Influence of sex steroid hormones on the neuroendocrine effects of evening light exposure in healthy adults

Carolina Guidolin [1, 2, [0009-0007-4959-2667]], Maydel Fernandez-Alonso [2 [0000-0002-3179-7476]], Johannes Zauner [1,2 [0000-0003-2171-4566]], Josef Trinkl [2, 3 [0009-0008-9064-4972]], Stephan Munkwitz [2 [0000-0002-9559-5566]], Manuel Spitschan [1, 2, 4, 5 [0000-0002-8572-9268]] *

[1] TUM School of Medicine and Health, Department Health and Sports Sciences, Chronobiology & Health, Technical University of Munich, Munich, Germany

[2] Max Planck Institute for Biological Cybernetics, Max Planck Research Group Translational Sensory & Circadian Neuroscience, Tübingen, Germany

[3] University Hospital and Faculty of Medicine, University of Tübingen, Tübingen, Germany

[4] TUM Institute for Advanced Study (TUM-IAS), Technical University of Munich, Garching, Germany

[5] TUMCREATE Ltd., Singapore, Singapore

* Corresponding author: Manuel Spitschan (manuel.spitschan@tum.de).

Abstract

Light regulates various non-visual functions in humans, including melatonin production, circadian phase shifting, alertness, and control of pupil size. These non-visual effects of light are mediated by melanopsin-containing intrinsically photosensitive retinal ganglion cells (ipRGCs). Individual sensitivity to non-visual light responses varies, with sex identified as contributing to these differences. Although the mechanisms behind sex differences in non-visual light effects are unclear, animal studies have shown that sex steroid hormones can modulate photic responses along the non-visual pathway. However, human studies are limited, often lacking documentation on participants' hormonal status and relying on between-subject designs, unable to capture hormonal fluctuations in naturally cycling individuals. Furthermore, the influence of exogenous hormones from contraceptive pill use remains largely unexplored. This within-subjects study addresses these gaps by examining how hormonal status modulates melanopsin sensitivity and evening melatonin suppression in healthy adults (aged 23–35). Three participant groups with distinct hormonal profiles will be recruited: naturally cycling individuals (NC group), individuals using monophasic combined oral contraceptives (MCOC group) and healthy male participants (HM group), with a maximum of 12 participants per group (n=36 total). Participants will undergo four evening experimental sessions across the menstrual cycle (NC group), the pill cycle (MCOC group), and an equivalent duration of 28 days (HM group). Each experimental session will span a total of 8 hours, starting seven hours prior and ending one hour after habitual bedtime, during which melanopsin sensitivity will be measured by state-of-the-art pupillometry, and light exposure to suppress melatonin will be delivered via a virtual reality head-mounted display (90 lux melanopic equivalent daylight illuminance [melanopic EDI]). Sex steroid hormone levels (estradiol, progesterone, testosterone, ethinylestradiol, synthetic progestins) will be measured through blood sampling. An additional dim-light (<1 lux melanopic EDI) control session to measure melatonin levels in the absence of light will be performed for each participant and randomised across the menstrual and pill cycle (NC and MCOC groups). This study will offer a first systematic account of how hormonal changes influence light sensitivity in healthy adults. Furthermore, our work will contribute to developing guidelines for including individuals with natural menstrual cycles in future circadian research, thereby encouraging more representative participant inclusion.

List of abbreviations

AUC	Area Under the Curve
AUDIT	Alcohol Use Disorders Identification Test
BF	Bayes factor
BFDA	Bayes Factor Design Analysis
BMI	Body Mass Index
CH	Confirmatory hypothesis
E2	Estradiol
E3G	Estrone-3-D-glucuronide
EDI	Equivalent daylight illuminance (e.g. melanopic EDI)
EDTA	Ethylenediaminetetraacetic acid
EE	Ethinylestradiol
ESS	Epworth Sleepiness Scale
FPR	False Positive Rate
HBT	Habitual bedtime
HM	Healthy male [participant]
I-PANAS-SF	International Positive and Negative Affect Schedule Short Form
ipRGCs	Intrinsically photosensitive retinal ganglion cells
KSS	Karolinska Sleepiness Scale
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry
LED	Light emitting diode
LH	Luteinising hormone
LMM	Liner Mixed Model
MCOC	Monophasic combined oral contraceptive
MCTQ	Munich Chronotype Questionnaire
NC	Naturally cycling
NSAIDs	Non-steroidal anti-inflammatory drug
P4	Progesterone

78	PeRBa	Perceived Research Burden Assessment
79	PIPR	Post Illumination Pupil Response
80	PMDD	Premenstrual Dysphoric Disorder
81	PSQI	Pittsburgh Sleep Quality Index
82	PSST	Premenstrual Symptoms Screening Tool
83	RSQ	Reproductive Status Questionnaire
84	SCN	Suprachiasmatic nucleus
85	SP	Synthetic progestin
86	T	Testosterone
87	TPR	True Positive Rate
88	VR	Virtual Reality

Introduction

Exposure to light regulates various physiological functions in humans, including the production of the hormone melatonin, the circadian timing system, alertness, and pupil size. These neuroendocrine, circadian, cognitive and pupillary effects of light exposure are termed "non-visual effects" and are mediated by the melanopsin-containing intrinsically photosensitive retinal ganglion cells (ipRGCs) in the retina (Blume et al., 2019). The contribution of melanopsin-containing ipRGCs to steady-state pupil size and melatonin suppression can be appreciated by looking at the spectral sensitivity of these responses, which is consistent with that of melanopsin (with peak at approximately 480 nm) (Spitschan, 2019).

There are significant individual variations in the effect of light exposure on people. For example, Phillips and colleagues (2019) demonstrated substantial differences in melatonin suppression to evening light exposure of almost a factor of 40: Some individuals showed 50% melatonin suppression near 10 lux while others required 400 lux to have the same response. These inter-individual differences may be due to state factors such as preceding light exposure, sometimes called photic history, and seasonal differences in physiology, behaviour or environmental exposures. Individual traits such as age, sex, chronotype, lens anatomy, pupil size, and genetic variability may also play a role (Chellappa, 2021; Spitschan & Santhi, 2022). Among these individual traits, sex has been analysed by a small number of studies as a potential factor explaining differences in non-visual effects response to light, with most studies focusing on melatonin suppression outcomes (for an overview of these studies and their methodologies, see Guidolin & Spitschan, 2024). Although conflicting results emerged from these studies, the two most recent and well-powered studies reported sex differences in melatonin suppression by light, with women exhibiting higher suppression at bright light intensities (Vidafar et al., 2024) as well as more moderate intensity (Fazlali et al., 2024) compared to men.

From a mechanistic perspective, these sex differences in non-visual effects of light could result from interactions between gonadal hormones, genetic sex, and environmental stimuli. Animal studies have demonstrated that gonadal hormones influence the response of the circadian system to photic inputs (Joye & Evans, 2022). For example, in female mice, recovery from simulated eastward jetlag is faster on the day of the proestrus compared to the day of the metestrus, and compared to males. This response is also mimicked by exogenous estradiol treatment, thus suggesting that gonadal hormones, and not just biological sex, influence photic processing along the non-visual circadian pathway (Joye & Evans, 2022). The site of action of gonadal hormones might be on the circadian pacemaker itself, the suprachiasmatic nucleus (SCN), and brain areas that project to the SCN. For example, in

mice, androgen receptors are densely localised in the SCN core itself, while estrogen receptors are sparsely distributed in the SCN, but are heavily expressed in areas that are direct (retina and intergeniculate leaflet) and indirect (dorsal raphe) afferents to the SCN (Mong et al., 2011).

Less is known about the effects of gonadal hormones on the non-visual effects of light in humans. This is due to the large sex data gap in circadian physiology and non-visual effects of light research (Spitschan et al., 2022). Along a woman's lifespan, major hormonal transition periods such as puberty, pregnancy, and perimenopause, as well as regular hormonal events such as the menstrual cycle, offer opportunities to study the role of sex steroids on brain functioning and behaviour. Similarly, the hormonal profile of individuals using oral contraceptives (the "pill cycle") also presents an opportunity to study the combined effect of exogenous sex hormones and inhibited endogenous sex hormones on the brain and behaviour (Hampson, 2023). These investigations are key to developing mechanistic models of how sex steroid hormones impact health and disease trajectories, and therefore essential for successful precision medicine interventions for both males and females (Barth et al., 2023).

Although sex differences in circadian physiology and non-visual effects of light have been reported (Spitschan et al., 2022), the role of sex steroid hormones in explaining these differences is unexplored. In our survey of the literature encompassing 11 published articles and one pre-printed article (Guidolin & Spitschan, 2024; *Supplementary materials*), we found that in many studies (n=5), participants' reproductive status was undocumented, and the menstrual or pill cycle phases were not accounted for. Furthermore, for participants using oral contraceptives, a lack of details on the specific pill formulation hinders investigations into the influence of exogenous hormones on non-visual effects of light of interest. To our knowledge, only very few studies have examined the effect of menstrual cycle phase on melatonin suppression. An earlier within-subjects study by Nathan and colleagues (1999) examined melatonin suppression across three menstrual cycle phases confirmed by menstrual onset and blood progesterone levels (follicular, luteal, and menstrual phases). While they found no significant effect, their small sample size (N=6) limited the study's power to detect an effect. Furthermore, despite including three menstrual phases, their phase classification approach was suboptimal, as no ovulation testing was performed, which is key to characterise hormonal fluctuations across the menstrual cycle (Schmalenberger et al., 2021). More recently, Vidafar and colleagues (2024) could not detect differences in melatonin suppression between the follicular and luteal phase across various illuminance levels (10 lux, 30 lux, 50 lux, 100 lux, 200 lux, 400 lux and 2000 lux at eye level). The number of participants in each phase differed for each light level (follicular phase n_{min}=4, n_{max}=16; luteal phase n_{min}=6, n_{max}=18). The authors also reported no effects of salivary estradiol or progesterone on melatonin suppression in

females, nor of testosterone in males. Similarly, Fazlali and colleagues (2024) observed no phase-related differences (follicular, n=11, versus luteal, n=13) in melatonin suppression under moderate light intensities (~150 lux melanopic EDI at eye level). Crucially, both studies were exploratory analyses of previously collected data, where the primary research question was unrelated to sex or menstrual cycle influences on melatonin suppression.

Specific methodological considerations must be incorporated starting in the study design phase to study how endogenous hormones across the menstrual cycle affect a physiological outcome of interest. It is recommended to treat the menstrual cycle as a within-person process, requiring a repeated-measures design covering at least one complete cycle (Schmalenberger et al., 2021). Furthermore, categorising cycle phases into “follicular” and “luteal” is discouraged, as each of these phases encompasses multiple hormonal changes (Schmalenberger et al., 2021), and sex steroid hormone levels should preferably be measured in blood, rather than saliva (Arslan et al., 2023; Barth et al., 2023). Finally, to correctly characterise menstrual cycle phase, it is essential that ovulation is confirmed, for example with at-home luteinising hormone (LH) urinary tests (Schmalenberger et al., 2021).

This work aims to assess the influence of endogenous and exogenous sex steroid hormones on the neuroendocrine effects of light exposure in healthy young adults in the evening, quantified by melatonin suppression by light. To this aim, we employ a repeated-measures within-subjects design with three participant groups: naturally cycling individuals (NC group), individuals taking monophasic combined oral contraceptives (MCOC group), and healthy male participants (HM group). To investigate the influence of both endogenous and exogenous sex hormones on melatonin suppression, we conduct dense sampling across the menstrual cycle in the NC group and pill cycle in the MCOC group. Finally, to investigate the site of action of endogenous and exogenous steroid hormones along the non-visual pathway, we assess whether these hormones influence the non-visual response to light at the retinal level (melanopsin sensitivity), quantified by measuring melanopsin-mediated pupil response.

Confirmatory Hypothesis 1. Based on previous literature, we hypothesise an effect of sex on melatonin suppression, with female participants exhibiting greater melatonin suppression by light compared to male participants (**CH1a**). Furthermore, since activation of melanopsin-containing ipRGCs causes melatonin suppression, we hypothesise an effect of sex on melanopsin sensitivity, whereby female participants exhibit larger melanopsin-mediated pupil responses than male participants (**CH1b**).

Confirmatory Hypothesis 2. We hypothesise a group effect on light-induced melatonin suppression, whereby melatonin suppression differs between naturally cycling individuals (NC group), individuals taking oral contraceptives (MCOC group), and healthy male participants

(HM group) (**CH2a**). Moreover, we hypothesise an effect of group (NC/MCOC/HM) on melanopsin-mediated pupil responses (**CH2b**).

Confirmatory Hypothesis 3. In naturally cycling individuals (NC group), we hypothesise an effect of the endogenous sex steroid hormones estradiol (E2) and progesterone (P4) on light-induced melatonin suppression (**CH3a**). We also expect an effect of E2 and P4 on melanopsin-mediated pupil responses (**CH3b**).

Confirmation Hypothesis 4. In individuals taking oral contraceptives (MCOC group), we hypothesise an effect of the exogenous sex hormones ethinylestradiol (EE) and synthetic progestin (SP) on light-induced melatonin suppression (**CH4a**). We also expect an effect of EE and SP on melanopsin-mediated pupil responses (**CH4b**).

A note on terminology

In this Registered Report, we refer to sex differences as differences between female participants (having reproductive organs consistent with the female biological sex) and male participants (having reproductive organs consistent with the male biological sex). The term "sex differences" is not interchangeable with "gender differences". Sex is determined by reproductive organs, genes, and hormones, while gender refers to socially and culturally defined characteristics (Tannenbaum et al., 2019; Torgrimson & Minson, 2005). We acknowledge that not every female participant in our study might identify as a woman, and not every male participant might identify as a man. While in this document we refer to female and male participants, during the study we will refer to participants with their preferred pronouns indicated at screening. A copy of the Sex and Gender Equity in Research (SAGER) Checklist (Heidari et al., 2016) for reporting sex and gender variables in human research is included in the *Supplementary materials*.

Methods

Ethics information

The study protocol will be approved by the Ethics Committee of the Technical University of Munich. During the in-person screening, participants will provide written informed consent before any experimental procedures. The study will be performed in line with the principles of the Declaration of Helsinki.

Sampling plan

Recruitment and screening

Participants will be recruited using several strategies, including fliers distributed around Tübingen or shared on social media. We will also use university and campus mailing lists and word of mouth to facilitate recruitment. The recruitment materials will redirect participants to an online screening form implemented on REDCap (Harris et al., 2009, 2019), which contains questionnaires assessing participants' eligibility for the study. Eligible individuals will then be invited to an in-person screening, where further eligibility criteria will be tested (for a detailed list of online and in-person inclusion and exclusion criteria, see Table 2). During the in-person screening, participants will be introduced to the laboratory space, instruments, and protocols, so they can familiarise themselves with the environment of the experimental sessions. Moreover, they will download the MyCap app (Harris et al., 2022), which will be used to fill in at-home questionnaires during the study (see *Continuous measurements*). Participants will be given the opportunity to ask any questions before providing written consent to participate in the study. All participant-facing information, including all questionnaires and study-related information, will be in English. Any verbal exchange between the researchers and the participants (in-person or via email) will also take place in English.

Participant compensation

Participants will be remunerated for their time. We consider “time spent participating in the study” as the time spent in the laboratory during experimental sessions, the number of weeks of keeping and logging a regular sleep/wake schedule, the time spent logging mood symptoms, the time spent performing at-home luteinising hormone (LH) tests (NC group), and the time spent logging oral contraceptive intake timing (MCOC group). This compensation strategy implies that different participants might be remunerated differently based on the participant group they belong to (NC group, MCOC group, or HM group). For example, the number of weeks of regular sleep/wake schedule might differ between participants due to our

scheduling strategy (see *Participants scheduling* below). Furthermore, to acknowledge the flexibility that NC group participants will adopt during the study period in terms of scheduling sessions sequentially, rather than establishing dates of experimental sessions at screening as for the other two groups, these participants will receive a bonus remuneration for each scheduled experimental session. Participants will also receive a bonus if they complete all experimental sessions and the discharge questionnaire (n=5). The compensation strategy for this study is summarised in *Supplementary materials* (Table S1).

Inclusion and exclusion criteria

We will recruit healthy participants aged between 23 and 35 years with good English language proficiency. The choice of this age group is due to the need for full ovulatory competence in the naturally cycling group, which is highest from the mid-twenties to mid-thirties of a female's lifespan (Hampson, 2020). Physical, mental, ocular and retinal health will be assessed during online screening and/or in-person screening (Table 2).

Domain	Criterion	Assessment method	Time of assessment
Age	≥23, ≤35 years old	Self-report	Screening (online questionnaire)
Ability to follow instructions in English	Good English proficiency	Self-report	Screening (online questionnaire)
Physical and mental health	Good physical and mental health	Self-report	Screening (online questionnaire)
Ocular and retinal health	Good ocular and retinal health	Self-report	Screening (online questionnaire)
Visual acuity	Normal or corrected-to-normal acuity	Landolt C chart or similar	Screening (in-person)
Colour vision	Normal colour vision	HRR colour vision test	Screening (in-person)

Table 1. Inclusion criteria applied to all participants.

We will exclude participants who currently have a diagnosis of chronic, neurological, and psychiatric disease. We will also exclude individuals who are underweight or obese, take any medication (excluding allergy medication), regularly use sleep aids and regularly consume nicotine, recreational drugs, and/or excessive alcohol. Furthermore, individuals with poor sleep quality, extreme early or late chronotypes and excessive daytime sleepiness will not be able to participate. We will also exclude individuals who have worked shifts or have travelled across time-zones (>1 hour) in the month prior to screening, according to moderate criteria in chronobiology studies (Yousefzadehfard et al., 2022). Individuals who have allergies to the foods served during the experimental sessions will be excluded. In terms of reproductive health and hormonal status, we will exclude intersex individuals and those with a record of bilateral oophorectomy or gender-affirming surgery, individuals who have a diagnosis of hormonal disorders, individuals with moderate to severe premenstrual dysphoric disorder (PMDD) symptoms, and individuals who take any exogenous hormones (except hormonal contraceptives). Furthermore, individuals who are currently pregnant or have been pregnant or breastfeeding in the last 12 months, and individuals who experience severe physical symptoms related to the menstrual cycle, will not be able to participate (Table 2).

Domain	**Criterion**	**Assessment method**	**Time of assessment**
Body Mass Index (BMI)	Underweight or obese, <18.5, >30 kg/m^2	Measured height and weight	Screening (in-person)
Medication use	Any regular or habitual use of systemic medications, except antiallergic medication	Self-report	Screening (online questionnaire)
Sleep aids use	Any use regular or habitual (≥1/week) of sleep aids	Self-report	Screening (online questionnaire)
Health condition	Diagnosis of any chronic, hormonal, ocular or neurological diseases or disorders	Self-report	Screening (online questionnaire)

Mental health	Diagnosis of psychiatric diseases or disorders (mood, anxiety, personality, post-traumatic stress, obsessive compulsive, panic, and eating disorders)	Self-report	Screening (online questionnaire)
Nicotine consumption	≥2 times/week	Self-report	Screening (online questionnaire)
Recreational drug use	Any regular use of recreational drugs (≥2 times/week)	Self-report	Screening (online questionnaire)
Alcohol abuse	Excessive alcohol use	AUDIT ≥8	Screening (online questionnaire)
Sleep quality	Poor sleep quality	PSQI >5	Screening (online questionnaire)
Sleep duration	Extremely short or extremely long sleep duration on workdays	Subjective sleep duration on workdays <6 or >10 hours (based on the MCTQ)	Screening (online questionnaire)
Chronotype	Extreme early or extreme late chronotype	MCTQ ≤2:00 MCTQ >5:00	Screening (online questionnaire)
Daytime sleepiness	Excessive daytime sleepiness	ESS >10	Screening (online questionnaire)
Shift work	No shift work in the past 1 month	Self-report	Screening (online questionnaire)

Time zone travel	No inter-time zone travel >1 hour in the past one month	Self-report	Screening (online questionnaire)
Food-related allergies	Allergy or intolerance for food ingredients in experimental meals	Self-report	Screening (online questionnaire)
Reproductive organs	Not consistent with either female or male biological sex	Modified RSQ	Screening (online questionnaire)
Reproductive system surgery	Record of bilateral oophorectomy or gender affirming surgery	Modified RSQ	Screening (online questionnaire)
Hormone intake	Intake of any exogenous hormones (with exception of contraceptive hormones)	Modified RSQ	Screening (online questionnaire)
Hormonal disorder	Diagnosis of hormonal disorder	Modified RSQ	Screening (online questionnaire)
Current or past pregnancy	Being currently pregnant or having had a pregnancy or having been breast-feeding in the last 12 months	Modified RSQ	Screening (online questionnaire)
Menstrual symptoms	Frequent severe physical symptoms related to menstrual cycle	Modified RSQ	Screening (online questionnaire)

Premenstrual dysphoric disorder	Moderate to severe PMDD symptoms	PSST	Screening (online questionnaire)

Table 2. Exclusion criteria. Abbreviations: BMI, Body Mass Index; AUDIT, Alcohol Use Disorders Identification Test (Saunders et al., 1993); PSQI, Pittsburgh Sleep Quality Index (Buysse et al., 1989); MCTQ, Munich Chronotype Questionnaire (Roenneberg et al., 2003); ESS, Epworth Sleepiness Scale (Johns, 1991); RSQ, Reproductive Status Questionnaire (Schmalenberger et al., 2021, modified in that questions covered by other questionnaires were removed); PSST, Premenstrual Symptoms Screening Tool (Steiner et al., 2003).

In addition to these general inclusion and exclusion criteria, additional criteria will be applied to individuals participating as part of the NC group and MCOC groups (Table 3).

Group	Domain	Criterion	Assessment method
NC group	**Menarche timing**	Menarche ≥3 years from experiment start	Modified RSQ
	Cycle information	Records (start dates) of at least the last six cycles	Modified RSQ
	Cycle information source	Records logged using: • Calendar (digital or paper) • Health app on smartphone • Cycle tracking app (to be confirmed by study team)	Modified RSQ
	Cycle regularity	Length of the last six cycles ≥26 days and ≤35 days	Modified RSQ

	Cycle variation	Variation in length of the last six cycles ≤6 days	Modified RSQ
	Last use of hormonal contraceptives	Last use of any hormonal contraceptives and copper intrauterine device (IUD) ≥6 menstrual cycles	Modified RSQ
MCOC group	**Pill type**	Monophasic combined oral contraceptive pill	Self-report
	Pill formulation	0.03 mg ethinylestradiol and 2 mg dienogest formulation (anti-androgenic)	Self-report
	Pill intake history	Stable pill intake regimen over the last six months	Self-report

Table 3. Inclusion criteria specific to the NC group and the MCOC group. Abbreviation: RSQ, Reproductive Status Questionnaire (Schmalenberger et al., 2021, modified in that questions covered by other questionnaires were removed).

Principles for adjusting the inclusion/exclusion criteria of the study

We are aware that our inclusion/exclusion criteria for the NC and MCOC groups may be stringent. For this reason, we cannot predict whether our target sample size for these two groups can be reached. In case recruitment of participants is not successful and n≥5 NC group participants have been excluded solely due to the criteria for cycle information, cycle regularity, cycle variation, and last use of hormonal contraceptives, we will adjust the inclusion criteria. Similarly, after n≥5 MCOC group individuals have been excluded solely due to the criteria for pill formulation and pill intake history, we will adjust the inclusion criteria. Strategies for criteria adjustments in both participant groups are shown in Table 4.

Group	**Domain**	**Old criterion**	**New criterion**

NC group	**Cycle information**	Records (start dates) of at least the last six cycles	Records (start dates) of at least the last three cycles
	Cycle regularity	Cycle length of the last six cycles ≥26 days and ≤35 days	Cycle length of the last three cycles ≥26 days and ≤35 days
	Cycle variation	Variation in length of the last six cycles ≤6 days	Variation in length of the last six cycles ≤8 days
	Last use of hormonal contraceptives	Last use of any hormonal contraceptives ≥6 menstrual cycles	Last use of any hormonal contraceptives ≥3 menstrual cycles
MCOC group	**Pill formulation**	0.03 mg ethinylestradiol and 2 mg dienogest formulation (anti-androgenic)	Any monophasic and anti-androgenic pill formulation
	Pill intake history	Stable pill intake regimen over the last six months	Stable pill intake regimen over the last three months

Table 4. Adjusted inclusion criteria for the MCOC group.

In case the inclusion and exclusion criteria are prohibitive and do not lead to successful recruitment of our sample size, we will review the criteria and make adjustments. Any such adjustments or protocol deviations will be reported as part of future publications.

Sample size

A summary of effect sizes calculated from existing literature on sex- and hormone-related differences in non-visual effects of light is reported in the *Supplementary materials* (Figure S1 and Figure S2), indicating a range of typical sample sizes between N=6 and N=55. For the present study, we used Bayes Factor Design Analysis (BFDA; Schönbrodt & Wagenmakers, 2018) for a fixed-n design to evaluate the expected evidence strength of our design with n=12

participants for each group. The choice of n=12 for each participant group is based on resources and time constraints of the current study, as we estimate that data collection for n=36 (i.e. N=12 participants for each of the three groups) will take approximately 18 months.

BFDA was conducted specifically for hypothesis CH3a, which tests the effect of endogenous hormones on melatonin suppression in the NC group. Since linear mixed-effects models (LMMs) represent the statistical modelling approach for all analyses, BFDA for LMMs was used to estimate the strength of evidence for the full model ($CH3a_{Full}$). First, due to the lack of reported effect sizes for the influence of endogenous hormones on melatonin suppression, we used a previous dataset of melatonin suppression using a comparable protocol (Fernandez-Alonso & Spitschan, 2025) to identify which effect sizes for E2 and P4 yielded the physiologically plausible melatonin suppression data. This dataset contains melatonin suppression values for n=27 participants (mean±1SD=49±27.7%). The results indicated that effect sizes for E2 and P4 ranging from -0.5 to 0.5 at 0.1 steps and intercept values from -0.9 to 0.9 at 0.3 steps yielded values within the 2.5^{th} and 97.5^{th} percentiles (1.4% and 90%) of our known melatonin suppression data. Once these effect sizes combinations (n=847) were identified, we performed BFDA simulations using the *lmBF* function from the *BayesFactor* R package (Morey, 2023, version 0.9.12-2), using a default prior distribution (Cauchy prior distribution with scale parameter $r = \sqrt{(2)}/2$) and performing 100 iterations over each of the 847 effect sizes combinations. The full model included melatonin suppression as the outcome, E2 and P4 as fixed effects, and random intercepts and slopes for E2 and P4 by participant. The null model included only a random intercept for each participant. The BF resulting from the ratio of the full model over the null model was then calculated for each iteration. For each parameter combination, we determined the true positive rate (TPR) as the proportion of BF>3 among the 100 iterations for that parameter combination.

Our results showed that higher absolute values (±0.5) of E2 and P4 yielded higher TPRs, while varying intercept values had a negligible impact on the TPR (*Supplementary materials Figure S3*), consistent with expectations since E2 and P4 were the predictors of interest. Out of n=847 parameter combinations, only n=157 yielded a TPR≥0.8. For E2, 51.2% of these values were generated from either an effect size of -0.5 or 0.5, while the remaining 49.8% were generated by smaller effect sizes. Similarly, for P4, 54.8% of the n=157 TPR≥0.8 values were generated from effect sizes of -0.5 or 0.5, while the remaining 49.8% were generated by smaller effect sizes. When simulating data under the null model (E2 and P4 effect sizes equal to 0), the average false positive rate (FPR) calculated under the seven possible intercept values was 0.12±0.027 (mean±1SD). These results highlight that if the true effect sizes of E2 and P4 (i.e. those collected in our study) are moderate to large, then there is a high probability

that the study will yield strong evidence for those effects, even with the limited sample size of n=12. On the other hand, for smaller effect sizes (e.g. ±0.3 or ±0.2), the current sample size might be insufficient to detect these effects. We believe that these findings justify the chosen sample size under the practical resource and time constraints. Visualisations of the BFDA results are reported in the *Supplementary materials (Figure* S3 and *S4).* Analysis code and files for the BFDA simulations are provided in the GitHub repository associated with this manuscript (https://github.com/tscnlab/GuidolinEtAl_PCIRRStage1_2025/tree/main/bfda).

Design

Participant scheduling

This study will use a repeated-measures within-subject design, with melatonin suppression and melanopsin sensitivity measured at four time points across the menstrual cycle (NC group), pill cycle (MCOC group) and across a corresponding time period of 28 days (HM group). Participants will also undergo an additional experimental session, in which melatonin concentration will be measured without light exposure (dim light control session). Experimental sessions will be scheduled so that a maximum of five participants can participate during the same evening. Participants will not interact during the experimental blocks, and each of them will sit in individual experimental booths, which are separated from each other by partition screens or walls. Two experimental sessions will be planned for each week, spaced no more than three days apart, excluding weekend days (i.e. Monday and Thursday or Tuesday and Friday). This frequency is chosen for feasibility due to staff and resource constraints. Furthermore, this strategy is expected to guarantee coverage of menstrual phases of the NC group, ensuring that a given menstrual phase falls within at least one of the two scheduled weekly sessions. Aligning the MCOC and HM groups to these same sessions as the NC group allows efficient data collection during the same evenings. In the following paragraphs, we outline scheduling principles for the experimental sessions for all three participant groups. Stopping criteria are also outlined in the *Supplementary materials (Table S2).*

Naturally cycling individuals

In order to capture different hormonal profiles across the menstrual cycle, we will schedule naturally cycling individuals to visit the laboratory during four menstrual phases with distinct hormonal levels: peri-ovulatory phase (rising E2, low P4), mid-luteal phase (high and stable E2 and P4), peri-menstrual phase (falling and low E2 and P4), and mid-follicular phase (slight rise in E2, low P4). The criteria to schedule the participants across the cycle are as follows (Figure 1A):

1. Peri-ovulatory phase (scheduled based on calendar method prediction and presence of at least one day of positive E2 test, and performed latest on day +1 from LH positive test, where day 0 indicates positive LH test)
2. Mid-luteal phase (day +6 to day +10 following LH test, where day 0 indicates positive LH test)
3. Peri-menstrual phase (day +12 to day +16 following positive LH test, where day 0 indicates positive LH test)
4. Mid-follicular phase (day +4 to day +8 of the next menstrual cycle, where cycle day +1 indicates the day of menses onset)

Compared to the recommended phasing criteria for the peri-ovulatory phase detailed in Schmalenberger et al. (2021) (-2 days prior to LH test to day +1 from positive LH test, where positive LH test is day 0), we choose to apply slightly less strict criteria for feasibility reasons. Relying solely on at-home LH testing would necessitate scheduling experimental sessions on very short notice, posing logistical challenges. To address this, we will use the Clearblue Advanced Digital ovulation tests (Clearblue, Swiss Precision Diagnostics, GmbH, Geneva, Switzerland; details in *Luteinising hormone (LH) testing and logging*). These tests detect both the rise in E2 preceding ovulation and the subsequent LH surge. This dual detection allows for greater flexibility in scheduling the peri-ovulatory session, as the E2-positive result typically precedes the LH-positive result by several days. However, because E2-positive duration is unpredictable and varies across individuals, we will combine calendar projections with hormone test results to schedule peri-ovulatory visits as follows:

- First, we will calculate the next expected menses onset based on the participant's average cycle length from the past six cycles.
- Then, we will estimate the day of the LH positive test by subtracting 14 days (i.e. the typical duration of the luteal phase) from this calculated next menses onset. The provisional peri-ovulatory session will be scheduled as close as possible to this predicted day. For instance, if the estimated LH-positive day is a Wednesday and a session is already scheduled for Thursday, participants will be provisionally invited for that Thursday.
- To confirm this session, participants must have at least one E2-positive test result.
- If an E2-positive result occurs >9 days before the provisionally scheduled session, the visit will be rescheduled earlier. This adjustment ensures the session does not occur more than one day after the LH-positive test in case the participant is having a shorter than expected follicular phase (i.e., not beyond day +1).

We believe this approach enables the reliable capture of rising and high E2 levels, which includes both the late follicular rise and the peri-ovulatory peak in E2. We will reassess this method after the first five participants complete all sessions to ensure the peri-ovulatory sampling window remains ≤5 days across individuals. Furthermore, compared to Schmalenberger and colleagues (2021), we extend the upper limit of the mid-follicular phase window from day +7 to day +8. This extension provides greater flexibility in scheduling mid-follicular sessions at a longer interval from the peri-menstrual visit.

Participant in-person screening will occur during the first menstrual cycle ("cycle 0") upon successful confirmation of eligibility and inclusion from the online screening. During the in-person screening, final eligibility and inclusion will be confirmed. The specific phase in the cycle at which participants undergo in-person screening will not be predetermined but will instead depend on availability. Between five and seven days from the in-person screening day, participants will complete the dim light control experimental session. This strategy will ensure that the dim light control will be performed at a random cycle phase for each participant. This dim light control session serves as a control for all experimental sessions with bright light exposure and prevents the need for dim light control sessions for each cycle phase, which we deem too burdensome on the participants. The next menstrual cycle ("cycle 1") will be the cycle in which experimental sessions with bright light exposure will be scheduled. Sampling will follow the same order for all participants and will start from the peri-ovulatory phase, as this phase can be predicted using E2 and LH tests, and end at the mid-follicular phase of the following cycle. This scheduling method implies that the experimental sessions will span two menstrual cycles: the experimental sessions for the peri-ovulatory and mid-luteal phases will occur during the first cycle (cycle 1), while the experimental session for the mid-follicular phase will occur during the following cycle ("cycle 2"). The experimental session for the peri-menstrual phase will fall either at the end of cycle 1 or at the beginning of cycle 2 (since the peri-menstrual phase spans between -3 days before menses onset to cycle day +2 menses onset, where cycle day +1 is the day of menses onset). A calendar visualisation of our scheduling system across two menstrual cycles is available in the *Supplementary materials (Figure S5*).

Participants will test for E2 and LH at home starting four days prior to their earliest estimated possible ovulation day. The earliest possible ovulation day is calculated by subtracting 14 days (i.e. typical luteal phase duration) from the shortest cycle length of their last six menstrual cycles. Once the test detects a rise in E2, the pre-scheduled peri-ovulatory experimental session will be confirmed. Participants will continue to test until obtaining a positive LH result to verify an ovulatory cycle. Following this, the mid-luteal session is

scheduled between days +6 to +10 and the peri-menstrual session between days +12 to +16, where day 0 is the LH-positive day. If the peri-menstrual session falls on day +1 of the cycle and the participant is experiencing dysmenorrhea (based on entries on *Pill intake and menses onset logging*, details in *Continuous measurements*) to an extent that makes participation unpleasant, the peri-menstrual experimental session will be rescheduled. Participants will also report the onset of their next menses (start date of cycle 2, entries on the *Pill intake and menses onset logging*), and the mid-follicular phase will be scheduled between days +4 to +8 days from menses onset (menses onset = day +1). Participants then repeat ovulation testing in cycle 2 to confirm it is ovulatory, as an anovulatory cycle could affect mid-follicular hormone levels (Hambridge et al., 2013). Finally, participants report the onset of their next menses (start of cycle 3) to complete cycle tracking.

Oral contraceptive (MCOC) group and males (HM) group

At screening, participants in the MCOC and HM groups will be allowed to choose one day of the week for the experimental session to take place out of a choice of two (e.g. either Mondays or Thursdays). Once a day of the week has been selected, they will come to the laboratory once a week for five consecutive weeks on this chosen day (one dim light control and four experimental sessions). The first experimental session will be scheduled at least five days from the in-person screening. For the MCOC group, since the pill cycle has a length of 28 days, we will be able to capture a whole pill cycle, including three experimental sessions during the active phase (i.e. oral contraceptive pill intake), and one during the inactive phase (i.e. placebo pill intake) (Figure 1B).

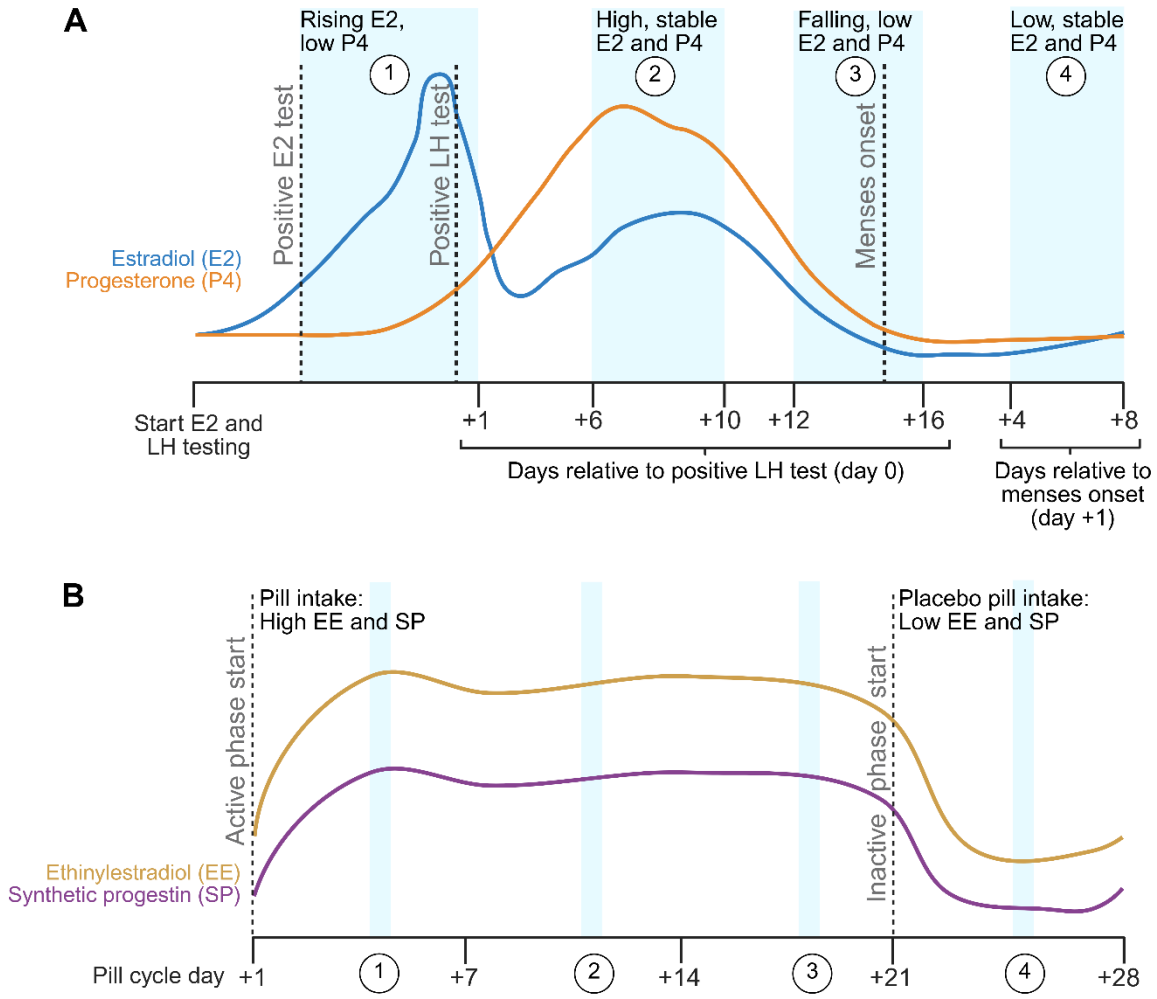

Figure 1: Bright light experimental sessions for NC and MCOC groups. **(A)** Example experimental sessions for an NC group participant. Experimental sessions will be scheduled starting from the peri-ovulatory phase (1). A provisional date for this session will be scheduled in advance based on the calendar method and confirmed by at least one day of a positive E2 test. The experimental session will take place latest on day +1 from a positive LH test. The day of the positive LH test will determine the date for the mid-luteal phase session (2) and the peri-menstrual session (3). Once menses onset (day +1 of cycle 2) is confirmed by the participant, the mid-follicular phase will be scheduled (4). **(B)** Example experimental sessions for an MCOC group participant. Experimental sessions will be scheduled starting at a random point in their pill cycle (here: starting during the first active phase week). Experimental sessions will be scheduled weekly on the same day during three weeks of the active phase (1, 2 and 3), and during the inactive phase (4). This exemplary figure was made using BioRender.com.

Randomisation of experimental sessions and blinding

This study comprises five experimental sessions: four experimental sessions with bright light exposure, and one dim light control session with no light exposure. The dim light control will always be the first experimental session for all groups. For the NC group, this will be performed at a random time point of menstrual cycle 0, as explained above (see *Naturally cycling individuals*). For the MCOC and HM groups, this will be performed as the first weekly experimental session out of the five sessions. The following four experimental sessions will be bright light exposure sessions (one for each target menstrual cycle and pill cycle phases for the NC and MCOC groups). While this system prevents the researcher from being blind to

the experimental condition (dim light control or bright light exposure), it ensures that the dim light control occurs at a random menstrual and pill cycle phase for each participant of the NC and MCOC groups. This is important because it ensures that a specific hormonal profile does not influence the dim light control session, which could influence the melatonin suppression calculated for the bright light exposure sessions.

Circadian stabilisation

Participants will be asked to follow a regular sleep/wake cycle schedule (16 hours wake, 8 hours sleep, with ±30 minutes deviation allowed) from at least five days prior to the first experimental session until the last experimental session. This sleep/wake schedule will be self-selected during the in-person screening. For the MCOC and HM groups, this one week of circadian stabilisation will be followed by five experimental weeks, with one experimental session for each week. For the NC group, there will be a variable time between the first experimental session (dim light control during cycle 0) and the bright light experimental sessions, depending on when the in-person screening during cycle 0 will take place. Participants will keep a regular sleep/wake cycle for at least five days prior to the dim light control session, and then again at least five days prior to the earliest expected peri-ovulatory experimental session. Adherence to this schedule will be monitored continuously using self-reported sleep and wake times (see *Morning sleep diary*), as well as a visual check of actimetry, light exposure and skin temperature data prior to each experimental session. On the morning after an experimental session, participants will be allowed to sleep two hours longer than their self-selected schedule. Stopping criteria for non-compliance with the selected sleep/wake schedule are outlined in the *Supplementary materials (Table S2).*

In-laboratory experimental procedure

Participants will arrive at the laboratory seven hours prior to their habitual bedtime (HBT). After compliance with their self-selected sleep-wake schedule is verified, they will first undergo alcohol and tetrahydrocannabinol (THC) screening using a urine multidrug test (Drug-Screen Multi 5, nal von minden GmbH, Moers, Germany) and a breath analyser (Breathalyzer, ACE Instruments GmbH, Freilassing, Germany), respectively. Then, participants will take turns to undergo pupillometry measurements with the retinaWISE instrument (Gibertoni et al., 2024) in a custom-made dark experimental tent. Light levels in the tent during stimulus delivery via this instrument will be kept to a minimum (<1 melanopic EDI lux). While waiting for their turn to undergo pupillometry measurements, participants will sit in a private, dimly lit booth (<10 melanopic EDI lux) and consume a previously selected meal from Huel (choice between Chicken Mushroom Pasta flavour and Thai Green Curry flavour; Hot & Savoury Pots, Huel Ltd., Tring, Hertfordshire, United Kingdom). The selected flavour will be kept constant

throughout the five experimental sessions, and participants will rinse their mouths with water after finishing the meal. After having completed the pupillometry measurements, participants will stay in their private booth and be allowed to read, write, draw, or listen to podcasts or audiobooks through in-ear headphones. Four hours prior to HBT, the dim light exposure lasting 120 minutes will start. After 90 minutes of dim light exposure, participants' pupils will be dilated using eye drops (Mydriaticum Stulln®, Tropicamide 5 mg/1 ml, Pharma Stulln GmbH, Stulln, Germany). Half an hour later, i.e. at two hours prior to HBT, a blood draw will be taken and immediately after, participants will wear a virtual reality (VR) headset to view either a bright light of ~90 lux melanopic EDI (n=4 experimental sessions) or a dark stimulus of <1 lux melanopic EDI (n=1 dim light control session). At the end of these two hours (i.e. at HBT), participants will remove the headsets, undergo a second blood draw, and sit in their experimental booth (<10 lux melanopic EDI). After one hour in this condition, the experimental session will be finished, and participants can return home. For this, given the late time in the evening, a taxi service will be offered to them.

Salivary samples for melatonin analysis will be collected every 30 minutes throughout the experimental session, starting five hours prior to HBT. From four hours to two hours before HBT, the sampling schedule will be organised in 30–minute "blocks" as follows (see Figure 1 inset A). Each block will start with a salivary sample, followed by three brief neurobehavioural questionnaires about participants' alertness, mood and visual discomfort. After this, participants will have a short break during which they will be able to drink water and use the bathroom. The time spent providing the salivary samples, filling in the questionnaires, and having free time for the break will be 10 minutes, meaning that the exact break length will depend on how long participants take for the first two tasks. After the break, participants will remain in their private room in dim light, free to read, write, listen to music and podcasts via in-ear headphones for 20 minutes. From two hours prior to HBT to HBT, the same structure of this 30–minute experimental block will be kept, with the only difference being that after the break, participants will wear the VR headset to be exposed to bright light (n=4 sessions), or the dim light control (n=1 session), for 20 minutes (Figure 1, inset B1 and B2, respectively). A total of four 20–minute bright light exposures (n=4 sessions) or dim light exposures (n=1 session) will be achieved during these two hours. During this time, participants will be able to listen to music or podcasts using in-ear headphones. For the hour in dim light after HBT, the sampling schedule will be identical to that prior to the light exposure. Throughout the evening, participants will have no access to screens or light sources other than those included in the experimental protocol. They will be asked to keep their phones in their bags. Furthermore, based on the title, the researcher will confirm that the content of any podcast, music or books used by participants is not of a sexual, soporific, or horror nature at the beginning of each

experimental session. If participants need to leave their private room during the break to use the bathrooms and/or drink, they will wear red-tinted glasses to limit light exposure. The light in the hallways and bathroom will be kept at <10 lux melanopic EDI.

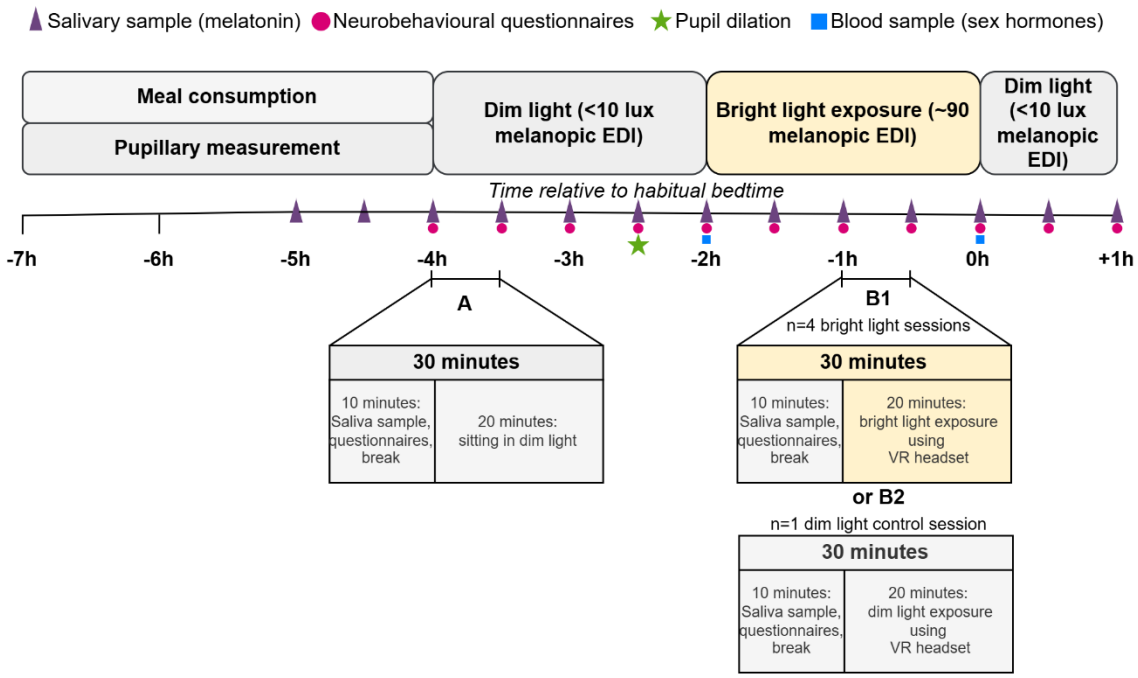

Figure 2. Experimental protocol for evening sessions. Participants will arrive in the laboratory 7 hours prior to HBT and end the experiment one hour after HBT. From 4 hours to 2 prior to HBT, the sampling schedule will be organised in 30–minute blocks, starting with the salivary samples, neurobehavioural questionnaires, a short break, followed by 20 minutes in dim light (inset A). The same sampling schedule will be kept during bright light exposure; however, here, the salivary samples, questionnaires, and breaks will be followed by a 20–minute delivery of the bright light stimulus via the VR headset (inset B1). Note that for the dim light control session (n=1), participants will wear the VR headset but receive a light stimulus of intensity <1 lux (inset B2).

Light exposure stimuli

Light exposure will be delivered at the vertical plane using a VR head-mounted display (HTC Vive Pro Eye, HTC, Taiwan). This device comprises a binocular display with an in-built eye tracking technology (Tobii eye tracker, Sweden) for real-time pupillometry and eye-movement measurements. The distance between the display and participants' eyes is variable as it depends on anatomical features. Furthermore, participants will need to adjust the height of the headset on the face and the interpupillary distance at the beginning of each experimental session as part of the in-built calibration procedure of the eye tracker. Light stimuli will be created using the licensed software Vizard 7 from WorldViz (Santa Barbara, CA, USA), which involves Python code for stimulus presentation. A detailed report on light stimuli delivery with this set-up and validation for melatonin suppression studies has been described by Fernandez-Alonso & Spitschan (2025). The light stimulus for the control condition (dim light) consists of a constant background, with intensity set to zero, yielding <0.01 lux melanopic EDI (mean±SD across five VR headsets: 0.0076±0.0046 lux melanopic EDI). The light

stimulus for the bright light experimental sessions consists of a constant full-field polychromatic white light, generated by setting the red, green and blue channels to equal values at 50% intensity, yielding ~90 lux melanopic EDI (mean±SD across five VR headsets: 90.94±0.28 lux melanopic EDI). Stimuli are presented dichoptically (to both eyes) using the VR headset, meaning that each VR lens will be set to approximately 90 lux melanopic EDI. Furthermore, the manufacturer reports a field of view of 110 degrees. To calibrate the VR headsets, light intensities expressed in melanopic EDI and photopic illuminance were measured for the left and right lenses of all five available VR head-mounted displays at gradually increasing intensity levels (from 0 to 1 at 0.1 steps), using a spectroradiometer (Jeti spectraval 1511-HiRes, Jena Technische Instrumente, Jena, Germany). Spectral power distributions for these measurements are available as CSV files in the GitHub repository associated to this manuscript (https://github.com/tscnlab/GuidolinEtAl_PCIRRStage1_2025/tree/main/vr_calibration). This calibration procedure showed that a 50% input intensity corresponded to approximately 90 lux melanopic EDI (see *Supplementary materials*, Table S3). We selected this light level for use in our experiment carefully, seeking an optimum point in the dose-response curve for melatonin suppression. We based our selection on the computational model developed by Giménez and colleagues (2022), which summarises empirical melatonin suppression data from 29 studies, and fits a simple heuristic prediction model for melatonin suppression as a function of melanopic EDI and pupil dilation. For our maximally achievable melanopic EDI of ~90 lux and a dilated pupil, the model predicts a melatonin suppression of 57.90%, which we consider to be a suitable response to avoid extrema of the dose response curves. A filled-in copy of the ENLIGHT Checklist (Spitschan et al., 2023) for reporting laboratory-based studies on the non-visual effects of light in humans is included in the *Supplemental materials*.

During the light exposure delivery, to ensure that participants keep their eyes open and were awake at all times, a visual stimulus (coloured dot, red, blue, green, and yellow, size set to 0.15 in arbitrary units) will be presented in the centre of the visual field, and participants will be prompted to press a button on a response pad every time it changed colour (intervals set to 30 seconds). While reaction times for this response will be collected, this test will only serve to ensure participants keep awake during the light exposure, and the data collected will not be used for the inferential analysis of this paper.

Pupillometry measurements

Pupillometry measurements to measure melanopsin-mediated pupil response (i.e. melanopsin sensitivity) will be performed with the retinaWISE binocular pupillometer (Gibertoni et al., 2024; Oculox Technologies, Muzzano, Switzerland). This device delivers light

stimuli using six calibrated light emitting diodes (LEDs) and can be used in combination with the retinaWISE software to target specific photoreceptors of interest, including melanopsin. Specific protocols delivering light stimuli of interest can be developed and delivered through this system, the measured outcome being the participant's pupil diameter throughout a given trial. Thus, the device enables delivery of silent substitution stimuli, which can change the activation of one or more photoreceptors while keeping other photoreceptors equally activated (Spitschan & Woelders, 2018). In the current study, we will create three protocols aimed at measuring 1) the post-illumination pupillary response (PIPR), and 2) LMS stimulation (silent substitution stimuli for joint modulation of L, M and S cones and constant activation of melanopsin), and 3) melanopsin stimulation (silent substitution stimuli for selective modulation of melanopsin and constant activation of L, M and S cones). The light stimuli will be delivered to the participants' right eye through a 10-30° field stop to achieve annular stimulation. For the PIPR protocol, six narrowband short-wavelength and six narrowband long-wavelength stimuli will be presented alternately. As part of the LMS-directed and melanopsin-directed sequences, the protocol will start with a 4-minute adaptation to the background spectrum to achieve light adaptation, followed by 16 pulses of LMS-directed and melanopsin-directed contrast, respectively. For all three protocols, the stimulus consists of a 3-second pulse of light windowed by a half cosine ramp (500ms on each side). Furthermore, the stimulus is accompanied by a 1-second pre-stimulus baseline and a 16-second post-stimulus measurement period.

The retinaWISE device will be placed in a black photography tent (Pop-Up Darkroom, ILFORD, Harman Technology Ltd., UK) to minimise external light exposure during the experiment. Before the start of pupillometry data collection, participants will be instructed on how to adjust the chair and chin rest to sit comfortably while stimuli will be delivered through the retinaWISE system. Then, participant alignment will follow using a test stimulus (green annular field, 523nm wavelength and 5% intensity, corresponding to 12.5 lux) to ensure that the participant can fully see the stimulus with their right eye. Once alignment is achieved (i.e. the device detects participants' right pupil, indicated by a red circle around the pupil's diameter), participants will sit back for a dark adaptation period of 10 minutes. Afterwards, they will place their chin on the chin rest, alignment will be checked again by the experimenter, and the three protocols will be run one after the other, in the following order: PIPR, LMS stimulation, and melanopsin stimulation. In between each stimulation, participants will be allowed to take a short break if needed, during which they can sit back but not leave the experimental tent. Throughout the experiment, participants will be asked to keep their gaze straight and limit head movements. To ensure this, an elastic headband attached to the device and strapped behind participants' heads will be used.

The retinaWISE device will be spectrally calibrated using a spectroradiometer (Jeti spectraval 1511-HiRes, Jena Technische Instrumente, Jena, Germany). The light from the retinaWISE device will be projected on a glass diffuser disc (diameter: 25.4 mm, DG10-220-A, Thorlabs, Ely, UK) positioned at the corneal plane.

Salivary samples

Salivary samples for melatonin analysis will be collected every 30 minutes, starting 5 hours prior to 1 hour after participants' HBT, using Salivettes (Sarstedt, Nümbrecht, Germany). Samples (≥1mL) will be centrifuged immediately after collection and stored at -20°C. The samples will be analysed for melatonin concentration using enzyme-linked immunoabsorbent assay (ELISA) by NovoLytiX GmbH, Switzerland, with typical specifications (limit of quantification 0.5–50 pg/mL, detection limit 0.5 pg/mL, mean intra-assay precision: 7.8%, mean inter-assay precision 10.0%).

Neurobehavioural tests

As part of each experimental block, participants will be filling in three questionnaires to measure their mood, visual comfort and alertness. These will be delivered using tablets available in each participant's experimental booth (i.e. one tablet for each participant). The data from these questionnaires will not be used for the confirmatory analyses presented here, but will be collected for exploratory analyses.

Mood

Participants will be asked to report their current mood using the International Positive and Negative Affect questionnaire Short Form (I-PANAS-SF; Thompson, 2007). In this questionnaire, participants answer the question "*To what extent do you currently feel:*" based on 10 mood components (*upset*, *hostile*, *alert*, *ashamed*, *inspired*, *nervous*, *determined*, *attentive*, *afraid*, *active*) using a five-point Likert-type item ("*Not at all"* – "*Extremely*").

Visual comfort

Participants will be asked to rate the current intensity of their visual (dis)comfort symptoms across various dimensions using a custom-made five-point discrete visual analogue scale ("*Please rate the intensity of these symptoms:*", ranging from "*None"* to "*Severe"*). The choice of visual discomfort dimensions included in this questionnaire is based on common visual discomfort symptoms based on the literature (*general discomfort*, *headache, blurry vision, double vision, eye strain, burning sensation in the eyes, watery eyes, increased sensitivity to light, nausea*).

Subjective sleepiness

Participants will rate their subjective sleepiness using the Karolinska Sleepiness Scale (KSS, Shahid et al., 2011) on a 10-point ordered category item "*Please rate your sleepiness in the last 5 minutes*", with response options: *"Extremely alert", "Very alert", "Alert", "Rather alert", "Neither alert nor sleepy",* "*Some signs of sleepiness", "Sleepy, but no effort to keep awake", "Sleepy, but some effort to keep awake", "Very sleepy, great effort to keep awake, fighting sleep",* "*Extremely sleepy, fighting sleep*".

Blood samples

Blood samples will be collected twice during each experimental night: once immediately before bright light exposure (2 hours prior to HBT) and once immediately before the end of bright light exposure (at HBT). The first blood sample will be used for analysis of sex steroid hormones. The second blood sample is collected for exploratory analyses looking at the effect of light exposure on metabolomics. During the dim-light control experimental session, blood samples will be collected at the same time points despite a lack of light exposure. At each time point, blood will be collected using two 2.7 ml ethylenediaminetetraacetic acid (EDTA) tubes. The samples will immediately be centrifuged at room temperature (10 minutes at 2750g) to separate plasma and serum. After centrifugation, the plasma from both EDTA tubes will be mixed and aliquoted into 200 µl Eppendorf tubes. These aliquots will then be snap-frozen in liquid nitrogen until the end of the experimental session and then stored at -80°C until analysis. Blood samples will be analysed for endogenous sex hormones (estradiol, progesterone and testosterone) using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) in all three groups. Furthermore, non-targeted LC-MS/MS will be used to detect synthetic hormones (ethinylestradiol and synthetic progestin) contained in the contraceptive pill. This analysis will be performed for all participants of the MCOC group, as well as for n=3 participants of the NC group and for n=3 participants of the HM group. As the concentration of exogenous hormones in the NC and HM groups is expected to be zero, analysis on a subset of participants (n=3/group as reported above) will serve as a control measure to ensure this. Steroids will be extracted from plasma using solid phase extraction and detected by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) using the UHPLC-system LC-40D (Shimadzu, Duisburg, Germany) and a QTrap6500+ (Sciex, Darmstadt, Germany).

Continuous measurements

This study involves daily or multi-day monitoring of a variety of behavioural and biological outcomes, including self-reported ones. For these, participants will use the MyCap app (Harris

et al., 2022) on their smartphones to fill in questionnaires, which integrates with REDCap (Harris et al., 2009, 2019) and will thus allow researchers to access the participants' responses. These continuous questionnaires consist of a morning sleep diary, an evening mood diary, an LH results log, a pill intake log, and a menses onset log.

Morning sleep diary

Every morning from the start of circadian stabilisation until the last experimental session, participants will be filling in a morning sleep diary upon wakening (Consensus Sleep Diary, core version; Carney et al., 2012). This questionnaire consists of nine items to assess their bed and sleep timing, sleep duration, number and duration of awakenings, sleep offset, time of getting out of bed, type of day (free/work day), and their subjective sleep quality (assessed using a five-point Likert-type item including "*Very poor*", "*Poor*", "*Fair*", "*Good",* and "*Very good*"). Responses to this questionnaire will be checked daily by the researcher to ensure compliance with the self-selected sleep/wake schedule. This data will not be used for the confirmatory analyses of this study, but might be used for exploratory analyses.

Evening mood diary

Every evening before going to sleep, participants will retrospectively report their mood using a modified version of the MoodZoom questionnaire (Tsanas et al., 2016). The original questionnaire is modified to replace the word "elated" with "happy", and to measure retrospective rather than current mood. This questionnaire asks participants to what extent the following words describe their mood for the past day: *anxious*, *happy*, *sad*, *angry*, *irritable*, and *energetic* using a seven-point ordered category scale ("*Not at all"*, "*Slightly*", "*Somewhat*", "*Moderately*", "*Quite a bit*", "*Very much so*", "*Extremely*"). This data will not be used for the confirmatory hypothesis here, but rather for exploratory analyses.

Luteinising hormone (LH) testing and logging

NC group participants will test for rising levels of urinary E2 and LH surge using at-home Clearblue Advanced Digital ovulation tests (Clearblue, Swiss Precision Diagnostics GmbH, Geneva, Switzerland). The Clearblue Advanced Digital Ovulation tests detect the rise in the estrone-3-D-glucuronide (E3G), a metabolite of estradiol, which precedes LH release from the pituitary gland. According to manufacturer data on n=87 individuals, 66% of users experienced positive E2 results for 0-4 days before obtaining a positive LH result, 25% for 5-9 days, 1% for more than 10 days, and 8% only got a positive E2 result, without getting a positive LH result (*Advanced Digital Ovulation Test*, 2016). When participants start testing, low levels of E3G and LH will return a negative result (black circle on the display monitor), and provide the device with a "baseline measurement". Once the test first detects a rise in E3G levels from baseline,

it returns a blinking smiley on the device's small display. Once the test detects LH above a threshold of 40 mIU/mL, the device's display returns a stable smiley which lasts 48 hours. During the in-person screening, participants will be trained in using these tests and provided with an infographic instruction sheet to refer to when using the tests at home. Participants will log their E3G and LH test results in a questionnaire on the MyCap app every day from the start of testing until a positive LH result is obtained, and add a picture or video (for the blinking smiley) of the result displayed on the monitor. They will perform this test in the morning after their longest sleep, as the first urine sample of the day and before to drinking any water or beverage.

Pill intake and menses onset logging

Participants of the MCOC group will be asked to take the contraceptive pill in the morning, directly after waking up, throughout the study. Every morning after intake, they will complete a short survey to log timing ("*Please indicate the time when you took the pill this morning"*) and type (active/inactive) of pill intake (*"Please indicate the type of pill you took today"*). In case participants have been taking the pill at another time of day until enrolment in the study, they will change this timing at least five days prior to the first experimental session (i.e. during circadian stabilisation). The NC group will be asked to log their menstrual cycle onset (for cycles 1, 2 and 3) as soon as menses start, using another short survey (adapted from Appendix 2, Schmalenberger et al., 2021). This survey will ask participants whether their menstrual bleeding has started and whether participants are experiencing dysmenorrhea.

Wrist actimetry

From the in-person screening, participants will wear a wrist actigraph (ActTrust2, Condor Instruments, São Paolo, Brazil) on their non-dominant wrist throughout the study. The sampling interval of the device will be set to 60 seconds. These data will be used to check compliance with the self-selected sleep/wake schedule throughout the experiment (see *Experimental procedure*). Participants will be instructed to remove the actigraph in case of contact sports and/or swimming. These data are also collected for future exploratory analyses.

Personal light logging

Starting from the in-person screening, participants will also wear a light logger at the chest level to measure their light exposure (ActLumus, Condor Instruments, São Paolo, Brazil). The sampling interval of the device will be set to 60 seconds. This device will be attached to a lanyard so that participants can wear it around their neck and measure light exposure at the chest level. This will serve to account for the effect of photic history on melatonin suppression, as it has been shown that the amount of light received during the day might alter the sensitivity

to the neuroendocrine effects of light (Smith et al., 2004). Photic history is not included as a predictor in our confirmatory hypotheses, but this data will be used for exploratory analyses. Participants will be instructed to remove the light logger in case of showering, contact sports and/or swimming, and to press the event marker on the device when taking it off or placing it back on. While button presses will not be used in our analysis, we will collect this data as it can be useful for future non-wear-related analyses.

Medicinal, dietary, and exercise restrictions during the study

Throughout the study period, participants should abstain from melatonin supplements (any form), sleep aids or supplements. Participants will also be asked to completely avoid recreational drugs and nicotine for the whole study duration. Furthermore, on the day of an experimental session, participants will be asked to avoid intake of antihistamines and non-steroidal anti-inflammatory drugs (NSAIDs). Note that, if needed, NC group participants will be allowed to use NSAIDs to alleviate dysmenorrhea on the day of the perimenstrual visit, but will log any intake. Alcohol consumption will not be allowed in the three days prior to an experimental session. Caffeinated drinks such as coffee, black tea, white tea, energy drinks, Yerba mate drinks, and cola drinks consumption will also be limited to a maximum of one cup to be consumed in the morning (cut-off time: noon). On the day of an experimental session, participants will be asked to avoid consumption of bananas, walnuts, sour cherries, citrus fruits, chocolate and any coloured candy or gum after 15:00, as these can interfere with the measurement of salivary melatonin. Furthermore, on days of scheduled experimental sessions, participants will be asked to avoid extreme physical activity, i.e. activity that is out of their routine, at least 4 hours before coming to the laboratory. Adherence to these instructions will be verbally checked on the day of the experimental session, prior to the start of the experiment.

Discharge questionnaire

Two days after their last experimental session, participants will be sent a link to a REDCap questionnaire. This questionnaire will first assess the burden participants experienced while taking part in the experiment (Perceived Research Burden Assessment, PeRBA, adapted for retrospective testing; Lingler et al., 2014). Then, participants will be asked open-ended questions about their experience in taking part in the study, with a focus on procedural aspects ("*How would you describe the overall effort required to participate in this study?*" and "*Were there any specific parts of the experiment that were particularly challenging to complete? This question refers to the time spent in the laboratory during the experimental sessions*"); convenience ("*How did the number of study visits work for you? Were they too many, too few, or about right?", "How did the scheduling of your visits affect your daily life or work/study*

routine?", "*How would you describe your communication with the research team (e.g., reminders, clarity, responsiveness?*", and "*Is there anything that could have made participation more convenient for you?*"); lifestyle impact ("*Were any of these study-related restrictions (e.g. regular sleep and wake times) more difficult or easier to manage because of your individual circumstances (e.g. hormonal status, lifestyle habits)?*", "*Do you feel that your personal background (e.g., being a naturally cycling individual, being on hormonal contraception, or being a male participant) influenced how burdensome the study was for you?"*, and "*Are there aspects of the study that you think would be harder or easier for people in a different group?*"); and perceived value and motivation ("*What motivated you to participate in this study?*", and *"Would you be willing to participate in a similar study in the future? Why or why not?).*

While we aim to capture the experience of the NC group in particular because of the flexible scheduling they will adopt during the study, we will collect responses from participants of all groups in order to have a comparison between groups. This data will be used for exploratory analyses.

Analysis plan

Data pre-processing and analysis will be performed in R Markdown notebooks in the R Statistical Software (R Core Team, latest version available at data analysis, reported at Stage 2).

Salivary melatonin pre-processing

Firstly, if present, mis-ordered melatonin samples will be chronologically re-ordered based on the timestamp of sample collection. Melatonin concentration in missing samples will be estimated using interpolation via the piecewise cubic hermite interpolating polynomial method, combined with forward and backward filling (function *pchip* from the package *pracma*, latest version available at data analysis). The area under the curve (AUC) for melatonin concentration during all experimental sessions will be calculated using Simpson's rule. Melatonin measurements where 40% of melatonin values are missing (i.e. two out five melatonin values during light exposure, or respective time in the dim light control) will be excluded, meaning that the AUC will not be calculated. Melatonin suppression for a given participant and for each experimental bright light session will be expressed as percentage suppression relative to the dim light experimental session:

$$\text{Melatonin suppression } (\%) = \left(1 - \frac{\text{AUC}_{\text{bright}}}{\text{AUC}_{\text{dim}}}\right) \cdot 100$$

Where AUC_{bright} and AUC_{dim} represent the area under the curve of the melatonin profile from two hours prior to HBT to HBT in the bright and dim light experimental sessions, respectively.

Sex steroid hormones analysis

As the bioassays for sex steroid hormones detection are still under development, key parameters such as the lower limit of quantification, upper limit of quantification, and limit of detection have not yet been fully established. These thresholds will be defined during the method validation phase, and will determine cut-off values for handling missing values and outliers. We will report these thresholds and data pre-processing details at Stage 2.

Pupillometry data pre-processing

For this analysis, we will calculate melanopsin-mediated pupil responses as the ratio of the melanopsin response amplitude to the LMS response amplitude. In other words, we will use pupil data from the melanopsin stimulation protocol and from the LMS stimulation protocol to identify changes in pupil size uniquely due to melanopsin and not to other photoreceptors. Data from the dim light control session will not be used for this analysis. The pre-processing pipeline will be the same for data obtained by both protocols, and include resampling of raw data to evenly spaced time points and filtering the raw data to remove physiologically impossible pupil values below 1 mm and above 10 mm, as well as blinks. Furthermore, missing values will be estimated using linear interpolation. For each trial, pupil size will be normalised to the baseline pupil size, i.e. to the mean pupil diameter one second before stimulus onset. Furthermore, the timing of each trial will be aligned to stimulus onset. Afterwards, for each trial, pupil size will be converted to a percentage change over time from baseline using the following formula:

$$\Delta\text{Pupil diameter(t) (\%)} = \frac{\text{Pupil diameter(t)}}{\text{Pupil diameter}_{\text{Baseline}}} \cdot 100$$

where the numerator *Pupil size* is the size of the measured pupil (i.e. right pupil, mm) in a given trial. Trials where >50% of baseline pupil size data is unavailable and trials where interpolated data >30% of total trial data will be excluded. Afterwards, all trials for a participant's given experimental session will be averaged. In order to quantify only melanopsin-related changes in pupil size, we will calculate melanopsin contribution to pupil responses as

$$\text{Melanopsin mediated pupil response} = \frac{\text{Melanopsin response amplitude}}{\text{LMS response amplitude}}$$

Where *Melanopsin response amplitude* is the area under the curve during stimulus presentation in the melanopsin stimulation protocol, and *LMS response amplitude* is the area under the curve during stimulus presentation for the LMS stimulation protocol.

NC group cycle phase: recategorisation

As described above, combined E3G and LH tests and menses onset will be used to ensure that experimental sessions are scheduled in the four menstrual phases of interest. To further validate that the experimental sessions of the NC group happened at the menstrual cycle phase of interest, two criteria will be applied in the following order (Criterion 1, counting method, and Criterion 2, relative hormone levels). For Criterion 1 (counting method), each phase will be characterised by a combination of backward and forward counting as follows (note that day 1 is menstrual onset of cycle 2, without day 0):

1. Peri-ovulatory phase: cycle day -20 to -10 before menstrual onset
2. Mid-luteal phase: cycle day -9 to -4 before menstrual onset
3. Peri-menstrual phase: cycle day -3 to +3 before menstrual onset
4. Mid-follicular phase: cycle day +4 to +8 from menstrual onset

These criteria are taken from Schmalenberger and colleagues (2021) and adapted for changes in menstrual phase definition used in the current study (outlined in *Participants scheduling – Naturally cycling individuals*). For Criterion 2 (relative hormone levels), relative E2 and P4 hormone levels will be used to re-categorise the peri-menstrual, peri-ovulatory and mid-luteal phases of the experimental sessions as follows (Schmalenberger et al., 2020):

- Peri-menstrual phase: peri-menstrual E2 < peri-ovulatory E2
- Mid-luteal phase: mid-luteal P4 > peri-ovulatory P4

If an experimental session meets the Criterion 1 (counting method), it will remain assigned to the original cycle phase and be included in the analyses. If an experimental session does not meet Criterion 1 for belonging to a specific phase, but it falls within another cycle phase investigated in the study based on the counting method proposed above, then it will be re-assigned to this phase. If this is not the case, compliance with Criterion 2 (relative hormone levels) will be checked. If Criterion 2 supports the original phase, this visit will remain assigned to the original menstrual phase. On the other hand, if Criterion 2 supports that the experimental visit was performed at another phase, then data from this visit will be reassigned to this phase. Experimental sessions that do not pass Criterion 1 and/or cannot be reassigned based on Criterion 1 or Criterion 2 will be excluded from the analyses.

Manipulation checks

In order to check that participant scheduling was successful in capturing different hormonal profiles, we will perform manipulation checks as a first step of data analysis. First, model diagnostics will be performed to ensure linearity of the predictors (residual plots), normality of residuals and random effects (using quantile-quantile plots), and homoscedasticity of residuals (residual plots). The manipulation checks will be performed as LMMs using the *lmer* function from the *lmerTest* package (latest version available at data analysis, Kuznetsova et al., 2017).

Manipulation check	**Type**	**Statistical model**
Verify different endogenous sex hormones levels in different menstrual phases (i.e. respective scheduled sessions): - Peri-ovulatory: high E2 and low P4, high T, low EE and SP - Mid-luteal: high E2 and P4, low T, low EE and SP - Peri-menstrual: falling E2 and P4, low T, low EE and SP - Mid-follicular: low E2 and P4, low T, low EE and SP	Within group (NC)	Five linear mixed models, with dependent variable being each of the hormones of interest (E2, P4, T, EE, P). Structure of linear mixed model: menstrual phase as categorical predictor (main effect) and participant ID as random factor.
Verify different exogenous hormones levels between active pill intake phases (n=3) and inactive pill intake phase (n=1) - Active phase: high EE and SP, low E2 ad P4 - Inactive phase: lower EE and SP and higher E2 and P4	Within group (MCOC)	Five linear mixed models, with dependent variable being each of hormones of interest (E2, P4, T, EE, P). Structure of linear mixed model: pill cycle phase as categorical predictor (main effect) and participant ID as random factor.

compared to active phase - Low T levels across the pill cycle		
Verify different endogenous and exogenous hormonal profiles between NC group, MCOC group, and HM group: - NC groups: high E2 and P4 levels, low T, no EE and SP - MCOC group: high EE and SP levels, low T, E2 and P4 - HM group: high T, low E2 and P4, no EE and SP	Between-groups	Five linear mixed models with dependent variable being each of the hormones of interest (E2, P4, T, EE, P). Structure of the linear mixed models: group as categorical predictor (main effect) and participant ID as random factor.

Table 5: Overview of manipulation checks.

Analytic plan for confirmatory analyses

Each confirmatory analysis will be analysed using an LMM. The outcome of each analysis will be assessed based on the resulting BF (see Table 7 and *Sample size* for interpretation of the evidence strength of each hypothesis based on BF values). First, model diagnostics will be performed to ensure linearity of the predictors (residual plots), normality of residuals and random effects (using quantile-quantile plots), and homoscedasticity of residuals (residual plots). We will use the following criteria to determine evidence for the full model CH_{Full} (Jeffreys, 1998): 1/3<BF<3 as anecdotal evidence for CH_{Full}, 3<BF<10 as substantial evidence for CH_{Full}, 10<BF<30 as strong evidence for CH_{Full}, and BF>100 as very strong evidence for CH_{Full}. Here, we report the full and null models for each confirmatory hypothesis according to the Wilkinson-Rogers notation (Wilkinson & Rogers, 1973).

Confirmatory Hypothesis 1. Based on previous literature, we hypothesise an effect of sex on melatonin suppression, with female participants exhibiting greater melatonin suppression by light compared to male participants (**CH1a**).

CH1a$_{Full}$: Melatonin suppression

= Sex + (1 | Participant ID)

CH1a$_{Null}$: Melatonin suppression

= (1 | Participant ID)

Furthermore, since activation of melanopsin-containing ipRGCs causes melatonin suppression, we hypothesise an effect of sex on melanopsin sensitivity, whereby female participants exhibit larger melanopsin-mediated pupil responses than male participants (**CH1b**).

CH1b$_{Full}$: Melanopsin-mediated pupil responses

= Sex + (1 | Participant ID)

CH1b$_{Null}$: Melanopsin-mediated pupil responses

= (1 | Participant ID)

Confirmatory Hypothesis 2. We hypothesise a group effect on light-induced melatonin suppression, whereby melatonin suppression differs between individuals in the NC group, the MCOC group, and the HM group. (**CH2a**).

CH2a$_{Full}$: Melatonin suppression

= Group + (1 | Participant ID)

CH2a$_{Null}$: Melatonin suppression

= (1 | Participant ID)

Moreover, we hypothesise an effect of group (NC/MCOC/HM) on melanopsin-mediated pupil responses (**CH2b**).

CH2b$_{Full}$: Melanopsin-mediated pupil responses

= Group + (1 | Participant ID)

CH2b$_{Null}$: Melanopsin-mediated pupil responses

= (1 | Participant ID)

Confirmatory Hypothesis 3. In naturally cycling individuals (NC group), we hypothesise an effect of the endogenous sex steroid hormones estradiol (E2) and progesterone (P4) on light-induced melatonin suppression (**CH3a**).

CH3a$_{Full}$: Melatonin suppression

= E2 levels + P4 levels + (E2 levels + P4 levels | Participant ID)

CH3a$_{Null}$: Melatonin suppression

= (1 | Participant ID)

We also hypothesise an effect of E2 and P4 on melanopsin-mediated pupil responses (**CH3b**).

CH3b$_{Full}$: Melanopsin-mediated pupil responses

= E2 levels + P4 levels + (E2 levels + P4 levels | Participant ID)

CH3b$_{Null}$: Melanopsin-mediated pupil responses

= (1 | Participant ID)

Confirmation Hypothesis 4. In the MCOC group, we hypothesise an effect of the exogenous sex hormones ethinylestradiol (EE) and synthetic progestin (SP) on light-induced melatonin suppression (**CH4a**).

CH4a$_{Full}$: Melatonin suppression

= EE levels + SP levels + (EE levels + SP levels | Participant ID)

CH4a$_{Null}$: Melatonin suppression

= (1 | Participant ID)

We also hypothesise an effect of EE and SP on melanopsin-mediated pupil responses (**CH4b**).

CH4b$_{Full}$: Melanopsin-mediated pupil responses

= EE levels + SP levels + (EE levels + SP levels | Participant ID)

CH4b$_{Null}$: Melanopsin-mediated pupil responses

= (1 | Participant ID)

Question	Hypothesis	Sampling plan	Analysis plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Confirmatory Hypothesis 1 (between-groups)						
Do female and male participants differ in the amount of melatonin suppression by evening light?	**CH1a**: We hypothesise an effect of sex on melatonin suppression outcomes, with female participants exhibiting greater melatonin suppression by bright light compared to male participants. Effect of interest: sex	Based on BFDA results of CH3a (N_{max} for each group = 12)	Linear mixed effects model for melatonin suppression Fixed effects: sex (categorical with two levels: female and male) Random effects: random intercept for participant ID (categorical with levels corresponding to participant number)	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH1a_{Full}$ 1<BF<10 = moderate evidence for $CH1a_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH1a_{Full}$ BF>30 = very strong evidence for $CH1a_{Full}$	Conclusive evidence for this hypothesis will identify that female participants do exhibit greater melatonin suppression by evening light compared to male participants.	Among the few studies that have looked at sex differences in melatonin suppression, the two most recent and most powered studies suggest that female participants had higher melatonin suppression by light at night compared to male participants. The result of our study will contribute to this growing evidence on potential sex differences and their directionality.
Is there a sex difference in light sensitivity earlier on along the non-visual pathway, at	**CH1b**: We hypothesise an effect of sex on melanopsin-mediated pupil	Based on BFDA results of CH3a (N_{max} for each group = 12)	Linear mixed effects model for melanopsin-mediated pupil responses	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH1b_{Full}$	Conclusive evidence for this hypothesis will identify that female participants have	The outcome of this analysis will inform whether sex differences in light sensitivity are

the photoreceptor level, measured by melanopsin-mediated pupil size?	responses, whereby female participants exhibit higher melanopsin activation than male participants. Effect of interest: sex		Fixed effects: sex (categorical with two levels: female and male) Random effects: participant ID (categorical with levels corresponding to participant number)	1<BF<10 = moderate evidence for $CH1b_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH1b_{Full}$ BF>30 = very strong evidence for $CH1b_{Full}$	higher melanopsin sensitivity compared to male participants.	present at the retinal level (quantified by melanopsin-mediated pupil responses). This analysis builds on the one performed in CH1a, to assess where, along the non-visual pathway, sex differences are to be observed. Refer to Table 7 for detailed interpretation.
Confirmatory Hypothesis 2 (between-groups)						
Do different hormonal profiles (NC group, MCOC group, HM group) influence the amount of melatonin suppression by evening light?	**CH2a**: We hypothesise a group effect on light-induced melatonin suppression, whereby melatonin suppression differs between naturally cycling individuals (NC group), individuals taking oral contraceptives (MCOC group), and healthy male participants (HM group). Effect of interest: group	Based on BFDA results of CH3a (N_{max} for each group = 12)	Linear mixed effects model for melatonin suppression Fixed effects: group (categorical with three levels: NC, MCOC and HM) Random effects: participant ID (categorical with levels corresponding to participant number)	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH2a_{Full}$ 1<BF<10 = moderate evidence for $CH2a_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH2a_{Full}$ BF>30 = very strong evidence for $CH2a_{Full}$	Conclusive evidence for this hypothesis will demonstrate that different hormonal profiles influence melatonin suppression.	To the best of our knowledge, no study has investigated melatonin suppression differences in the three participant groups included in this study. The outcomes of this analysis will provide the first account of the role of three different hormonal profiles in melatonin suppression.

Do different hormonal profiles (NC group, MCOC group, HM group) influence melanopsin-mediated pupil responses?	**CH2b**: We expect that a group effect (NC/MCOC/HM) on melanopsin-mediated pupil responses. Effect of interest: group	Based on BFDA results of CH3a (N_{max} for each group = 12)	Linear mixed effects model for melatonin suppression Fixed effects: group (categorical with three levels: NC, MCOC and HM) Random effects: participant ID (categorical with levels corresponding to participant number)	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH2b_{Full}$ 1<BF<10 = moderate evidence for $CH2b_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH2b_{Full}$ BF>30 = very strong evidence for $CH2b_{Full}$	Conclusive evidence for this hypothesis will demonstrate that different hormonal profiles influence melanopsin sensitivity.	The outcome of this analysis will inform whether group differences in light sensitivity are present at the retinal level (quantified by melanopsin-mediated pupil responses). This analysis builds on the one performed in CH2a, to assess where, along the non-visual pathway, group differences are to be observed. Refer to Table 7 for detailed interpretation.
Confirmatory Hypothesis 3 (within-group, NC group)						
Do endogenous sex steroid hormones influence melatonin suppression across the menstrual cycle?	**CH3a**: In naturally cycling individuals (NC group), we hypothesise an effect of the endogenous sex steroid hormones estradiol (E2) and progesterone (P4) on light-induced melatonin suppression.	Fixed-n Bayes Factor Design Analysis (N_{max} = 12)	Linear mixed effects model for melatonin suppression Fixed effects: E2 levels (continuous) and P4 level (continuous) Random effects: intercept and slopes for E2 and P4 levels by participant ID	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH3a_{Full}$ 1<BF<10 = moderate evidence for $CH3a_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH3a_{Full}$ BF>30 = very strong evidence for $CH3a_{Full}$	Conclusive evidence for this hypothesis will demonstrate that absolute E2 and P4 levels influence melatonin suppression in naturally cycling individuals.	To our knowledge, only one between-subject study indicated that salivary E2 and P4 did not have an effect on melatonin suppression in female participants. Thus, the outcome of this within-subject analysis will inform whether levels of endogenous sex

	Effect of interest: E2 levels and P4 levels		(categorical with levels corresponding to participant number)			steroid hormones across the menstrual cycle, reliably measured in plasma rather than saliva, influence melatonin suppression.
Do endogenous sex steroid hormones across the menstrual cycle influence light sensitivity at the retinal level, as quantified by melanopsin-mediated pupil responses?	**CH3b:** In naturally cycling individuals, we hypothesise an effect of endogenous sex steroid hormones E2 and P4 on melanopsin-mediated pupil responses. Effect of interest: E2 levels, and P4 levels	Based on BFDA results of CH3a (N_{max} = 12)	Linear mixed effects model for melanopsin activation Fixed effects: E2 levels (continuous) and P4 level (continuous) Random effects: intercept and slopes for E2 and P4 levels by participant ID (categorical with levels corresponding to participant number)	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH3b_{Full}$ 1<BF<10 = moderate evidence for $CH3b_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH3b_{Full}$ BF>30 = very strong evidence for $CH3b_{Full}$	Conclusive evidence for this hypothesis will demonstrate that absolute E2 and P4 levels influence melanopsin-mediated pupil responses in naturally cycling individuals.	The outcome of this analysis will inform whether endogenous sex hormones across the menstrual cycle influence light sensitivity at the retinal level (quantified by melanopsin-mediated pupil responses). This analysis builds on the one performed in CH3a to assess where, along the non-visual pathway, endogenous hormones influence light sensitivity. Refer to Table 7 for detailed interpretation.
Confirmatory Hypothesis 4 (within-group, MCOC group)						

Do exogenous sex steroid hormones influence melatonin suppression across the pill cycle?	**CH4a:** In individuals taking oral contraceptives (MCOC group), we hypothesise an effect of the exogenous sex hormones ethinylestradiol (EE) and synthetic progestin (SP) on light-induced melatonin suppression Effect of interest: EE levels and SP levels	Based on BFDA results of CH3a (N_{max} = 12)	Linear mixed effects model for melatonin suppression Fixed effects: EE levels (continuous) and SP levels (continuous) Random effects: intercept and slopes for EE and SP levels by participant ID (categorical with levels corresponding to participant number)	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH4a_{Full}$ 1<BF<10 = moderate evidence for $CH4a_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH4a_{Full}$ BF>30 = very strong evidence for $CH4a_{Full}$	Conclusive evidence for this hypothesis will demonstrate that absolute exogenous hormone levels (EE and SP) influence melatonin suppression in individuals taking the monophasic combined oral contraceptive pill.	To our knowledge, no study has investigated the effects of exogenous sex hormones in the oral contraceptive pill on melatonin suppression. Thus, the outcome of this analysis will provide a first account of whether levels of exogenous sex steroid hormones across the pill cycle influence melatonin suppression.
Do exogenous sex steroid hormones influence light sensitivity at the retinal level (quantified by melanopsin-mediated pupil response) across the pill cycle?	**CH4b**: In individuals taking oral contraceptives, we expect an effect of EE and SP on melanopsin-mediated pupil responses. Effect of interest: EE levels and SP levels	Based on BFDA results of CH3a (N_{max} = 12)	Linear mixed effects model for melanopsin-mediated pupil responses Fixed effects: EE levels (continuous) and SP levels (continuous) Random effects: intercept and slopes for EE and SP levels by participant ID (categorical with levels corresponding to	<u>Inconclusive evidence</u> 1<BF<3 = anecdotal evidence for $CH4b_{Full}$ 1<BF<10 = moderate evidence for $CH4b_{Full}$ <u>Conclusive evidence</u> 10<BF<30 = strong evidence for $CH4b_{Full}$ BF>30 = very strong evidence for $CH4b_{Full}$	Conclusive evidence for this hypothesis will demonstrate that absolute exogenous hormone levels (EE and SP) influence melanopsin-mediated pupil responses in individuals taking the monophasic combined oral contraceptive pill.	Adding to the CH4a analysis, this analysis will inform on the site of action of exogenous hormones on light sensitivity along the non-visual pathway. This analysis builds on the one performed in CH4a to assess where, along the non-visual pathway, exogenous hormones influence light

			participant number)			sensitivity. Refer to Table 7 for detailed interpretation.

Table 6. Design table.

Outcome interpretation matrix for hormonal effects of neuroendocrine (melatonin suppression) and pupillary (melanopsin sensitivity) outcomes

For each Confirmatory hypothesis of this study, we analyse the influence of a given parameter of interest (e.g. sex, group, hormone levels) on melatonin suppression as well as melanopsin activation. The rationale for performing these two analyses in parallel is that looking at melanopsin function could provide mechanistic insights into whether any sex, group, or hormone effect on melatonin suppression already exists at the retinal level. The table below adds details on how we will interpret and compare results from these models looking at light sensitivity at different sites along the non-visual pathway (neuroendocrine: melatonin suppression, pupillary: melanopsin sensitivity). Here, we present an example for interpreting results from CH3a and CH3b.

NC group		**Effect of E2 and P4 levels on melatonin suppression**	
		Conclusive evidence	Lack of evidence
Effect of E2 and P4 levels on melanopsin sensitivity	Conclusive evidence	E2 and P4 levels influence melatonin suppression and this effect is already observable at the level of the retina.	E2 and P4 levels influence melanopsin function, but this is not mirrored in an effect on melatonin suppression. The differences in melanopsin function might be compensated somewhere in the non-visual pathway between melanopsin and the pineal gland.
	Lack of evidence	E2 and P4 levels influence melatonin suppression, but this effect cannot be seen at the retinal level. E2 and P4 might directly or indirectly act at the post-retinal level to influence melatonin suppression.	There is no influence of E2 and P4 on melatonin suppression and melanopsin function. An influence might be observed at other locations of the non-visual pathway.

Table 7. Details on interpretations and comparisons of models where melatonin suppression and melanopsin activation are the dependent variables.

References

Advanced Digital Ovulation Test: Typically identifies 4 or more Fertile Days. (2016, April 20). https://www.clearblue.com/ovulation-tests/advanced-digital

Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *The International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Arslan, R. C., Blake, K., Botzet, L. J., Bürkner, P.-C., DeBruine, L., Fiers, T., Grebe, N., Hahn, A., Jones, B. C., Marcinkowska, U. M., Mumford, S. L., Penke, L., Roney, J. R., Schisterman, E. F., & Stern, J. (2023). Not within spitting distance: Salivary immunoassays of estradiol have subpar validity for predicting cycle phase. *Psychoneuroendocrinology*, *149*, 105994. https://doi.org/10.1016/j.psyneuen.2022.105994

Barth, C., Crestol, A., Lange, A.-M. G. de, & Galea, L. A. M. (2023). Sex steroids and the female brain across the lifespan: Insights into risk of depression and Alzheimer's disease. *The Lancet Diabetes & Endocrinology*, *11*(12), 926–941. https://doi.org/10.1016/S2213-8587(23)00224-3

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

Bland, M. J., & Altman, D. G. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Blume, C., Garbazza, C., & Spitschan, M. (2019). Effects of light on human circadian rhythms, sleep and mood. *Somnologie*, *23*(3), 147–156. https://doi.org/10.1007/s11818-019-00215-x

Brainard, G. C., Hanifin, J. P., Greeson, J. M., Byrne, B., Glickman, G., Gerner, E., & Rollag, M. D. (2001). Action Spectrum for Melatonin Regulation in Humans: Evidence for a Novel Circadian Photoreceptor. *The Journal of Neuroscience*, *21*(16), 6405–6412. https://doi.org/10.1523/JNEUROSCI.21-16-06405.2001

Brown, T. M., Brainard, G. C., Cajochen, C., Czeisler, C. A., Hanifin, J. P., Lockley, S. W., Lucas, R. J., Münch, M., O'Hagan, J. B., Peirson, S. N., Price, L. L. A., Roenneberg, T., Schlangen, L. J. M., Skene, D. J., Spitschan, M., Vetter, C., Zee, P. C., & Jr, K. P. W. (2022). Recommendations for daytime, evening, and nighttime indoor light exposure to best support physiology, sleep, and wakefulness in healthy adults. *PLOS Biology*, *20*(3), e3001571. https://doi.org/10.1371/journal.pbio.3001571

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The Consensus Sleep Diary: Standardizing Prospective Sleep Self-Monitoring. *Sleep*, *35*(2), 287–302. https://doi.org/10.5665/sleep.1642

Chellappa, S. L. (2021). Individual differences in light sensitivity affect sleep and circadian rhythms. *Sleep*, *44*(2), zsaa214. https://doi.org/10.1093/sleep/zsaa214

Fazlali, F., Lazar, R. R., Yahya, F., Stefani, O., Spitschan, M., & Cajochen, C. (2024). *Sex and seasonal variations in melatonin suppression, and alerting response to light* (p. 2024.10.18.619012). bioRxiv. https://doi.org/10.1101/2024.10.18.619012

Fernandez-Alonso, M., & Spitschan, M. (2025). Standardized and Calibrated Light Stimuli via Head-Mounted Displays for Investigating the Nonvisual Effects of Light. *Journal of Pineal Research*, *77*(4), e70051. https://doi.org/10.1111/jpi.70051

Gibertoni, G., Hromov, A., Piffaretti, F., & Geiser, M. H. (2024). Development of an Innovative Pupillometer Able to Selectively Stimulate the Eye's Fundus Photoreceptor Cells. *Diagnostics*, *14*(17), Article 17. https://doi.org/10.3390/diagnostics14171940

Giménez, M. C., Stefani, O., Cajochen, C., Lang, D., Deuring, G., & Schlangen, L. J. M. (2022). Predicting melatonin suppression by light in humans: Unifying photoreceptor-based equivalent daylight illuminances, spectral composition, timing and duration of light exposure. *Journal of Pineal Research*, *72*(2), e12786. https://doi.org/10.1111/jpi.12786

Guidolin, C., & Spitschan, M. (2024). *Sex-related and hormone-related differences in non-visual responses to light exposure—Comparison—ORKG*. https://orkg.org/comparison/R733094

Guidolin, C., Zauner, J., Hartmeyer, S. L., & Spitschan, M. (2025). *Collecting, detecting and handling non-wear intervals in longitudinal light exposure data* (p. 2024.12.23.627604). bioRxiv. https://doi.org/10.1101/2024.12.23.627604

Hambridge, H. L., Mumford, S. L., Mattison, D. R., Ye, A., Pollack, A. Z., Bloom, M. S., Mendola, P., Lynch, K. L., Wactawski-Wende, J., & Schisterman, E. F. (2013). The influence of sporadic anovulation on hormone levels in ovulatory cycles. *Human Reproduction (Oxford, England)*, *28*(6), 1687–1694. https://doi.org/10.1093/humrep/det090

Hampson, E. (2020). A brief guide to the menstrual cycle and oral contraceptive use for researchers in behavioral endocrinology. *Hormones and Behavior*, *119*, 104655. https://doi.org/10.1016/j.yhbeh.2019.104655

Hampson, E. (2023). Oral contraceptives in the central nervous system: Basic pharmacology, methodological considerations, and current state of the field. *Frontiers in Neuroendocrinology*, *68*, 101040. https://doi.org/10.1016/j.yfrne.2022.101040

Harris, P. A., Swafford, J., Serdoz, E. S., Eidenmuller, J., Delacqua, G., Jagtap, V., Taylor, R. J., Gelbard, A., Cheng, A. C., & Duda, S. N. (2022). MyCap: A flexible and configurable platform for mobilizing the participant voice. *JAMIA Open*, *5*(2), ooac047. https://doi.org/10.1093/jamiaopen/ooac047

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Heidari, S., Babor, T. F., De Castro, P., Tort, S., & Curno, M. (2016). Sex and Gender Equity in Research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*, *1*(1), 2. https://doi.org/10.1186/s41073-016-0007-6

J. Nathan, P., R. Norman, T., & D. Burrows, G. (1999). Effect of the menstrual cycle stage on the melatonin suppression by dim white light. *Psychoneuroendocrinology*, *24*(2), 193–200. https://doi.org/10.1016/S0306-4530(98)00075-4

Jeffreys, H. (1998). *The Theory of Probability*. OUP Oxford.

Johns, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, *14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

Joye, D. A. M., & Evans, J. A. (2022). Sex differences in daily timekeeping and circadian clock circuits. *Seminars in Cell & Developmental Biology*, *126*, 45–55. https://doi.org/10.1016/j.semcdb.2021.04.026

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lingler, J. H., Schmidt, K., Gentry, A., Hu, L., & Terhorst, L. (2014). Perceived Research Burden Assessment (PeRBA): Instrument Development and Psychometric Evaluation. *Journal of Empirical Research on Human Research Ethics : JERHRE*, *9*(4), 46–49. https://doi.org/10.1177/1556264614545037

Mong, J. A., Baker, F. C., Mahoney, M. M., Paul, K. N., Schwartz, M. D., Semba, K., & Silver, R. (2011). Sleep, Rhythms, and the Endocrine Brain: Influence of Sex and Gonadal

Hormones. *Journal of Neuroscience*, *31*(45), 16107–16116. https://doi.org/10.1523/JNEUROSCI.4175-11.2011

Morey, R. (2023). *BayesFactor: Computation of Bayes Factor for Common Designs* (Version 0.9.12-4.7) [Computer software]. https://cran.r-project.org/web/packages/BayesFactor/index.html

Phillips, A. J. K., Vidafar, P., Burns, A. C., McGlashan, E. M., Anderson, C., Rajaratnam, S. M. W., Lockley, S. W., & Cain, S. W. (2019). High sensitivity and interindividual variability in the response of the human circadian system to evening light. *Proceedings of the National Academy of Sciences*, *116*(24), 12019–12024. https://doi.org/10.1073/pnas.1901824116

R Core Team. (2025). *R Core Team (2025). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria.* [Computer software]. <https://www.R-project.org/>.

Roenneberg, T., Wirz-Justice, A., & Merrow, M. (2003). Life between clocks: Daily temporal patterns of human chronotypes. *Journal of Biological Rhythms*, *18*(1), 80–90. https://doi.org/10.1177/0748730402239679

Saunders, J. B., Aasland, O. G., Babor, T. F., De La Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction*, *88*(6), 791–804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x

Schmalenberger, K. M., Eisenlohr-Moul, T. A., Jarczok, M. N., Eckstein, M., Schneider, E., Brenner, I. G., Duffy, K., Schweizer, S., Kiesner, J., Thayer, J. F., & Ditzen, B. (2020). Menstrual Cycle Changes in Vagally-Mediated Heart Rate Variability Are Associated with Progesterone: Evidence from Two Within-Person Studies. *Journal of Clinical Medicine*, *9*(3), Article 3. https://doi.org/10.3390/jcm9030617

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

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. M. (2011). Karolinska Sleepiness Scale (KSS). In A. Shahid, K. Wilkinson, S. Marcu, & C. M. Shapiro (Eds.), *STOP, THAT and*

One Hundred Other Sleep Scales (pp. 209–210). Springer New York. https://doi.org/10.1007/978-1-4419-9893-4_47

Smith, K. A., Schoen, M. W., & Czeisler, C. A. (2004). Adaptation of human pineal melatonin suppression by recent photic history. *Journal of Clinical Endocrinology and Metabolism*, *89*(7), 3610–3614. Scopus. https://doi.org/10.1210/jc.2003-032100

Spitschan, M. (2019). Melanopsin contributions to non-visual and visual function. *Current Opinion in Behavioral Sciences*, *30*, 67–72. https://doi.org/10.1016/j.cobeha.2019.06.004

Spitschan, M., Kervezee, L., Lok, R., McGlashan, E., Najjar, R. P., Allen, A. E., Andersen, M., Bará, S., Blattner, P., Blume, C., Boivin, D. B., Bonmatí-Carrión, M.-Á., Broszio, K., Brown, T. M., Chellappa, S. L., Duffy, J. F., Eto, T., Flynn-Evans, E., Fotios, S., … Zeitzer, J. (2023). ENLIGHT: A consensus checklist for reporting laboratory-based studies on the non-visual effects of light in humans. *eBioMedicine*, *98*. https://doi.org/10.1016/j.ebiom.2023.104889

Spitschan, M., & Santhi, N. (2022). Individual differences and diversity in human physiological responses to light. *eBioMedicine*, *75*. https://doi.org/10.1016/j.ebiom.2021.103640

Spitschan, M., Santhi, N., Ahluwalia, A., Fischer, D., Hunt, L., Karp, N. A., Lévi, F., Pineda-Torra, I., Vidafar, P., & White, R. (2022). Sex differences and sex bias in human circadian and sleep physiology research. *eLife*, *11*, e65419. https://doi.org/10.7554/eLife.65419

Spitschan, M., & Woelders, T. (2018). The Method of Silent Substitution for Examining Melanopsin Contributions to Pupil Control. *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.00941

Steiner, M., Macdougall, M., & Brown, E. (2003). The premenstrual symptoms screening tool (PSST) for clinicians. *Archives of Women's Mental Health*, *6*(3), 203–209. https://doi.org/10.1007/s00737-003-0018-4

Tannenbaum, C., Ellis, R. P., Eyssel, F., Zou, J., & Schiebinger, L. (2019). Sex and gender analysis improves science and engineering. *Nature*, *575*(7781), 137–146. https://doi.org/10.1038/s41586-019-1657-6

Thompson, E. R. (2007). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. https://doi.org/10.1177/0022022106297301

Torgrimson, B. N., & Minson, C. T. (2005). Sex and gender: What is the difference? *Journal of Applied Physiology*, *99*(3), 785–787. https://doi.org/10.1152/japplphysiol.00376.2005

Tsanas, A., Saunders, K. E. A., Bilderbeck, A. C., Palmius, N., Osipov, M., Clifford, G. D., Goodwin, G. M., & De Vos, M. (2016). Daily longitudinal self-monitoring of mood

variability in bipolar disorder and borderline personality disorder. *Journal of Affective Disorders*, *205*, 225–233. https://doi.org/10.1016/j.jad.2016.06.065

Vidafar, P., McGlashan, E. M., Burns, A. C., Anderson, C., Shechter, A., Lockley, S. W., Phillips, A. J. K., & Cain, S. W. (2024). Greater sensitivity of the circadian system of women to bright light, but not dim-to-moderate light. *Journal of Pineal Research*, *76*(2), e12936. https://doi.org/10.1111/jpi.12936

Wilkinson, G. N., & Rogers, C. E. (1973). Symbolic Description of Factorial Models for Analysis of Variance. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *22*(3), 392–399. https://doi.org/10.2307/2346786

Yousefzadehfard, Y., Wechsler, B., & DeLorenzo, C. (2022). Human circadian rhythm studies: Practical guidelines for inclusion/exclusion criteria and protocol. *Neurobiology of Sleep and Circadian Rhythms*, *13*, 100080. https://doi.org/10.1016/j.nbscr.2022.100080