interquartile range (25th–75th percentile), the red line denotes the mean, and the grey line the median. Whiskers show the 5th–95th percentile range under both images.

Approximately 75% of the cosine amplitudes are negative (M-shaped cosine functions) suggesting a perimenstrual decrease in reported abilities. 25% show the

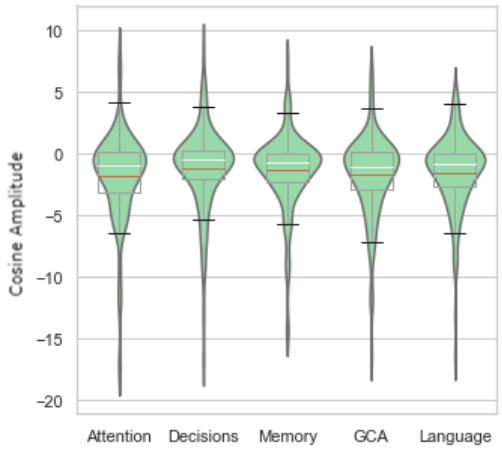


Figure 5.6: Distribution of cosine amplitudes for the 5 derived Self-Assessed Cognitive Ability variables. The rectangles within the violin plots show the interquartile range (25th–75th percentile), the red line denotes the mean, and the grey line the median. Whiskers show the 5th–95th percentile range under both images.

opposite cyclical change (W-shaped). Note that this pattern is also seen with physical and affective symptoms, but that the trajectory is reversed due to the reverse-scoring system. All means lie below the medians, showing that the distributions are negatively skewed (there are more outliers with a pronounced M-shape than a W-shape). Although in the box plots *Concentration* and *Cognitive Control* demonstrated different profiles to the other symptoms, this is not reflected in their cosine amplitudes.

Figure 5.6 shows the cosine amplitude distributions for the derived variables. *Decisions* has the narrowest distribution (i.e., the smallest interquartile range) and *Attention* has the widest distribution. Again, all means lie below the medians, meaning that the distributions are negatively skewed. Moreover, all inter-quartiles ranges lie below the axis (i.e., approximately 75% of subject-symptoms have negative, or M-shaped, cosine functions). All 5 variables have a similar distribution.

Figure 5.7 shows the correlations for physical, affective and cognitive cosine amplitudes. Derived variables are used here instead of the 15 individual items for easy

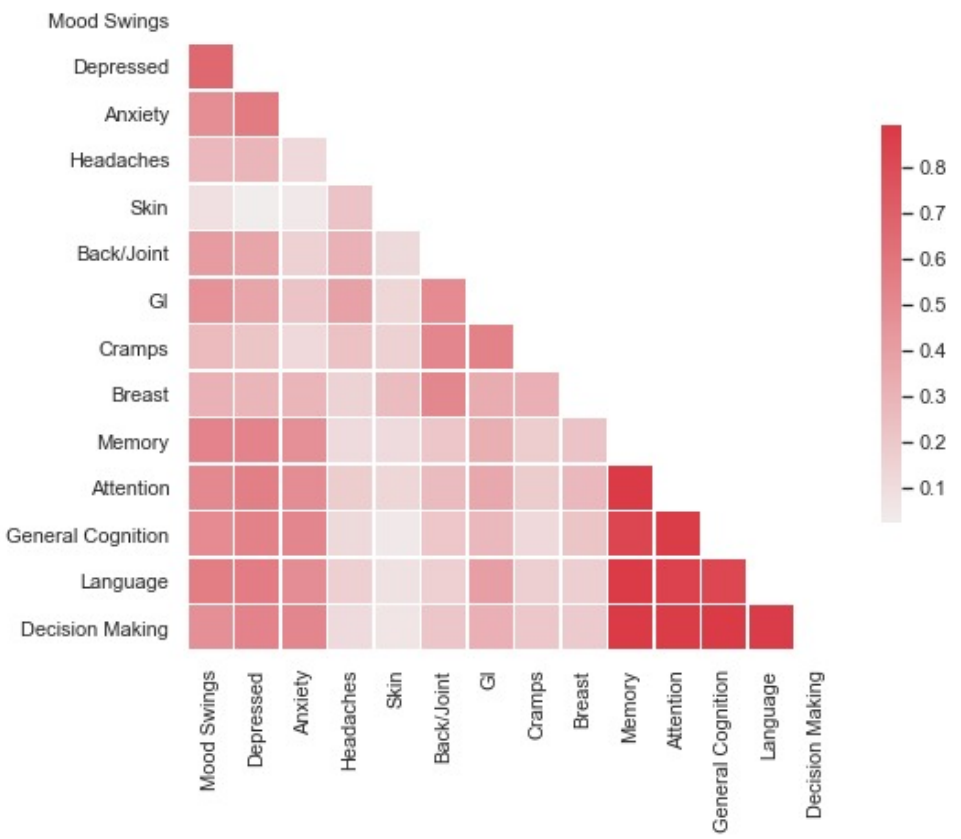


Figure 5.7: Correlation heat map of physical, affective and cognitive cosine amplitudes

comparison with the other symptom changes.

Within cluster, derived cognitive variables show the highest shared variance (ranging from 67%-72%), followed by affective symptoms (23%-43%). As seen in the previous chapter, physical symptoms show weak to negligible correlations amongst themselves, excluding a few exceptions like *Cramps* and *GI* (28% shared variance). Of particular interest are the moderate correlations between affective and cognitive symptoms all of which fall in the range of 22%-31% shared variance. Physical coefficients as a whole, correlate weakly with affective or cognitive symptoms. An exception is the *GI* coefficient, which shares 16% of the variance with *Language* and 21% with *Mood*

Swings.

Pairwise correlations for the 15 individual SACAs show ranges of 31%-74% shared variance between items. As expected, items show higher shared variance within group since they are tapping into the same construct. A clear example is *Decisiveness*, *Good Decisions* and *Decision Confidence* (67%-72%). Not surprisingly, items which fall under the GCA construct like *Think Clearly* and *Cognitive Control* show strong correlations with a number of variables from other constructs. Interestingly, however, *Creativity*, although classified as a GCA, is the item least correlated with any other SACA (31%-53%). It also correlates negligibly with physical and affective symptoms.

5.4 Multilevel Analyses

5.4.1 Overview

Multilevel analyses were run on three different models, each one building on the next, to confirm whether fitting a more complex model results in a better explanation of the random variance. All multilevel analyses were conducted using the Stata GLLAMM (Generalized Linear Latent and Mixed Models) program with adaptive quadrature.

5.4.2 Findings

The 1st level model looked at between-participants mean differences, including the intercept as random effect. All fifteen variables had mean differences that were significantly different than zero, as suggested by the exclusion of zero from the 95% CI and confirmed by the Likelihood -ratio test ($p<0.001$). Intraclass Correlation Coefficients (ICC) are shown, in order of increasing magnitude, in table 5.1 below. The ICCs show the proportion of variance attributed to mean-level differences between participants. The higher the ICC (*Spatial memory, Multitasking* and *Semantic Memory)*, the more stable the variable is within-person. The lower the ICC *(Concentration* and *Cognitive Control),* the more variance attributed to variations between within-individual replications.

The 2nd level analysis included the intercept as a random effect as well as the fixed effect of cosine. All variables showed significant fixed effects indicating that introducing cyclicity improves the model. Table 5.1 shows the fixed effects for all 15 variables. Results do not show a wide range of variability in fixed effects: lowest is

Table 5.1: ICCs (model 1), fixed (model 2) and random (model 3) effects for each of the 15 Self-Assessed Cognitive Ability variables

Ability	**ICC**	**% Var Fixed**	**% Var Random**
Concentration	0.16	0.8	3.2
Cognitive Control	0.19	1.1	3.5
Speaking	0.20	0.8	3.2
Attention to Present	0.21	1.0	4.3
Decision Confidence	0.21	0.4	3.0
Comprehension	0.22	0.9	3.4
Think Clearly	0.22	1.2	4.5
Decisiveness	0.22	0.4	3.2
Learnability	0.23	0.8	3.4
Good Decisions	0.23	0.4	3.8
Writing	0.24	0.7	3.9
Creativity	0.27	0.4	2.2
Semantic Memory	0.30	0.7	3.7
Multitasking	0.30	0.9	4.0
Spatial Memory	0.31	0.6	3.7

ICCs show the proportion of variance attributed to mean-level differences between participants, with higher scores indicating higher between-subject variability as opposed to within-person variability. The effect size quantifies the proportion of within-person variance explainable by the particular model. Model 1 had a random intercept and does not account for slope. All variables had mean differences that were significantly different than zero ($p<0.001$). Model 2 had a fixed slope and a random intercept. All variables showed significant fixed effects ($p<0.001$) indicating that introducing cyclicity improved the model. Model 3 had a random slope and a random intercept. All variables showed significant random effects ($p<0.001$) indicating that accounting for individual differences improved the model.

0.4% for *Creativity, Good Decisions, Decision Confidence, Decisiveness* and highest is 1.2% for *Think Clearly*.

In the 3rd level, the analysis was run with both the intercept and slope as random factors. All variables showed significant random effects (exclusion of zero from the 95% CI). Table 5.1 shows the random effects for all 15 variables. Results range from 2.2% (*Creativity*) up to 4.5% (*Think Clearly*). The more complex model, with random individual-specific cosine amplitudes, explains approximately four times the variance that the 2nd level model does.

Table 5.2 presents the ICCs, fixed and random effects for each of the five derived variables. Working with derived variables, rather than all 15 question items, makes it easier to compare cognitive changes with physical and affective symptom ones. For the ICCs all five variables had mean level differences that were significantly different than zero. Apart from *Memory* which had the highest ICC (0.36), the rest showed similar proportions of variance. Physical and affective ICCs (see Table 4.1 in Chapter 4) showed more variation (range of 0.2–0.44) than cognitive ones (0.26–0.36).

All five derived variables showed significant fixed effects, endorsing a more complex model. While the effect sizes for physical variables differed dramatically, ranging from 1% to the much higher 14.5% (refer to Table 4.2), effect sizes were very similar for cognitive variables, with the exception of *Decision Making*, which was lower than the rest (0.6%). Cognitive fixed effect sizes are more in line with affective symptoms, which were also on the lower end (ranging from 1.6% to 3.9%).

A similar trend is seen with the random effects. Effect sizes for cognitive variables do not show much variability (4.5%-5.5%). They are in the same range as affective symptoms (3.8%-6.6%). Physical symptom effect sizes, on the other hand,

Table 5.2: ICCs (model 1), fixed (model 2) and random (model 3) effects for each of the 5 derived variables

Ability	**ICC**	**% Var Fixed**	**% Var Random**
Attention	0.26	1.3	5.4
GCA	0.27	1.3	4.9
Language	0.27	1.2	4.9
Decision Making	0.28	0.6	4.5
Memory	0.36	1.1	5.5

ICCs show the proportion of variance attributed to mean-level differences between participants, with higher scores indicating higher between-subject variability as opposed to within-person variability. The effect size quantifies the proportion of within-person variance explainable by the particular model. Model 1 had a random intercept and does not account for slope. All variables had mean differences that were significantly different than zero ($p<0.001$). Model 2 had a fixed slope and a random intercept. All variables showed significant fixed effects ($p<0.001$) indicating that introducing cyclicity improved the model. Model 3 had a random slope and a random intercept. All variables showed significant random effects ($p<0.001$) indicating that accounting for individual differences improved the model.

varied from 2.8% to the much higher 24.1% (refer to Table 4.2). Random effects were significant for all variables, indicating once again the importance of considering individual differences when looking at menstrual cycle effects. For the derived cognitive variables, the random model explains approximately four times the variance that the fixed model does. This increase is similar to that shown by affective symptoms (see Table 4.2). Physical symptoms, on the other hand, had random effect sizes approximately double those of the fixed (see Table 4.2).

5.5 Structural Equation Modelling I: Cognition

5.5.1 Overview

The objective of this analysis is to identify the model of best fit for self-assessed cognition as tested using the metacognitive evaluation measure DOME. By determining the optimal model, the relationship between self-assessed cognitive abilities (SACAs) can be observed. Because study interest is in menstrual change in abilities, analysis will be carried out on the saved cosine amplitudes and not the raw scores. Since the questionnaire items have not been previously fitted onto a model, a data led EFA (exploratory factor analysis) will be the starting point. The factor structure suggested by EFA will then be tested using CFA (confirmatory factor analysis). It is hypothesized that the EFA suggested model will be a five-factor correlated model because theory and preliminary testing in Chapter 2 showed the different constructs of metacognition to be distinct yet related. The predicted model is presented in Figure 5.8. Three questions are expected to load onto each factor, according to which aspect of metacognition they pertain to.

5.5.2 EFA

For step one of finding a model of best fit, an EFA was carried out to see what model configuration would be suggested based on the data. Analysis was carried out on the saved cosine amplitudes and not the raw scores.

5.5.2.1 Assumption tests

The Doornik-Hansen test rejected normality (χ^2 (30) =358.052, p<0.001), so Principal Axis Factoring was chosen as the factor extraction method instead of Maximum

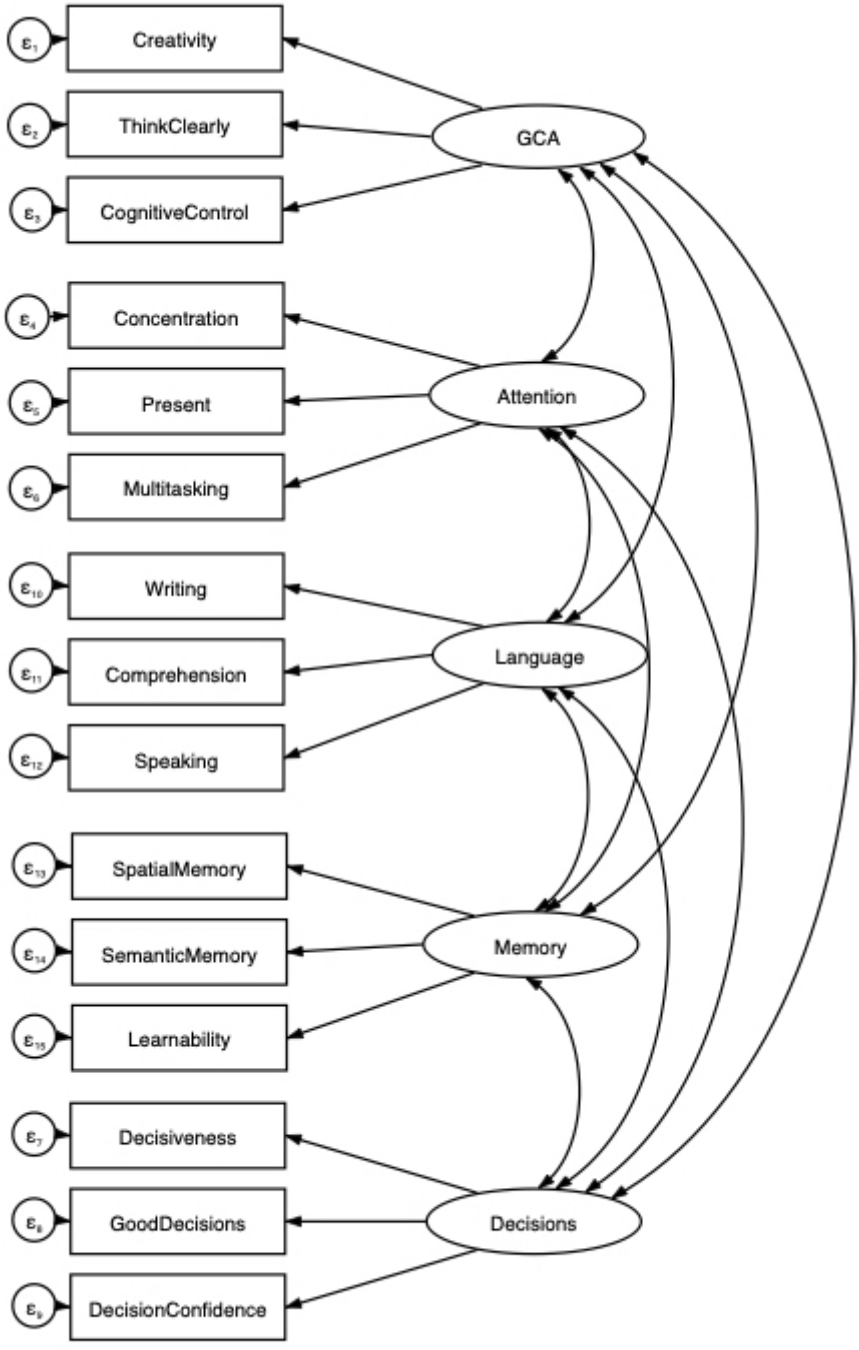


Figure 5.8: Predicted five-factor correlated model

Likelihood. The following assumptions were met: a higher case to variable ratio than the required minimum of 5:1 (8:1), a higher KMO than the required 0.6 (.956) and a significant Bartlett's test ($\chi^2(105) = 2214.27$, $p<.001$). The requirement of a determinant greater than 0.00001 was not met (3.180E-9), so one should proceed with caution as this suggests possible multicollinearity between variables. All communalities were above .2 (minimum loading was 0.74).

5.5.2.2 EFA Findings

Using the Kaiser (K1) criterion a 1-factor solution was identified based on the eigenvalue greater than 1 rule. This first factor accounted for 76.3% of the total variance. A rotated factor matrix could not be produced, since only one factor was extracted. The scree plot (Figure 5.9) demonstrated an "elbow" at two components, also favoring a one-factor solution. The model suggested by the EFA was therefore not a five-factor model, but a one-factor solution with all 15 variables loading onto a single latent construct.

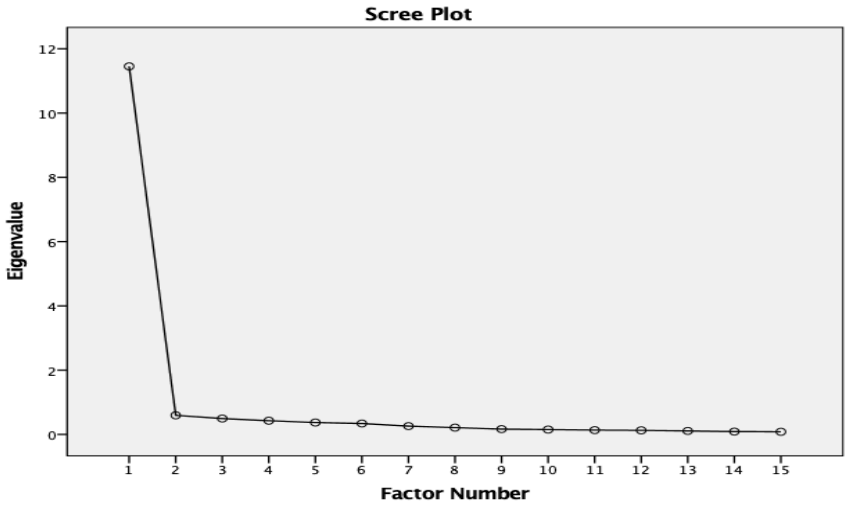


Figure 5.9: Scree plot for all 15 SACAs

5.5.3 CFA

For step two of finding a model of best fit, the one factor model suggested by EFA was tested by CFA. Analysis was carried out on the saved cosine amplitudes and not the raw scores.

5.5.3.1 Assumption tests

The multivariate normality assumption failed (χ^2 (30)=358.052, p<0.001), and so the more robust version of Maximum Likelihood with the Satorra-Bentler correction was used.

5.5.3.2 CFA Findings

The EFA model tested was a data driven one-factor model with all 15 variables loading onto one latent factor. Results presented in Table 5.3 and 5.4 show that the model performed moderately well. In order to see whether a better model fit could be found the initially predicted five-factor model with correlations among all five latent factors was also tested. This model performed only moderately well and did not show improvements over the one-factor model. This was surprising since theory had informed factor and variable loading structure and a priori hypothesis had predicted this model fitting well.

Table 5.3: Goodness of fit scores for five-factor and one-factor models

Model	χ^2	*df*	*p-value*	*CFI*	*AIC*	*RMSEA*	*SRMR*
One-Factor	148.815	90	<0.001	0.954	7756	0.074	0.033
Five-Factor Correlated	115.301	80	0.006	0.972	7716	0.061	0.029

CFI = comparative fit index, AIC = Akaike's information criterion, RMSEA = root mean square error of approximation, SRMR = standardized root means square residual

Table 5.4: Assessing goodness of fit for five-factor and one-factor models

Model	*p-value*	*CFI*	RMSEA		*SRMR*		*Model Fit*
	>.05	>0.95	<0.06	<0.08	<0.06	<0.08	*Decision*
One-Factor	✗	✓	✗	✓	✓	✓	Moderate Fit
Five-Factor Correlated	✗	✓	✗	✓	✓	✓	Moderate Fit

Model fit assessment based on how many fit index cut-off criteria are met. If the minimum threshold for all tests is met, models are described as having a "good fit", models with three out of four a "moderate fit" and those with two or less "a poor fit". Note: CFI = comparative fit index, RMSEA = root mean square error of approximation, SRMR = standardized root means square

In the five-factor model all five factors showed a high level of reliability: Attention (α=.92), Decisions (α=.94), Memory (α=.92), Language (α=.91), and GCA (α=.87). The one-factor model composed of 15 cognitive items also showed a high level of reliability (α=.97).

5.5.4 Aggregating variables

The "moderate" models fit all tests except for the chi-square (Table 5.4), a test notoriously sensitive to sample size (Gatignon 2010; Singh et al., 2016). The most complex model, the five-factor correlated (Image 5.8), had a case to free parameter ratio of 120:40, coming in at much lower than the accepted sample size of 5:1 minimum. In these situations, Bagozzi and Edwards (1998) proposed reducing the number of path coefficients by collapsing items within the same sub-scale in a technique known as parceling. By aggregating two or more items through averaging, you are working with composites of items instead of individuals, simplifying complex models while keeping the structure intact (Rocha & Chelladurai, 2012). Collapsing the items also helps solve the problem of multicollinearity seen in the EFA.

The three items making up each latent variable were collapsed, with their raw scores averaged and a coefficient extracted for each of the newly derived variables:

Attention, Decisions, Memory, Language and GCA (General Cognitive Ability). The model was therefore reduced from 15 items to 5. Pairwise coefficients show high shared variance between all variables (lowest shared - 67%). An exploratory factor analysis was carried out to determine whether all variables should load onto one latent factor or more.

5.5.4.1 Aggregated variables EFA: Assumptions and findings

The Doornik-Hansen test rejected the null hypothesis of multivariate normality of the data (χ^2 (10)=101.730, p<0.001). Principle Axis Factoring was therefore chosen as a factor extraction method in favour of Maximum Likelihood for this analysis. The case to variable ratio was higher than the required 5:1 (24:1) and the determinant greater than the required 0.00001 (0.006). The Kaiser-Meyer-Olkin test came in at .9 and Bartlett's was significant (χ^2(10)=803.7, p=<0.001). The diagonals of the anti-image correlation matrix were all above .5 and communalities all well over .2. Factor analysis assumptions were marked as met.

Using the Kaiser (K1) criterion a 1-factor solution was identified based on the "eigenvalue greater than 1 rule. This first factor accounted for 89.6% of the total variance. A rotated factor matrix could not be produced since only one factor was extracted.

The scree plot (Figure 5.10) demonstrated an "elbow" at two components, hence also favouring a one-factor solution. The one-factor model composed of the aggregated variables *Attention, Decisions, GCA, Memory and Language* showed a high level of reliability (α=.97). PA was not carried out since both the Kaiser and scree plot support the unidimensionality of the cognitive model.

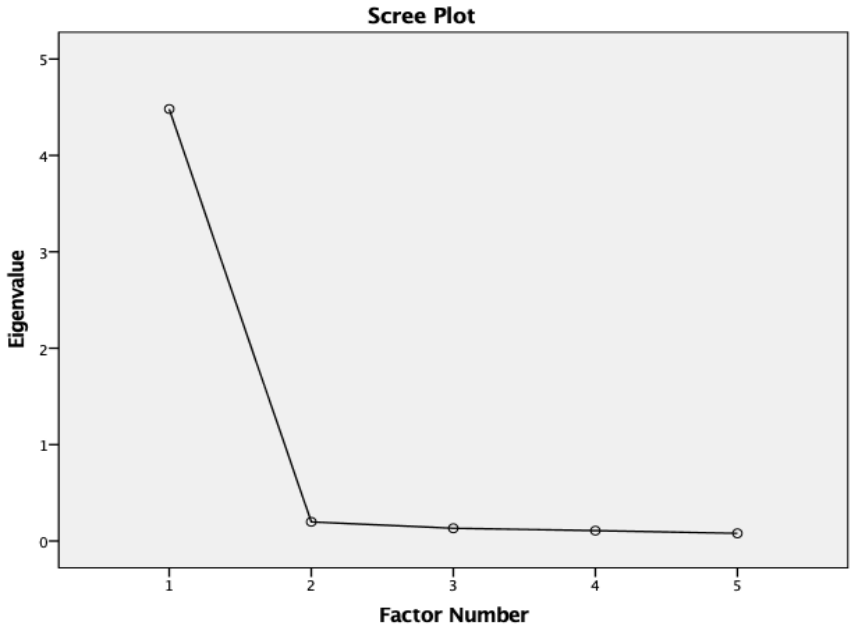


Figure 5.10: Aggregated variables scree plot

5.5.4.2 Aggregated variables CFA: Assumptions and findings

CFA was carried out on the one-factor set-up identified through EFA. The multivariate normality assumption failed (χ^2 (10)=101.730, p<0.001), and so the more robust version of Maximum Likelihood with the Satorra-Bentler correction was used.

Fit indices for the model show that it fits the data well (χ^2=7.813, *df*=5, p= 0.167; CFI=0.99; RMSEA=0.069; AIC=2384; SRMR=0.013). A comparison of this model's fit with previous models can be seen in Table 5.5.

Table 5.5: Comparing goodness of fit for collapsed model with previous models

Model	*p-value* >.05	*CFI* >0.95	RMSEA <0.06	<0.08	*SRMR* <0.06	<0.08	*Model Fit Decision*
Five-Factor Correlated	✗	✓	✗	✓	✓	✓	Moderate Fit
One-Factor All Items	✗	✓	✗	✓	✓	✓	Moderate Fit
One-Factor Derived	✓	✓	✗	✓	✓	✓	Good Fit

5.5.5 Cognition model of best fit

The best fitting model is a collapsed model, with five variables loading onto one factor. While the parceling of variables does remove the problem of multicollinearity, loadings as seen in Figure 5.11 are higher than expected. Correlation coefficients between cognitive variables are all above .9. In comparison, physical correlation coefficients range from .56-.76, and affective ones from .70-.87 (see Figure 4.10). Plausible explanations for such factor loadings among cognitive variables are given in the discussion.

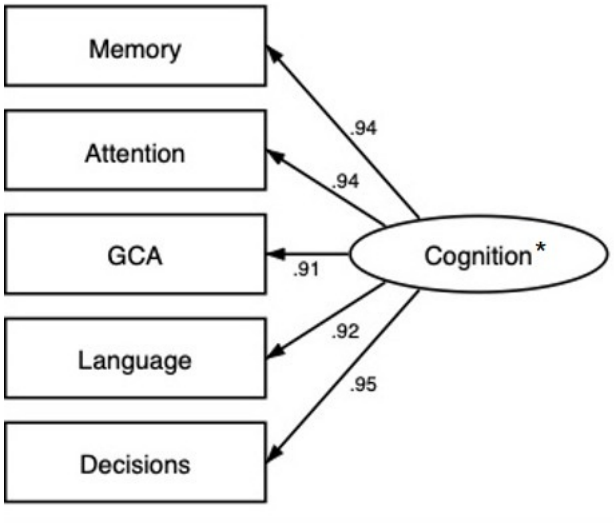


Figure 5.11: Cognition model of best fit. Cognition is self-assessed.

5.6 Structural Equation Modelling II: Cognition, Physical, Affective

5.6.1 Outline

The next series of SEMs will model the underlying relationship between self-assessed cognitive abilities and the other symptoms, physical and affective. Since no similar model exists, the model configuration to be tested will be provided by an EFA. The best fitting cognition model - a one-factor model - was established in the previous section. Similarly, we have also modelled the relationship between physical, affective and headache symptoms, where a three-factor model was confirmed in Chapter 4. Based on this model as well as literature, one can hypothesise that symptoms will load onto four distinct, yet related, latent factors. A correlated four-factor model made up of the factors Physical, Affective, Headaches and Cognition will therefore be tested and compared to the model suggested by the EFA, provided that they have different configurations.

5.6.2 EFA

For step one of finding a model of best fit, an EFA was carried out on all cognitive, physical and affective variables to see what model configuration would be suggested based on the data. As in the previous section, analysis was carried out on the saved cosine amplitudes and not the raw scores.

5.6.2.1 EFA assumptions

The Doornik-Hansen test rejected normality (χ^2 (26) = 157.442, p<0.001) and so Principal Axis Factoring was chosen as the factor extraction method instead of Maximum Likelihood. The following assumptions were met: a higher case to variable ratio than the required minimum of 5:1 (9:1), a higher KMO than the required 0.6 (.89) and a significant Bartlett's test (χ^2 (78) = 1208.3, p<0.001). The requirement of a determinant greater than 0.00001 was met (2.455E-5). All communalities were above .2 (minimum loading was 0.302 for *Headache*), so no variables were removed from the analysis.

5.6.2.2 EFA Findings

Using the Kaiser (K1) criterion a two-factor model was suggested with the first factor accounting for 48.5% of the variance and the second factor 16.4%; cumulative total of 64.9%. All other presented factors were discarded since they had eigenvalues below 1. Varimax rotation was used to rotate loadings and make the readings clearer to interpret. Items were sorted by size and those loading at lower than .5 were suppressed. The Rotated Component Matrix (Table 5.6 below) shows all affective and cognitive variables loading onto factor one and all physical variables and headache loading onto factor two. The scree plot (Figure 5.12) also favored a two-factor solution, by demonstrating an "elbow" at three components. PA was not carried out since the Eigenvalue and scree plot tests were in agreement.

Table 5.6: Rotated Component Matrix suggesting Two-Factor Model

Variable	Component	
	1	2
Depressed	.624	
Anxiety	.637	
MoodSwings	.565	
GI		.737
BackJoint		.809
Cramps		.727
Breast		.613
Headache		.545
Attention	.916	
Decisions	.933	
Memory	.925	
GCA	.922	
Language	.921	

Extraction Method: Principal Axis Factoring. Rotation Method: Varimax. Loadings blank at <.5

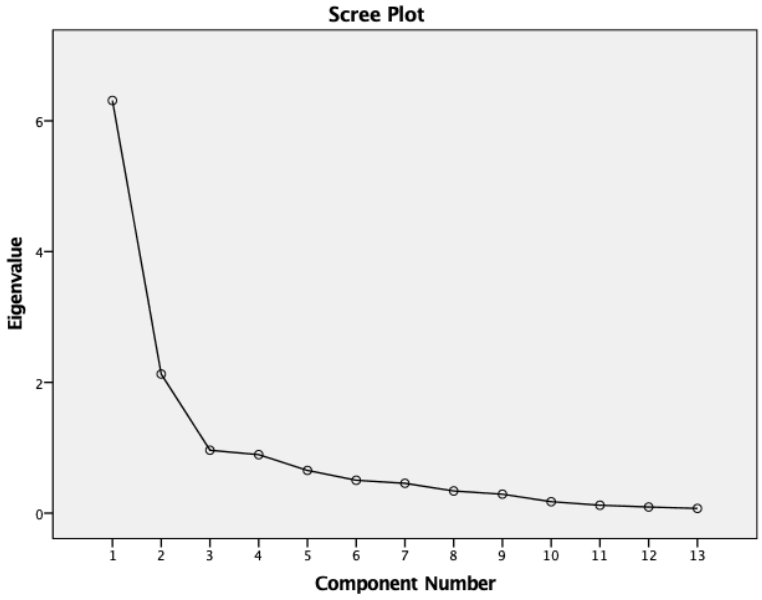


Figure 5.12: Scree plot for cognitive, physical and affective changes

5.6.3 CFA

The CFA model to be tested was a two-factor correlated model suggested by the EFA with all affective and cognitive variables forming one construct and all physical and headache symptoms forming another.

5.6.3.1 CFA assumptions

The Doornik-Hansen test rejected normality of the data (χ^2 (26) = 157.442, p<0.001) and so a Maximum Likelihood estimation method with robust standard errors was employed together with the Satorra-Bentler correction.

5.6.3.2 CFA findings

The EFA suggested two-factor model, did not fit the data well. Results are presented in Table 5.7. As seen in Table 5.8 the criteria were met by none of the tests. Following this, a second model was tested - the hypothesized, four-factor correlated model. This model demonstrated a good fit, meeting all test criteria. The four-factor model composed of three affective items (α=.82), four physical items (α=.70), five cognitive items (α=.97), and one item for headaches shows a high level of reliability.

One could attempt further refinements to improve the model fit at this point (aiming to meet the more stringent SRMR cut-off), but it is not advisable to arbitrarily

Table 5.7: Goodness of fit scores for the two-factor and four-factor models

Model	χ^2	*df*	*p-value*	*CFI*	*AIC*	*RMSEA*	*SRMR*
Two-Factor Correlated	138.134	64	<0.001	0.913	2251	0.099	0.094
Four-Factor Correlated	76.177	60	0.078	0.981	2181	0.048	0.052

CFI = comparative fit index, AIC = Akaike's information criterion, RMSEA = root mean square error of approximation, SRMR = standardized root means square residual

Table 5.8: Assessing goodness of fit for two-factor and four-factor models

Model	*p-value* >.05	*CFI* >0.95	RMSEA <0.06	<0.08	*SRMR* <0.06	<0.08	*Model Fit Decision*
Two-Factor Correlated	✗	✗	✗	✗	✗	✗	Poor Fit
Four-Factor Correlated	✓	✓	✓	✓	✗	✓	Good Fit

increase model complexity. As with most complex models, the four-factor model already has a high number of free parameters (31). This brings the sample size to free parameter ratio (3.8:1) below the suggested 5:1 (Bentler & Chou, 1987), which is not ideal. It is important to remember, however, that this is a general rule of thumb. Recent simulation research, suggests, in fact, that the N:q ratio is an oversimplification and that size requirements depend on a number of factors including normality of the data, model complexity, estimation methods used as well as loading magnitude (see reviews by Hancock, 2013; Jackson, 2007; Lee, Cai, & MacCallum, 2012). Sideridis et al., (2014), as an example, found that a sample size of 50-70 would be enough for a model of functional brain connectivity, provided that it involved at least four latent variables.

5.6.4 Model of best fit

A visual representation of the final model of best fit is presented in Figure 5.13.

When modelling the relationship between self-assessed cognition, and physical, affective and headache symptoms in the menstrual cycle, the best-fit model is a four-factor model with correlations between all latent factors. Presented in order of decreasing magnitude, latent constructs share the following variances: Cognition and Affective 50%, Physical and Affective 32%, Headaches and Physical 22%, Physical and

Cognition 14%, Headaches and Affective 13%, and Headaches and Cognition 3%. All constructs are significantly correlated except for Headaches and Cognition (p=0.108).

As far as menstrual cycle symptom change is concerned, it is clear that symptoms are related. Of particular interest is the strong link between affective and metacognitive changes, as well as the correlation between physical symptoms and headaches, which was not significant in Kiesner et al. (2016). Despite this, symptoms still retain an element of distinctiveness. This can be seen in the differing loading coefficients for the different constructs, which are much higher for cognitive and affective symptoms than for physical symptoms.

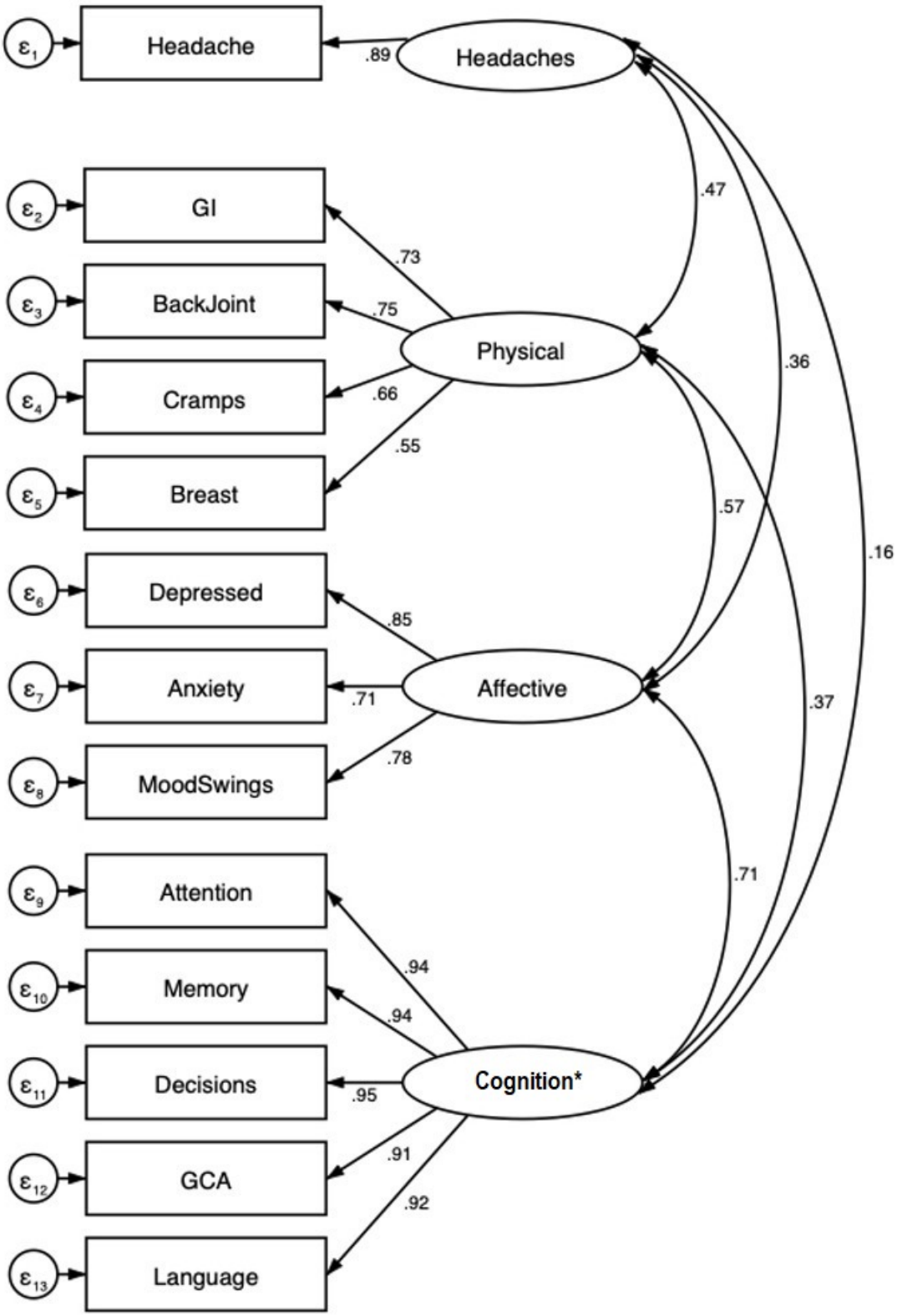


Figure 5.13: Model of best fit for headaches, physical, affective and cognitive constructs. Cognition is self-assessed.

5.7 Chapter summary

This chapter looked at metacognitive patterns in the context of the menstrual cycle; specifically individual differences in manifestations and covariations in self-assessed cognitive abilities. The majority of analyses were carried out on five main constructs of cognition: Attention, Decision Making, Language, Memory and GCA as well as the fifteen lower-level items tapping into those constructs. Results demonstrated metacognitive cyclicity, lack of variation between cognitive domains, and variation in the strength and direction of change cyclicity across individuals. Metacognitive changes were examined against other menstrual complaints, with factor structures indicating distinct but significantly correlated cognitive, headache, affective and physical constructs. The results presented here will be discussed in further detail in the discussion (Chapter 6).

Chapter 6

Discussion

6.1 Chapter Overview

The present work sought to investigate metacognitive fluctuations with respect to the menstrual cycle in young, naturally-cycling women using daily self-assessments of multiple cognitive abilities and analysed within and across individuals across multiple menstrual cycles. Results confirmed the existence of metacognitive variations in time with the menstrual cycle and revealed significant individual differences and covariation between cognitive abilities and physical and affective symptoms. The scope of this chapter is to analyse the findings presented in the previous two chapters in light of existing literature (Section 6.2). The outcome should be a comprehensive understanding of the metacognitive profile in the context of the cycle, to bridge the current knowledge gap in this unexplored domain. Insight is provided into menstrual metacognitive patterns in terms of direction and magnitude, individual differences, biases, groupings with other symptoms, specificity of modality, and possible underlying pathways. Replicability is tested by comparing findings in the physical and affective domains with those presented in Kiesner et al., (2016) where the same methodology was used. Section 6.3 lists the study's limitations, and the implications of the findings are outlined in section 6.4. Concluding remarks are presented in the last part of the chapter in Section 6.5.

6.2 Analysis of results

6.2.1 Metacognition oscillates in time to the menstrual cycle

The contention over whether the menstrual cycle has an effect on cognition has already been dissected (see Chapter 2). Among inconsistent and contradictory results, the most recent comprehensive reviews suggest no consistent association (Sundström-Poromma & Gingnell, 2014; 2018). In the present study, metacognition shows clear cyclicity in relation to the menstrual cycle. All reported abilities follow the same trajectory – a rapid peak after cycle onset, followed by a gentle decline throughout the cycle, with a sudden drop at the start of menstruation. At the beginning of the second cycle the slope rapidly increases and repeats the pattern. This M-shaped trend signifies mid-cycle improvement and perimenstrual decline in reported cognitive ability; a result demonstrated by the overall negative amplitude coefficients showcased by all variables.

A link between metacognition and the menstrual cycle in the absence of a conclusive effect of the cycle on cognition (Sundström-Poromma & Gingnell, 2014; 2018) proffers two important points. First, that metacognition may – at least to some degree – operate separately from cognition. This has already been evidenced in other populations; e.g., in cancer patients undergoing chemotherapy, where perceived cognitive impairment is reported as a main symptom but objective cognitive performance is not affected (Diane & Tallman, 2015; Rosa et al., 2021; Oerlemans et al., 2022). The claim for a dissociation is substantiated by both neural and computation evidence (Shimamura, 2000; Timmermans et al., 2012; Fleming et al., 2012; Vaccaro & Fleming, 2018). Secondly, the menstrual cycle's impact on metacognition suggests that objective cognitive testing alone may not provide a comprehensive understanding of the intricate interplay between the menstrual cycle and cognitive processes. Metacognitive testing

taps into dimensions that are inaccessible to cognitive tests such as personal beliefs, attitudes, cultural taboos, social norms, and concurrent symptom effects – all which shape how women experience and manage their menstrual cycle (Kowalski & Chapple, 2000; Morgan & Rapkin, 2002, Marvan & Cortes-Iniestra, 2008; Feinberg, 2020). Similarly, DSM-V listed PMDD symptoms such as '*subjective* difficulty in concentrating' and '*subjective* sense of feeling out of control' suggest the existence of factors that extend beyond objective performance. In a world where objective testing is considered the 'ground truth', it is imperative to recognise the complexity of processes like the menstrual cycle, where controlled testing may have limited external validity and may not be reflective of the real experience. This is especially important when considering that findings have real life implications; such as informing policies like menstrual leave and in education and awareness campaigns focused on reducing the stigma surrounding menstruation.

6.2.2 Distinct but correlated

To elucidate the causal mechanisms underlying menstrual metacognitive ability, it is necessary to examine symptom comorbidity with other physical and psychological symptoms. Although the existence of concomitance alone may not establish causation, the lack of an association would certainly disconfirm it. Existing research has already established co-presence between metacognitive symptoms and physical, affective and behavioural complaints when examining the factor structures of PMS and PMDD (Teng et al., 2004; Wang et al., 2007). The present study, however, is the first to show the covariation of metacognitive cyclical *change* with other symptoms, demonstrating that symptoms do not simply covary across time, but are also synchronized with the menstrual cycle.

Results show that cognitive, affective, physical and headache variables load

onto four respective individual factors which correlate significantly between themselves —indicating that the latent factors are distinct yet interrelated. Headaches and cognitive changes are exceptions to this pattern as they did not exhibit a significant relationship between their cyclicity, suggesting different causal pathways for the two. This is not surprising, considering that menstrual headaches have consistently been shown to follow distinct patterns to other menstrual symptoms as discussed in detail in section 6.2.7 below (Kiesner & Martin, 2013; Kiesner, 2016).

The correlation between physical and cognitive changes is weak, indicating that they are somewhat related, but share only a small amount (14%) of their variance. On the other hand, affective and cognitive changes are highly correlated, with 50% of their variance being shared. The strength of this relationship is supported by other results in the present study, including similar loading coefficients, similar within symptom-type homogeneity (section 6.2.3) and similar intraindividual variance patterns (section 6.2.5). Based on these findings it is therefore more likely that affective and subjective cognitive changes share common causes (e.g., hormonal changes, sociocultural effects), whereas the relationship between physical and reported cognitive abilities may be influenced by other factors.

6.2.3 Domain-general or domain-specific

Literature is divided as to whether metacognition is domain-general – where results on one variable is a good indicator of results on another (Ais et al., 2016; Faivre et al., 2018; McCurdy et al., 2013; Song et al., 2011) – or domain-specific, where metacognitive processes are particular to a certain domain or area of knowledge and success in one is not predictive of the other (Baird et al., 2013; Fitzgerald et al., 2017; Garfinkel et al., 2016; Morales et al., 2018). A recurrent finding throughout this present study is

the lack of heterogeneity in the magnitude and direction of change across the menstrual cycle demonstrated by the different self-assessed cognitive profiles. This is seen in the similar time-plot trajectories, raw score means and spread, and cosine amplitude scores and distributions. The exceptions are *Concentration* and *Cognitive Control,* in that their raw scores show a mean and median at the 50 mark as opposed to the other variables, which all have means higher than 50. Their cosine amplitude profiles, however, match the others' – with all variables demonstrating both positive and negative coefficients across individuals (both premenstrual and midcycle increases), suggestive of the same type of cyclical change. As such, cognitive ability cosine amplitudes demonstrate high correlation with 67-72% of the variance shared.

This low discrimination between cognitive changes hints at a shared metacognitive process system for: general cognition, attention, memory, decision-making and language, the constructs that were tested here. This conclusion is supported by SEM findings where parcelled cognitive variables all loaded onto one factor, and factor loadings were high (upwards of .91). An argument can therefore be made for a shared etiological pathway for menstrual cognitive complaints and improvements – biological or otherwise. This is especially the case when one compares reported cognitive changes to physical and affective symptoms which show a much higher rate of variability amongst themselves (Kiesner et al., 2011; Kiesner et al., 2016).

It is important to stress that the findings in this study do not refute the possibility of domain-specific systems. For example, it is possible that domain-sharing occurs across different types of cognition as long as one is judging self-competence (as was tested here) but that a different system is utilized when judging others, or that the same domain is accessed for all judgements of performance (as tested here) but that a different one is employed for judgements of agency (David et al., 2012). Indeed, a number of

authors believe in the co-existence of both domain-general and domain-specific systems (Fleming et al., 2014; Morales et al., 2018; Valk et al., 2016; Fielder et al., 2019).

6.2.4 Metacognitive bias

When making metacognitive judgements, a positive metacognitive bias is the tendency to give overly high confidence ratings to one's performance (Fleming & Lau, 2014; Charles et al., 2020; Dunn et al., 2021). The higher the overestimation of ability, the lower the metacognitive sensitivity, i.e., the extent to which the judgement accurately predicts task performance (Siedlecka et al., 2016). Findings from this study suggest possible systematic metacognitive bias for reported menstrual cognitive abilities based on a tendency to report high scores across almost all cognitive items. For the majority, item score means and interquartile were positively mean biased (M = 52.65, SD = 1.2; t (28) = 8.5, $p<0.001$) with an overall positive skewness indicated by a higher mean than median. This means that users reported heightened cognitive ability, rather than the expected neutral ability, on most days. The cause of the observed metacognitive bias is unknown. One possibility is that people overestimate their cognitive abilities due to a frame-of-reference shift where symptom free menstrual days appear better than they actually are. A similar phenomenon is seen in patients who overestimate how healthy they were pre-illness due to such response shift effects (Lindberg et al., 2017; Hinz et al., 2021; Ham et al., 2020). Another possibility is fulfillment of pre-existing menstrual beliefs. Attitudes, stereotypes and expectations have repeatedly been associated with biases in retrospective reporting of menstrual abilities (Marvan & Cortes-Iniestra, 2001; Kowalski & Chapple, 2000; Morgan & Rapkin, 2002; Feinberg, 2020).

Alternatively, one cannot discount effects external to metacognitive bias. For example, positively skewed scores might be a result of careless response. Cognitive

abilities were self-scored on a 100-point Visual Analogue Scale (VAS), with the anchors 'Worse than usual' and 'Better than usual' at the minima and maxima. The slider starting point was at the middle of the scale. It is possible that users swiped right (towards 'Better than usual') because it was more convenient. This is plausible when considering that the majority of users submitted their data through mobile phones and were likely right-handed. Participants also couldn't progress unless the sliders were moved, justifying deviation from the 50-point 'Same as usual' starting mark. Additionally, the VAS, being a continuous scale, typically discourages systematic biases (e.g., end-aversion and social desirability bias) since it is harder to estimate which value is expected on the scale continuum (Hasson & Arnetz, 2005; Klimek et al., 2017). This is particularly the case when number markers are hidden, as in this study. With this phenomenon however, one would expect to see the effect across all variables. With *Concentration* and *Cognitive control* demonstrating means around the 50-point mark and showing significant differences to all variables ($\chi^2(14) = 762$, $p = <0.001$), consistent systematic bias due to carelessness becomes doubtful.

6.2.5 Individual variation

One of the long-standing issues in the field of menstrual research is the lack of deep understanding of heterogeneity in symptom changes across the menstrual cycle. This is due to the tendency of studies to overlook individual changes and instead focus on average effects (reviewed in Sundstrom Poromaa and Gingnell, 2014, 2018; Schmalenberger et al., 2021). The neglect of individual differences can lead to an underestimation of effect sizes and the acceptance of false null hypotheses, particularly when dealing with small sample sizes. Because of this, not testing for individual differences has been cited as the primary contributor to the conflicting and inconclusive results observed in menstrual cycle research (Schmalenberger et al., 2021). One strength of the present

study lies in its use of multilevel modelling to incorporate both average effects (fixed effects) and individual differences in response to the cycle (random effects). Results indicate that for all symptoms and abilities, there is a contribution of average effects by the menstrual cycle, but a significant model improvement when factoring in individual differences. For physical changes, the impact of individual differences in cyclicity is approximately twice as large as the impact of average effects. The same improvement is four-fold for cognitive and affective changes. These findings highlight just how heterogeneous menstrual symptom changes are, and align with research indicating differences in the timing, duration and intensity of menstrual symptoms experienced (Shoep et al., 2019; Kiesner, 2011; Lustyk et al., 2008; Kiesner, 2017) as well as differences in elicited treatment response (Halbreich et al., 2006; Halbreich, 2008; Freeman et al., 2011). Possible contributors to individual variations include differences in: hormone levels at the same point in the cycle (Sundstrom Poromaa & Gingnell, 2014), hormone sensitivity – e.g., due to varying amounts of neurotransmitters and different genotype (Kiesner, 2017; Jacobs & D'Esposito, 2011; Smith et al., 2014), metacognitive sensitivity (Fleming, 2014) and menstrual beliefs; the latter based on individual attitudes, past experiences, and cultural factors (Brooks-Gunn & Ruble, 1986; Obermeyer & Sievert, 2007; Chrisler & Gorman, 2019). The results of this study demonstrate the exciting potential for within-person individual variation in metacognition, highlighting the need for further research to fully identify and understand these variations.

Findings also offer valuable insights into the potential relationships between variables as determined by the degree of explained random variation. While these relationships, and the possible explanations for them, have been discussed in detail in other sections (see 6.2.2, 6.2.7), comments related to individual differences in these associations will be made here. Psychological symptoms and cognitive abilities share extremely similar variance patterns highlighting once again the possibility that the causal

mechanisms of the two are linked to some extent. Both symptom groups show the same level of improvement when factoring in within-person changes and a relatively low fraction of variance attributable to both models when compared to physical symptoms (M=1.25±0.4; M=4.62±0.7 for fixed and random respectively). An exception is mood swings whose variance pattern is more in line with physical than psychological symptoms. Interestingly, headaches show the opposite pattern, aligning more strongly with psychological than physical symptoms in terms of individual variation. Literature shows that physical symptoms (specifically abdominal cramps) are in fact more strongly associated with mood swings than with other affective symptoms (Kiesner et al., 2016) and headaches correlate more with depression/anxiety symptoms than with physical symptoms, even when different headache patterns (mid-cycle vs. perimenstrual) are displayed (Kiesner, 2016; Kiesner & Martin). This relationship can be attributed to the temporal nature of these symptoms. Mood swings are transient and fleeting, making them more similar to some physical symptoms like cramps than to psychological symptoms like depressed affect. On the other hand, the link between headaches and anxiety/depressed symptoms may be due to their relative stability across hours or days.

6.2.6 Underlying pathways

Although the current research is not enough to determine the causal mechanisms underlying metacognition and the menstrual cycle, findings indicate a number of plausible pathways that warrant further investigation.

6.2.6.1 Direct hormonal influence

The first possibility is a direct effect of steroids. Results demonstrating mid-cycle improvement and pre-menstrual worsening in ability show a similar trajectory to estro-

gen fluctuations from menstruation to ovulation. Metacognition sensitivity to gonadal steroids is well established (Hauser et al., 2017; Livermore et al., 2021).

6.2.6.2 Metacognitive beliefs

Alternatively, the present study's result of perimenstrual worsening could be a reflection of the generally negative, global and personal, attitudes and beliefs surrounding menstruation (Delaney et al., 1976; Sommer, 1992). In their Self-Regulatory Executive Function model, Wells & Matthews (2009) hypothesise that distorted external beliefs are internalised as destructive metacognitions (self-beliefs) which are perpetrated through maladaptive cognitive coping practices such as hyper-vigilance and rumination and which in turn have an effect on the perception of menstrual symptom presence and severity. Evidence supports the role of metacognitive beliefs in menstrual conditions, including premenstrual increases in self-focused attention and rumination (Oxley, 1998; Craner et al., 2016; Li et al., 2020) and the efficacy of cognitive therapy in treating PMS and PMDD. (Lustyk et al. 2009; Kleinstäuber et al., 2012; Kancheva et al., 2021). The current study suggests that future research should directly assess metacognitive beliefs and control for their potential influence.

6.2.6.3 Symptom mediation

Findings of cognitive changes co-varying with physical and affective symptom changes provide support for symptom-mediated menstrual metacognitive routes. Of the physical changes, the one that correlates most strongly with cognitive change is gastrointestinal discomfort. This is not surprising, considering the intricate bi-directional interactions within the gut-brain axis (reviewed in Martin et al., 2018). Gut tissue has been linked to cognitive dysfunction through multiple pathophysiological mechanisms such as: the direct activation of the vagus nerve through infection (Moreno-Valladares et al., 2022),

inflammation via gut microbiota (Rundek et al., 2021), and through the production of gut-derived neurotransmitters (Rea et al., 2016).

Cognitive and affective changes demonstrate a strong relationship (50% shared variance). A large amount of literature in the field attests to the idea that at least some portion of this shared variance is due to an affective mediating factor. Affective stressors have been shown to alter the immune system via the production of proinflammatory cytokines which cross the blood-brain barrier and interfere with high-level cognition and metacognition (Peters et al., 2019; Kindler et al., 2020; Bourgognon, & Cavanagh, 2020; Muscat & Barrientos, 2021). Evidence of symptom mediation can be seen in studies where subjective memory complaints (SMCs) are consistently related to anxiety and depressive symptoms, but hardly any link is found between SMCs and objective cognition (Ford et al., 2004; Weber & Mapstone, 2009; Reid, 2006; Pullens et al., 2010; Jansen, 2013; Kuba et al., 2019). It is also seen in studies where menstrual cognitive effects in PMS/PMDD populations are notoriously inconclusive – except for when studying tasks with a cognitive-emotional basis (reviewed in Le et al., 2020; Eggert et al., 2017; Slyepchenko et al., 2017).

Going in the other direction, there is the possibility of an unexplored pathway whereby metacognition mediates the effect of hormones on affective symptoms. The metacognitive processing (monitoring and regulation) of both cognition and emotion occurs in the same brain regions i.e., the amygdala, PFC, ACC and insula (Sacher et al., 2013; Jonkman et al., 2021; Landin et al., 2019) and so a hormonal effect on metacognitive emotion regulation may influence affective symptoms. Behavioral evidence has shown that women who struggle with emotional regulation may experience more severe menstrual symptoms, and that the interplay between sex hormones and emotion regulation strategies can influence negative affect (Mahon et a., 2015; Manikandan et

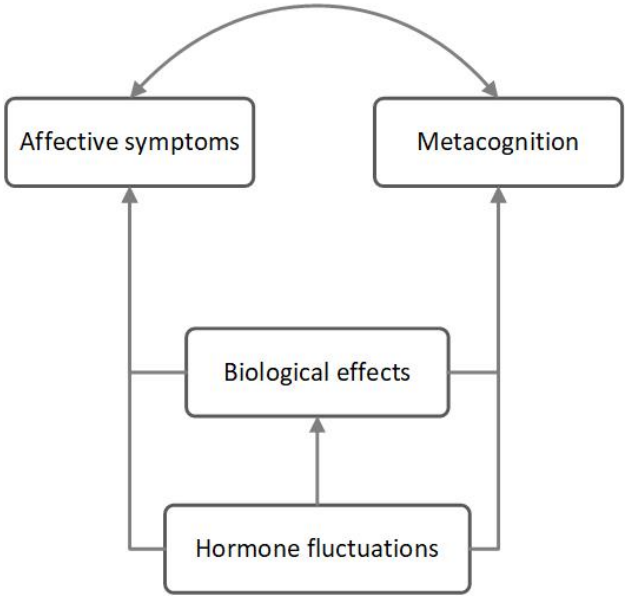


Figure 6.1: Metacognition as a mediating factor for affective symptoms. Schematic diagram showcasing the potential role of metacognition as a factor mediating either direct hormonal influence or else a secondary biological effect (e.g., inflammation) on affective symptoms. The pathway between affective symptoms and metacognition is posited as being bidirectional.

al., 2016; Graham et al., 2017). While this thesis does not directly test for this hypothesis it does provide a foundation for future research by demonstrating two key findings: 1) an influence of the menstrual cycle on metacognition and 2) an association between cognitive and affective changes. Thus, metacognitive emotion regulation may be an important factor to consider when studying emotional experiences and menstrual symptoms in women, with this avenue being an opportune area for future investigation. A schematic highlight both affective symptoms and metacognitive changes as mediating factors and the bidirectional pathway between them is presented in Figure 6.1.

Symptom-mediated pathways can also arise from the mere experience of a symptom influencing others, as proposed by the physical distress model (Kiesner et al., 2016; van Iersel et al., 2016). Physical pain/discomfort has been shown to affect

mood and metacognitive judgments (Riva et al., 2011; Allen et al., 2016; Hauser et al., 2017). Alternatively, psychological symptoms may possibly cause physical symptoms to feel more intense (Kiesner, 2017). Evidence for the physical distress model could be demonstrated via stronger correlations between psychological symptoms and physical symptoms that cause the greatest physical discomfort (Kiesner & Pastor, 2010). Findings from this study support this hypothesis, as the strongest associations with psychological variables are demonstrated by gastrointestinal discomfort and pain in the back and joints, and the weakest is with skin, which arguably demonstrates less severe physical discomfort. However, no significant association was found between headaches and cognition, casting doubt on the physical distress model as a cause of cognitive changes. It is important to note that these findings are limited to the domain of metacognition and do not necessarily reflect the complex interplay between physical and affective symptoms in influencing each other psychologically.

Finally, reported changes in cognition may be due to behavioral and developmental changes that are not directly determined by the menstrual cycle but are long-term consequences of menstrual symptom stressors. Disrupted participation and engagement in important developmental activities have been found to contribute to maladaptive psychological adjustments and coping practices as well as destructive metacognitions (Kiesner, 2017; Wells & Matthew, 2009). The effect of changes in lifestyle and social engagement on the menstrual experience are well established (Phelan et al., 2020; Robb et al., 2020; Lebar et al., 2022).

While the causal mechanisms discussed in this section are presented as separate entities, reality is likely much more complex. Outcomes are expected to be a consequence of numerous cause-and-effect events, feedback loops, and interconnected pathways that interact with one another in multiple directions.

6.2.7 Replicability of findings

This present study replicated Kiesner et al.'s (2016) methodological design, specifically examining whether findings from the original study regarding physical and affective cyclical patterns could be reproduced in a sample from a different nationality. Results show that a number of findings were successfully replicated.

In relation to cyclical patterns of symptoms, as in the original study, all physical and affective variables showed a general group tendency towards reduced symptoms during the mid-cycle phase and increased symptom reporting during the pre-menstrual phase. Despite this collective tendency however, the present study corroborates original findings that 1) variations exist between symptoms in terms of magnitude and direction and 2) that these symptoms can be grouped based on patterns of occurrence. Specifically, in both studies *Depressed, Anxiety* and *Headache* have the highest percentages of negative cosine amplitudes across individuals and *Breast* and *Cramps* have exclusively positive amplitudes. This suggests that some symptoms peak only during the perimenstrual phase, and some, like headaches and mood changes, demonstrate both premenstrual and midcycle increases across participants. This finding is significant because it suggests that these two types of symptoms may have different underlying causes or mechanisms.

With regards to individual variation, the present study confirms the findings of the primary research, which showed that individual differences in within-person cyclicity account for around twice the amount of variance as compared to the average effects. The effect sizes for both studies were similar and symptoms followed the same order of magnitude in both studies with *Anxiety* on the low end of the spectrum and *Breast* showing the highest effect sizes. These results reinforce the idea that there is considerable variation among individuals in physical and affective symptom change, and

highlight the need to factor in individual differences in cyclical patterns, rather than solely relying on average levels.

Finally, the present study supports the correlated three-factor model in Kiesner et al. (2016), which contends that headaches, physical symptoms, and affective symptoms are best represented as distinct but interrelated factors. Consistent with previous research, the present findings demonstrate that affective symptoms exhibit greater internal consistency compared to physical symptoms, as reflected in their loading coefficients. This may be attributed to the differential neurological underpinnings of these symptom domains, with physical symptoms being rooted in diverse tissues and affective symptoms having a more limited neurological basis (Kiesner & Pastore, 2010). Results underscore the validity of the framework proposed by Kiesner et al. (2016) and provide further insights into the nature of physical and affective symptoms.

Some results were not reproduced in the current study. To begin with, the residual correlation between observed variables *Cramps* and *Mood Swings,* as suggested by the modification indices in the original paper, did not improve the model, though an acceptable fit was still produced. Rather, outside of the latent constructs, *Mood Swings* tended to be most highly correlated with *GI* as discussed in section 6.2.5 above. While the specific symptoms differ, the claim for a link based on symptom temporality still holds. This paper proposes that in addition to menstrual cramps (van Iersel et al., 2016; Kiesner et al., 2016), any physical symptom of acute and transitory nature can potentially act as an explanatory model for mood swings.

Another replicability issue was the variable *Skin,* which did not successfully fit the model, in contrast to data presented by Kiesner et al., (2016). *Skin* did not load onto the three-factor solution and had to be removed as it did not correlate with any other variable except *Breast* and *Headache*. There are several reasons that could explain

why *Skin* might act differently to the other physical variables. Firstly, the effects of the menstrual cycle on skin may be more complex and dynamic, for example causing separate changes in oil production, hydration levels, and inflammation (Stoll et al., 2001; Arora et al., 2010; Raghunath et al., 2015). Therefore, it is likely that skin shows more between person variation, as demonstrated by the higher ICC values. Some people may experience an increase in acne breakouts, while others may experience dryness, oiliness, or sensitivity (Shah et al., 2001). Secondly, skin, like headaches, exhibits small magnitudes compared to other symptoms and is more stable in terms of within-cycle change (also supported by the high ICC values; see table 4.1). This is because skin changes are not particular to the menstrual cycle and are highly influenced by external factors (e.g., environmental stressors). Third and finally, like breast pain, skin shows a different temporal profile to symptoms such as cramps, with a premenstrual increase rather than perimenstrual increase, indicated by the time-plot (see figure 4.2). More research is needed to determine whether skin would provide a better model fit if loaded onto its own latent factor.

One other notable discrepancy between the current replication and the original investigation was observed in the correlation between physical and affective symptoms. Specifically, the current study revealed a stronger correlation between these symptom domains, with a shared variance of 32%, compared to the original study's shared variance of 18%. This finding hints at a stronger underlying link between physical and affective symptoms than previously hypothesized, potentially implicating shared physiological or psychological mechanisms underlying these symptom domains.

Finally, the replication exhibits an association between headaches and physical symptoms, which was not found in the original study. This finding challenges the primary assumption of independence between the two constructs and the claim of no

shared pathways between them. What is clear in both studies is that headaches operate differently from other physical symptoms. In the current research, when the headache variable was tested as part of a two-factor model, loading onto the physical latent construct, the model did not fit well. This suggests that although the constructs are correlated, the different symptom types are distinct and require separate latent constructs. More research looking into the relationship between physical symptoms and headache menstrual changes are needed.

Taken together findings support the original conclusion that individual differences exist in cyclical changes of various physical and emotional symptoms, and that their associations depend on symptom specificity (Kiesner et al., 2016). The replicability of the results strengthens the credibility and validity of the findings, decreasing the likelihood of chance occurrence and increasing the confidence in their applicability to other contexts. Care however, should still be taken with potential generalizability, as discussed in detail in section 6.3.

6.3 Limitations

This study made significant effort to uphold high standards in menstrual cycle research. Measures taken included: accounting for intraindividual differences, measuring across more than one cycle, collecting data daily rather than periodically, and having a large sample size. A number of study limitations however still need to be acknowledged.

One constraint of this study pertains to the limited scope of data collection, which only spanned over two menstrual cycles. This decision was driven by pragmatic considerations to optimise participant retention. Although this duration is enough to establish the presence or absence of menstrual cycle effects on metacognition, a longer-term data collection scheme covering multiple cycles would allow for more statistically rigorous intra-individual outcomes and would help mitigate the potential confounding influence of random, non-cyclic effects. It is therefore recommended that future studies aim to extend data collection beyond two menstrual cycles, if practicable.

Also driven by pragmatic considerations and real-world constrains, was the decision to not incorporate any neurophysiological assessments (e.g., hormonal or inflammation measures). Although these would have helped verify cycle phases and augment the understanding of underlying mechanisms, their use did not pay off against researcher and participant burden and their omission did not compromise results. Likewise, measures of participants' menstrual beliefs would have provided insight into possible underlying pathways related to sociocultural factors, but this was deemed beyond the scope of the present research.

This study may be subject to sample bias, the direction of which is unclear. Potentially, women with higher levels of symptoms may have been more inclined to

participate, leading to an overestimation of the effect being studied. Conversely, underrepresentation may result from exclusion criteria. Exclusions included pregnant and breastfeeding women, menopausal women, and those using hormonal treatments. The latter group is particularly important to consider, as those experiencing severe menstrual symptoms, whether diagnosed or not, are often prescribed hormonal treatments (Pearlstein & Steiner, 2012; Vannuccini et al., 2021). Excluding these women may have resulted in an underestimation of the effect of symptoms. Further, given that only female university students in their second or third year were recruited, there is likely a selection bias that favors younger women, excluding other age groups who may be more vulnerable to metacognitive deficits. Caution is therefore warranted in interpreting the findings of this study, as they may not fully reflect the experiences of all women with menstrual symptoms. More replication studies are needed to ensure that findings can be generalized to the populations which were not represented in the current study.

Despite significant efforts made to reduce response bias (through question randomization and thorough explanation), complete mitigation of such factors was not achieved. All metacognitive, physical, and affective symptoms and abilities exhibited a similar phenomenon whereby their amplitudes decreased during the second cycle. This could possibly suggest either inflated attention to symptoms at the beginning of the study levelling out in time or else study fatigue whereby participants' responses became less accurate and consistent over time. A visual inspection of the ability/symptom average time-plots shows that change in time occurs consistently across both cycles and is not random, suggesting that effect cannot be attributed solely to bias. Further, the findings from this study demonstrate a strong replication of the correlations and individual differences reported by Kiesner et al. (2016), thereby supporting the presence of the hypothesised effect.

Finally, in this study, data was analyzed using cosine functions, which are mathematically simple and well-suited for the task due to the cyclical nature of the data. However, using a single fixed frequency to model the data with cosine waves may oversimplify or exclude important information. This includes harmonics of symptom values (e.g., variations at twice the menstrual cycle frequency), shifts in phase (e.g., symptoms that get worse before they get better during the cycle), or even periodic but non-smooth variations (e.g., symptoms that occur at only specific points in the cycle). Although visual inspection of individual subject trajectories suggested that symptom harmonics and phase shifts were absent, periodic but non-smooth variations could not be ruled out. In spite of this, the simplicity of the cosine model provides a good trade-off for any lost information.

6.4 Implications

6.4.1 Contributions to the field

The present thesis contributes to existing menstrual research in a number of ways. Since the discovery of steroid hormone receptors in the brain (Pfaff, 1968; Eichenbaum & Otto, 1992; Bixo et al., 1995), one of the main questions has always been whether the menstrual cycle has an effect on cognition. Findings from this thesis contribute to the debate by tackling metacognition—an area which has been mostly overlooked. Metacognition is cognitive introspection (thinking about thinking), occurring constantly and at a higher-level to most cognitive processes. The metacognitive processes of monitoring and regulation are highly vulnerable to influences such as global and individual beliefs, metacognitive biases, mood, and lack of self-awareness (Wells & Simons, 2009; Spada et al., 2010; David, 2012; Szczepanik et al., 2021) while also being susceptible to hormonal influences (Barrau-Sastre et al., 2022; Graham et al., 2018; Hodgetts et al., 2015; Kara et al., 2022). This makes metacognition more reflective of the complex dynamics of the menstrual cycle than standardized cognitive tests. By demonstrating an effect of the menstrual cycle on metacognition, findings from the present thesis contribute to the discussion in two ways. Firstly, evidence for an effect of the cycle on metacognition, in the absence of a conclusive one for objective cognition (Sundström-Poromma & Gingnell, 2014; 2018) provides support for claims that metacognition operates independently from cognition to some extent (Shimamura, 2000; Timmermans et al., 2012; Fleming et al., 2012; Vaccaro & Fleming, 2018). This suggests that the effects of metacognition on cognitive processes should be studied separately. Secondly results indicate that objective testing on its own may not fully capture the complexity of the menstrual cycle; which goes beyond biological and physiological

aspects and is subject to social, cultural and personal factors that vary across individuals and populations (Sommer, 1992; Roberts, 2004; Karlsson et al., 2014; Blake et al., 2022). Studies that do not account for these dimensions (most research in the field) have limited external validity and are likely not reflective of and cannot comment on real-life experiences. This might shed light on why menstrual cognitive studies are so notoriously inconclusive and contradictory (Souza et al., 2012; Sundström-Poromma & Gingnell, 2014; 2018; Le et al., 2020).

Findings also make a valuable contribution to the literature by providing a clearer understanding of menstrual metacognitive patterns in young, naturally-cycling women. This is especially significant since this is the first time menstrual metacognitive change has been systematically documented. Clear insight is provided into menstrual metacognitive patterns in terms of direction and magnitude. The current research provides evidence that metacognition is not specific to certain domains, but rather a more general ability that applies to various cognitive processes. Several studies have shown that metacognition seems to operate similarly across different areas of thinking such as attention, memory, and decision-making (Ais et al., 2016; Faivre et al., 2018; McCurdy et al., 2013; Song et al., 2011). This implies that metacognition is not highly selective and can be applied to different domains with good predictive power. Findings also suggest potential systemic metacognitive bias for reported abilities. Although this thesis attributes the bias to a fulfillment of pre-existing menstrual beliefs (Marvan & Cortes-Iniestra, 2001; Morgan & Rapkin, 2002; Feinberg, 2020) or a frame-of-reference shift (Lindberg et al., 2017; Hinz et al., 2021; Ham et al., 2020), further research is necessary to explore this phenomenon in greater depth.

The results of this study demonstrate how individual variations in cyclical change are consistently significant and hold greater importance than average levels of

change. Failure to account for individual differences can lead to assumptions of homogeneity in group studies that when violated, result in the underestimation of effect sizes and the acceptance of false null hypotheses. This is especially the case if sensitive populations (e.g., subjects with PMS, PMDD etc.) are not previously identified and the sample sizes are small. Schmalenberger et al., (2021), cite lack of accountability for individual differences as the single biggest contributor towards conflicting and inconclusive results in menstrual cycle research. There is need for further research to fully identify and understand these differences in metacognitive changes.

This thesis also makes a direct contribution to the current discussion regarding underlying menstrual mechanisms. Headaches and cognitive changes are not significantly linked suggesting different underlying pathways. Affective and cognitive changes share a high percentage of variance – a relationship long supported by past literature (Ford et al., 2004; Weber & Mapstone, 2009; Reid, 2006; Eggert et al., 2017; Slyepchenko et al., 2017; Le et al., 2020) – suggesting possible similar underlying causes. Results from this study provide support for both somatic effects (such as hormonal fluctuations), and cognitive factors (such as beliefs and expectations surrounding the menstrual cycle) as possible underlying causes (see section 6.2.6). Finally, physical symptoms and cognitive changes are only mildly correlated suggesting a relationship influenced by other factors. Co-variation evidence further provides support for symptom-mediated pathways. In particular, gastrointestinal discomfort correlates with cognitive change, in line with literature linking the two via multiple pathophysiological mechanisms (Moreno-Valladares et al., 2022; Rundek et al., 2021; Rea et al., 2016). Affective stressors may alter the immune system through increased production of circulating proinflammatory cytokines (Dias et al., 2022) which cross the blood-brain barrier (Stenfors et al., 2017) and disrupt high level cognition and metacognition (Peters et al., 2019; Kindler et al., 2020; Bourgognon, & Cavanagh, 2020; Muscat & Barrientos,

2021).

Although not directly tested in this thesis, this study offers a basis for investigation into an unexplored pathway in which metacognition plays a role in mediating a hormonal effect on affective symptoms. This claim is supported by two noteworthy findings uncovered in the research: the influence of the menstrual cycle on metacognition, and a correlation between cognitive alterations and affective changes. This pathway is possible since the same brain regions (such as amygdala, PFC, ACC, and insula) process both cognitive and emotional metacognitive monitoring and regulation (Sacher et al., 2013; Jonkman et al., 2021; Landin et al., 2019). Emotional metacognitive regulation sensitivity to hormonal changes can therefore result in affective symptom changes. This hypothesis is supported by evidence from behavioral studies (Mahon et al., 2015; Manikandan et al., 2016; Graham et al., 2017).

Finally, findings from this thesis serve as a confirmation and validation of previous studies (Kiesner et al., 2016). Results confirm a group tendency towards perimenstrual increase for all physical and affective symptoms as well as within-individual changes in magnitude and direction across symptom changes and across individuals. The replication of results highlights the robustness of the research findings and increases confidence in the validity of the results. Furthermore, by focusing on the Maltese context, the present study affirms that findings can be generalized to other populations. Results which were not replicated encourage revision or expansion of previous hypotheses. Specifically, this study challenges earlier studies that found no link between headaches and physical symptoms (Kiesner, 2016) and also suggests a stronger underlying link between physical and affective symptoms than was previously found.

6.4.2 Clinical implications

The results of this thesis have potential clinical implications, particularly in relation to the diagnosis and treatment of menstrual disorders.

Currently, the classification of menstrual conditions is based on symptom presence with very little understanding of underlying causes. Knowing condition etiology and causal factors provides a more accurate and specific classification that allows targeted treatment. In the case of PMDD, metacognitive symptoms such as ”subjective sense of being overwhelmed”, “feeling out of control”, and ”subjective difficulty concentrating or focusing” are currently placed in a group of ‘behaviour’ symptoms based on symptom presence. Behaviour symptoms are considered non-essential accessory PMDD symptoms and their presence is not required for diagnosis (DSM-5; APA, 2013). Recent studies have suggested that metacognitive symptoms should be viewed as main symptoms of PMDD on the grounds that they load onto the same factor as affective and physical symptoms, which suggests common underlying pathways for all symptoms (Teng et al., 2004; Wang et al., 2007). The present study supports this by providing evidence of symptom associations between cognitive, affective, and physical changes. Headaches and cognitive ability changes demonstrated no significant association, suggesting no common underlying mechanism and possibly distinct etiological pathways in menstrual conditions.

Acknowledging symptom heterogeneity can also enhance differential diagnosis and foster targeted treatment. For example, PMDD is often thought of as a singular disorder with one pathophysiology. However, recent findings of heterogeneity within the population have raised the question of possible distinct subtypes of PMDD (Pearlstein et al., 2005; Halbreich, 2008; Eisenlohr-Moul et al., 2019). The present study ascertains that individuals demonstrate differences in direction and magnitude for a range

of symptoms. Given the current prescribing practices of women being given the same hormonal treatment to manage menstrual conditions without consideration for individual symptoms (Makuch et al., 2013; Peterson et al., 2023), it is imperative to conduct further research to comprehensively characterise and comprehend the inter-individual variations in these symptoms (Pearlstein & Steiner, 2008).

Another potential clinical implication is the application of the Daily Online Metacognitive Evaluation (DOME) questionnaire to other domains of research. Significant effort was dedicated to the development of DOME, in order to satisfy a gap in the field for a questionnaire which can measure daily self-assessed cognitive ability, i.e., metacognition. There is, to the author's knowledge, no existing tool that can measure daily self-assessed metacognitive knowledge across a number of domains while also focusing on abilities rather than deficits. Results from preliminary pilot testing attest to DOME's validity and reliability as a metacognitive measure, as well as its sensitivity to metacognitive change. Validation results based on the current study (comprised of 120 participants over a period of 60 days) will be reported in future research. Outside of the menstrual cycle, this tool would be well applied to areas where a discrepancy exists between clinical neuropsychological evaluations and subjective reports of cognitive function such as: bipolar disorder, major depressive disorder, epilepsy, fibromyalgia and cancer (Torrent, 2012; Srisurapanont, 2015; Rodin, 2012; Demant et al., 2015; Srisurapanont et al., 2017; Samarasekera et al., 2015). To this end, DOME may be employed to monitor metacognitive experiences in patients or to compare metacognition in patients as opposed to healthy populations. For example, DOME can be utilized in significant areas examining the potential of subjective cognitive impairments as biomarkers for diseases like Alzheimers (Lin et al., 2019; Jessen et al., 2020; Rabin et al., 2017).

6.4.3 Social policy

Findings from this study show that metacognitive ability fluctuates in time with the menstrual cycle, with a decrease in ability reported during the premenstrual phase. While more research is needed to fully understand the relationship between metacognition and the menstrual cycle, these findings could potentially inform social policy such as menstrual leave. Menstrual leave allows women to take time off work without having to use sick days or vacation time, which can be a significant benefit for those who experience regular menstrual disruptions. Statistics show that PMS symptoms affect approximately 75% of the female population, and in 20% of cases interferences are severe (Rapkin & Winer, 2004; Kiesner, 2009; Houston et al., 2006). In terms of metacognition, disruptions can include difficulty concentrating, a perpetual experience of brain cloudiness, and ineffective decision-making. In such cases, allowing for menstrual leave may be a reasonable accommodation to ensure that individuals are not penalized for temporary changes in metacognitive function. As of this year, 2023, Spain has become the first European country to introduce this policy, joining Japan, South Korea, Indonesia, and Taiwan (Le Monde, 2023; Politico, 2023). MPs in Malta have called for a national discussion, with some Maltese companies adopting the measure (Times of Malta, 2023; Malta Independent, 2023). It is therefore fitting that this study has focused on the Maltese population, aligning with the ongoing national debate on menstrual leave policy. Opponents argue that rather than promoting gender equality, menstrual leave reinforces gender stereotypes and can lead to discrimination against women. Ultimately, the aim of menstrual leave policies is to offer assistance and protection to women who suffer from severe menstrual disruptions. Evidence from the present study of the existence of individual differences in symptom severity, provides a strong argument in favour of a social policy that protects the more vulnerable of these women.

In addition, awareness of changes in metacognition across the menstrual cycle could also help organizations to better support female employees. For example, organizations may consider offering additional training or resources to help employees improve their metacognitive abilities during the premenstrual phase, or adjust workloads to account for potential changes in metacognitive ability. In the field of education and public health, present study findings could help inform campaigns aimed at reducing menstrual stigma through increased awareness and understanding of menstrual health.

6.5 Chapter summary

Chapter 6 focused on analyzing results from the present study in the context of existing literature. The aim was to convey novel insights into the nature of metacognition in the context of the menstrual cycle, an area not yet systematically explored. The chapter examined menstrual metacognitive patterns in terms of direction, magnitude, biases and modality as well as individual differences in within-person ability change. The potential links to other menstrual symptoms and the possible underlying mechanisms joining them together were also discussed. Finally, the chapter looked into the replicability of findings, and discussed study limitations and implications. Concluding comments based on the outcomes are made in the subsequent chapter (Chapter 7).

Chapter 7

Conclusion

The aim of this thesis was to investigate the effects of the menstrual cycle on metacognition in healthy, naturally-cycling women. The novel questionnaire DOME was developed to extract insights into the nature of metacognition in the context of the cycle in terms of: magnitude and direction of change, individual differences in within-person symptoms, and covariation with other menstrual symptoms.

This research is relevant because it addresses a gap in the field with regards to the role of metacognition in the menstrual cycle. Metacognition is critical in this context as it can possibly explain subjective cognitive impairments that are frequently reported in the absence of conclusive objective cognitive performance changes (Moos, 1969; Farage et al., 2008; Souza et al., 2012; Sundström-Poromma & Gingnell, 2014; 2018). Metacognition has been shown to play a role in other conditions where a dissociation between objective and subjective measurement has been documented, including cancer patients and fibromyalgia. However, its impact in the context of the menstrual cycle remains unexplored (Hoven et al., 2019; Lenzo et al., 2020; Oerlemans et al., 2022).

Novel findings show a clear effect of the menstrual cycle on metacognition as evidenced by self-assessed cognitive changes that showed cyclical fluctuations corresponding to different cycle days. All reported metacognitive abilities showed a perimenstrual decline and a mid-cycle improvement, suggestive of either a hormonal (possibly estrogen) effect (Walker, 1997; Reed & Carr, 2015) or an influence stemming from negative menstrual beliefs (Oxley, 1998; Craner et al., 2016; Li et al., 2020). More research is needed to differentiate between the two possible causes. Findings show that significant individual differences exist in both the magnitude and direction of metacognitive cyclical change, underlining the need for future studies to look beyond the average level of change.

Amongst themselves, metacognitive domains demonstrate very little variability in terms of overall magnitude and direction, implying that results from one domain could predict results in another. This provides support for metacognition being domain-general, at least as far as self-perceived judgements are involved (Fitzgerald et al., 2017; Garfinkel et al., 2016; Morales et al., 2018). By extension, this thesis argues for a shared etiological pathway for all menstrual metacognitive complaints or improvements.

Through the exploration of associations between menstrual symptom changes, several distinct symptom clusters were identified. Cognitive abilities, physical, affective and headache symptoms presented as four individual yet correlated factors. The findings highlight strong correlations between cognitive and affective changes suggesting that the underlying mechanisms of affective and cognitive symptom types may be shared. It is also possible that the shared variance between affective and cognitive symptoms may be attributed to an affective mediating factor. For example, affective stressors can impact the immune system, producing proinflammatory cytokines that affect high-level cognition and metacognition (Peters et al., 2019; Kindler et al., 2020; Muscat & Barrientos, 2021). Alternatively, findings from the present research highlight the potential role of metacognition in mediating affective symptoms, possibly through hormonal influences on metacognitive emotion regulation (Mahon et al., 2015; Manikandan et al., 2016; Graham et al., 2017). In contrast to cognitive and affective symptoms, physical and cognitive changes show weak correlation, suggesting separate influences. Among physical changes, gastrointestinal discomfort has the strongest correlation with cognitive change, likely due to interactions in the gut-brain axis (Martin et al., 2018). Interestingly, there is no relationship between headaches and cognitive changes, indicating distinct causal pathways. This may be due to different menstrual patterns associated with headaches compared to other symptoms (Martin et al., 2007; Kiesner et al., 2013).

Specific findings from this research serve as a replication of Kiesner, et al., (2016). Results confirm a group tendency towards a perimenstrual increase in physical and affective symptoms, as well as significant within-individual changes for both symptom types. The findings also suggest two distinct temporal characterisations of symptoms seen in Kiesner et al., (2016). Firstly, symptoms can be categorized based on their duration, with some being transient and fleeting (e.g., mood swings, cramps), while others, such as headaches and anxiety/depressed symptoms, show relative stability over hours or days (van Iersel et al., 2016; Kiesner et al., 2016). Secondly, symptoms can be characterized by their phase. Certain symptoms, like breast tenderness and cramps, show only one type of cyclical phase, increasing before menstruation. On the other hand, other symptoms such as headaches display heterogeneity, with some participants experiencing an increase in symptoms during the perimenstrual phase, while others observe an increase during the midcycle phase. These classifications are significant because different symptom groupings may have different underlying causes or mechanisms. Some results from Kiesner et al., (2016) were not reproduced in the present study. The current findings challenge the idea of independence between headaches and physical symptoms reported by Kiesner et al. (2016) and question the assertion that no shared pathways exist between these domains. Furthermore, findings suggest a stronger underlying link between physical and affective symptoms than previously demonstrated, potentially implicating shared physiological or psychological mechanisms. This aligns with previous research in the field (Teng et al., 2004; Wang et al., 2007).

By exploring metacognition, this research offers valuable insights beyond the limitations of traditional cognitive measures. This is because metacognition captures dimensions that standardized tests cannot access, including personal beliefs, attitudes, cultural norms, and symptom influences (Kowalski & Chapple, 2000; Morgan & Rap-

kin, 2002; Marvan & Cortes-Iniestra, 2008; Feinberg, 2020). All these factors shape and influence women's experiences of their symptoms, meaning that objective studies have limited external validity and may not accurately reflect or provide meaningful commentary on real-life experiences. Recognizing the crucial role of metacognition, given the complexities of the menstrual cycle, becomes imperative when addressing whether the menstrual cycle impacts cognition. This is particularly significant when considering the social implications of these findings in informing policies and reducing stigma surrounding menstruation.

The study's novel findings also have clinical implications for diagnosing and treating menstrual conditions. Covariation suggests that metacognitive symptoms should be reclassified as primary symptoms, along with physical and affective symptoms, in PMDD (see Section 6.4.2). Moreover, the absence of a common underlying mechanism between changes in headaches and cognitive ability indicates distinct etiological pathways. Individual differences in symptoms (Section 6.2.5) highlight the importance of personalized treatment, including tailored hormonal treatment prescriptions (Makuch et al., 2013; Peterson et al., 2023), necessitating further research to comprehensively comprehend these variations (Pearlstein & Steiner, 2008). Additionally, the study highlights the potential application of the DOME questionnaire in other areas like cancer research, where subjective reports of cognitive disruption in the absence of objective results are well established (Liou, 2019; Rodin & Ahles, 2012).

This study contributes the first systematic documentation of metacognitive changes in the context of the menstrual cycle. The main focus was on addressing three main research questions: 1) Does metacognition oscillate in time to the menstrual cycle? 2) Do metacognitive changes covary with other symptoms? 3) Do individual differences exist in these metacognitive changes? By addressing these pivotal ques-

tions, a solid foundation has been established for further exploration in this field. Moving forward, future studies can delve into the underlying mechanisms that drive these metacognitive shifts, investigating the influence of cultural and societal factors, and exploring the potential efficacy of interventions in enhancing metacognitive functioning across various phases of the menstrual cycle. This continued research will not only enhance our understanding of the cycle's influence on cognition but also contribute to the development of targeted interventions, personalized treatments, and policy changes aimed at improving the well-being of individuals navigating the complexities of the menstrual cycle.

References

Ais, J., Zylberberg, A., Barttfeld, P., et al. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition 146:* 377–386

Aldahadha, B. (2021). Metacognition, mindfulness attention awareness, and their relationships with depression and anxiety. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *39*(2), 183-200.

Allen, A.M., McRae-Clark, A.L., Carlson, S., Saladin, M.E., Gray, K.M., Wetherington, C. L., McKee, S.A., & Allen, S.S. (2016). Determining menstrual phase in human biobehavioral research: a review with recommendations. *Exp. Clin. Psychopharm. 24,* 1–11

Allen, M., Frank, D., Schwarzkopf, D. S., Fardo, F., Winston, J. S., Hauser, T. U., & Rees, G. (2016). Unexpected arousal modulates the influence of sensory noise on confidence. *Elife*, *5*, e18103.

American College of Obstetricians and Gynecologists. (2022). Premenstrual syndrome: clinical management guidelines for obstetrician–gynecologists. *ACOG Practical Bulletin15,* 1 – 9

Anderson, J. C. & Gerbing, D.W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49*, 155–173

Andreano, J. M., and Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learn. Mem. 16,* 248–266

Annunziata, Maria Antonietta & Muzzatti, B & Giovannini, Lorena & Lucchini, G. (2011). Cognitive functioning self-assessment scale (CFSS): Preliminary psychometric data. *Psychology, health & medicine. 17.* 207-12.

Armbruster, D., Grage, T., Kirschbaum, C., & Strobel, A. (2018). Processing emotions: Effects of menstrual cycle phase and premenstrual symptoms on the startle reflex, facial EMG

and heart rate. *Behavioral Brain Research 351,* 178–187

Arnold, A. P. (2009). The organizational-activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. *Hormones and behavior*, *55*(5), 570–578

Arnsten A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat. Rev. Neurosci*. 10, 410–422

Arnsten, A. F. (1998). The biology of being frazzled. *Science*, *280*(5370), 1711-1712

Arora, M. K., Seth, S., & Dayal, S. (2010). The relationship of lipid profile and menstrual cycle with acne vulgaris. *Clinical Biochemistry, 43,* 1415–1420

Asher, A., & Myers, J.S. (2015). The Effect of Cancer Treatment on Cognitive Function. *Clinical Advances in Hematology & Oncology, 13* (7), 441-450

Azcoitia, I, Yague, J.G, & Garcia-Segura, L.M. (2011). Estradiol synthesis within the human brain. *Neuroscience, 191*, 139-147

Azevedo, R. (2020). Reflections on the field of metacognition: issues, challenges, and opportunities. *Metacognition Learning 15***,** 91–98

Bagozzi, R.P., Edwards J.R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods 1*(1):45-87

Baird B, Smallwood J, Gorgolewski KJ, et al. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. Journal of Neuroscience 33(42): 16657–16665

Baker, L. (2010). *Metacognition.* In International Encyclopedia of Education (Third Edition) 204-210. Elsevier

Banks, W. A. (2012). Brain meets body: the blood-brain barrier as an endocrine interface. *Endocrinology*, *153*(9), 4111–4119

Bannbers, E., Gingnell, M., Engman, J., Morell, A., Comasco, E., Kask, K., et al. (2012). The effect of premenstrual dysphoric disorder and menstrual cycle phase on brain activity during response inhibition. *J. Affect. Disord. 142,* 347–350

Barrau-Sastre, P., Birulés, I., Verdaguer-Rodríguez, M., López-Carrilero, R., Ferrer-Quintero, M., García-Mieres, H., ... & Spanish Metacognition Group. (2022). Influence of Menstrual Cycle Length and Age at Menarche on Symptoms, Cognition, Social Cognition, and Metacognition in Patients with First-Episode Psychosis. *Women*, *2*(2), 135-146.

Barrientos, M., Tapia, L., Silva, J. R., & Reyes, G. (2020). Biological stress reactivity and introspective sensitivity: An exploratory study. *Frontiers in psychology*, *11*, 543.

Barth, C., Steele, C.J., Mueller, K, Rekkas, V.P., Arélin, K., Pampel, A., Burmann, I., Kratzsch, J., Villringer, A., & Sacher J. (2016). In-vivo Dynamics of the Human Hippocampus across the Menstrual Cycle. *Sci Rep. 7,* 32833

Batmaz, S., Altinoz, A. E., & Sonkurt, H. O. (2021). Cognitive attentional syndrome and metacognitive beliefs as potential treatment targets for metacognitive therapy in bipolar disorder. *World Journal of Psychiatry*, *11*(9), 589.

Baumeister, R. F., Finkenauer, C., & Vohs, K. D. (2001). "Bad is stronger than good". *Review of General Psychology, 5* (4), 323–370.

Becker, J. B., & Hu, M. (2008). Sex differences in drug abuse. *Front. Neuroendocrinol. 29*, 36–47

Bellon, E., Fias, W., Ansari, D., & De Smedt, B. (2020). The neural basis of metacognitive monitoring during arithmetic in the developing brain. *Human Brain Mapping, 41*(16), 4562-4573

Beltz, A.M. & Moser, J.S. (2020). Ovarian hormones: a long overlooked but critical contributor to cognitive brain structures and function. *Annals of the New York Academy of Sciences,* 14255–

Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research, 16,* 78–117

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246

Bercik, P., Park, A. J., Sinclair, D., Khoshdel, A., Lu, J., Huang, X., ... & Verdu, E. F. (2011). The anxiolytic effect of Bifidobacterium longum NCC3001 involves vagal pathways for gut–brain communication. *Neurogastroenterology & Motility*, *23*(12), 1132-1139.

Berenbaum, S.A., & Beltz, A.M. (2011). Sexual differentiation of human behavior: Effects of prenatal and pubertal organizational hormones. *Frontiers in Neuroendocrinology, 32*(2), 183-200

Bernal, A., & Paolieri, D. (2022). The influence of estradiol and progesterone on neurocognition during three phases of the menstrual cycle: Modulating factors. *Behavioral Brain Research, 417,* 113593.

Biegon, A., Alia-Klein, N., Alexoff, D.L., Fowler, J., Kim, W., Logan, J., Pareto, D., Preston-Campbell, R., Wang, G. (2020). Relationship of estrogen synthesis capacity in the brain with obesity and self-control in men and women. *PNAS, 117* (37)

Bixo, M., Backstrom, T., Winblad, B., Andersson, A. (1995). Estradiol and testosterone in specific regions of the human female brain in different endocrine states. J. *Steroid Biochem. Molec. Biol, 55(3), 297–303*

Blake, K. R., McCartney, M., & Arslan, R. C. (2022). Menstrual cycle and hormonal contraception effects on self-efficacy, assertiveness, regulatory focus, optimism, impulsiveness, and risk-taking. *Journal of Experimental Social Psychology*, *103*, 104382

Blumenthal, P. D. & Edelman, A. Hormonal (2008). Contraception. *Obstetrics & Gynaecology, 112,* (3), 670-684

Bolger, N. & Laurenceau, J.P. (2013). *Intensive Longitudinal Methods. An introduction to Diary and Experience Sampling Research.* New York: Guilford Press

Borkowski, J. G., Chan, L. K. & Muthukrishna, N. (2020). A process-oriented model of metacognition: links between motivation and executive functioning. In (Gregory Schraw & James C. Impara) *Issues in the Measurement of Metacognition* 1–42 (Buros Institute of Mental Measurements)

Bouman, A., Heineman, M. J., & Faas, M. M. (2005). Sex hormones and the immune response in humans. *Human reproduction update*, *11*(4), 411-423

Bourgognon, J. M., & Cavanagh, J. (2020). The role of cytokines in modulating learning and memory and brain plasticity. *Brain and neuroscience advances*, *4*, 2398212820979802.

Broadbent, D.E., Cooper, P.F., FitzGerald, P., & Parkes, K.R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*, 1–16.

Brooks-Gunn, J., & Ruble, D. N. (1986). Men's and women's attitudes and beliefs about the menstrual cycle. *Sex Roles*, *14*(5), 287-299.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York: Guilford Press

Brown, R. G., & Fernie, B. A. (2015). Metacognitions, anxiety, and distress related to motor fluctuations in Parkinson's disease. *Journal of Psychosomatic Research*, *78*(2), 143-148.

Browne M.W., Cudeck R. (1993). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research. 21(2)*:230-258

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition, 2*(1), 16

Bull, J.R., Rowland, S.P., Berglund-Scherwitzl, E.B., Scherwitzl, R., Gemzell-Danielsson, K., & Harper, J., (2019). Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *npj Digital Med. 2*, 1, 83-

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., & Robinson, E. S., et al. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci. 14,* 365–376

Carroll, H. A., Lustyk, M. K. B., & Larimer, M. E. (2015). The relationship between alcohol consumption and menstrual cycle: a review of the literature. *Archives of women's mental health*, *18*(6), 773-781.

Chakravarty, E.F., & Sammaritano, L.R. (2019). Pregnancy and Autoimmune Disease, Reproductive and Hormonal Issues. *Dubois' Lupus Erythematosus and Related Syndromes, 9,* 499-519

Chang, Y. T., & Chen, Y. C. (2009). Study of menstrual attitudes and distress among postmenarcheal female students in Hualien County. *Journal of Nursing Research*, *17*(1), 20-29.

Charles, L., Chardin, C., & Haggard, P. (2020). Evidence for metacognitive bias in perception of voluntary action. *Cognition,* 194(), 104041-

Chen, J.R., Yan, Y.T., Wang, T.J., Chen, LJ., Wang, Y.J., & Tseng, G.F. (2009). Gonadal hormones modulate the dendritic spine densities of primary cortical pyramidal neurons in adult female rat. *Cerebral Cortex, 19*(11), 2719–2727

Chrisler, J. C. (2007). PMS as a culture-bound syndrome. In J. C. Chrisler, C. Golden, & P. D. Rozee (Eds.), Lectures on the psychology of women (pp. 154–171). Boston: McGraw Hill.

Chrisler, J. C. (2008). 2007 Presidential address: Fear of losing control: Power, perfectionism, and the psychology of women. Psychology of Women Quarterly, 32, 1–12

Chrisler, J. C. (2011). Leaks, lumps, and lines: Stigma and women's bodies. *Psychology of women quarterly*, *35*(2), 202-214.

Chrisler, J. C., & Gorman, J. A. (2019). The menstrual cycle: Attitudes and behavioral concomitants. In *Routledge International handbook of women's sexual and reproductive health* (pp. 55-69). Routledge

Clarke, A. E., & Ruble, D. N. (1978). Young adolescents' beliefs concerning menstruation. *Child Development*, 231-234

Clos, M., Bunzeck, N. & Sommer, T. Dopamine is a double-edged sword: dopaminergic modulation enhances memory retrieval performance but impairs metacognition. *Neuropsychopharmacol* **44**, 555–563 (2019).

Cobo, J., Nieto, L., Ochoa, S. et al. (2016). Insight and gender in schizophrenia and other psychoses. *Psychiatry Res 243*:268–277.

Cole, J. H., & Cole, W. (1992). *Women Pilots of World War II*. University of Utah Press.

Colzato, L. S., & Hommel, B. (2014). Effects of estrogen on higher-order cognitive functions in unstressed human females may depend on individual variation in dopamine baseline levels. *Frontiers in Neuroscience*, *8*(8 APR), 8–11

Colzato, L. S., Pratt, J., & Hommel, B. (2012). Estrogen modulates inhibition of return in healthy human females. *Neuropsychologia*, *50*(1), 98–103.

Colzato, L.S., Hertsig, G., van den Wildenberg, W.P.M., & Hommel, B. (2010). Estrogen modulates inhibitory control in healthy human females: Evidence from the stop-signal paradigm. *Neuroscience, 167,* 709–715

Cook, S. A., Salmon, P., Dunn, G., Holcombe, C., Cornford, P., & Fisher, P. (2015b). The association of metacognitive beliefs with emotional distress after diagnosis of cancer. *Health Psychology*, *34*(3), 207.

Cook, S. A., Salmon, P., Dunn, G., Holcombe, C., Cornford, P., & Fisher, P. (2015a). A prospective study of the association of metacognitive beliefs and processes with persistent emotional distress after diagnosis of cancer. *Cognitive therapy and research*, *39*(1), 51-60

Cornish, I.M. (2000). Factor structure of the Everyday Memory Questionnaire. *British Journal of Psychology, 91,* 427–438

Costello, A. B & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation: 10* (7)

Craner, J. R., Sigmon, S. T., & Young, M. A. (2016). Self-focused attention and symptoms across menstrual cycle phases in women with and without premenstrual disorders. *Cognitive Therapy and Research, 40*(1), 118–127

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16, 297-334

Cupini, L. M., Corbelli, I., & Sarchelli, P. (2021). Menstrual migraine: what it is and does it matter?. *Journal of Neurology*, *268*(7), 2355-2363

David, A. S., Bedford, N., Wiffen, B.& Gilleen, J. (2012) Failures of metacognition and lack of insight in neuropsychiatric disorders. ***Phil. Trans. R. Soc. B,*** *367,* 1379–1390

De Bondt, T., Pullens, P., Van Hecke, W., Jacquemyn, Y., and Parizel, P.M., (2016). Reproducibility of hormone-driven regional grey matter volume changes in women using SPM8 and SPM12. *Brain structure & function 221,* 4631-4641

De Bondt, T., Smeets, D., Pullens, P., Van Hecke, W., Jacquemyn, Y., & Parizel, P.M. (2015). Stability of resting state networks in the female brain during hormonal changes and their relation to premenstrual symptoms. *Brain Res 1624*:275–285

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature neuroscience*, *16*(1), 105-110.

Demir, O., Sal, H., & Comba, C. (2021). Triangle of COVID, anxiety and menstrual cycle. *Journal of Obstetrics and Gynaecology*, *41*(8), 1257-1261.

de Rooij, A., Van Der Leeden, M., Roorda, L. D., Steultjens, M. P., & Dekker, J. (2013). Predictors of outcome of multidisciplinary treatment in chronic widespread pain: an observational study. *BMC Musculoskeletal Disorders*, *14*(1), 1-11.

De Rosa, N., Della Corte, L., Giannattasio, A. (2021). Cancer-related cognitive impairment (CRCI), depression and quality of life in gynecological cancer patients: a prospective study. *Arch Gynecol Obstet 303*, 1581–1588

Delaney, J., Lupton, M. J., and Toth, E. (1976).*The Curse: A Cultural History of Menstruation*, Dutton, New York

Deuster, P. A., Powell-Dunford, N., Crago, M. S., & Cuda, A. S. (2011). Menstrual and oral contraceptive use patterns among deployed military women by race and ethnicity. *Women & Health*, *51*(1), 41-54

Dias, S. P., Brouwer, M. C., Boelen, A., & van de Beek, D. (2022). Cerebrospinal fluid sex steroid hormones in bacterial meningitis. *Medicine*, *101*(36), e30452.

Diekhof, E.K. & Ratnayake, M. (2016). Menstrual cycle phase modulates reward sensitivity and performance monitoring in young women: preliminary fMRI evidence. *Neuropsy-*

chologia 84: 70–80

Dietrich, T., Krings, T., Neulen, J., Willmes, K., Erberich, S., Thron, A., & Sturm, W. (2001). Effects of blood estrogen level on cortical activation patterns during cognitive activation as measured by functional MRI. *NeuroImage, 13,* 425–432

Ding, Lin & Velicer, Wayne & Harlow, Lisa. (1995). Effects of Estimation Methods, Number of Indicators Per Factor, and Improper Solutions on Structural Equation Modeling Fit Indices. *Structural Equation Modeling: A Multidisciplinary Journal. 2.* 119-143

Dixon, R.A., Hultsch, D.F., & Hertzog, C. (1988). The Metamemory in Adulthood (MIA) questionnaire. *Psychopharmacology Bulletin*, *24,* 671–688

Doornik, J. A., & Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics 70:* 927–939

Dragin, N., Nancy, P., Villegas, J., Roussin, R., Le Panse, R., & Berrih-Aknin, S. (2017). Balance between estrogens and proinflammatory cytokines regulates chemokine production involved in thymic germinal center formation. *Scientific reports*, *7*(1), 1-13.

Drogos, L. L., Rubin, L. H., Geller, S. E., Banuvar, S., Shulman, L. P., & Maki, P. M. (2013). Objective cognitive performance is related to subjective memory complaints in midlife women with moderate to severe vasomotor symptoms. *Menopause (New York, NY)*, *20*(12).

Drueke, B., Gauggel, S., Weise, L., Forkmann, T., & Mainz, V. (2022). Metacognitive judgements and abilities in patients with affective disorders. *Current Psychology*, 1-13.

Dubol, M., Epperson, N., Sacher, J., Pletzer, B., Derntl, B., Lanzenberger, R., Sundström-Poromaa, I., & Comasco, E. (2021). Neuroimaging the menstrual cycle: A multimodal systematic review. *Frontiers in Neuroendocrinology,60*

Dunn, T. L., Gaspar, C., McLean, D., Koehler, D. J., & Risko, E. F. (2021). Distributed Metacognition: Increased Bias and Deficits in Metacognitive Sensitivity When Retrieving Information from the Internet. *Technology, Mind, and Behavior, 2*(3).

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*(1), 6–25.

Eggert, L., Kleinstäuber, M., Hiller, W., & Witthöft, M. (2017). Emotional interference and attentional processing in premenstrual syndrome. *Journal of behavior therapy and experimental psychiatry*, *54*, 77-87

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of personality and social psychology*, *84*(1), 5

Eichenbaum, H., & Otto, T. (1992). The hippocampus - what does it do? *Behav Neural* Biol. 57, 2–36

Eisenlohr-Moul, T. A., Rubinow, D. R., Schiller, C. E., Johnson, J. L., Leserman, J., & Girdler, S. S. (2016). Histories of abuse predict stronger within-person covariation of ovarian steroids and mood symptoms in women with menstrually related mood disorder. *Psychoneuroendocrinology, 67,* 142–152.

Eisenlohr-Moul, T.A., Rubinow, D.R., Schiller, C.E., Johnson, J.L., Leserman, J., Girdler, S.S., 2016. Histories of abuse predict stronger within-person covariation of ovarian steroids and mood symptoms in women with menstrually related mood disorder. Psychoneuroendocrinology. 67, 142–152.

Englander-Golden, P., Whitmore, M. R., and Dienstbier, R. A. (1978). Menstrual cycle as a focus of study and self-reports of moods and behaviors. Motivat. Emot. 2: 75-86

Engman, J., Linnman, C., Van Dijk, K.R.A., & Milad, M.R. (2016). Amygdala subnuclei resting-state functional connectivity sex and estrogen differences. *Psychoneuroendocrinology 63*:34–42

Epperson, C. N., Haga, K., Mason, G. F., Sellers, E., Gueorguieva, R., Zhang, W., ... & Krystal, J. H. (2002). Cortical γ-aminobutyric acid levels across the menstrual cycle in healthy women and those with premenstrual dysphoric disorder: a proton magnetic resonance spectroscopy study. *Archives of general psychiatry*, *59*(9), 851-858.

Erbil, N., Felek, N., & Karakaşlı, E. (2015). The relationship between attitudes towards menarche and current attitudes towards menstruation of women: A comparative study. *Journal of Human Sciences*, *12*(2), 1120-1130

Faivre, N., Filevich, E., Solovey, G., et al. (2018). Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *The Journal of Neuroscience*

*38(*2): 263–277

Falkenstein, E., Tillmann, H. C., Christ, M., Feuring, M., & Wehling, M. (2000). Multiple actions of steroid hormones--a focus on rapid, nongenomic effects. *Pharmacological reviews*, *52*(4), 513–556.

Farage, M. A., Neill, S., & MacLean, A. B. (2009). Physiological changes associated with the menstrual cycle: a review. *Obstetrical & gynecological survey*, *64*(1), 58-72.

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

Farias, S. T., Mungas, D., Reed, B. R., Cahn-Weiner, D., Jagust, W., Baynes, K., & Decarli, C. (2008). The measurement of everyday cognition (ECog): scale development and psychometric properties. *Neuropsychology*, *22*(4), 531–544

Fehring, R.J., Schneider, M., & Raviele, K., (2006). Variability in the phases of the menstrual cycle. *J. Obstet. Gynecol. Neonatal Nurs. 35,* 376–384

Feinberg, E. I. (2020). *Self-Objectification and Menstruation the Role of Negative Attitudes and Emotions About Menstruation on the Experience of Physical Menstrual Symptoms* (Doctoral dissertation, The New School)

Ferrer-Quintero, M., Fernández, D., López-Carrilero, R. *et al.* (2022). Males and females with first episode psychosis present distinct profiles of social cognition and metacognition. *Eur Arch Psychiatry Clin Neurosci* 272, 1169–1181

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and cognition*, *9*(2), 288-307.

Fiedler, K., Ackerman, R., & Scarampi, C. (2019). Metacognition: Monitoring and controlling one's own knowledge, reasoning and decisions. In R.J. Sternberg & J. Funke (Eds.). *The Psychology of Human Thought: An Introduction (pp. 89-111).* Heidelberg: Heidelberg University Publishing

Fisher, P. L., Byrne, A., & Salmon, P. (2017). Metacognitive therapy for emotional distress in adult cancer survivors: a case series. *Cognitive therapy and research*, *41*(6), 891-901.

Fisher, P. L., Cook, S. A., & Noble, A. (2016). Clinical utility of the Metacognitions Questionnaire 30 in people with epilepsy. *Epilepsy & Behavior*, *57*, 185-191

Fitzgerald, L.M., Arvaneh, M. and Dockree, P.M. (2017). Domain-specific and domain-general processes underlying metacognitive judgments. *Consciousness and Cognition 49*: 264–277

Flashman, L. A., & McAllister, T. W. (2002). Lack of awareness and its impact in traumatic brain injury. *NeuroRehabilitation*, *17*(4), 285–296

Flavell J. H. (1979). Metacognitive aspects of problem solving In Resnick L. B. (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Erlbaum.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in human neuroscience*, *8*, 443.

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *Journal of Neuroscience*, *32*(18), 6117-6125.

Fleming, S. M.& Dolan, R. J. (2014). The neural basis of metacognitive ability. *Phil. Trans. R. Soc. B 367,* 1338–1349

Fleming, S.M. & Dolan, R.J. (2012). Metacognition: computation, biology and function. *The Royal Society, 367* (1594), 1280-1286

Fleur, D.S., Bredeweg, B. & van den Bos, W. (2021). Metacognition: ideas and insights from neuro- and educational sciences. *npj Sci. Learn. 6,* 13

Flynn, D., van Schaik, P., & van Wersch, A. (2004). A Comparison of Multi-Item Likert and Visual Analogue Scales for the Assessment of Transactionally Defined Coping Function. *European Journal of Psychological Assessment, 20*(1), 49-58

Fontana, A., & Pontari, B. (1994). Menstrual-related perceptual changes in women with premenstrual syndrome: Factors to consider in treatment. *Counselling Psychology Quar-*

terly, *7*(4), 399-406

Ford, N., Slade, P., & Butler, G. (2004). An absence of evidence linking perceived memory problems to the menopause. *British journal of general practice*, *54*(503), 434-438.

Fuentes, N., & Silveyra, P. (2019). Estrogen receptor signaling mechanisms. *Advances in protein chemistry and structural biology*, *116*, 135–170

Funke, F. & Reips, U. (2012). Why Semantic Differentials in Web-Based Research Should Be Made from Visual Analogue Scales and Not from 5-Point Scales. *Field Methods. 24*(3), 310-327

García-Gómez, E., González-Pedrajo, B., & Camacho-Arroyo, I. (2013). Role of sex steroid hormones in bacterial-host interactions. *BioMed research international*, *2013*.

García-Mieres, H., Usall, J., Feixas, G. & Ochoa, S. (2020). Gender differences in the complexity of personal identity in psychosis. *Schizophr Res 222:*467–469

Garfinkel, S.N., Manassei, M.F., Hamilton-Fletcher G, et al. (2016) Interoceptive dimensions across cardiac and respiratory axes. *Philosophical Transactions of the Royal Society B: Biological Sciences 371*(1708): 20160014

Gasbarri, A., Pompili, A., d'Onofrio, A., Cifariello, A., Tavares, M. C., & Tomaz, C. (2008). Working memory for emotional facial expressions: Role of the estrogen in young women. *Psychoneuroendocrinology*, *33*(7), 964–972

Gaskin, C.J & Happell, B. (2014). On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use, *International Journal of Nursing Studies, 51* (3): 511-521

Gatignon, H. (2010). *Confirmatory Factor Analysis. In Statistical Analysis of Management Data* (pp. 59-122). New York, NY: Springer.

Gehlert, S., Song, I., Chang, C.-H., Hartlage, S., (2009). The prevalence of premenstrual dysphoric disorder in a randomly selected group of urban and rural women. *Psychol. Med. 39,* 129–136

Genazzani, A. R., Lucchesi, A., Stomati, M., Catarsi, S., Genazzani, A. D., Criscuolo, M., et. al. (1997). Effects of sex steroid hormones on the neuroendocrine system. *Eur. J. Contracept. Reprod. Heal. Care* 2, 63–69.

Gennari, L., Merlotti, D., & Nuti, R. (2011). Aromatase activity and bone loss. *Advances in Clinical Chemistry, 54*, 129-164

Ghiasi, A., Keramat, A., & Mollaahmadi, L. (2018). The relationship between attitudes toward menstruation and perimenstrual symptoms among female students of Shahroud University of Medical Sciences, Northeast Iran. *Shiraz E-Medical Journal*, *19*(8).

Gilewski, M.J., & Zelinski, E.M. (1988). Memory Functioning Questionnaire (MFQ). *Psychopharmacology Bulletin*, *24,* 665–670.

Gilewski, M.J., Zelinski, E.M., & Schaie, K.W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging, 5*, 482–490

Golub, S. (1992). Periods: From menarche to menopause. Newbury Park: Sage

Gorsuch, R.L. (1997). Exploratory Factor Analysis: Its Role in Item Analysis, *Journal of Personality Assessment, 68*:3, 532-560

Gurvich, C., & Thomas, N. (2021). Hormones and Cognition. *Brain sciences*, *11*(3), 318.

Ham, R.M.T., Broering, J.M., Cooperberg, M.R., Carroll, P., & Wilson, L.S. (2020). Understanding the major factors affecting response shift effects on health-related quality of life: What the then-test measures in a longitudinal prostate cancer registry. *Clinical Genitourinary Cancer, 18*: e21–e27.

Hampl, R., Bičíková, M. & Sosvorová, L. (2015). Hormones and the blood-brain barrier. *Hormone Molecular Biology and Clinical Investigation, 21*(3), 159-164

Hampson, E. (2020). A brief guide to the menstrual cycle and oral contraceptive use for researchers in behavioral endocrinology. *Horm. Behav. 119,* 104655

Hampson, E. & Morley, E.E. (2013). Estradiol concentrations and working memory performance in women of reproductive age. *Psychoneuroendocrinology, 38*, 2897–2904

Hancock, G. R., & Mueller, R. O. (Eds.). (2013). *Quantitative methods in education and the behavioral sciences: Issues, research, and teaching. Structural equation modeling: A second course (2nd ed.).* IAP Information Age Publishing.

Handy, A. B., Greenfield, S. F., Yonkers, K. A., & Payne, L. A. (2022). Psychiatric symptoms across the menstrual cycle in adult women: A comprehensive review. *Harvard Review of Psychiatry*.

Hao, J., Rapp, P.R., Janssen, W.G. (2007). Interactive effects of age and estrogen on cognition and pyramidal neurons in monkey prefrontal cortex. *Proc Natl Acad Sci U S A, 104*(27), 11465–11470

Harrington, D. (2009). *Confirmatory factor analysis.* New York: Oxford University Press.

Hatta, T., & Nagaya, K. (2009). Menstrual Cycle Phase Effects on Memory and Stroop Task Performance. *Arch. Sex. Behav., 38,* 821–827

Hauser, T. U., Allen, M., Purg, N., Moutoussis, M., Rees, G., & Dolan, R. J. (2017). Noradrenaline blockade specifically enhances metacognitive performance. *Elife*, *6*, e24901.

Hausmann, M., Slabbekoorn, D., Van Goozen, S. H., Cohen-Kettenis, P. T., & Gunturkun, O. (2000). Sex hormones affect spatial abilities during the menstrual cycle. *Behav. Neurosci.* 114, 1245–1250

Hawkins, S. M., & Matzuk, M. M. (2008). The menstrual cycle: basic biology. *Annals of the New York Academy of Sciences*, *1135*, 10–18

Hayes, J. P. & Kaszniak, A. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology Review. 15*. 105-130

Henderson, V. W. (2018). Progesterone and human cognition. *Climacteric*, *21*(4), 333-340

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational & Psychological*

Measurement, 66(3), 393-416

Hermans, E. J., Henckens, M. J., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in neurosciences*, *37*(6), 304-314.

Hertzog, C., Dixon, R. A., Hultsch, D.F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5*(2), 215–227

Heyes, C., Bang, D., Shea, N., Frith, C.D, & Fleming, S.M. (2020). Knowing Ourselves Together: The Cultural Origins of Metacognition. *Trends in Cognitive Sciences, 24* (5), 349-362

Hjelmervik, H., Hausmann, M. & Osnes, B. 2014. Resting states are resting traits: an fMRI study of sex differences and menstrual cycle effects in resting state cognitive control networks. *PLoS One 9*: e103492

Hjelmervik, H., Westerhausen, R., Osnes, B., Endresen, C. B., Hugdahl, K., Hausmann, M., & Specht, K. (2012). Language lateralization and cognitive control across the menstrual cycle assessed with a dichotic-listening paradigm. *Psychoneuroendocrinology*, *37*(11), 1866-1875.

Hill, S. E., & Durante, K. M. (2009). Do women feel worse to look their best? Testing the relationship between self-esteem and fertility status across the menstrual cycle. *Personality and Social Psychology Bulletin*, *35*(12), 1592-1601

Hill, N. L., Mogle, J., Wion, R., Munoz, E., DePasquale, N., Yevchak, A. M., & Parisi, J. M. (2016). Subjective cognitive impairment and affective symptoms: a systematic review. *The Gerontologist*, *56*(6)

Hinz, A., Zenger, M., Leuteritz, K., Mehnert-Theuerkauf, A., & Petrowski, K. (2021). Do patients suffering from chronic diseases retrospectively overestimate how healthy they were before they fell ill?. *International journal of clinical and health psychology: IJCHP*, 21(3), 100230.

Hodgetts, S., Hausmann, M., & Weis, S. (2015). High estradiol levels improve false memory rates and meta-memory in highly schizotypal women. *Psychiatry Research, 229* (3),

708-714

Hodgetts, S.L. (2016). *The neuromodulatory properties of gonadal steroid hormones with regard to individual differences in cognition and brain organization* (Unpublished doctoral thesis). Department of Psychology, Durham University, UK

Holst, A. S., Jacques-Aviñó, C., Berenguera, A., Pinzón-Sanabria, D., Valls-Llobet, C., Munrós-Feliu, J., ... & Medina-Perucha, L. (2022). Experiences of menstrual inequity and menstrual health among women and people who menstruate in the Barcelona area (Spain): a qualitative study. *Reproductive health*, *19*(1), 1-16

Hoven, M., Lebreton, M., Engelmann, J.B. *et al.* (2019). Abnormalities of confidence in psychiatry: an overview and future perspectives. *Transl Psychiatry* **9**, 268

Howard, Matt. (2015). A Review of Exploratory Factor Analysis (EFA) Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? *International Journal of Human-Computer Interaction.* 32(1)

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55

Hu, X., Liu, Z., Chen, W., Zheng, J., Su, N., Wang, W. … Luo, L. (2017). Individual differences in the accuracy of judgments of learning are related to the gray matter volume and functional connectivity of the left mid-insula. *Frontiers in Human Neuroscience*, 11, 399.

Hyde, J.S. (2016). Sex and cognition: Gender and cognitive functions. *Curr. Opin.Neurobiol., 38,* 53–56

IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp

Irwin, M. R., & Cole, S. W. (2011). Reciprocal regulation of the neural and innate immune systems. *Nature Reviews Immunology*, *11*(9), 625-632

Itriyeva, K. (2022). The normal menstrual cycle. *Current Problems in Pediatric and Adolescent Health Care*, 101183

Jacobs, E., & D'Esposito, M. (2011). Estrogen shapes dopamine-dependent cognitive processes: implications for women's health. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *31*(14), 5286–5293.

Jackson, D.L. (2003). Revisiting Sample Size and Number of Parameter Estimates: Some Support for the N:q Hypothesis, *Structural Equation Modeling: A Multidisciplinary Journal,* 10:1, 128-141

Jackson, D. L. (2007). The effect of the number of observations per parameter in misspecified confirmatory factor analytic models. *Structural Equation Modeling, 14,* 48-76

Jackson, D.L., Gillaspy, A. & Purc-Stephenson, R. (2009). Reporting Practices in Confirmatory Factor Analysis: An Overview and Some Recommendations. *Psychological methods. 14*. 6-23

Janelsins, M. C., Lei, L., Netherby-Winslow, C., Kleckner, A. S., Kerns, S., Gilmore, N., ... & Culakova, E. (2022). Relationships between cytokines and cognitive function from pre- to post-chemotherapy in patients with breast cancer. *Journal of Neuroimmunology*, *362*, 577769

Jansen, C. E. (2013, November). Cognitive changes associated with cancer and cancer therapy: Patient assessment and education. In *Seminars in Oncology Nursing* (Vol. 29, No. 4, pp. 270-279). WB Saunders.

Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., ... & Wagner, M. (2020). The characterisation of subjective cognitive decline. *The Lancet Neurology*, *19*(3), 271-278.

Jichuan Wang, Xiaoqian Wang. (2019). *Structural Equation Modeling: Applications Using Mplus, 2nd Edition*. Wiley

Joensson, M., Romer Thomsen, K., Andersen, L.M., Gross, J., Mouridsen, K., Sandberg, K., Ostergaard, L., Lou, H., (2015). Dopamine activates conscious self-monitoring through medial prefrontal cortex. *Human Brain Mapping, 36*(5), 1866- 1877

Johnston-Robledo, I., & Chrisler, J. C. (2020). The menstrual mark: Menstruation as social stigma. *The Palgrave Handbook of Critical Menstruation Studies*, 181-199.

Jöreskog, K. G., & Sörbom, D. (1981). LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods (Research Report 8–8). Uppsala, Sweden: University of Uppsala, Department of Statistics

Jonkman, L. E., Fathy, Y. Y., Berendse, H. W., Schoonheim, M. M., & Van De Berg, W. D. J. (2021). Structural network topology and microstructural alterations of the anterior insula associate with cognitive and affective impairment in Parkinson's disease. *Scientific Reports, 11*(1), 1-10

Kancheva Landolt, N., & Ivanov, K. (2021). cognitive behavioral therapy-a primary mode for premenstrual syndrome management: systematic literature review. *Psychology, Health & Medicine*, *26*(10), 1282-1293.

Kara, O., Kaymaz, N. & Uzun, M.E. (2022). The effect of hyperandrogenism and obesity on mindfulness and metacognition in adolescents with polycystic ovary syndrome. *Arch Womens Ment Health* **25**, 911–921

Karlsson, T. S., Marions, L. B., & Edlund, M. G. (2014). Heavy menstrual bleeding significantly affects quality of life. *Acta obstetricia et gynecologica Scandinavica*, *93*(1), 52-57.

Keenan, P. A., Ezzat, W. H., Ginsburg, K., Moore, G. J. (2001). Prefrontal cortex as the site of estrogen's effect on cognition. *Psychoneuroendocrinology, 26*(6), 577-590

Kenny, D.A. (2011, September 6). Terminology and Basics of SEM. *Basics.* `http://davidakenny.net/cm/basics.htm`

Kiesner, J. (2009). Physical characteristics of the menstrual cycle and premenstrual depressive symptoms. *Psychological Science*, *20*(6), 763–770

Kiesner, J. (2011). One woman's low is another woman's high: Paradoxical effects of the menstrual cycle. *Psychoneuroendocrinology,* 36, 68–76

Kiesner, J. (2017). The Menstrual Cycle-Response and Developmental Affective-Risk Model: A multilevel and integrative model of influence*, Psychol. Rev.* 124 (2017) 215-244

Kiesner, J., & Martin, V. T. (2013). Mid-cycle headaches and their relationship to different patterns of premenstrual stress symptoms. *Headache: The Journal of Head and Face*

Pain, *53*(6), 935-946.

Kiesner, J., & Pastore, M. (2010). Day-to-day co-variations of psychological and physical symptoms of the menstrual cycle: Insights to individual differences in steroid reactivity. *Psychoneuroendocrinology*, *35*(3), 350-363.

Kiesner, J., Mendle, J., Eisenlohr-Moul, T. A., & Pastore, M. (2016). Cyclical Symptom Change Across the Menstrual Cycle: Attributional, Affective, and Physical Symptoms *Clinical Psychological Science*

Kindler, J., Lim, C. K., Weickert, C. S., Boerrigter, D., Galletly, C., Liu, D., ... & Weickert, T. W. (2020). Dysregulation of kynurenine metabolism is related to proinflammatory cytokines, attention, and prefrontal cortex volume in schizophrenia. *Molecular psychiatry*, *25*(11), 2860-2872

King, S. (2021). Menstrual leave: Good intention, poor solution. In *Aligning Perspectives in Gender Mainstreaming* (pp. 151-176). Springer, Cham.

Kissling, E. A. (1996). Bleeding out loud: Communication about menstruation. *Feminism & Psychology*, *6*(4), 481-504

Kleinstauber, M., 2019. Are there temporal subtypes of premenstrual dysphoric disorder? Using group-based trajectory modeling to identify individual differences in symptom change. Psychol. Med. 50, 964–972

Kleinstäuber, M., Witthöft, M., & Hiller, W. (2012). Cognitive-behavioral and pharmacological interventions for premenstrual syndrome or premenstrual dysphoric disorder: a meta-analysis. *Journal of clinical psychology in medical settings*, *19*(3), 308-319.

Klimek, L., Bergmann, K.C., Biedermann, T., Hellings, P. & Jung, K. (2017). Visual analogue scales (VAS): Measuring instruments for the documentation of symptoms and therapy monitoring in cases of allergic rhinitis in everyday health care. *Allegro J Int. 26*(1), 16-24

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling.* Guilford Publications.

Kline, R. B. (2011). *Methodology in the Social Sciences. Principles and practice of structural equation modeling (3rd ed.).* Guilford Press.

Ko, Y., & Lau H. (2012). A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. ***Phil. Trans. R. Soc. B,*** *367,* 1401–1411

Kollmann, J., Gollwitzer, M., Spada, M. M., & Fernie, B. A. (2016). The association between metacognitions and the impact of Fibromyalgia in a German sample. *Journal of psychosomatic research*, *83*, 1-9

Kowalski, R. M., & Chapple, T. (2000). The social stigma of menstruation: Fact or fiction? *Psychology of Women Quarterly*, *24*(1), 74-80

Kraft, B., Jonassen, R., Stiles, T. C., & Landrø, N. I. (2017). Dysfunctional metacognitive beliefs are associated with decreased executive control. *Frontiers in psychology*, *8*, 593.

Kralik, J., Lee , J., Rosenbloom, P., Jackson, P., Epstein, S., Romero, O., Sanz , R., Larue, O., Schmidtke, H., Lee, S., & McGreggor, K. (2020). Metacognition for a common model of cognition. *Procedia of Computer Science* 145, 730-739

Kuba, K., Weißflog, G., Götze, H., García-Torres, F., Mehnert, A., & Esser, P. (2019). The relationship between acceptance, fatigue, and subjective cognitive impairment in hematologic cancer survivors. *International Journal of Clinical and Health Psychology*, *19*(2), 97-106.

Kulakaç, Ö., Oncel, S., Fırat, M. Z., & Akcan, A. (2008). Menstrual Attitude Questionnaire: validity and reliability study. *Turkiye Klinikleri J Gynecol Obst*, *18*(6), 347-56

Kunimoto C., Miller, J. & Pashler, K. (2001). Confidence and Accuracy of Near-Threshold Discrimination Responses. *Conscious Cogn. 10*(3), 294–340

Kuppens, K., Neels, H., van Wilgen, C. P., Roussel, N., Heyrman, A., Lambrecht, L., ... & Nijs, J. (2015). Do illness perceptions in patients with fibromyalgia differ across countries? A comparative study. *Myopain*, *23*(1-2), 13-20.

Landin, J. D., Palac, M., Carter, J. M., Dzumaga, Y., Santerre-Anderson, J. L., Fernandez, G. M., ... & Werner, D. F. (2019). General anesthetic exposure in adolescent rats causes

persistent maladaptations in cognitive and affective behaviors and neuroplasticity. *Neuropharmacology, 150,* 153-163

Larki, M., Salavati, A., & Azmoude, E. (2022). The Relationship Between Menstrual Patterns and Menstrual Attitude Dimensions among Women of Reproductive Age: Cross-sectional survey. *Sultan Qaboos University Medical Journal*, *22*(2), 241.

Lau, H. C. & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences, 103*(49), 18763–18768

Le, J., Thomas, N., & Gurvich, C. (2020). Cognition, The Menstrual Cycle, and Premenstrual Disorders: A Review. *Brain Sciences, 10*(4), *198–*

LeBlanc, E. S., Neiss, M. B., Carello, P. E., Samuels, M. H., & Janowsky, J. S. (2007). Hot flashes and estrogen therapy do not influence cognition in early menopausal women. *Menopause*, *14*(2), 191-202.

Lebar, V., Laganà, A. S., Chiantera, V., Kunič, T., & Lukanović, D. (2022). The Effect of COVID-19 on the menstrual cycle: a systematic review. *Journal of Clinical Medicine*, *11*(13), 3800.

Lee, A. L., Goldstein, R. S., & Brooks, D. (2017). Chronic pain in people with chronic obstructive pulmonary disease: prevalence, clinical and psychological implications. *Chronic Obstructive Pulmonary Diseases*, *4*(3), 194

Lee, T., Cai, L., & MacCallum, R. C. (2012). *Power analysis for tests of structural equation models. In R. H. Hoyle (Ed.), Handbook of structural equation modeling (p. 181–194).* The Guilford Press.

Leeners, B., Kruger, T. H. C., Geraedts, K., Tronci, E., Mancini, T., Ille, F., … Hengartner, M. P. (2017). Lack of Associations between Female Hormone Levels and Visuospatial Working Memory, Divided Attention and Cognitive Bias across Two Consecutive Menstrual Cycles. *Frontiers in Behavioral Neuroscience*, *11*(July), 1–10

Lenzo, V., Sardella, A., Martino, G., & Quattropani, M. C. (2020). A systematic review of metacognitive beliefs in chronic medical conditions. *Frontiers in psychology*, *10*, 2875.

Levin, E. R., Vitek, W. S., & Hammes, S. R. (2018). Estrogens, progestins, and the female reproductive tract. *Goodman & Gilman's: The Pharmacological Basis of Therapeutics; Brunton, LL, Hilal-Dandan, R., Knollman, BC, Eds*, 803-833.

Levitt, R. B., & Barnack-Tavlaris, J. L. (2020). Addressing menstruation in the workplace: the menstrual leave debate. *The Palgrave handbook of critical menstruation studies*, 561-575

Lewicka, M., Czapinski, J., & Peeters, G. (1992). "Positive-negative asymmetry or "When the heart needs a reason"". *European Journal of Social Psychology*

Lin, Y., Shan, P. Y., Jiang, W. J., Sheng, C., & Ma, L. (2019). Subjective cognitive decline: preclinical manifestation of Alzheimer's disease. *Neurological Sciences*, *40*, 41-49

Lindberg, P., Netter, P., Koller, M., Steinger, B., & Klinkhammer-Schalke M. (2017). Breast cancer survivors'recollection of their quality of life: Identifying determinants of recall bias in a longitudinal population-based trial. *PLoS One.;*12

Liou, K. T., Ahles, T. A., Garland, S. N., Li, Q. S., Bao, T., Li, Y., Root, J. C., & Mao, J. J. (2019). The Relationship Between Insomnia and Cognitive Impairment in Breast Cancer Survivors. *JNCI cancer spectrum*, *3*(3), pkz041

Lisofsky, N., Mårtensson, J., Eckert, A., Lindenberger, U., Gallinat, J., & Kühn, S. (2015). Hippocampal volume and functional connectivity changes during the female menstrual cycle. *NeuroImage, 118*, 154–162

Livermore, J. J., Holmes, C. L., Moga, G., Adamatzky, K., Critchley, H. D., Garfinkel, S. N., & Campbell-Meiklejohn, D. (2021). Serotonergic effects on interoception. *bioRxiv*, 2020-08

Lockwood, J., Daley, D., Townsend, E., & Sayal, K. (2017). Impulsivity and self-harm in adolescence: a systematic review. *European child & adolescent psychiatry*, *26*(4), 387-402

Luca, M., Chattipakorn, S. C., Sriwichaiin, S., & Luca, A. (2020). Cognitive-behavioral correlates of dysbiosis: A review. *International journal of molecular sciences*, *21*(14), 4834

Lustyk, M. K. B., Gerrish, W. G., Shaver, S., & Keys, S. L. (2009). Cognitive-behavioral therapy for premenstrual syndrome and premenstrual dysphoric disorder: a systematic review. *Archives of women's mental health*, *12*(2), 85-96.

Lustyk, M. K. B., Widman, L., Paschane, A., & Ecker, E. (2008). Stress, quality of life and physical activity in women with varying degrees of premenstrual symptomatology. *Women & health*, *39*(3), 35-44.

Lysaker, P.H., Gagen, E., & Wright, A. (2019). Metacognitive deficits predict impaired insight in schizophrenia across symptom profiles: a latent class analysis. *Schizophr Bull 45*:48–56

MacCallum, R. C. (1995). Model specification: Procedures, strategies, and related issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications.* Thousand Oaks, CA: Sage

Maccoby, E.E. & Jacklin, C.N. Sex differences in intellectual functioning. In Proceedings of the Invitational Conference on Testing Problems, 28th Oct 1972; *New York, NY, USA*; 1972. pp. 37–55

McDonald, Roderick & Ho, Moon-Ho. (2002). Principles and Practice in Reporting Structural Equation Analyses. *Psychological methods. 7*. 64-82

Maher-Edwards, L., Fernie, B. A., Murphy, G., Wells, A., & Spada, M. M. (2011). Metacognitions and negative emotions as predictors of symptom severity in chronic fatigue syndrome. *Journal of psychosomatic research*, *70*(4), 311-317.

Mahon, J. N., Rohan, K. J., Nillni, Y. I., & Zvolensky, M. J. (2015). The role of perceived control over anxiety in prospective symptom reports across the menstrual cycle. *Archives of women's mental health*, *18*(2), 239-246.

Maki, P. M., Rich, J. B., & Rosenbaum, R. S. (2002). Implicit memory varies across the menstrual cycle: estrogen effects in young women. *Neuropsychologia* 40, 518–529

Manikandan, S., Nillni, Y. I., Zvolensky, M. J., Rohan, K. J., Carkeek, K. R., & Leyro, T. M. (2016). The role of emotion regulation in the experience of menstrual symptoms and perceived control over anxiety-related events across the menstrual cycle. *Archives of women's mental health*, *19*(6), 1109-1117.

Manski, R., Grindlay, K., Burns, B., Holt, K., & Grossman, D. (2014). Reproductive health access among deployed US servicewomen: a qualitative study. *Military Medicine*, *179*(6), 645-652.

Mäntylä T., Rönnlund M., Kliegel M. (2010). Components of executive functioning in metamemory. *Appl. Neuropsychol. 17,* 289–298

Marrocco, J. & McEwen, B.S. (2016). Sex in the brain: hormones and sex differences. *Dialogues in clinical neuroscience, 18*(4), 373–383

Martin, C. R., Osadchiy, V., Kalani, A., & Mayer, E. A. (2018). The brain-gut-microbiome axis. *Cellular and molecular gastroenterology and hepatology, 6*(2), 133-148.

Marván, M. L., & Cortés-Iniestra, S. (2008). *Menstruación: qué es y qué no*. Editorial Pax México.

Marván, Maria & Cortés-Iniestra, Sandra. (2001). Women's beliefs about the prevalence of premenstrual syndrome and biases in recall of premenstrual changes. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association. 20.* 276-80.

Mccurdy, L.Y., Maniscalco, B., Metcalfe, J., et al. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *Journal of Neuroscience 33*(5): 1897–1906

McDonald, Roderick & Ho, Moon-Ho. (2002). Principles and Practice in Reporting Structural Equation Analyses. *Psychological methods. 7*. 64-82

McEwen, B. S. & Alves S. E. (1999). Estrogen actions in the central nervous system. *Endocr. Rev. 20*, 279–307

Meeker, T.J., Veldhuijzen, D.S., Keaser, M.L., Gullapalli, R.P., & Greenspan, J.D. (2020). Menstrual Cycle Variations in Gray Matter Volume, White Matter Volume and Functional Connectivity: Critical Impact on Parietal Lobe. *Frontiers in Neuroscience, 14,*

Menard, C. S., & Harlan, R. E. (1993). Up-regulation of androgen receptor immunoreactivity in the rat brain by androgenic-anabolic steroids. *Brain Res. 622*, 226–236.

Merriman, J. D., Aouizerat, B. E., Cataldo, J. K., Dunn, L., Cooper, B. A., West, C., ... & Miaskowski, C. (2014). Association between an interleukin 1 receptor, type I promoter polymorphism and self-reported attentional function in women with breast cancer. *Cytokine*, *65*(2), 192-201.

Merskin, D. (1999). Adolescence, advertising, and the ideology of menstruation. *Sex Roles*, *40*(11), 941-957

Meza-Moreno, F. I., Pimienta-Alcaraz, M. J., Vázquez-Valdez, M. F., Balderas-Sánchez, C. P., & Toledo-Fernández, A. (2021). Self-reported executive function, and not performance-based measures, strongly associates with symptoms of premenstrual syndrome/premenstrual dysphoric disorder. *Salud mental*, *44*(2), 83-90.

Mitchell, E. S., & Woods, N. F. (2011). Cognitive symptoms during the menopausal transition and early postmenopause. *Climacteric*, *14*(2), 252-261.

Mitsushima, D. (2010). Sex steroids and acetylcholine release in the hippocampus. *Vitam. Horm. 82,* 263–277

Moglia, M.L., Nguyen, H.V., Chyjek, K., Chen, K.T., & Casta˜no, P.M., (2016). Evaluation of smartphone menstrual cycle tracking applications using an adapted APPLICATIONS scoring system. *Obstet. Gynecol. 127*, 1153–1160

Moores, T. T., Chang, J. C. J., & Smith, D. K. (2006). Clarifying the role of self-efficacy and metacognition as predictors of performance: Construct development and test. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, *37*(2-3), 125-132

Moos, R. H., Kopell, B. S., Melges, F. T., Yalom, I. D., Lunde, D. T., Clayton, R. B., and Hamburg, D. A. (1969). Fluctuations in symptoms and moods during the menstrual cycle. J. Psychosom. Res. 13: 37-44.

Morales, J., Lau, H. and Fleming, S.M. (2018). Domain-specific patterns of activity support metacognition in human prefrontal cortex. *Journal of Neuroscience 38*(14): 3534–3546

Moreno-Valladares, M., Moncho-Amor, V., Bernal-Simon, I., Agirre-Iturrioz, E., Álvarez-Satta, M., & Matheu, A. (2022). Norovirus Intestinal Infection and Lewy Body Disease in an Older Patient with Acute Cognitive Impairment. *International Journal of Molecular*

Sciences, 23(15), 8376.

Morgan, M., & Rapkin, A. (2002). Cognitive flexibility, reaction time, and attention in women with premenstrual dysphoric disorder. *The Journal of Gender-specific Medicine: JGSM: the Official Journal of the Partnership for Women's Health at Columbia*, *5*(3), 28-36

Moritz, S., Klein, J.P., Desler, T. et al. (2017). Neurocognitive deficits in schizophrenia. Are we making mountains out of molehills? *Psychol Med 47*:2602–2612.

Moses-Payne, M. E., Rollwage, M., Fleming, S. M., & Roiser, J. P. (2019). Postdecision Evidence Integration and Depressive Symptoms. *Frontiers in psychiatry*, *10*, 639.

Moulds, M. L., Yap, C. S., Kerr, E., Williams, A. D., & Kandris, E. (2010). Metacognitive beliefs increase vulnerability to rumination. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *24*(3), 351-364.

Moulton, V. R. (2018). Sex Hormones in Acquired Immunity and Autoimmune Disease. *Frontiers in Immunology, 9, 2279-*

Muharam, R., Agiananda, F., Budiman, Y. F., Harahap, J. S., Prabowo, K. A., Azyati, M., ... & Sumapraja, K. (2022). Menstrual cycle changes and mental health states of women hospitalized due to COVID-19. *PloS one*, *17*(6), e0270658.

Munro, B.H. (2004). *Statistical methods for health care research.* Philadelphia: Lippincott

Mumenthaler, M. S., O'Hara, R. U. T. H., Taylor, J. L., Friedman, L. E. A. H., & Yesavage, J. A. (2001). Influence of the menstrual cycle on flight simulator performance after alcohol ingestion. *Journal of studies on alcohol*, *62*(4), 422-433.

Muscat, S. M., & Barrientos, R. M. (2021). The Perfect Cytokine Storm: How Peripheral Immune Challenges Impact Brain Plasticity & Memory Function in Aging. *Brain Plasticity*, (Preprint), 1-14

Mutlu, H. H., Bilican, F. I., Mutlu, H. H., & Gumus, M. (2018). A comparison of metacognitive factors among patients with cancer and the control group. *Psycho-Oncology*, *27*(4), 1277-1283.

Myers, J.S. (2010). The possible role of cytokines in chemotherapy-induced cognitive deficits. *Adv Exp Med Biol.* 678: 119–123.

Nelson, T. O. (2001). Metamemory, psychology of. In *International Encyclopedia of the Social & Behavioral Sciences* (eds Smelser, N. J. & Baltes, P. B.) 9733–9738

Nelson, T., & Narens, L. (1990). Metamemory: a theoretical framework and new findings. *The psychology of learning and motivation* (pp. 125–173). New York: Academic press.

Nunnally, J.C. (1987). *Psychometric theory.* New York: McGraw-Hill

Nunnally, J.C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.).* New York, NY: McGraw-Hill.

Obermeyer, C. M., & Sievert, L. L. (2007). Cross-cultural comparisons: midlife, aging, and menopause. *Menopause*, *14*(4), 663-667

O'Connor, D.W., Pollitt, P.A., Roth, M., Brook, P.B., & Reiss, B.B. (1990). Memory complaints and impairment in normal, depressed, and demented elderly persons identified in a community survey. *Archives of General Psychiatry*, *47*, 224–227

Oerlemans, S., Schagen, S.B., van den Hurk, C.J. (2022). Self-perceived cognitive functioning and quality of life among cancer survivors: results from the PROFILES registry. *J Cancer Surviv*.,*16*, 303–313

Ortega-González, M., Aguirre-Acevedo, D. C., & Jiménez-Benítez, M. (2021). Metacognitions on Pain, Functionality and Intensity of Pain in Women with Fibromyalgia. *International Journal of Psychological Research*, *14*(1), 78-90

Overgaard, M. & Sandberg, K (2012). Metacognition: computation, neurobiology and function || Kinds of access: different methods for report reveal different kinds of metacognitive access. *Philosophical Transactions: Biological Sciences, 367*(1594), 1287–1296

Owens, S. A., Eisenlohr-Moul, T. A., & Prinstein, M. J. (2020). Understanding when and why some adolescent girls attempt suicide: an emerging framework integrating menstrual cycle fluctuations in risk. *Child development perspectives*, *14*(2), 116-123.

Oxley, T. (1998). Menstrual management: An exploratory study. Feminism & Psychology, 8, 185–191

Pahlevan, M., Besharat, M. A., Borjali, A., & Farahani, H. (2019, November). A hypothetical model of pain perception in patients with chronic pain: The predictive role of unconscious, emotional, behavioral and meta-cognitive factors. In *Annales Médico-psychologiques, revue psychiatrique* (Vol. 177, No. 9, pp. 875-883). Elsevier Masson.

Paige, K. E. (1971). Effects of oral contraceptives on affective fluctuations associated with the menstrual cycle. *Psychosomatic Medicine*

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology*, 422.

Papageorgiou, C., & Wells, A. (2001). Metacognitive beliefs about rumination in recurrent major depression. *Cognitive and Behavioral Practice*, *8*(2), 160-164.

Parlee, M. B. (1974). Stereotypic beliefs about menstruation: a methodological note on the Moos Menstrual Distress Questionnaire and some new data. *Psychosomatic medicine*

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed Strategies for Learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*(6), 1239–1252.

Pearlstein, T., Yonkers, K. A., Fayyad, R., & Gillespie, J. A. (2005). Pretreatment pattern of symptom expression in premenstrual dysphoric disorder. *Journal of affective disorders*, *85*(3), 275-282

Peper, J.S., van den Heuvel, M.P., Mandl, R.C.W.W., Pol, H.E.H.H., & van Honk, J. (2011). Sex steroids and connectivity in the human brain: a review of neuroimaging studies. *Psychoneuroendocrinology 36*(8):1101–1113

Peters, A. T., Ren, X., Bessette, K. L., Goldstein, B. I., West, A. E., Langenecker, S. A., & Pandey, G. N. (2019). Interplay between pro-inflammatory cytokines, childhood trauma, and executive function in depressed adolescents. *Journal of Psychiatric Research*, *114*, 1-10.

Petersen, N., Touroutoglou, A., Andreano, J.M., and Cahill, L. (2015). Oral contraceptive pill use is associated with localized decreases in cortical thickness. *Hum Brain Mapp 36* 2644-2654

Peugh, J.L. (2010). A practical guide to multilevel modelling. *Journal of School Psychology 48*, 85–112

Pfaff D. W. (1968). Autoradiographic localization of radioactivity in rat brain after injection of tritiated sex hormones. *Science (New York, N.Y.)*, *161*(3848), 1355–1356

Pfaff, D.W. & Keiner M. (1973). Atlas of estradiol-concentrating cells in the central nervous system of the female rat. *J Comp Neurol*. 151, 121–58

Pfaffl, M. W., Lange, I. G., & Meyer, H. H. D. (2003). The gastrointestinal tract as target of steroid hormone action: Quantification of steroid receptor mRNA expression (AR, ERα, ERβ and PR) in 10 bovine gastrointestinal tract compartments by kinetic RT-PCR. *The Journal of steroid biochemistry and molecular biology*, *84*(2-3), 159-166.

Phakiti, A. (2018). Confirmatory Factor Analysis and Structural Equation Modeling. *The Palgrave Handbook of Applied Linguistics Research Methodology*, 459–500

Phelan, A. L. (2020). COVID-19 immunity passports and vaccination certificates: scientific, equitable, and legal challenges. *The lancet*, *395*(10237), 1595-1598

Pletzer, B., Crone, J.S., Kronbichler, M., & Kerschbaum, H. (2016). Menstrual cycle and hormonal contraceptive-dependent changes in intrinsic connectivity of resting-state brain networks correspond to behavioral changes due to hormonal status. *Brain Connect 6*(7):572–585

Pletzer, B., Harris, T., & Hidalgo-Lopez, E. (2018). Subcortical structural changes along the menstrual cycle: beyond the hippocampus. *Scientific reports, 8*(1), 16042

Pletzer, B., Harris, T.A., & Scheuringer, A. (2019). The cycling brain: menstrual cycle related fluctuations in hippocampal and fronto-striatal activation and connectivity during cognitive tasks. *Neuropsychopharmacology 44:* 1867–1875

Pletzer, B., Steinbeisser, J., van Laak, L., & Harris, T. (2019). Beyond Biological Sex: Interactive Effects of Gender Role and Sex Hormones on Spatial Abilities. *Front. Neurosci.,13,* 675

Portney, G. (2015). *Foundations of Clinical Res: Applications to Practice.* 3rd Edition. Watkins P. Davis Company

Potter, J., Bouyer, J., Trussell, J., & Moreau, C. (2009) Premenstrual syndrome prevalence and fluctuation over time: Results from a French population-based survey. *J. Women's Health, 18,* 31–39

Protopopescu, X., Butler, T., Pan, H., Root, J., Altemus, M., Polanecsky, M., McEwen, B., Silbersweig, D., & Stern, E. (2008). Hippocampal structural changes across the menstrual cycle. *Hippocampus, 18*(10), 985–988

Pullens, M. J., De Vries, J., & Roukema, J. A. (2010). Subjective cognitive dysfunction in breast cancer patients: a systematic review. *Psycho-oncology*, *19*(11), 1127-1138

Purewal, R., & Fisher, P. L. (2018). The contribution of illness perceptions and metacognitive beliefs to anxiety and depression in adults with diabetes. *Diabetes research and clinical practice*, *136*, 16-22.

Qiu, L., Su, Jie., Yinmei, Ni., Bai, Yang et al. (2018). The neural system of metacognition accompanying decision-making in the prefrontal cortex. *Plos Biol. 16* (4)

Qin, S., Hermans, E. J., Van Marle, H. J., Luo, J., & Fernández, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological psychiatry*, *66*(1), 25-32.

Quattropani, M. C., Lenzo, V., Fries, W., & Belvedere, A. (2014). Dysfunctional metacognitive beliefs and gastrointestinal disorders. Beyond an 'organic'/'functional'categorization in the clinical practice. *Mediterranean Journal of Clinical Psychology*, *2*(1)

Quattropani, M. C., Lenzo, V., Mucciardi, M., & Toffle, M. E. (2016). Metacognition as predictor of emotional distress in cancer patients. *Life Span and Disability*, *19*(2), 221-239

Rabiepoor, S., Valizadeh, R., & Barjasteh, S. (2017). Study of menstrual attitudes and knowledge among postmenarcheal students. *International Journal of Pediatrics*, *5*(5), 4991-5001

Rabin, L. A., Smart, C. M., & Amariglio, R. E. (2017). Subjective cognitive decline in preclinical Alzheimer's disease. *Annual review of clinical psychology*, *13*, 369-396

Raghunath, R. S., Venables, Z. C., & Millington, G. W. M. (2015). The menstrual cycle and the skin. *Clinical and experimental dermatology*, *40*(2), 111-115

Ranieri, J., Guerra, F., & Di Giacomo, D. (2020). Role of metacognition thinking and psychological traits in breast cancer survivorship. *Behavioral Sciences*, *10*(9), 135

Rastelli, D., Robinson, A., Lagomarsino, V. N., Matthews, L. T., Hassan, R., Perez, K., ... & Rao, M. (2022). Diminished androgen levels are linked to irritable bowel syndrome and cause bowel dysfunction in mice. *The Journal of clinical investigation*, *132*(2).

Rea, K., Dinan, T. G., & Cryan, J. F. (2016). The microbiome: a key regulator of stress and neuroinflammation. *Neurobiology of stress*, *4*, 23-33.

Reed, B. G., & Carr, B. R. (2015). *The normal menstrual cycle and the control of ovulation.* South Dartmouth (MA) Press

Rees, M. (1995). The age of menarche. *ORGYN: Organon's magazine on women & health*, (4), 2–4

Reid, L. M., & MacLullich, A. M. (2006). Subjective memory complaints and cognitive impairment in older people. *Dementia and geriatric cognitive disorders*, *22*(5-6), 471-485

Reilly, J., & Kremer, J. (1999). A qualitative investigation of women's perceptions of premenstrual syndrome: Implications for general practitioners. *Br. J. Gen. Pract., 49,* 783–786

Reinholdt-Dunne, M. L., Blicher, A., Nordahl, H., Normann, N., Esbjørn, B. H., & Wells, A. (2019). Modeling the relationships between metacognitive beliefs, attention control and symptoms in children with and without anxiety disorders: A test of the S-REF model. *Frontiers in psychology*, *10*, 1205.

Reyes, G., Silva, J. R., Jaramillo, K., Rehbein, L., & Sackur, J. (2015). Self-knowledge dim-out: stress impairs metacognitive accuracy. *PLoS One*, *10*(8), e0132320.

Richardson, J. T. (1992). The menstrual cycle, cognition, and paramenstrual symptomatology. In *Cognition and the menstrual cycle* (pp. 1-38). Springer, New York, NY

Riva, P., Wirth, J. H., & Williams, K. D. (2011). The consequences of pain: The social and physical pain overlap on psychological responses. *European Journal of Social Psychology, 41*(6), 681-687

Robb, C. E., De Jager, C. A., Ahmadi-Abhari, S., Giannakopoulou, P., Udeh-Momoh, C., McKeand, J., ... & Middleton, L. (2020). Associations of social isolation with anxiety and depression during the early COVID-19 pandemic: a survey of older adults in London, UK. *Frontiers in Psychiatry*, *11*, 591120.

Roberts, B., Eisenlohr-Moul, T., Martel, M.M. (2018). Reproductive steroids and ADHD symptoms across the menstrual cycle. *Psychoneuroendocrinology. 88*, 105–114

Roberts, T.-A. (2004). Female trouble: The Menstrual Self-evaluation Scale and women's self-objectification. Psychology of Women Quarterly, 28, 22–26

Rocha, C. & Chelladurai, P. (2012). Item Parcels in Structural Equation Modeling: an Applied Study in Sport Management. *International Journal of Psychology and Behavioral Sciences. 2*. 46-53

Rodin, G. & Ahles,T.A. (2012). Accumulating evidence for the effect of chemotherapy on cognition. J. *Clin.Oncol. 30,* 3568–3569

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning*, *7*(3), 151-173

Rosenberg, L., & Park, S. (2002). Verbal and spatial functions across the menstrual cycle in healthy young women. *Psychoneuroendocrinology* 27, 835–841

Rouault, M., Seow, T., Gillan, C.M., & Fleming, S.M. (2018). Psychiatric Symptom Dimensions Are Associated with Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry, 84,* 6, 443-451

Rozin, P., & Royzman, E. B. (2001). "Negativity bias, negativity dominance, and contagion". *Personality and Social Psychology Review. 5* (4), 296–320

Ruble, D. N. (1977). Premenstrual symptoms: A reinterpretation. Science 197: 291-292.

Ruiz-Cortés, Z.T. (2012). Gonadal Sex Steroids: Production, Action and Interactions in Mammals. In (Ed.), Steroids - From Physiology to Clinical Medicine. IntechOpen.

Rundek, T., Roy, S., Hornig, M., Cheung, Y. K., Gardener, H., DeRosa, J., ... & Sacco, R. L. (2021). Gut permeability and cognitive decline: A pilot investigation in the Northern Manhattan Study. *Brain, behavior, & immunity-health*, *12*, 100214.

Ryu, A., & Kim, T. H. (2015). Premenstrual syndrome: A mini review. *Maturitas*, *82*(4), 436-440.

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,* 374.

Salas-Sender, M., López-Carrilero, R., Barajas, A. et al. (2019). Gender differences in response to metacognitive training in people with first-episode psychosis. *J Consult Clin Psychol* 88:516–525

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (p. 399–419). Sage Publications, Inc.

Saykin, A.J., Wishart, H.A., Rabin, L.A., et al. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. *Neurology*, *67*(5), 834-842

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. *Psychoneu-*

roendocrinology, *123*, 104895

Schmalenberger, K.M., Eisenlohr-Moul, T.A., Würth, L., Schneider, E., Thayer, J.F., Ditzen, B., & Jarczok, M.N., (2019). A systematic review and meta-analysis of withinperson changes in cardiac vagal activity across the menstrual cycle: implications for female health and future studies. *J. Clin. Med. 8,* 1946

Schmidt, P.J., Nieman, L.K., Danaceau, M.A., Adams, L.F., Rubinow, D.R. (1998). Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *N. Engl. J. Med. 338*, 209–216

Schmitt, T.A. (2011). Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment 29*(4), 304-321

Schmitz, T.W., Rowley, H.A., Kawahara, T.N. & Johnson, S.C. (2006). Neural correlates of self-evaluative accuracy after traumatic brain injury. *Neuropsychologia, 44*(5), 762–773.

Schoep, M. E., Nieboer, T. E., van der Zanden, M., Braat, D. D., & Nap, A. W. (2019). The impact of menstrual symptoms on everyday life: a survey among 42,879 women. *American journal of obstetrics and gynecology*, *220*(6), 569-e1

Schoning, S., Engelien, H., Kugel, S., Schafer, H., Schiffbauer, P., Zwitserlood, E., Pletziger, E., Beizai, P., Kersting, A., Ohrmann, P., Greb, R.R., Lehmann, W. Heindel, W., Arolt, V., & Konrad, C. (2007). Functional anatomy of visuo-spatial working memory during mental rotation is influenced by sex, menstrual cycle, and sex steroid hormones. *Neuropsychologia, 45*, 3203-3214

Schraw G., Dennison R. S. (1994). Assessing metacognitive awareness. *Contemp. Educ. Psychol. 19*, 460–475

Schütze, R., Rees, C., Smith, A., Slater, H., & O'Sullivan, P. (2020). Metacognition, perseverative thinking, and pain catastrophizing: a moderated-mediation analysis. *European Journal of Pain*, *24*(1), 223-233.

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology, Vol. 1. Attitudes and social cognition* (pp. 203–229). American Psychological Association.

Scott, R.B., Dienes, Z., Barret, A.B., Bor, B., & Seth, A.K. (2014). Blind Insight: Metacognitive Discrimination Despite Chance Task Performance. *Psychological Science, 25*(12) 2199–2208

Sellers, R., Wells, A., & Morrison, A. P. (2018). Are experiences of psychosis associated with unhelpful metacognitive coping strategies? A systematic review of the evidence. *Clinical psychology & psychotherapy*, *25*(1), 31-49.

Shah, M. G., & Maibach, H. I. (2001). Estrogen and skin: an overview. *American journal of clinical dermatology*, *2*, 143-150

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701–703

Sherwin, B. B. (2012). Estrogen and cognitive functioning in women: lessons we have learned. *Behav. Neurosci*. 126, 123–127

Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory: A study of the feeling-of knowing phenomenon in amnesic patients. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 12*(3), 452–460

Shimamura, A.P. (2000). Toward a Cognitive Neuroscience of Metacognition. *Consciousness & Cognition, 9(2), 313–323.*

Shirazi, T. N., Levenberg, K., Cunningham, H., Self, H., Dawood, K., Cárdenas, R., ... & Puts, D. A. (2021). Relationships between ovarian hormone concentrations and mental rotations performance in naturally-cycling women. *Hormones and Behavior*, *127*, 104886

Siedleckal, M., Paulewicz, B., Wierzchon, M. (2016). But I Was So Sure! Metacognitive Judgments Are Less Accurate Given Prospectively than Retrospectively. *Front. Psychol., 7,* 218

Sideridis, G., Simos, P., Papanicolaou, A., & Fletcher, J. (2014). Using Structural Equation Modeling to Assess Functional Connectivity in the Brain Power and Sample Size Considerations. *Educational and Psychological Measurement, 74(5):* 733–758.

Singh, K., Junnarkar, M., & Kaur, J. (2016). Measures of Positive Psychology, Develompent and Validation. Berlin: Spring

Slyepchenko, A., Lokuge, S., Nicholls, B., Steiner, M., Hall, G. B., Soares, C. N., & Frey, B. N. (2017). Subtle persistent working memory and selective attention deficits in women with premenstrual syndrome. *Psychiatry Research*, *249*, 354-362.

Smith, C. T., Wallace, D. L., Dang, L. C., Aarts, E., Jagust, W. J., D'Esposito, M., & Boettiger, C. A. (2016). Modulation of impulsivity and reward sensitivity in intertemporal choice by striatal and midbrain dopamine synthesis in healthy adults. *Journal of neurophysiology*, *115*(3), 1146-1156

Smith, G., Della Sala, S., Logie, R.H. & Maylor, E.A. (2000). Prospective and Retrospective Memory in Normal Aging and Dementia: A Questionnaire Study. *Memory, 8*, 311-321

Soares, C. N., Cohen, L. S., Otto, M. W., & Harlow, B. L. (2001). Characteristics of women with premenstrual dysphoric disorder (PMDD) who did or did not report history of depression: A preliminary report from the Harvard Study of Moods and Cycles. *Journal of Women's Health and Gender-Based Medicine, 10,* 873–878

Solem, S., Thunes, S. S., Hjemdal, O., Hagen, R., & Wells, A. (2015). A metacognitive perspective on mindfulness: An empirical investigation. *BMC psychology*, *3*(1), 1-10.

Solís-Ortiz, S. & Corsi-Cabrera, M. (2008). Sustained attention is favored by progesterone during early luteal phase and visuo-spatial memory by estrogens during ovulatory phase in young women. *Psychoneuroendocrinology, 33*, 989–998

Solis-Ortiz, S., Guevara, M.A. & Corsi-Cabrera, M. (2004). Performance in a test demanding prefrontal function is favored by early luteal phase progesterone: An electroencephalographic study. *Psychoneuroendocrinology, 29*, 1047–1057

Sommer, B. (1992). Cognitive performance and the menstrual cycle. In *Cognition and the menstrual cycle* (pp. 39-66). Springer, New York, NY

Song, C., Kanai, R., Fleming, S.M., Weil, R.S., Schwarzkopf, D.S., & Rees, G. (2011). Relating interindividual differences in metacognitive performance on different perceptual tasks. *Conscious Cogn. 20:* 1787–1792

Song, J.H.H, Loyal, S. & Lond, B. (2021) Metacognitive Awareness Scale, Domain Specific (MCAS-DS): Assessing Metacognitive Awareness During Raven's Progressive Matrices. *Front. Psychol., 11* 10.3389/fpsyg.2020.607577

Sotomayor-Zárate, R., Cruz, G., M Renard, G., Espinosa, P., & D Ramirez, V. (2014). Sex hormones and brain dopamine functions. *Central Nervous System Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Central Nervous System Agents)*, *14*(2), 62-71

Souza, E. G. V., Ramos, M. G., Hara, C., Stumpf, B. P., & Rocha, F. L. (2012). Neuropsychological performance and menstrual cycle: a literature review. *Trends in Psychiatry and Psychotherapy*, *34*, 5-12.

Spada, M. M., Gay, H., Nikčevic, A. V., Fernie, B. A., & Caselli, G. (2016). Meta-cognitive beliefs about worry and pain catastrophising as mediators between neuroticism and pain behavior. *Clinical Psychologist*, *20*(3), 138-146.

Spada, M. M., Georgiou, G. A., & Wells, A. (2010). The relationship among metacognitions, attentional control, and state anxiety. *Cognitive behavior therapy, 39*(1), 64–71.

Spada, M. M., Mohiyeddini, C., and Wells, A. (2008a). Measuring metacognitions associated with emotional distress: factor structure and predictive validity of the metacognitions questionnaire 30. Pers. Individ. Dif. 45, 238–242.

Spada, M. M., Nikčević, A. V., Moneta, G. B., and Wells, A. (2008b). Metacognition, perceived stress, and negative emotion. Pers. Individ. Dif. 44, 1172–1181

Spencer, D., Nietfeld, J. L., Cao, L., & Difrancesca, D. (2021). Exploring the interplay between attributions and metacognitive monitoring ability in a post-secondary classroom. *The Journal of Experimental Education*, 1-17

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and self-regulated learning constructs. Educational Research and Evaluation,10(2), 117–139

StataCorp. (2019). Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.

Steiger, J. H. & Lind, C. (1980). "Statistically based tests for the number of common factors", *Annual meeting of the Psychometric Society*, Iowa City, IA

Streiner, D. L. (2003) Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80,* 99-103

Steiner, M., Pearlstein, T., Cohen, L.S., Endicott, J., Kornstein, S.G., Roberts, C., Roberts, D.L., & Yonkers, K. (2006). *Journal of Women's Health,* 57-69

Stenfors, C. U., Jonsdottir, I. H., Hanson, L. M., & Theorell, T. (2017). Associations between systemic pro-inflammatory markers, cognitive function and cognitive complaints in a population-based sample of working adults. *Journal of psychosomatic research*, *96*, 49-59

Stern, Z., Katz-Navon, T. & Naveh, E, (2008). The Influence of Situational Learning Orientation, Autonomy and Voice on Error Making: The Case of Resident Physicians (September, 25 2008). *Management Science, Vol. 54,* pp. 1553-1564

Streiner, D.L. (2006). Research Methods in Psychiatry Building a Better Model: An Introduction to Structural Equation Modelling. *Can J Psychiatry, 51:* 317-324

Stoll, S., Shalita, A. R., Webster, G. F., Kaplan, R., Danesh, S., & Penstein, A. (2001). The effect of the menstrual cycle on acne. *Journal of the American Academy of Dermatology, 45,* 957–960

Stuart, M. J., Singhal, G., & Baune, B. T. (2015). Systematic review of the neurobiological relevance of chemokines to psychiatric disorders. *Frontiers in cellular neuroscience*, *9*, 357

Stumpf, W.E. (1968). Estradiol-concentrating neurons: topography in the hypothalamus by dry-mount autoradiography. *Science.162,* 1001–3.

Sun, X., Zhu, C., & So, S. H. W. (2017). Dysfunctional metacognition across psychopathologies: a meta-analytic review. *European Psychiatry*, *45*, 139-153.

Sundstrom-Poromaa, I. (2018). The Menstrual Cycle Influences Emotion but Has Limited Effect on Cognitive Function. *Vitam. Horm., 107,* 349–376

Sundström-Poromaa, I., Comasco, E., Sumner, R., & Luders, E. (2020). Progesterone–Friend or foe? *Frontiers in Neuroendocrinology*, *59*, 100856.

Sundström, Poromaa, I. & Gingnell, M. (2014). Menstrual cycle influence on cognitive function and emotion processing from a reproductive perspective. *Front. Neurosci, 8*, 8

Tabachnick, B.G & Fidell, L.S. (2001). *Using Multivariate Statistics, Fourth Edition.* Needham Heights, MA: Allyn & Bacon. ISBN 0-321-05677-9. Hardcover

Taipa, R., das Neves, S. P., Sousa, A. L., Fernandes, J., Pinto, C., Correia, A. P., ... & Sousa, N. (2019). Proinflammatory and anti-inflammatory cytokines in the CSF of patients with Alzheimer's disease and their correlation with cognitive decline. *Neurobiology of Aging*, *76*, 125-132

Tan, D. A., Haththotuwa, R., & Fraser, I. S. (2017). Cultural aspects and mythologies surrounding menstruation and abnormal uterine bleeding. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *40*, 121-133

Tavakol M, Dennick R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55

Teng, C. T., Vieira Filho, A. H. G., Artes, R., Gorenstein, C., Andrade, L. H., & Wang, Y. P. (2005). Premenstrual dysphoric symptoms amongst Brazilian college students: factor structure and methodological appraisal. *European archives of psychiatry and clinical neuroscience*, *255*(1), 51-56.

Thimm, M., Weis, S. & M. Hausmann, M. (2014). Menstrual cycle effects on selective attention and its underlying cortical networks. *Neuroscience 258:* 307–317

Thomas, A., Duran, A., & Abraham, B. P. (2020). Correlation of menstrual distress to severity of gastrointestinal symptoms in inflammatory bowel disease patients. *Indian Journal of Gastroenterology*, *39*(5), 514-520.

Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association

Timmermans B., Schilbach L., Pasquali A.& Cleeremans A.2012. Higher-order thoughts in action: consciousness as an unconscious re-description process. *Phil. Trans. R. Soc. B, 367,* 1412–

Toffoletto, S., Lanzenberger, R., Gingnell, M., Sundström-Poromaa, I., & Erika, C. (2014). Emotional and cognitive functional imaging of estrogen and progesterone effects in the female human brain: A systematic review. *Psychoneuroendocrinology, 50()*, 28–52

Troyer, A.K., & Rich, J.B. (2002). Psychometric properties of a new metamemory questionnaire for older adults. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *57,* 19–27

Tsang, S., Royse, C.F., & Terkawi, A.S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth. 11,* 80-89

Vaccaro, A.G., & Fleming, S.M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and Neuroscience Advances, 2()*

Valk S.L., Bernhardt, B.C., Böckler A, et al. (2016) Substrates of metacognition on perception and metacognition on higher-order cognition relate to different subsystems of the mentalizing network. *Human Brain Mapping 37*(10): 3388–3399.

van Iersel, K. C., Kiesner, J., Pastore, M., & Scholte, R. H. (2016). The impact of menstrual cycle-related physical symptoms on daily activities and psychological wellness among adolescent girls. *Journal of adolescence*, *49*, 81-90.

Vandenberg, S.G., & Kuse, A.R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Percept. Mot. Ski., 47*, 599–604

Von Ah, D. & Tallman, E.F. (2015). Perceived Cognitive Function in Breast Cancer Survivors: Evaluating Relationships with Objective Cognitive Performance and Other Symptoms Using the Functional Assessment of Cancer Therapy—Cognitive Function Instrument. *Journal of Pain and Symptom Management, 49*, 4, 697-706

Walker, A. E. (1997). *The menstrual cycle*. Psychology Press

Wang, Y. P., Teng, C. T., Vieira Filho, A. H. G., Gorenstein, C., & Andrade, L. H. (2007). Dimensionality of the premenstrual syndrome: confirmatory factor analysis of premenstrual dysphoric symptoms among college students. *Brazilian Journal of Medical and Biological Research*, *40*(5), 639-647.

Weber, M., & Mapstone, M. (2009). Memory complaints and memory performance in the menopausal transition. *Menopause*, *16*(4), 694-700.

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain: a journal of neurology*, *97*(4), 709–728.

Wells A, Simons M (2009) Metacognitive therapy. John Wiley & Sons, Ltd.

Wells, A. (1995). Meta-cognition and worry: A cognitive model of generalized anxiety disorder. *Behavioral and cognitive psychotherapy*, *23*(3), 301-320.

Wells, A. (2001). Metacognitive beliefs about rumination in recurrent major depression. *Cognitive and Behavioral Practice*, *8*(2), 160-164

Wells, A. (2013). Advances in metacognitive therapy. *International Journal of Cognitive Therapy*, *6*(2), 186-201.

Wells, A., & Carter, K. E. (2009). Maladaptive thought control strategies in generalized anxiety disorder, major depressive disorder, and nonpatient groups and relationships with trait anxiety. *International Journal of Cognitive Therapy*, *2*(3), 224-234.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: properties of the MCQ-30. *Behavior research and therapy*, *42*(4), 385-396.

Wells, A., & Matthews, G. (1996). Modelling cognition in emotional disorder: The S-REF model. *Behavior Research Therap*y, *34* (11), 881-888

Whitfield, K., (2005). An investigation of the influence of gender and menstrual cycle phase on metacognitive judgements of efficacy. Masters thesis. Simon Fraser University, Canada

Winne, P. H. (2018). Cognition and metacognition within self-regulated learning. In D. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 254–270). New York: Routledge.

Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Lawrence Erlbaum Associates Publishers.

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction*, *35*(4-5), 356-373.

Xu, M., Chen, D., Li, H., Wang, H., & Yang, L. Z. (2022). The Cycling Brain in the Workplace: Does Workload Modulate the Menstrual Cycle Effect on Cognition? Frontiers in behavioral neuroscience, 16

Yama, K., Asari, Y., Ono, A., Machida, M., & Miura, J. (2020). Plasma Interleukin-10 Levels Are Altered in Women with Severe Premenstrual Syndrome: A Preliminary Study. *Women's Health Reports*, *1*(1), 73-79

Yeung, D. Y., Tang, C. S. K., & Lee, A. (2005). Psychosocial and cultural factors influencing expectations of menarche: A study on Chinese premenarcheal teenage girls. *Journal of Adolescent Research*, *20*(1), 118-135.

Yılmaz, A. E., Gençöz, T., & Wells, A. (2011). The temporal precedence of metacognition in the development of anxiety and depression symptoms in the context of life-stress: A prospective study. *Journal of Anxiety Disorders*, *25*(3), 389-396.

Yoshida, T., Molton, I. R., Jensen, M. P., Nakamura, T., Arimura, T., Kubo, C., & Hosoi, M. (2012). Cognitions, metacognitions, and chronic pain. *Rehabilitation Psychology*, *57*(3), 207.

Zaka, M., & Mahmood, K. T. (2012). Pre-menstrual syndrome-a review. *Journal of Pharmaceutical Sciences and Research*, *4*(1), 1684.

Zarbo, C., Brugnera, A., Compare, A., Malandrino, C., Candeloro, I., Secomandi, R., Betto, E., Bellia, A., Rabboni, M., Bondi, E., & Frigerio, L. (2022). Psychological and cognitive

factors implicated in pain experience in women with endometriosis. *Health Care for Women International*.

Zarbo, C., Brugnera, A., Compare, A., Secomandi, R., Candeloro, I., Malandrino, C., ... & Frigerio, L. (2019). Negative metacognitive beliefs predict sexual distress over and above pain in women with endometriosis. *Archives of Women's Mental Health*, *22*(5), 575-582.

Zare, H., Abazarian Tehrani, M., & Alipour, A. (2013). The effect of menstrual cycle on metamemory, everyday memory and prospective memory in women aged 18-45 years. *The Iranian Journal of Obstetrics, Gynecology and Infertility*, *15*(41), 1-8.

Zendehdel, M., & Elyasi, F. (2018). Biopsychosocial etiology of premenstrual syndrome: A narrative review. *Journal of family medicine and primary care*, *7*(2), 346.

Ziadni, M. S., Sturgeon, J. A., & Darnall, B. D. (2018). The relationship between negative metacognitive thoughts, pain catastrophizing and adjustment to chronic pain. *European Journal of Pain*, *22*(4), 756-762.

Appendices

Appendix A

Questionnaire for Main Study

General Questions

Kindly answer the following questions:

1. Participant Code

2. On what day are you starting DOME?

(This question only needs to be answered once per cycle)

First day of menstruation

Second day of menstruation

3. The responses provided are in relation to:

Today

Yesterday

4. Did you have menstrual bleeding or spotting?

Yes

No

5. If yes, how many times did you change your pad/tampon during the day?

6. If a major event occurred today that you think affected your responses, please mention it (example car crash, holiday, fallout)

In the last 24 hours how would you rate these measures RELATIVE to how you usually feel. Your:

7. Ability to think clearly and efficiently

Worse than usual——————Better than usual

8. Cognitive Control (i.e., sense of feeling in control versus being mentally overwhelmed)

Worse than usual——————Better than usual

9. Creativity Level

Worse than usual——————Better than usual

10. Ability to concentrate and avoid distractions

Worse than usual——————Better than usual

11. Ability to multitask

Worse than usual——————Better than usual

12. Ability to pay attention to what is happening in the present

Worse than usual——————Better than usual

13. Ability to be decisive

Worse than usual——————Better than usual

14. Ability to make good decisions

Worse than usual——————Better than usual

15. Confidence in your decisions

Worse than usual——————Better than usual

16. Ability to remember information you know (e.g. a phone number, fact you had learnt)

Worse than usual——————Better than usual

17. Ability to learn new information

Worse than usual——————Better than usual

18. Ability to remember where you put an item (e.g. phone, bag, remote)

Worse than usual——————Better than usual

19. Ability to write clear instructions for others to follow

Worse than usual——————Better than usual

20. Ability to comprehend a long explanation or talk

Worse than usual——————Better than usual

21. Ability to give a long explanation or talk

Worse than usual——————Better than usual

<u>In the last 24 hours:</u>

22. Did you have trouble sleeping?

Not at all	A little	A moderate amount	Quite a bit	A lot

23. Did you have cramping or pain in the lower abdomen?

Not at all	A little	A moderate amount	Quite a bit	A lot

24. Did you experience joint or back pain?

Not at all	A little	A moderate amount	Quite a bit	A lot

25. Did you suffer from bloating or water retention?

Not at all	A little	A moderate amount	Quite a bit	A lot

26. Did you experience breast swelling or breast pain?

Not at all	A little	A moderate amount	Quite a bit	A lot

27. Did you experience an increase in pimples or oily skin?

Not at all	A little	A moderate amount	Quite a bit	A lot

28. Did you have less tolerance towards noise and light?

Not at all	A little	A moderate amount	Quite a bit	A lot

29. Did you have gastrointestinal or stomach complaints/pains?

Not at all	A little	A moderate amount	Quite a bit	A lot

30. How was your appetite?

Not at all hungry	A little hungry	Average hungry	Quite hungry	Almost always hungry

31. Did you have headache pain?

Not at all	Mild	Moderate	Severe	Excruciating

32. Did you experience an increase in bowel movements?

Not at all	A little	A moderate amount	Quite a bit	Very much

33. Did you feel anxious?

Not at all	A little	A moderate amount	Quite a bit	Very much

34. Did you feel tense?

Not at all	A little	A moderate amount	Quite a bit	Very much

35. Did you have sudden mood swings?

Not at all	A little	A moderate amount	Quite a bit	Very much

36. Did you feel irritable?

Not at all	A little	A moderate amount	Quite a bit	Very much

37. Did you feel sad?

Not at all	A little	A moderate amount	Quite a bit	Very much

38. Did you feel depressed?

Not at all	A little	A moderate amount	Quite a bit	Very much

39. Did you experience any Cognitive Fog (i.e., sense of confusion or lack of mental clarity)?

Not at all	A little	A moderate amount	Quite a bit	Very much

40. Did you experience any Cognitive Fatigue (i.e., feeling of mental fatigue or exhaustion)?

Not at all	A little	A moderate amount	Quite a bit	Very much

Appendix B

Consent Form

TITLE OF PROJECT:

The Mapping of Self-assessed Cognitive Symptoms throughout the Menstrual Cycle via DOME; an instrument for metacognition

By signing the following you confirm that you want to participate in this study.

You acknowledge that you have received enough information about the study, what it involves and the intended use of any data that you supply.

You also acknowledge that you understand that you are free to withdraw from the study without having to give a reason for withdrawing and without any adverse result of any kind.

No monetary payment will be provided for this study.

Signed ..... Date .....

(NAME IN BLOCK LETTERS) .....

Appendix C

Menstrual Cycle Questionnaire

Kindly fill in the following. You are free to leave out anything you do not feel comfortable filling in.

Code: _____

Age: _____

Country of birth: _____

How old were you when you started having menstrual periods? Age: _____
If you cannot remember your exact age, were you:

- ☐ Younger than 10
- ☐ 10-12 years old
- ☐ 13-15 years old
- ☐ 16 or older
- ☐ Don't know

Are your periods regular? (predictable within one week)

- ☐ Yes
- ☐ No

How many days are there between the start of one period and the start of the next on average?

- ☐ Less than 21 days
- ☐ 22-24 days
- ☐ 25-28 days
- ☐ 29-32 days
- ☐ 33 – 35 days
- ☐ More than 36 days

How many days of bleeding do you usually have each period? (Bleeding for which you needed a tampon or sanitary pad, NOT discharge for which you needed a panty liner only.) _____ days

How heavy is your menstrual flow usually?

- ☐ Light
- ☐ Moderate
- ☐ Heavy (clots/flooding)

Do you have any irregular bleeding between cycles?

- ☐ Yes - always
- ☐ Yes - sometimes
- ☐ No

Are you currently using any hormonal contraception? (This includes pills, injections, patches, the implant and MIRENA coil.)

- ☐ No
- ☐ Yes: (please specify) _____

Have you used any other hormonal contraception in the past? (This includes pills, injections, patches, the implant and MIRENA coil.)

- ☐ No
- ☐ Yes: (please specify) _____

Have you ever been pregnant?

- ☐ No
- ☐ Yes

Have you ever been diagnosed with any medical conditions?

☐ No

☐ Yes: (please specify) _____

Have you ever been diagnosed with any psychological/psychiatric conditions?

☐ No

☐ Yes: (please specify) _____

Have you ever been diagnosed with:

☐ Endometriosis

☐ Polycystic ovaries

☐ Pelvic inflammatory disease / Pelvic infection

☐ Uterine fibroids

☐ Blocked tubes

☐ Other menstrual cycle related problem (please specify): _____

Do you smoke?

☐ Never

☐ Occasionally

☐ Regularly

Do you exercise?

☐ Never

☐ Occasionally

☐ Regularly

Appendix D

Information Sheet

Title of project: The Mapping of Self-Assessed Cognitive Symptoms throughout the Menstrual Cycle via DOME, an instrument for metacognition

Investigator: Nicole Caruana

UM Ethics ID Number: 796:14/02/19-Nicole Caruana

Contact email: nicole.caruana.10@um.edu.mt

Dear potential participant, I would like to invite you to take part in this project which is part of my PhD in Cognitive Science at the University of Malta.

Project Overview

This research will involve data collection from approximately 150 women using an in-house questionnaire that measures changes in metacognition as a function of the menstrual cycle. The instrument – DOME (Daily Online Metacognitive Evaluation) comprises 40 items and is an online self-assessment of cognitive, physical and affective symptoms. The scope of the study is to map the reported symptoms to the menstrual cycle.

Eligibility Requirements:

Female participants over the age of 18 are requested for this study. People who have: gone through menopause, are pregnant, are taking oral contraceptives or other hormonal drugs, or do not have regular periods are not eligible to take part in the study.

What you will need to do and time commitment:

The whole questionnaire will take around six minutes and has to be filled in daily over a period of two months. As a participant you will be required to log in using the given link, and enter your answers to all 40 questions. The questionnaire cannot be submitted unless all questions are completed. You will be asked to answer the questions at the end of the day, after you reflect on your memory, attention and decision-making for that day.

On a given day if you feel that, for example, your attention was better than usual, drag the marker towards the 'better than usual' anchor. If you feel it was worse than usual,

drag the marker towards the 'worse than usual' anchor. On days when you think that there was no change from usual, leave the marker in the middle of the line. There is no right or wrong answer. Try not to overthink your responses and just go with what you 'feel'.

Additional information:

Participation in this project does not involve any known risks or discomforts. It is important to note that your participation is entirely voluntary, and you have the right to withdraw from the study at any time, even after agreeing to participate. Your personal information will remain confidential and will not be included in the research report. If you decide to take part in this research, please carefully read the consent form and sign it. Should you have any questions or concerns, please feel free to contact me or my supervisor using the provided contact details.

Regards

Nicole Caruana

PhD Cognitive Science

Professor Ian Thornton

Cognitive Science Department

ian.thornton@um.edu.mt

Appendix E

Project Description

Title of project: The Mapping of Self-Assessed Cognitive Symptoms throughout the Menstrual Cycle via DOME, an instrument for metacognition

Investigator: Nicole Caruana. Student Number: 05015892

Supervisors: Professor Ian Thornton (University of Malta); Professor Jeff Kiesner (University of Padua)

Project Overview

The menstrual cycle is a biological phenomenon with social implications. It influences a woman's personal identity, behaviour and perception of ability, and in turn shapes culturally shared beliefs and expectations about the role of women in society (Richardson, 1992). Women regularly report cognitive complaints which they attribute to the cycle (Farage, Osborn, & MacLe, 2008) but neurophysiological studies have indicated that cognition is not in fact affected (Sundsrom Poromaa & Gingnell, 2014). One possibility for the disparity between perceived and actual functioning could be that the menstrual cycle influences accurate monitoring of one's own cognitive ability (Whitfield, 2005).

This research project will monitor the metacognitive experience of women across two cycles, measuring daily self-assessment of cognitive ability. Metacognitive experiences are conscious cognitive experiences that pertain to intellectual enterprise, whose main purpose is to tell you how well you are doing (Flavell, 1999). An example would be making a judgement on whether you will be able to remember a given piece of information. The primary aims of this study are to determine whether changes are cyclical with respect to the menstrual cycle, the extent of individual differences in ability changes, and to examine the covariance of cognitive symptom changes, both among themselves and in relation to other symptoms. A systematic link between metacognition and the menstrual cycle and insights into metacognitive patterns in the context of the cycle will inform underlying causal mechanisms and pathways, and have implications in the diagnoses and treatment of menstrual conditions.

Data collection from approximately 150 women, will take place using a custom designed questionnaire that measures changes in metacognition. The instrument – DOME (Daily Online Metacognitive Evaluation) comprises 40 items and is an online self-assessment measure of cognitive abilities and physical and affective symptoms. Female participants over the age of 18 will be required for this study. People who have gone through menopause, are pregnant, are taking oral contraceptives or other hormonal drugs, or do not have regular periods are not eligible to take part. The whole

questionnaire takes around six minutes and will be be filled in daily over a period of two months. Participants log in using a given link. The questionnaire cannot be submitted unless all questions are completed.

There are no known risks associated with this project. Participation is completely voluntary, with no monetary rewards. No names or other identifying information will be included in the research report.

Appendix F

Raw Correlations

Table A1: Correlation scores for physical and affective cosine amplitudes

	Mood Swings	Depressed	Anxiety	Headaches	Skin	Joint Back Pain	GI	Cramps	Breast
Mood Swings	1.000	0.663	0.481	0.273	0.086	0.411	0.465	0.268	0.312
Depressed	0.663	1.000	0.575	0.295	0.023	0.371	0.368	0.219	0.293
Anxiety	0.481	0.575	1.000	0.128	0.045	0.162	0.227	0.120	0.293
Headaches	0.273	0.295	0.128	1.000	0.230	0.310	0.387	0.236	0.155
Skin	0.086	0.023	0.045	0.230	1.000	0.117	0.142	0.159	0.256
Joint Back Pain	0.411	0.371	0.162	0.310	0.117	1.000	0.498	0.517	0.510
GI	0.465	0.368	0.227	0.387	0.142	0.498	1.000	0.540	0.339
Cramps	0.268	0.219	0.120	0.236	0.159	0.517	0.540	1.000	0.317
Breast	0.312	0.293	0.293	0.155	0.256	0.510	0.339	0.317	1.000

Table A2: Correlation scores for physical, affective and cognitive cosine amplitudes

	Mood Swings	Depressed	Anxiety	Headaches	Skin	Joint Back Pain	GI	Cramps	Breast	Memory	Attention	General Cognition	Language	Decision Making
Mood Swings	1.000	0.663	0.481	0.273	0.086	0.411	0.465	0.268	0.312	0.530	0.504	0.502	0.563	0.471
Depressed	0.663	1.000	0.575	0.295	0.023	0.371	0.368	0.219	0.293	0.535	0.556	0.543	0.564	0.532
Anxiety	0.481	0.575	1.000	0.128	0.045	0.162	0.227	0.120	0.293	0.469	0.482	0.517	0.483	0.522
Headaches	0.273	0.295	0.128	1.000	0.230	0.310	0.387	0.236	0.155	0.106	0.178	0.118	0.167	0.106
Skin	0.086	0.023	0.045	0.230	1.000	0.117	0.142	0.159	0.256	0.110	0.139	0.050	0.083	0.059
Joint Back Pain	0.411	0.371	0.162	0.310	0.117	1.000	0.498	0.517	0.510	0.214	0.262	0.208	0.167	0.213
GI	0.465	0.368	0.227	0.387	0.142	0.498	1.000	0.540	0.339	0.328	0.356	0.275	0.403	0.319
Cramps	0.268	0.219	0.120	0.236	0.159	0.517	0.540	1.000	0.317	0.180	0.189	0.126	0.172	0.206
Breast	0.312	0.293	0.293	0.155	0.256	0.510	0.339	0.317	1.000	0.226	0.279	0.218	0.172	0.194
Memory	0.530	0.535	0.469	0.106	0.110	0.214	0.328	0.180	0.226	1.000	0.890	0.832	0.884	0.891
Attention	0.504	0.556	0.482	0.178	0.139	0.262	0.356	0.189	0.279	0.890	1.000	0.880	0.855	0.882
General Cognition	0.502	0.543	0.517	0.118	0.050	0.208	0.275	0.126	0.218	0.832	0.880	1.000	0.822	0.893
Language	0.563	0.564	0.483	0.167	0.083	0.167	0.403	0.172	0.172	0.884	0.855	0.822	1.000	0.875
Decision Making	0.471	0.532	0.522	0.106	0.059	0.213	0.319	0.206	0.194	0.891	0.882	0.893	0.875	1.000