ORCA – Online Research @ Cardiff

Journal Pre-proof

Ovulatory shift, hormonal changes, and no effects on incentivized decision-making

Miloš Fišar, Lubomír Cingl, Tommaso Reggiani,
Eva Kundtová-Klocová, Radek Kundt, Jan Krátký,
Katarína Kostolanská, Petra Bencúrová, Marie Kudličková Pešková,
Klára Marečková
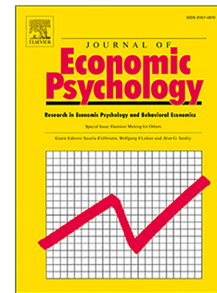
Please cite this article as: M. Fišar, L. Cingl, T. Reggiani et al., Ovulatory shift, hormonal changes, and no effects on incentivized decision-making. *Journal of Economic Psychology* (2023), doi: https://doi.org/10.1016/j.joep.2023.102656.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Ovulatory Shift, Hormonal Changes, and No Effects on Incentivized Decision-Making[*]

Miloš Fišar [†,1,2], Lubomír Cingl[3], Tommaso Reggiani[4,1,5], Eva Kundtová-Klocová[6], Radek Kundt[6], Jan Krátký[6], Katarína Kostolanská[7], Petra Bencúrová[7,8], Marie Kudličková Pešková[8], and Klára Marečková[8]

[1] *MUEEL Masaryk University Experimental Economics Laboratory, Masaryk University, Lipová 41a, 60200 Brno, Czech Republic.*
[2] *Competence Center for Experimental Research, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria.*
[3] *Faculty of Business Administration, Prague University of Economics and Business, Nám. W. Churchilla 4, 130 67 Prague, Czech Republic.*
[4] *Cardiff Business School, Cardiff University, Aberconway Building, Colum Dr, Cardiff CF10 3EU, United Kingdom.*
[5] *IZA, Schaumburg-Lippe-Straße 5–9, 53113 Bonn, Germany.*
[6] *LEVYNA Laboratory for the Experimental Research of Religion, Masaryk University, Arna Nováka 1, 602 00 Brno, Czech Republic.*
[7] *Department of Biochemistry, Faculty of Medicine, Masaryk University, Brno, Czech Republic*
[8] *Brain and Mind Research, Central European Institute of Technology (CEITEC), Masaryk University, Kamenice 753/5, 625 00 Brno, Czech Republic.*

July 29, 2023

Abstract

Employing an incentivized controlled lab experiment, we investigate the effects of ovulatory shift on salient behavioral outcomes related to i) risk preferences, ii) rule violation, and iii) exploratory attitude. As evolutionary psychology suggests, these outcomes may play an important role in economic decision-making and represent behavioral aspects that may systematically vary over the menstrual cycle to increase the reproductive success. Exploiting a within-subjects design, 124 naturally cycling females participated in experimental sessions during their ovulation and menstruation, the phases between which the difference in the investigated behavior should be the largest. In each session, hormonal samples for cortisol, estradiol, and testosterone were collected. The group of women was also contrasted against an auxiliary reference group composed of 47 males, who are not subject to hormonal variations of this nature. Our results reveal no systematic behavioral differences between the ovulation and menstruation phases.

Keywords: hormonal cycle; ovulatory shift hypothesis; risk preferences; rule violation; exploratory attitude.
JEL Classification: C91, D81, J16. **PsycINFO Classification:** 2260, 2970.

1 Introduction

Understanding the degree to which human preferences are influenced by physiological factors is becoming increasingly important in both economics and psychology, including the impact of one of the most significant natural changes in hormonal levels in humans that occur across the menstrual cycle. In this paper, we bridge the economic and psychological literatures by applying incentivized economic tasks to measure behavioral changes inspired by evolutionary psychological theory.

Mainly related to the explanation of gender gaps in preferences, a growing body of economic literature has examined this relationship because most life-changing events happen between the ages of 15 and 50, which coincides with the time when women can have children (Croson and Gneezy, 2009). The phases of the menstrual cycle (MC) have been found to affect economic behavior, such as competitiveness (Buser, 2012b; Wozniak et al., 2014), social preferences (Buser, 2012a), loss aversion (Lazzaro et al., 2016), or competitive bidding (Chen et al., 2013; Schipper, 2014; Pearson and Schipper, 2013). Conceptually, these papers have explored the biological underpinnings of human behavior, focusing on the role of selected hormones due to their large and predictable fluctuation across the MC. However, with one exception, there is typically no deeper theory of why such changes should occur. When formulating the mechanisms that explain why hormones should affect behavior, these studies simply base their predictions on findings derived from medicine, psychology, and/or behavioral (neuro-)endocrinology.[1] This approach may clearly be problematic as the absence of theory is arguably one of the roots of the replication crisis (Muthukrishna and Henrich, 2019).

In psychology, on the other hand, a prominent evolutionary theory labeled as the ovulatory shift hypothesis (OSH) suggests a reason why there should be behavioral changes across the MC (Gangestad and Simpson, 2000; Gangestad and Thornhill, 2008). This theory is based on the observation, that in many mammalian species, including humans' close cousin the chimpanzee (*Pan troglodytes*), the period just preceding ovulation is associated with remarkable behavioral and body changes, such as increases in female sexual behavior, pheromone release, and changes in physical appearance, which are meant to increase the chance of attracting a male for successful mating and the conception of offspring (e.g., Deschner et al., 2003; Stallmann and Froehlich, 2000). Women, however, do not develop any easily perceptible external body cues around ovulation, making it effectively concealed (Gangestad and Thornhill, 2008). The conclusion that there are no cues of a high-fertility period in women has been noted as premature because a lack of explicit signs does not necessarily imply that more subtle changes are absent (Haselton and Gildersleeve, 2011).

Evolutionary theorists expect female behavior to change throughout the menstrual cycle to best fit the reproductive processes with behavioral patterns to maximize chances of survival and further reproduction. The OSH proposes several broad empirical predictions. A prominent example, the good-genes OSH suggests that ovulation in women in a relationship is connected with a temporary shift in preferences from long-term relationships with men who have cues of prosociality to short-term extra-pair relationships with men that have cues of strong genetic qualities. This way, women

---

[1]The exception is Pearson and Schipper (2013) who formulate a theory similar to the ovulatory shift hypothesis (see below for details) as an *ad-hoc* explanation of their findings. Chen et al. (2013) base their predictions on the medical and psychological literature and cite Richardson (1992), who claims that most menstruating women tend to "experience a variety of physical, psychological and behavioral changes during the period between ovulation and menstruation." Schipper (2014) bases predictions on a handbook of behavioral endocrinology (Nelson, 2011) and takes the MC as a proxy for hormonal fluctuation. Lazzaro et al. (2016) begin with a list of behaviors that are affected by the MC beyond those related to reproductive function and continue with brain-imaging studies, focusing on the reward centers that should imply changes in economic behavior. Buser (2012b) argues that if the MC changes behavior, this will represent evidence that the gender gap in competitiveness is partially caused by biological factors. Wozniak et al. (2014) focus on behavioral gender differences, so they also discuss differences in hormones and ground the related mechanisms in behavioral neuro-endocrinology – that estrogen affects the density of serotonin-binding receptors in the brain, including areas associated with the reward circuit.

would have children with good genes who will be cared for by reliable fathers in a stable and secure prosocial environment (Jones et al., 2019; Gangestad and Thornhill, 2008; Pillsworth and Haselton, 2006; Larson et al., 2012). Other predictions cover lower attraction toward male relatives (Lieberman et al., 2011), changes in what women find attractive in men (Gildersleeve et al., 2014), changes in women's own attractiveness (Haselton and Gildersleeve, 2011), changes in sexual desire (Roney and Simmons, 2013), and changes in intra-sexual competitiveness (Lucas and Koff, 2013). The main suspected underlying mechanisms for those changes are primarily the sex hormones estradiol and progesterone, which potentially affect women's libido (Gangestad and Dinh, 2022).

Empirical evidence in favor of the OSH predictions is rather complicated. Generally, older studies report mostly support for the predictions (e.g., Miller et al., 2007; Haselton and Gildersleeve, 2011; Gildersleeve et al., 2014; Lieberman et al., 2011), while the more recent methodologically advanced studies often find null effects on the part of several OSH predictions (e.g., Schleifenbaum et al., 2022; Holzleitner et al., 2022; Mei et al., 2022; Roney et al., 2023). The proposed reasons for such differences include methodological inconsistencies, such as having small samples and, thus, under-powered statistical tests; measuring the outcome variables as self-reported preferences; vague definitions; not pre-registering data analysis and thus allowing many possibilities for ad-hoc hypotheses testing; or assessing the phase of the menstrual cycle via self-reports and not hormonal analysis (Welling and Burriss, 2019; Gangestad et al., 2016; Kiesner et al., 2020; Jones et al., 2019). A prominent example of methodological issues is a large diary study by Arslan et al. (2021b), in which the authors initially find support for the predictions of the good-genes ovulatory shift hypothesis but subsequently claim the original findings to be weak and inconsistent (Arslan et al., 2021a).

The OSH predictions regarding heightened general sexual desire, sexual behavior, and own attractiveness ratings seem to have strong empirical support (Schleifenbaum et al., 2021; Boudesseul et al., 2019; Arslan et al., 2021b; Marcinkowska et al., 2022). Regarding intra-sexual competitiveness, the literature also seems to find support for OSH-based changes: Eisenbruch and Roney (2016) find that women in fertile phase demand higher shares in the Ultimatum game, in accordance with Lucas and Koff (2013). However, in large samples, Blake et al. (2022) find ovulation to be related to higher perceived self-efficacy, but not assertiveness, and Blake (2022) concludes that fertile women are motivated to gain status through prestige rather than dominance.

The prediction of an increased interest in extra-pair relationships, especially when primary partners are less attractive, seems to be contested (see reviews in Larson et al., 2012; Gangestad and Dinh, 2021). Among the new, methodologically sound studies, Dinh et al. (2022) find a correlation between progesterone levels and extra-pair interest after controlling for male partner attractiveness. However, Thomas et al. (2021) do not find any stronger desires for non-binding relationships, even after controlling for relationship status. Also Arslan et al. (2021b) find an increase in both extra- and in-pair desire, and no partner-retention moderation. Relatedly, no larger inbreeding avoidance was found in Holzleitner et al. (2022), in contrast to Lieberman et al. (2011).

Much less support is provided for the good-genes OSH. Initial studies found women during ovulation to have stronger preference for male traits such as facial and vocal masculinity, expressions of dominant behavior, and body scent (reviews in Haselton and Gildersleeve, 2011; Gildersleeve et al., 2014). The recent methodologically advanced literature finds mostly null effects of ovulation on women's preferences for men's traits such as facial and body masculinity, a beard, and facial symmetry; as well as on men's ability to identify or respond to the fertile phase (Marcinkowska et al., 2018; Jones et al., 2018; Dixson et al., 2018; Jünger et al., 2018; Stern et al., 2020, 2021; Schleifenbaum et al., 2022; Mei et al., 2022; Roney et al., 2023), although the discussion is still ongoing (Gangestad et al., 2019; Gangestad and Dinh, 2021).

In this paper, we investigate whether and how the predictions of the OSH apply to behaviors

beyond mate selection, specifically economic behavior in three relevant domains: risk preferences, rule violation, and exploratory behavior. If a woman truly searches for extra-pair experiences during the fertile part of the menstrual cycle, this shift may extend to increased risk-taking, a higher propensity to break the rules, and more exploratory behavior. To test these three hypotheses, we employ a controlled preregistered economic experiment with incentivized choices in three tasks that have been established in the economic literature to measure the mentioned traits: a Bomb Risk Elicitation Task that measures financial risk attitudes (Crosetto and Filippin, 2013), a rule-violation game based on Hruschka et al. (2014), and an exploration-exploitation task based on Lenow et al. (2017).

In a within-subject design, 124 naturally cycling women participated once during menstruation and once during the ovulation phase in a counterbalanced order. We measured estradiol and testosterone levels from their saliva samples. Because ovulation is difficult to self-assess, the subjects were encouraged to self-administer a standard home luteinizing hormone (LH) test of ovulation before coming to this type of session. As an auxiliary reference group, 47 men also participated twice, with a two-week interval between sessions. To avoid potential confounding effects caused by short and/or long-term stress and anxiety, we also control for salivary and hair cortisol and State and Trait Anxiety (Spielberger et al., 1983). Our main result is that we observe no differences between decisions made during the ovulatory phase and menstruation for any of the tasks.

Several factors led us to expect a difference between ovulation and menstruation in terms of not only reproductive but also financial behavior. Mainly, we assumed that the ovulatory shift effects would manifest through hormonal pathways rather than fertility *per se* and influence the underlying neural structures that play a role in broader behaviors and thus include both domains, financial decision-making and mate choice (Hampson, 2020; Gangestad et al., 2016; Roney and Simmons, 2013; Dinh et al., 2022), as in the literature investigating the effects of oxytocin on social behavior (Zak et al., 2004; Fehr, 2009). We expected estradiol to be the leading direct hormonal pathway because its peak concentration occurs just before ovulation and drops only after ovulation. Estradiol has been shown to affect behavior in animal studies (e.g., the overall cost/benefit ratio in decision-making, Uban et al., 2012), and other research shows estradiol's control over dopamine activity (Jacobs and D'Esposito, 2011).[2] Importantly, several studies suggest such general effects: Pearson and Schipper (2013) show a change in bidding behavior among women during ovulation as compared to all other phases of the cycle; Eisenbruch and Roney (2016) find different behavior in the Ultimatum game for ovulating women, and Coenjaerts et al. (2021) report significant sex-specific effects on the part of exogenously administered estradiol during the Ultimatum game.

The menstruation phase represents a natural baseline for comparison with ovulation because this phase is characterized by both (i) a low risk of conception and (ii) low levels of ovarian hormones, including estradiol. The behavioral change in this phase due to OSH should thus be minimal.[3] The existence of an overt physical marker also makes menstruation easy for the participants to reliably identify, therefore, it is easy for researchers to measure (Hampson, 2020). Related economic studies examine the "high" and "low" hormone phases across the entire MC, as measured based on the self-reported position in the MC (Buser, 2012b,a; Chen et al., 2013; Pearson and Schipper, 2013;

---

[2]In addition to estradiol, luteinizing hormone and follicle-stimulating hormone also peak during ovulation. Progesterone, testosterone, and often even body temperature also change across the MC but in a different pattern (Hampson, 2020). Note that fluctuations of hormones across the MC may also affect neural pathways indirectly (e.g., through interaction with neurotransmitter systems (Barth et al., 2015), brain center volume (Zsido et al., 2022), and the activation of and connectivity between brain centers (Pletzer et al., 2019). For more details on the role of estrogens, with estradiol being a member of this family, in human physiology and cognition, see Galea et al. (2017) or Hall (2019).

[3]The luteal phase also involves a low risk of conception, but it is characterized by high levels of progesterone and thus may be relevant to other studies that focus on progesterone-related outcomes. However, we focus on estradiol (Hall, 2019).

Drichoutis and Nayga, 2014; Wozniak et al., 2014) or saliva or blood samples,[4] without specifically contrasting the two phases.

We contribute to the literature in several ways. First, to the best of our knowledge, no other study has investigated the effect of menstruation or ovulation on rule-violating behavior using an incentivized task. Only stated preference methods, with a rather vague methodology, have been used concerning female promiscuity and its punishment (Cashdan, 1995; Muggleton et al., 2019; Welling and Burriss, 2019). We use a standardized task belonging to the class of mind-games following Fischbacher and Föllmi-Heusi (2013). Games of this class contain a simple explicit rule that concerns the private reporting of a random event unobservable by the experimenter, such as a dice roll or a coin toss, and participants are financially motivated to break this rule. Rule-breaking is undetectable on an individual level, but due to the law of large numbers, it is statistically detectable on the aggregate level. Such games are well established in the literature (Abeler et al., 2019; Gerlach et al., 2019) and seem to have external validity (Cohn et al., 2015; Cohn and Maréchal, 2018; Dai et al., 2018; Hanna and Wang, 2017). We used a task based on Hruschka et al. (2014), which requires subjects to repeatedly roll dice with only two colors on their sides. Before each roll, subjects must choose one of these two colors. If the subsequent roll shows this color on the dice, they are supposed to allocate a token on their own account but if the other color appears, the token is lost to the experimenter. This task has the advantage of increased privacy over the standard cheating games, because even if a subject is directly observed during a dice roll, it can never be established with certainty that the subject cheated. This task also allows for a more precise measurement on an individual level as there are more decisions per person.

Our second contribution is that we are careful concerning the crucial methodology issues associated with this strand of the literature. First, we analyze the results with respect to MC phase in free-cycling women not only based on the self-reports but also by using two hormones, estradiol and testosterone. Second, we encouraged the use of LH home-kits and queried the current position in the MC, as recommended by Schmalenberger et al. (2021) and Blake et al. (2016), because self-reporting, which is frequently used e.g., in the diary studies (Arslan et al., 2021b) does not seem to deliver satisfactory estimates (Jones et al., 2019). We use a recommended within-subject design as a between-subject design would require a prohibitively large number of observations due to the likelihood of large genetic variation in the studied preferences (Gildersleeve et al., 2014). Because most of the studies showing positive results are under-powered (Gangestad et al., 2016), with our three main conditions, within-subject measurements, an expected effect size of $D = 0.25$ and 124 female subjects, we achieve a power of $\beta = 0.74$ for a one-sided t-test, as we have clear predictions regarding the directions of the effects. In this way, we can ensure that we measure a null effect or at least that the effects are smaller than what is economically relevant.

Third, although we strongly encouraged participants to attend only during the two mentioned phases of the MC and we maintained a strict protocol to ensure that the hormonal measurements were accurate, we have a remarkably low rate of consistency of hormonal change with the reported phases. During ovulation, both estradiol and testosterone should increase as compared to during menstruation, but in our sample, both hormones increased in only 36 of 124 female subjects; only estradiol increased in 68 subjects and only testosterone in 75 subjects. This inconsistency may be due to subjects' consciously or unconsciously misreporting the true phase of the cycle or due to the low reliability of salivary estradiol measurement via ELISA kits, as suggested by Arslan et al. (2023). To address this issue, we re-ran the main analysis twice more: once only for women with an increase

---

[4]Lazzaro et al. (2016) assessed behavioral data on women who attended the experiment four times during their regular natural cycle (menstruation, mid-follicular, luteal, and ovulation), which was validated through an analysis of blood samples and contrasting the various phases with one another. Schipper (2014) measured testosterone, estradiol, progesterone, and cortisol before and after experimental sessions and did not focus on a specific phase of the cycle but, rather, on hormonal levels and their changes due to participation in the experiment.

in estradiol during ovulation and another time only for women with an increase in testosterone. We find no effects on any of the three behavioral measures. Overall, this highlights the complicated nature of the empirical investigation of the female MC, as measuring hormones is more complex than it might seem.

Our study also speaks to the recent suggestion in the psychological literature that risk-taking, rule-violations, and status-seeking behavior, in general, can be affected by baseline levels of cortisol and testosterone, which has been labeled the dual-hormone hypothesis (Lee et al., 2015; Dekkers et al., 2019; Mehta et al., 2015; Denson et al., 2013). To investigate it, we use salivary cortisol and testosterone as controls in our analysis, in addition to the estradiol, and run a separate set of regressions in which we interact testosterone with cortisol. A recent meta-analysis (Dekkers et al., 2019) provides only tentative evidence for the dual-hormone hypothesis, which might be due to publication bias but also due to variation in the long-run exposure to stress (Stalder et al., 2017; Mayer et al., 2018). To address this, we also measure hair cortisol as a proxy for long-run stress, as salivary cortisol may indicate short-run stress. To capture the subjective effects of long-run and short-run stress, we further measure State and Trait Anxiety (Spielberger et al., 1983).

Related studies investigating the ovulatory cycle's effects on risk-taking and exploratory behavior, including frequently cited papers on the effects of the MC, predominantly use stated preferences (e.g., asking women if in the last 24 hours they have engaged in behaviors that would be considered risky, as well as how open they would be to a new partner), and correlate them with self-reported MC phase (Chavanne and Gallup Jr., 1998; Bröder and Hohmann, 2003; Kurath and Mata, 2018). Recent studies measure risk preferences using financially incentivized procedures and conclude on no effects of the MC (Lazzaro et al., 2016; Drichoutis and Nayga, 2014), or no significant correlations between risk preference and testosterone for freecycling women (Schipper, 2023). We use incentivized tasks that are easy to understand and operate, the Bomb Risk Elicitation Task (BRET, Crosetto and Filippin, 2013) and the foraging task (Lenow et al., 2017). In the BRET, the subjects observe a square containing 100 fields, and they are informed that under one of them, a bomb is hidden. The fields are automatically and slowly uncovered one by one, and the subjects only decide when to stop. With each uncovered field, subjects receive a monetary reward, but if the bomb field is uncovered, they receive nothing. The BRET allows for a precise estimation of risk preferences and is independent of loss aversion, as loss aversion may be affected by the MC (Lazzaro et al., 2016; Durante et al., 2020). The foraging task requires subjects to decide whether to continue harvesting apples from a current tree that, over time, produces fewer and fewer apples or try a new tree, which has uncertain productivity, instead. This task resembles a measure of ambiguity with a specific framing adapted to exploration. In this study, we explore gender differences in exploration as the first in the literature.

With robust null results for several alternative specifications regarding all three measures of behavior, our study adds another piece of evidence to the growing body of literature evaluating the predictions of the ovulatory shift hypothesis and, more broadly, to the literature studying the effects of hormones on human decision-making. Note, however, that, due to our focus on general behavioral patterns not specific to the context of human mating, and a lack of clear support for the studied behaviors unequivocally increasing individual fitness, our observations do not allow drawing any conclusions concerning the validity of the OSH. Without clear evidence that any given behavioral fluctuation leads to adaptive benefits, it is reasonable to believe that changes in behavior caused by hormonal fluctuations may be considered a byproduct of different functions.

2 Experimental Design

Our experimental design focuses on the effects of natural hormonal fluctuations related to the menstrual cycle on a set of behavioral outcomes. To address this conjuncture, we designed a within-subjects laboratory experiment with three behavioral tasks and pre-registered it. The collection of behavioral outcomes is accompanied by the collection of biological samples (saliva and hair) to properly measure the concentrations of two hormones.[5]

We invited female and male participants to participate in the experiment in separate sessions. The female subjects participate in two experimental sessions: one during ovulation and one during menstruation (about two weeks apart). The within-subjects design allows us to capture intra-subject behavioral changes during the two periods, minimizing the noise in the hormonal measures. By design, we balanced the proportion of female subjects participating in the first session during ovulation or menstruation. Male subjects (the control group) also participated in two experimental sessions two weeks apart. The male population is mainly addressed to control for systematic learning effects between the first and second experimental sessions.

2.1 Behavioral tasks

The experiment incorporates behavioral tasks to address three domains of behavior/decision-making that have been found to be influenced by gender differences in the existing literature. We focus on rule-breaking behavior (Hruschka et al., 2014), exploration and the exploitation of resources (Lenow et al., 2017), and risk-taking behavior (Crosetto and Filippin, 2013). We also consider and measure individuals' cognitive skills (Frederick, 2005) and state-trait anxiety (Spielberger et al., 1983). In the instructions, behavioral tasks were coded with letters "A"-"D" and the task order was randomly determined on an individual level.

Risk: Bomb Risk Elicitation Task (BRET) To elicit the risk attitudes of the participants, we implemented the Bomb Risk Elicitation Task, which was developed by Crosetto and Filippin (2013). Compared to the traditional multiple price-list approaches, such as that of Holt and Laury (2005), the BRET offers a series of analytical advantages. To our purposes, the most prominent is that, due to its intuitive procedure, this tool (i) requires minimal numeracy skills to be understood and used by the subjects and thus avoids potential gender differences and (ii) minimizes potential anchoring effects stemming from the use of the same task twice in a within-subjects design. The intuitive procedure thus minimizes potentially significant gender-math stereotyping effects (e.g., Quinn and Spencer, 2001; Kiefer and Sekaquaptewa, 2007) which may systematically affect and bias the measurement of the individual degree of risk aversion in the female population. Indeed, the related literature does not find gender-specific patterns when using this task (Crosetto and Filippin, 2016; Filippin and Crosetto, 2016).

The participant is presented with a 10x10 matrix of boxes on a screen, and a bomb is hidden under one of the 100 boxes. We use the continuous version of the task, in which a new box is opened each second and the participants are asked to stop the uncovering anytime they want. They are unaware of whether the bomb is located in a box that has already been opened when deciding to stop. Participants are paid for each opened box but lose the reward for this task if the bomb is located in a box that was opened. A risk-neutral person will open exactly 50 boxes, those who

---

[5]Deviations to the pre-analysis plan: we intended to recruit 150 women, but after exerting maximum effort, we were able to attract only 138, out of whom 124 completed both sessions. We also intended to measure progesterone, which we did, but by mistake, the sampling procedure used an improper type of plastic tube that interacted with the progesterone. Thus, the measurements were not reliable. Therefore, we discarded them altogether, and to determine the phase of the cycle, we used only estradiol and testosterone.

are risk-averse will open less than 50 boxes, and those who seek risk will open more than 50. The location of the bomb is typically revealed to the participant once they decide to stop. However, we keep the results private from participants after their first session, as we do not want any learning effects to influence their behavior in the second session.

Rule Violation: Dice Roll Task   To observe rule-violating (often interpreted as cheating) behavior, we implemented a dice roll task modelled after Hruschka et al. (2014). In this task, the participants use a bi-colored dice (three sides are white, and the other three sides are blue (Supplementary Section B.3.4)). Each person is asked to think about one of the colors and then roll the dice. After the dice roll, the participant is supposed to determine the result. If the color outcome is in line with their mental selection, they choose "option A" and receive a point; if there is a mismatch, they choose "option B," generating no points. As there is no control over the actual dice roll, it is impossible to identify rule-violation on an individual level. However, given the repetition involved (> 30 dice rolls), a statistical probability of rule-violation versus truth-telling exists. We did not report exactly how many rolls there were, but we kept the number constant across subjects and sessions, mainly to avoid an anchoring effect because of using the measure twice in a within-subjects design. In this design choice, we differ from previous papers, which have always announced the total number of decisions. In contrast to Fischbacher and Föllmi-Heusi (2013), who used only one decision per subject, we can measure the departure from the predicted 50/50 split on an individual level by presenting multiple decisions to our subjects. However, we focus only on the comparison across the two main groups of measurements. The unknown number of decisions should also decrease the demand effect that some participants may feel when presented with a clear but undetectable option to cheat (Eckerd et al., 2021; Zizzo, 2010). For the dice roll paradigm, the experimental evidence when reports are associated with gains is mixed; Clot et al. (2014) found that females cheat more often than males, while Conrads et al. (2017) detected the opposite pattern. Muehlheusser et al. (2015) and Ezquerra et al. (2018) did not observe systematic gender differences (see Abeler et al., 2019 for a comprehensive meta-study on the dice roll paradigm).

Exploration and Exploitation: Foraging Task   We adapted the paradigm created by Lenow et al. (2017) to measure the exploration versus the exploitation of resources. The incentivized task concerns the harvesting of imaginary apples. The participant has two decisions to make: (i) harvest apples from the same tree (*exploit resources*) or (ii) move to another tree in the orchard (*explore resources*). Each tree has a stochastic depletion rate, and the longer the participants stay with a given tree, the less efficient the harvest is. On the other hand, exploring the orchard is time consuming, and constantly transferring to the next tree would not be efficient. To the best of our knowledge, no studies on gender differences have been conducted in the context of the foraging task.

Cognitive Reflection Test (CRT)   Cognitive abilities are measured using the original set of three questions created by Frederick (2005) and a selected questions from the extension created by Miklánek and Zajíček (2020). We provide nine questions in total to the participants and limit the completion time to 9 minutes. The participants are paid for each correctly answered question. The CRT was used only at one of the two sessions, and it is utilized as an individual-specific control variable.

Full instructions for each task, as they were presented to the participants, are provided in Supplementary Section B. We also provide screenshots of the tasks (upplementary Section C) and the post-experiment questionnaire (upplementary Section D).

3 Procedures and Methods

Our interdisciplinary research incorporates methods derived from two fields. The behavioral data are collected using the standard methodology and guidelines of a laboratory experiment in economics (i.e., adhering to no deception, complete anonymity, blind protocol, and monetary incentives). This is accompanied by collecting and analyzing biological samples (saliva and hair) to determine the basal concentration of the selected hormones (salivary cortisol, estradiol, and testosterone, as well as cortisol in hair samples). The experiment was approved by the Research Ethics Committee of Masaryk University and preregistered at the Open Science Framework repository (https://osf.io/bc673/).

3.1 Laboratory Experiment

The experiment was conducted in the Czech Republic at Masaryk University Experimental Economics Laboratory (MUEEL) and the Laboratory of Experimental Economics at the Prague University of Economics (LEE), from November 2018 to June 2020. The subjects were recruited through Hroot (Bock et al., 2014) at MUEEL and ORSEE (Greiner, 2015) at LEE. We incorporated within-subjects data collection, and subjects were invited for two consecutive sessions, which were about 2 weeks apart. Free-cycling female subjects participated during their ovulation and menstruation phases. Male subjects were invited to participate as a control group, and they participated in two sessions 2 weeks apart to control for the time effects in the within-subjects design. The participants were provided with a summary of information regarding informed consent during the recruitment stage, during which it was emphasized to the female subjects that they have to participate in the first session during either ovulation or menstruation. The menstruation phase was assessed based on individual self-observations and self-reporting. The costs for ovulation tests, if used by the participants, were fully compensated for based on receipts. The experiment was programmed in zTree (Fischbacher, 2007), and the instructions were delivered in the Czech language (see Supplementary Section B).

In total, 157 sessions (only about 2.18 people participated in each session on average, and most sessions had only one participant at a time) were conducted. 124 female and 47 male subjects[6] participated in both sessions of the experiment. We required participating women to attend only during ovulation or menstruation. Because this is difficult to predict, they were encouraged to use a standard home ovulation kit to assess their luteinizing hormone levels, which they would then be reimbursed for, and they were free to report their willingness to attend to the research assistant on the night before the session. Most sessions had, therefore, only one participant. Due to the collection of biological samples, we imposed strict restriction criteria on recruitment, as well as the organization of the sessions. Each session was conducted early in the morning (8:00 am to 9:30 am), and only female research assistants were present in the laboratory for all sessions.[7] The participants had to adhere to the following rules to be allowed to participate: (i) no eating before the session, (ii)

---

[6]In total, 138 women and 58 men participated in the first session of the experiment, and 124 women and 47 men participated in both sessions. The attrition rate was 10.1% for women and 19% for men. One subject was removed from the session due to a lack of compliance with the rules; twenty-two did not appear for the second session, without providing any reason. When comparing the observable characteristics of the non-returning women with those of the returning women, we do not observe any systematic differences, except that the former were slightly more risk averse. The non-returning-women sub-sample is balanced with regard to their MCs (6 out of 13 participated during ovulation).

[7]During the sessions with men, only a male assistant was present. Interactions with the other gender may affect behavior (Booth and Nolen, 2012; Aries, 1996; Haselton and Gildersleeve, 2011; Cigarini et al., 2020), and we decided to keep this dimension constant. We note that, as the real world includes many interactions with other genders, avoiding such imposes a certain limitation on our research. However, random variation of the gender of the assistant and/or whether a session includes single-sex or mixed-sex participants would further increase the noise in our measurements and thus decrease the statistical power of our tests.

no tooth brushing or smoking at least 45 minutes before the session, (iii) no hormonal medication, and (iv) no dentist appointments around the date of the session. Female subjects could participate only if they had a free cycle and took no contraceptives.

For all participants (female and male alike), the procedure for each session was structured into several phases as follows:

I. *Check-in.* At the first session, the subject is checked in, receives a random ID number (to maintain the participant's anonymity between sessions and payments), and signs the informed-consent form (see Supplementary Section E).

II. *Biological sample collection.* For saliva collection, the participants receive a plastic collection container, a plastic tube, and bottled water to support salivation. They then provide a saliva sample by passively drooling into the plastic container. The participants are also provided with a muesli bar after saliva collection to avoid a potential hunger effect on behavioral outcomes.

For the hair sampling, participants are guided by the research assistant to a separate room in the laboratory, where a strand of their hair is collected. If possible, a hair lock was collected from the back of the head, about 2 cm above the hairline. The hair was then wrapped in tin foil to preserve the composition of the hair strands, following the procedure of Kostolanska et al. (2019).

III. *Experiment.* After all biological samples are collected, the research assistant distributes written instructions for behavioral tasks (see Supplementary Section B) and begins the experiment remotely on the subjects' computers.

IV. *Payments.* The subjects are paid at the end of the two sessions. Only one of the three behavioral tasks is randomly selected for payment in each session. After the first session, however, they receive only the show-up fee (150 CZK $\approx$ 5.80 EUR) to motivate them to participate in the second session and minimize attrition. The participants receive, at the end of the second session, a show-up fee and payment for the two behavioral tasks (approximately 86 CZK $\approx$ 3.50 EUR).

3.2 Hormonal Analysis

The hormones were analyzed using the Enzyme-linked immunosorbent assay (ELISA). The procedure for saliva sampling and hair collection, storage, and analysis are described below.

3.2.1 Saliva sample collection and storage

5 to 10 ml of saliva were collected from each participant via drooling into a plastic container. Samples were immediately transferred into 12-ml plastic tubes for biofluid storage and stored at -70°C until further processing. Tubes were labeled with a five-digit code assigned to the individual participants alongside the collection block label (A or B) based on the sample collection timing to allow the simultaneous analysis of paired samples. Samples were stored in the freezer for a maximum of 1.5 years (storage length from 1 month to 1.5 years). Researchers performing the analysis of hormones in saliva samples were blinded. That is, they did not know the participant's identity or the stage of the menstrual cycle in which the samples were collected.

3.2.2 ELISA analysis and data processing

Enzyme-linked immunosorbent assay was used to determine the concentrations of cortisol (ng/l), estradiol (pg/ml), and testosterone (pg/l) in the saliva samples. Direct saliva ELISA kits (AL PCO) were used to determine the concentrations of cortisol and testosterone, while the ELISA kit

10

manufactured by Abnova was used for estradiol assessment. On the day of the analysis, saliva samples were thawed at room temperature, and the appropriate amount of sample was taken for analyses of individual hormones and processed, following the manufacturer's instructions. The remaining saliva fraction was stored at -70°C. To prevent freeze-thaw cycles, hormones from each sample set (corresponding to one ELISA plate) were quantified in one session. Each set contained paired samples (blocks A and B) from 20 individuals for analyses using Alpco kits or from 19 individuals analyzed using Abnova kits (the reduction of samples was due to a higher number of assay controls required). Samples were analyzed in duplicate. Absorbance was measured via a SunriseTM microplate reader (TECAN) at 450 nm (AL PCO) or 405 nm (Abnova). Regression curves and final concentrations were calculated using online software (https://www.elisaanalysis.com/) with a four-parameter logistic regression fitting for cortisol assessment and a five-parameter logistic regression for estradiol and testosterone.

3.3 Hair sample collection and analysis

The hair samples obtained from the participants were analyzed via solid-phase extraction liquid chromatography, coupled with mass spectrometry, to determine the hair cortisol concentration (pg/mg). The hair cortisol concentration (HCC) represents the long-term level of cortisol in an organism, independent of the circadian and pulse rhythms, despite the fact that the mechanism of cortisol incorporation into the hair shaft has not yet been fully elucidated. The HCC was determined in 160 subjects (124 females/36 males) and covered the first three centimeters of hair cut from the scalp. Such segmentation led to the revelation of the HCC about the three prior consecutive months, as average hair growth is approximately 1 cm per month.

4 Results

4.1 Descriptive statistics

Summary statistics are reported in Table 1. In both groups, the average age is approximately 22.5 years, with the vast majority of subjects being Czech and Slovak nationals and only a residual share from other countries. Regarding the female sample, following the precise rationale of the experimental design, 50% of the participants took part in the first session during their reported ovulation phase, and the other 50% took part during the menstruation phase. This ensured the balance of the two main menstrual phases (ovulation versus menstruation) over the two experimental sessions and canceled out potential order effects on measurement. When considering the means of the key variables across the two groups, we observe that the means for the BRET are close to the mid-point of the choice set. We do not observe any significant gender gap in risk attitudes, which is consistent with previous studies using the dynamic BRET task, as compared to multiple-price measures (Crosetto and Filippin, 2013, 2016; Filippin and Crosetto, 2016; Crosetto et al., 2014). The within-subjects differences in the behavioral measures (marked as Δ) are constructed as level differences between the ovulation and menstruation phases for women and between Session 1 and Session 2 for men. The provided summary statistics clearly indicate that, for the full sample of women, we observe only negligible differences, which, in the standard parametric and non-parametric tests, are not significantly different from zero. In the remaining analysis, we also focus only on exploratory behavior, as it is complementary to exploitative behavior given the construction of the task.

Table 1: Descriptives

	Females	Males
Age	22.33 (3.08)	22.95 (2.17)
country CZ:	57%	70%
country SK:	33%	24%
country Other:	10%	6%
CRT score	3.37 (1.94)	4.93 (2.12)
Ovulation - session 1st	50%	.
Risk measure	45.85 (15.62)	44.37 (12.53)
Rule violation measure	19.41 (4.12)	21.55 (5.88)
Exploratory measure	10.44 (4.97)	11.38 (3.88)
Δ Risk measure	1.11 (19.60)	5.63 (12.94)
Δ Rule-violation measure	-0.11 (3.86)	-0.63 (3.90)
Δ Exploratory measure	0.40 (3.63)	0.04 (2.38)
Subjects	124	47

Note: Mean values with standard deviations in parentheses. Within-subject Δ measures are computed as (ovulation − menstruation) for females, (session 1 − session 2) for males.

4.2 Consistency of hormonal measures and cycle phase

Hormonal levels are expected to be higher during ovulation and lower during menstruation. Given this assumption, we should be able to identify a positive Δ at the intra-individual level by contrasting the individual hormonal levels during ovulation against those during menstruation. For the female sample, this relationship holds for 68 subjects when we focus on estradiol, 75 subjects regarding testosterone, and 69 subjects regarding cortisol. For these subjects, the intra-individual differences in the hormonal levels are always highly statistically significant (signed-MWU test, $p<0.001$).

4.3 Models

We begin our analysis by assuming that all participating women reported their cycle phases correctly, as they were strongly encouraged to do. Nevertheless, as subjects may consciously or unconsciously misreport their cycle phases, we perform several robustness checks in which, instead of the dummy variable indicating the ovulation phase, we use (i) only women with a larger value of estradiol during their reported ovulation phase;(ii) only women with a larger value of testosterone during reported ovulation; and (iii) two counting days method, in which ovulation is identified as being the midpoint of the cycle calculated once from the number of days elapsed from the last day of menstruation and another time from the number of days until the expected first day of the next menstruation.

Given the within-subjects design of the experimental setup and the sequential exposure to multiple tasks in each experimental session, our regression analysis relies on panel linear mixed models with random effects to account for potential intra-individual dependencies over sessions (West et al., 2006).[8]

This parametric approach allows us to control for potential ordering effects induced by the sequential administration of multiple tasks and the individual-specific hormonal levels measured through our repeated samplings. Given the richness of the experimental design, a conventional non-parametric approach, in which females' outcomes registered under ovulation are simply contrasted against those

---

[8]Supplementary Section, Table A.16 reports further OLS estimates based on a between-subjects assessment considering only the observations gathered in the first session for the female group. The outcome of this analysis is fully compatible with the results delivered by the within-subjects analyses.

Table 2: Risk Behavior

Outcome: BRET	(1) Pooled	(2) Females	(3) Females	(4) Females	(5) Females	(6) Males	(7) Males	(8) Males
Female	1.2719							
	(2.2682)							
Female*Ovulation	0.6672							
	(1.5425)							
Ovulation		0.6439						
		(1.6664)						
Estradiol			0.0000	0.0000	0.0000	-0.0011	-0.0010	-0.0010
			(0.0001)	(0.0001)	(0.0001)	(0.0012)	(0.0009)	(0.0009)
Testosterone			-0.0018	-0.0010	-0.0006	0.0169*	0.0078	0.0088
			(0.0051)	(0.0050)	(0.0050)	(0.0074)	(0.0050)	(0.0048)
Cortisol			0.0369			-0.0260		
			(0.0470)			(0.0832)		
HairCortisol_1			-0.2995			0.0125		
			(0.2475)			(0.1487)		
Y1SCORE				-0.0740			-0.0570	
				(0.0936)			(0.1540)	
Y2SCORE					-0.0954			-0.4048*
					(0.1171)			(0.1758)
CONST.	47.6106***	48.3701***	49.8243***	51.3087***	52.9431***	52.9425***	49.6343***	64.3375***
	(3.3763)	(3.1609)	(4.3085)	(4.3956)	(5.8933)	(8.3530)	(9.3614)	(9.8989)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
ll	-1388.6888	-1021.038	-1019.9091	-1020.7778	-1020.7561	-278.24431	-359.64865	-357.20622
$Wald - \chi^2$	28.31	19.68	22.07	20.12	20.23	14.21	16.26	21.97
$p > \chi^2$	0.0001	0.0014	0.0048	0.0053	0.0051	0.0765	0.0228	0.0026
Obs.	342	248	248	248	248	72	94	94

Note: Estimates from panel linear mixed models with random effects accounting for potential intra-individual dependencies over sessions. Controls include the score in the CRT, the order of sessions, and the order of the game during a session. Standard errors are in parentheses. The outcome variable is the number of opened boxes and it empirically spans from 0 to 100. *** $p<0.001$, ** $p<0.01$, * $p<0.05$. The decrease in observations in Model (6) can be attributed to the constraint of collecting hair samples, particularly for subjects with hair length shorter than 3 cm, where hair cortisol measurements are not feasible.

observed under the menstrual phase, would reveal only spurious differences potentially influenced by spillover generated by other factors, such as ordering effects. Symmetric considerations apply to the male group when simply comparing outcomes observed in the first session against those observed in the second experimental session.

The structure of regression analysis models is the same across all three behavioral outcomes. Model (1) includes a pooled sample of both men and women, the female dummy, and the interaction of the dummy indicating the ovulation phase with the female dummy. Models (2) to (5) only focus on the female sample. In Model (2), we begin our analysis by simply regressing the outcome variable on our primary variable of interest, the dummy for the (reported) ovulation phase. This should reveal intra-subject behavior changes during the ovulation phase as compared to the menstruation phase. Model (3) replicates the same analysis but now omits the ovulation dummy and adds further controls targeting individual hormonal levels (i.e., estradiol, testosterone, and salivary and hair cortisol. In this model, we assume that the intensity of the treatment effect of ovulation goes through a hormone-related change. Models (4) and (5) keep controlling for estradiol and testosterone, but instead of the cortisol measures, subjectively perceived anxiety is added, once in the short term and another time in the long term. Models (6) to (8) perform a similar function as Models (3) to (5) but

with men. In all models, following design-based considerations, we control for a parsimonious array of controls–the order of sessions, the order of the specific tasks during a session, and the individual CRT score (a proxy for individual-specific cognitive abilities)–but do not include the coefficients, as they are not of primary interest.[9]

4.4   Risk behaviour

Table 2 addresses risk-taking behavior. Following the ovulatory shift hypothesis, we would expect females to engage in more risky behavior during ovulation, rather than during the menstruation phase. The outcome variable is represented by the number of boxes opened by the subject out of a maximum of 100. The subject gains an additional reward from each opened box but loses all rewards if the bomb is located in an opened box.

The estimated coefficients of Model (1) with the pooled sample do not reveal any significant differences from zero, meaning that there is no gender difference in risk-taking and, more importantly, no difference between the ovulation and menstruation phases in this regard. In Model (2), the coefficient associated with the ovulation phase also does not reveal any statistically significant effect. Model (3) replicates the same analysis, now adding a further control targeting the individual hormonal levels, instead of the *ovulation* condition. The results of this specification are in line with the previous results. The coefficients associated with the analyzed hormones do not reveal any statistically significant effect. Models (4) and (5) exchange cortisol measures with state and trait anxiety, again without any of the coefficients being significantly different from zero. Again, all the coefficients are insignificant.

Models (6), (7), and (8) focus on the male group, mirroring the analyses conducted with the female sample. Risk-taking appears to be mildly statistically positively associated with the level of testosterone, but this is true only in one specification out of three, which we consider insignificant.

4.5   Rule-violations

Table 3 addresses rule-violating (cheating) behavior and mirrors the analyses conducted in the previous section. According to the ovulatory shift hypothesis, we would expect females to show a comparatively higher rate of violating rules (i.e., over-reporting the frequency of outcomes favorable for the self) during the ovulation session than during the menstruation session.

In the pooled sample, we observe a small gender effect, in which females report about two tokens fewer than males. Regarding the female-only Models (Models 2 to 5), subjects in their ovulatory phase do not violate rules significantly more often than subjects during the menstruation phase if only hormones are taken into account and short- and long-term anxiety levels are controlled for. The various coefficients associated with the ovulation dummy are relatively small in terms of their magnitude and never statistically significant at any conventional level. In Models (7) and (8), for men, the effect of testosterone turns out to be significant and negative.

4.6   Exploration

According to the ovulatory shift model, exploratory attitudes will be more pronounced during the ovulation phase than the menstruation phase. Table 4 addresses exploratory behavior. Following

---

[9]Given the characteristics of the experimental design, we use an indicator variable *session2* to capture the differential effect observed in the second session. This control variable filters any difference generated by a repetition effect (Berger and Baumeister, 2017; Charness et al., 2012). In addition, we also control for the potential ordering effect generated by the random sequence implemented to administer the various tasks during the experimental sessions (Belton and Sugden, 2018).

the paradigm established by Lenow et al. (2017), exploratory behavior is captured by the number of virtual trees visited by the subjects.

In the pooled Model (1), we see that women tended to explore a little less than men, but there was no effect on the part of ovulation. In line with the patterns observed for the previous outcomes, all models from (2) to (5) consistently show no differential effect on the part of the ovulation-phase dummy. Similarly, all the control variables that refer to individuals' hormonal levels appear to be orthogonal to the observed incentivized behavior. No relevant effects on exploratory attitude are detected in the male groups (Models 6 to 8), with only a statistically weak and positive effect on the part of the cortisol measure.

4.7 Dual-Hormone Hypothesis

The dual-hormone hypothesis posits that basal testosterone may affect risk-taking but only for low-cortisol individuals. However, a recent meta-analysis provides evidence that the effect may be small or, potentially, even simply a result of publication bias and calls for more studies (Dekkers et al., 2019). Because we measure both testosterone and cortisol, we can also easily test for this. Table 5, showing the levels of the hormones, and Table 6, showing the logs of the hormonal measures, show no effect on the part of cortisol, testosterone, or their interaction on any of the decision-making tasks. This result does not provide any support for the dual-hormone hypothesis.

4.8 Robustness checks

In Supplementary Section A, we also report on a number of robustness checks. First, in order to reduce the influence of outliers, we take logs of the hormonal measures and rerun the entire analysis for the three behavioral measures. We report the results in Tables A.1 to A.3. The only difference we find is that, for men's rule-violating behavior, the coefficient of testosterone is more statistically significant than without the log specification. Next, we restrict the sample to only females with a positive estradiol increase in the ovulation phase as compared to menstruation, which occurred in 68 subjects. Tables A.4 to A.6 qualitatively show the same results as the main analysis. Third, we redo the analysis for the subsample of women with a positive increase of testosterone under ovulation, which amounted to 75 women. Tables A.7 to A.9, again, show null results. Last, we use the counting-days method of identifying women who attended the session around the day when ovulation should occur. Here, we use the reported date of the last day of menstruation and focus on the women attending the reported ovulation session 12–16 days afterward (forward counting). Another time, we used the number of days until the next expected first day of menstruation (similar to backward counting), though we did not validate menstruation actually occurring on that day. Tables A.10 to A.15 provide the same null results as all of the other analyses.

5 Discussion

Overall, in a controlled laboratory experiment with incentivized measures of three behavioral outcomes that play a crucial role in economic decision-making, we find no systematic changes in decision-making, as would be implied if the ovulatory-shift hypothesis holds for financial behavior. We first analyze the three outcomes for the entire sample of participants (124 females and 47 males), assuming that all women truthfully attended the session during the ovulation and menstruation phases as reported, and control for the hormonal measures and psychological anxiety. We also do not find any support for the dual-hormone hypothesis. However, there are several caveats that restrict the potential conclusions derived from the null results.

Interestingly, we find a low consistency of a positive increase in estradiol (68 females) and testosterone

15

Table 3: Rule-violations

Outcome: CHEAT	(1) Pooled	(2) Females	(3) Females	(4) Females	(5) Females	(6) Males	(7) Males	(8) Males
Female	-2.0789**							
	(0.7874)							
Female*Ovulation	-0.0298							
	(0.3397)							
Ovulation		0.0166						
		(0.3332)						
Estradiol			0.0000	0.0000	0.0000	0.0000	0.0002	0.0001
			(0.0000)	(0.0000)	(0.0000)	(0.0004)	(0.0004)	(0.0004)
Testosterone			0.0006	0.0003	0.0005	-0.0041	-0.0044*	-0.0043*
			(0.0013)	(0.0013)	(0.0013)	(0.0031)	(0.0020)	(0.0020)
Cortisol			-0.0083			0.0521		
			(0.0120)			(0.0303)		
HairCortisol_1			0.0686			-0.0769		
			(0.0734)			(0.0856)		
Y1SCORE				-0.0185			0.0098	
				(0.0224)			(0.0596)	
Y2SCORE					-0.0405			-0.1154
					(0.0348)			(0.1010)
CONST.	20.8577***	18.2778***	17.9375***	18.7857***	19.8620***	21.9139***	21.0532***	26.2672***
	(1.0499)	(0.7324)	(1.1218)	(1.1057)	(1.6773)	(4.1344)	(3.5992)	(5.1813)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
ll	-957.67849	-671.31039	-669.74711	-670.1639	-669.82621	-211.85555	-273.46791	-272.83971
$Wald-\chi^2$	21.06	16.69	20.16	19.28	20.03	8.24	9.27	10.55
$p > \chi^2$	0.0018	0.0051	0.0098	0.0074	0.0055	0.4106	0.2340	0.1596
Obs.	342	248	248	248	248	72	94	94

Note: Estimates from panel linear mixed models with random effects accounting for potential intra-individual dependencies over sessions. Controls include the score in the CRT, the order of sessions, and the order of the game during a session. Standard errors are in parentheses. The outcome variable is the number of tokens devoted to self and it empirically spans from 9 to 32. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

(75 females) with the reported ovulation phase compared to the menstruation phase. We suggest two explanations for this and run two sets of robustness checks where we restrict the analysis only to the women with a consistent increase in either of the two hormones. We find no statistically significant difference in any of the three behavioral outcomes in any of the specifications.

We note that we strongly encouraged our female participants to self-administer the home ovulation tests and come only if they tested positive,[10] but we did not strictly require the tests before the start of the session. Therefore, because estimating ovulation is rather complicated, there may have

[10]We do acknowledge the fact that this feature of our recruitment protocol may represent a potential source of "purely cognitive" (following the taxonomy proposed by Zizzo, 2010) experimental demand effect, offering mild cues related to true experimental objectives. By design (individual and anonymized decision-making; sessions conducted by RAs and not by the researchers involved in this study), we exclude the interference of the stronger form of "social" experimental demand effect that–in addition to its cognitive dimension–is magnified by the perceived social pressure that the experimenter explicitly or implicitly places on a subject through instructions and cues (Zizzo, 2010). Focusing on risk- and cheating-related behavioural tasks, De Quidt et al. (2018) present a technique for assessing the robustness of laboratory results to the experimental demand effect. This methodology hinges on a model in which subjects respond to their beliefs about the researcher's objectives. The experimental demand effect is then induced in a structured way to measure its impact. Bounds are obtained by manipulating those beliefs by exogenously varying the intensity of the EDE. They estimated bounds average 0.13 standard deviations, suggesting that a conventional latent experimental demand effect is likely to be modest in its magnitude.

Table 4: Exploration

Outcome: Exploration (Exploitation)	(1) Pooled	(2) Females	(3) Females	(4) Females	(5) Females	(6) Males	(7) Males	(8) Males
Female	-1.7115*							
	(0.8016)							
Female*Ovulation	0.3305							
	(0.2940)							
Ovulation		0.3126						
		(0.3174)						
Estradiol			-0.0000	-0.0000	-0.0000	0.0001	0.0002	0.0002
			(0.0000)	(0.0000)	(0.0000)	(0.0003)	(0.0002)	(0.0002)
Testosterone			-0.0001	-0.0003	-0.0003	-0.0005	0.0010	0.0010
			(0.0015)	(0.0015)	(0.0015)	(0.0019)	(0.0013)	(0.0013)
Cortisol			-0.0159			0.0445*		
			(0.0131)			(0.0177)		
HairCortisol_1			0.0581			0.0379		
			(0.0925)			(0.0516)		
Y1SCORE				-0.0044			-0.0284	
				(0.0236)			(0.0376)	
Y2SCORE					-0.0214			-0.0025
					(0.0443)			(0.0659)
CONST.	13.3954***	11.5893***	12.2705***	12.1648***	12.9163***	12.8225***	13.1482***	12.2634***
	(1.0587)	(0.8857)	(1.3424)	(1.2597)	(2.0934)	(2.1630)	(2.1904)	(3.1925)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
ll	-937.76846	-696.55732	-695.00358	-695.98701	-695.88714	-175.03337	-230.38849	-230.67098
$Wald - \chi^2$	14.74	14.42	12.45	12.64	15.49	9.80	9.14	9.16
$p > \chi^2$	0.0224	0.0471	0.0715	0.0868	0.0814	0.0043	0.2003	0.2413
Obs.	342	248	248	248	248	72	94	94

Note: Estimates from panel linear mixed models with random effects accounting for potential intra-individual dependencies over sessions. Controls include the score in the CRT, the order of sessions, and the order of the game during a session. Standard errors are in parentheses. The outcome variable empirically ranges between 1 and 29 virtual trees visited. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

been some participants who misreported the phase and did not attend the session during ovulation. However, we do our best to address this with our robustness checks, focusing on hormonal measures while assuming hormones to be the mechanism through which OSH functions.

Second, a recent paper highlights the low reliability of salivary estradiol for the identification of cycle phase (Arslan et al., 2023). Thus, there may be a reason for being cautious if our estradiol measurements and the associated results are correct. On the other hand, the testosterone (and cortisol) measures obtained from salivary samples should be reliable, according to Arslan et al. (2023). Therefore, as we performed the second robustness check with only those women who showed an increase in testosterone, we did our best to address this issue.

Third, it is entirely possible that the ovulatory shift hypothesis is restricted to the domain of sexually reproductive behavior and does not affect broader behaviors, including financially motivated behavior, as we assumed in our study. However, this conclusion also matters for the economically relevant behaviors identified in our games and delivers important information to policymakers and managers alike. Moreover, it has been suggested that the ovulatory shift should differ with respect to relationship status because single women do not have anyone to cheat on (Pillsworth and Haselton, 2006; Stern et al., 2020). We did not ask about relationship status, and thus, we cannot

17

Table 5: Dual-Hormone Hypothesis (pooled sample, levels)

	(1) BRET	(2) CHEAT	(3) EXPLOR
Testosterone	-0.0081	-0.0038	-0.0022
	(0.0109)	(0.0030)	(0.0027)
Cortisol	-0.0079	-0.0078	-0.0057
	(0.0515)	(0.0149)	(0.0139)
Testosterone*Cortisol	0.0002	0.0000	0.0000
	(0.0002)	(0.0001)	(0.0000)
Female	1.6962	-2.2661**	-1.5607
	(2.1876)	(0.7778)	(0.7975)
CONST.	47.9801***	21.6220***	13.7291***
	(4.2320)	(1.2869)	(1.2724)
Controls	✓	✓	✓
ll	-1387.956	-956.56006	-937.9765
$Wald - \chi^2$	29.75	23.45	14.30
$p > \chi^2$	0.0002	0.0028	0.0742
Obs.	342	342	342

Note: Estimates from panel linear mixed models with random effects accounting for potential intra-individual dependencies over sessions. Controls include the score in the CRT, the order of sessions, and the order of the game during a session. Standard errors are in parentheses. The outcome variable empirically ranges between 1 and 29 virtual trees visited. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

Table 6: Dual-Hormone Hypothesis (pooled sample, logs)

	(1) BRET	(2) CHEAT	(3) EXPLOR
LOG_Testosterone	-1.5506	-1.6584	-1.3564
	(4.8870)	(1.3565)	(1.2382)
LOG_Cortisol	-0.8019	-2.2514	-2.0275
	(6.2409)	(1.7815)	(1.6435)
LOG_Testosterone*LOG_Cortisol	0.4283	0.4029	0.4174
	(1.3330)	(0.3712)	(0.3390)
Female	1.5736	-2.2998**	-1.4467
	(2.2655)	(0.7922)	(0.8162)
CONST.	50.5478*	30.1115***	19.9146**
	(22.9051)	(6.5346)	(6.0624)
Controls	✓	✓	✓
ll	-1388.5067	-956.23623	-937.46374
$Wald - \chi^2$	28.69	24.14	15.40
$p > \chi^2$	0.0004	0.0022	0.0518
Obs.	342	342	342

Note: Estimates from panel linear mixed models with random effects accounting for potential intra-individual dependencies over sessions. Controls include the score in the CRT, the order of sessions, and the order of the game during a session. Standard errors are in parentheses. The outcome variable empirically ranges between 1 and 29 virtual trees visited. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

report whether it matters, but the empirical evidence thus far does not suggest that its importance (Thomas et al., 2021; Stern et al., 2020). Therefore, our null results do not directly contradict the validity of the OSH predictions, only if the ovulatory shift hypothesis potentially extends to financial behavior.

We position our study within the literature on the empirical evidence for extended predictions of the ovulatory-shift hypothesis and, more generally, on the effects of hormonal variation on decision-making. However, a full comparison with the related studies is rather difficult due to the large methodological differences and a low level of consistency between the self-reported cycle phase and hormonal measurements. Our results contribute to the ongoing discussion of the incorporation of biomedical technologies and instruments into applied economic research (Smith, 2023). With respect to the investigation of the hormonal pathways of the potential effect, we note that our study focuses on, arguably, one of the largest naturally occurring variations in hormonal levels in humans. Still, this change occurs endogenously and is not an external manipulation administered in a randomized manner, such as in randomized control trials (RCTs), which are considered to be the gold standard in experimentation (Kvarven et al., 2020). Studies using RCTs, however, also frequently show null effects on the part of exogenously administered hormones on human behavior, be they oral contraceptives, estrogen, or testosterone (Zethraeus et al., 2009; Ranehill et al., 2018; Nadler et al., 2021).

References

Abeler, J., Nosenzo, D., and Raymond, C. (2019). Preferences for Truth-Telling. *Econometrica*, 87(4):1115–1153.

Aries, E. (1996). *Men and women in interaction: Reconsidering the differences*. Oxford University Press.

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., and Sunni, L. (2023). Not within spitting distance : salivary immunoassays of estradiol have subpar validity for cycle phase. *Psychoneuroendocrinology*, 149(105994):1–11.

Arslan, R. C., Driebe, J. C., Stern, J., Gerlach, T. M., and Penke, L. (2021a). The evidence for good genes ovulatory shifts in Arslan et al. (2018) is mixed and uncertain. *Journal of Personality and Social Psychology*, 121(2):441–446.

Arslan, R. C., Schilling, K. M., Gerlach, T. M., and Penke, L. (2021b). Using 26,000 diary entries to show ovulatory changes in sexual desire and behavior. *Journal of Personality and Social Psychology*, 121(2):410–431.

Barth, C., Villringer, A., and Sacher, J. (2015). Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. *Frontiers in Neuroscience*, 9(February):1–20.

Belton, C. A. and Sugden, R. (2018). Attention and novelty: An experimental investigation of order effects in multiple valuation tasks. *Journal of Economic Psychology*, 67:103–115.

Berger, R. and Baumeister, B. (2017). Repetition effects in laboratory experiments. *Social Dilemmas, Institutions, and the Evolution of Cooperation*, page 547.

Blake, K. R. (2022). When fertile, women seek status via prestige but not dominance. *Proceedings of the National Academy of Sciences of the United States of America*, 119(46):1–8.

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

Blake, K. R., McCartney, M., and Arslan, R. C. (2022). Menstrual cycle and hormonal contraception effects on self-efficacy, assertiveness, regulatory focus, optimism, impulsiveness, and risk-taking. *Journal of Experimental Social Psychology*, 103(November 2021):104382.

Bock, O., Baetge, I., and Nicklisch, A. (2014). Hroot: Hamburg Registration and Organization Online Tool. *European Economic Review*.

Booth, A. L. and Nolen, P. (2012). Gender differences in risk behaviour: Does nurture matter? *Economic Journal*, 122(558):56–78.

Boudesseul, J., Gildersleeve, K. A., Haselton, M. G., and Bègue, L. (2019). Do women expose themselves to more health-related risks in certain phases of the menstrual cycle? A meta-analytic review. *Neuroscience and Biobehavioral Reviews*, 107:505–524.

Bröder, A. and Hohmann, N. (2003). Variations in risk taking behavior over the menstrual cycle. *Evolution and Human Behavior*, 24(6):391–398.

Buser, T. (2012a). Digit ratios, the menstrual cycle and social preferences. *Games and Economic Behavior*, 76(2):457–470.

Buser, T. (2012b). The impact of the menstrual cycle and hormonal contraceptives on competitiveness. *Journal of Economic Behavior & Organization*, 83(1):1–10.

Cashdan, E. (1995). Hormones, Sex, and Status in Women. *Hormones and Behavior*, 29:354–366.

Charness, G., Gneezy, U., and Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of economic behavior & organization*, 81(1):1–8.

Chavanne, T. J. and Gallup Jr., G. G. (1998). Variation in Risk Taking Behavior Among Female College Students as a Function of the Menstrual Cycle. *Evolution and Human Behavior*, 19:27–32.

Chen, Y., Katuščák, P., and Ozdenoren, E. (2013). Why cant a woman bid more like a man? *Games and Economic Behavior*, 77(1):181–213.

Cigarini, A., Vicens, J., and Perello, J. (2020). Gender-based pairings influence cooperative expectations and behaviours. *Scientific Reports*, 10(1):1–10.

Clot, S., Grolleau, G., and Ibanez, L. (2014). Smug alert! exploring self-licensing behavior in a cheating game. *Economics Letters*, 123(2):191–194.

Coenjaerts, M., Pape, F., Santoso, V., Grau, F., Stoffel-Wagner, B., Philipsen, A., Schultz, J., Hurlemann, R., and Scheele, D. (2021). Sex differences in economic decision-making: Exogenous estradiol has opposing effects on fairness framing in women and men. *European Neuropsychopharmacology*, 50(September):46–54.

Cohn, A. and Maréchal, M. A. (2018). Laboratory Measure of Cheating Predicts School Misconduct. *The Economic Journal*, 128(615):2743–2754.

Cohn, A., Maréchal, M. A., and Noll, T. (2015). Bad boys: the effect of criminal identity on dishonesty. *Review of Economic Studies*, 82:1289–1308.

Conrads, J., Ellenberger, M., Irlenbusch, B., Ohms, E. N., Rilke, R. M., and Walkowitz, G. (2017). Team goal incentives and individual lying behavior. Technical report, WHU-Otto Beisheim School of Management.

Crosetto, P. and Filippin, A. (2013). The "bomb" risk elicitation task. *Journal of Risk and Uncertainty*, 47(1):31–65.

Crosetto, P. and Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, 19(3):613–641.

Crosetto, P., Filippin, A., and Heider, J. (2014). A Study of Outcome Reporting Bias Using Gender Differences in Risk Attitudes. *CESifo Economic Studies*, pages 1–24.

Croson, R. T. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

Dai, Z., Galeotti, F., and Villeval, M. C. (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Management Science*, 64(3):1081–1100.

De Quidt, J., Haushofer, J., and Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, 108(11):3266–3302.

Dekkers, T. J., van Rentergem, J. A., Meijer, B., Popma, A., Wagemaker, E., and Huizenga, H. M. (2019). A meta-analytical evaluation of the dual-hormone hypothesis: Does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy? *Neuroscience and Biobehavioral Reviews*, 96(September 2017):250–271.

21

Denson, T. F., Mehta, P. H., and Ho Tan, D. (2013). Endogenous testosterone and cortisol jointly influence reactive aggression in women. *Psychoneuroendocrinology*, 38:416–424.

Deschner, T., Heistermann, M., Hodges, K., and Boesch, C. (2003). Timing and probability of ovulation in relation to sex skin swelling in wild West African chimpanzees, Pan troglodytes verus. *Animal Behaviour*, 66(3):551–560.

Dinh, T., Emery, M., and Gangestad, S. W. (2022). Hormonal influences on women's extra-pair sexual interests : The moderating impact of partner attractiveness. *Evolution and Human Behavior*, 43(5):442–454.

Dixson, B. J. W., Lee, A. J., Blake, K. R., Jasienska, G., and Marcinkowska, U. M. (2018). Hormones and Behavior Women ' s preferences for men ' s beards show no relation to their ovarian cycle phase and sex hormone levels. *Hormones and Behavior*, 97(November 2017):137–144.

Drichoutis, A. C. and Nayga, R. M. (2014). Do risk and time preferences have biological roots? *Southern Economic Journal*, 82(May 2014):140528150321007.

Durante, K. M., Griskevicius, V., and Ulu, S. S. (2020). The effect of fertility on loss aversion. *Journal of Business Research*, 120(June 2018):599–607.

Eckerd, S., DuHadway, S., Bendoly, E., Carter, C. R., and Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, 67(2):261–275.

Eisenbruch, A. B. and Roney, J. R. (2016). Conception Risk and the Ultimatum Game: When Fertility is High, Women Demand More. *Personality and Individual Differences*, 98:272–274.

Ezquerra, L., Kolev, G. I., and Rodriguez-Lara, I. (2018). Gender differences in cheating: Loss vs. gain framing. *Economics Letters*, 163:46–49.

Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7(2-3):235–266.

Filippin, A. and Crosetto, P. (2016). A Reconsideration of Gender Differences in Risk Attitudes. *Management Science*, 62(11):3138–3160.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fischbacher, U. and Föllmi-Heusi, F. (2013). Lies in disguise-an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4):25–42.

Galea, L. A., Frick, K. M., Hampson, E., Sohrabji, F., and Choleris, E. (2017). Why estrogens matter for behavior and brain health. *Neuroscience and Biobehavioral Reviews*, 76:363–379.

Gangestad, S. W. and Dinh, T. (2021). Robust Evidence for Moderation of Ovulatory Shifts by Partner Attractiveness in Arslan et al.'s (2020) Data. *Journal of Personality and Social Psychology*, 121(2):432–440.

Gangestad, S. W. and Dinh, T. (2022). Women's Estrus and Extended Sexuality: Reflections on Empirical Patterns and Fundamental Theoretical Issues. *Frontiers in Psychology*, 13(June).

Gangestad, S. W., Dinh, T., Grebe, N. M., Del Giudice, M., and Emery Thompson, M. (2019). Psychological cycle shifts redux, once again: response to Stern et al., Roney, Jones et al., and Higham. *Evolution and Human Behavior*, 40(6):537–542.

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

Gangestad, S. W. and Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences*, 23(4):573–587.

Gangestad, S. W. and Thornhill, R. (2008). Human oestrus. *Proceedings of the Royal Society B: Biological Sciences*, 275(1638):991–1000.

Gerlach, P., Teodorescu, K., and Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, 145(1):1–44.

Gildersleeve, K., Haselton, M. G., and Fales, M. R. (2014). Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological bulletin*, 140(5):1205–1259.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*.

Hall, J. E. (2019). Neuroendocrine Control of the Menstrual Cycle. In *Yen and Jaffe's Reproductive Endocrinology*, pages 149–166.e5. Elsevier, eighth edi edition.

Hampson, E. (2020). A brief guide to the menstrual cycle and oral contraceptive use for researchers in behavioral endocrinology. *Hormones and Behavior*, 119(December 2019):104655.

Hanna, R. and Wang, S.-y. (2017). Dishonesty and Selection into Public Service: Evidence from India. *American Economic Journal: Economic Policy*, 9(3):262–290.

Haselton, M. G. and Gildersleeve, K. (2011). Can men detect ovulation? *Current Directions in Psychological Science*, 20(2):87–92.

Holt, C. A. and Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review*, 95(3):902–912.

Holzleitner, I. J., Driebe, J. C., Arslan, R. C., Hahn, A. C., Lee, A. J., Shea, K. J. O., Gerlach, T. M., Penke, L., Jones, B. C., and Debruine, L. M. (2022). No increased inbreeding avoidance during the ovulatory phase of the menstrual cycle. *Evolutionary Human Sciences*, 4(e47):1–19.

Hruschka, D., Efferson, C., Jiang, T., Falletta-Cowden, A., Sigurdsson, S., McNamara, R., Sands, M., Munira, S., Slingerland, E., and Henrich, J. (2014). Impartial institutions, pathogen stress and the expanding social network. *Human Nature*, 25(4):567–579.

Jacobs, E. and D'Esposito, M. (2011). Estrogen shapes dopamine-dependent cognitive processes: Implications for women's health. *Journal of Neuroscience*, 31(14):5286–5293.

Jones, B. C., Hahn, A. C., and DeBruine, L. M. (2019). Ovulation, Sex Hormones, and Women's Mating Psychology. *Trends in Cognitive Sciences*, 23(1):51–62.

Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., Fasolt, V., Morrison, D., Lee, A. J., Holzleitner, I. J., O'Shea, K. J., Roberts, S. C., Little, A. C., and DeBruine, L. M. (2018). No Compelling Evidence that Preferences for Facial Masculinity Track Changes in Women's Hormonal Status. *Psychological Science*, 29(6):996–1005.

Jünger, J., Motta-Mena, N. V., Cardenas, R., Bailey, D., Rosenfield, K. A., Schild, C., Penke, L., and Puts, D. A. (2018). Do women's preferences for masculine voices shift across the ovulatory cycle? *Hormones and Behavior*, 106(October):122–134.

Kiefer, A. K. and Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, 43(5):825–832.

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

Kostolanska, K., Novotna, L., Taborska, E., and Pes, O. (2019). Online solid-phase extraction liquid chromatography–mass spectrometry of hair cortisol using a surrogate analyte. *Chemical Papers*, 73(1):151–158.

Kurath, J. and Mata, R. (2018). Individual differences in risk taking and endogeneous levels of testosterone, estradiol, and cortisol: A systematic literature search and three independent meta-analyses. *Neuroscience and Biobehavioral Reviews*, 90:428–446.

Kvarven, A., Strømland, E., and Johannesson, M. (2020). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour*, 4(4):423–434.

Larson, C. M., Pillsworth, E. G., and Haselton, M. G. (2012). Ovulatory Shifts in Women's Attractions to Primary Partners and Other Men: Further Evidence of the Importance of Primary Partner Sexual Attractiveness. *PLoS ONE*, 7(9):e44456.

Lazzaro, S. C., Rutledge, R. B., Burghart, D. R., and Glimcher, P. W. (2016). The Impact of Menstrual Cycle Phase on Economic Choice and Rationality. *PLOS ONE*, 11(1):e0144080.

Lee, J. J., Gino, F., Jin, E. S., Rice, L. K., and Josephs, R. A. (2015). Hormones and ethics: Understanding the biological basis of unethical conduct. *Journal of Experimental Psychology: General*, 144(5):891–897.

Lenow, J. K., Constantino, S. M., Daw, N. D., and Phelps, E. A. (2017). Chronic and acute stress promote overexploitation in serial decision making. *Journal of Neuroscience*, 37(23):5681–5689.

Lieberman, D., Pillsworth, E. G., and Haselton, M. G. (2011). Kin affiliation across the ovulatory cycle: Females avoid fathers when fertile. *Psychological Science*, 22(1):13–18.

Lucas, M. and Koff, E. (2013). How conception risk affects competition and cooperation with attractive women and men. *Evolution and Human Behavior*, 34(1):16–22.

Marcinkowska, U. M., Galbarczyk, A., and Jasienska, G. (2018). La donna è mobile? Lack of cyclical shifts in facial symmetry, and facial and body masculinity preferences—A hormone based study. *Psychoneuroendocrinology*, 88(July 2017):47–53.

Marcinkowska, U. M., Shirazi, T., Mijas, M., Roney, J. R., Marcinkowska, U. M., Shirazi, T., Mijas, M., James, R., Marcinkowska, U. M., Shirazi, T., and Mijas, M. (2022). Hormonal Underpinnings of the Variation in Sexual Desire , Arousal and Activity Throughout the Menstrual Cycle – A Multifaceted Approach Hormonal Underpinnings of the Variation in Sexual Desire , Arousal and Activity. *The Journal of Sex Research*, 00(00):1–7.

Mayer, S. E., Lopez-duran, N. L., Sen, S., and Abelson, J. L. (2018). Chronic stress, hair cortisol and depression: A prospective and longitudinal study of medical internship. *Psychoneuroendocrinology*, 92:57–65.

Mehta, P. H., Welker, K. M., Zilioli, S., and Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology*.

Mei, M., Grillot, R. L., Abbey, C. K., Thompson, M. E., Roney, J. R., and Roney, J. R. (2022). Does scent attractiveness reveal women ' s ovulatory timing ? Evidence from signal detection analyses and endocrine predictors of odour attractiveness. *Proceedings of the Royal Society B*, 289(20220026):1–9.

Miklánek, T. and Zajíček, M. (2020). Personal traits and trading in an experimental asset market. *Journal of Behavioral and Experimental Economics*, 86:101538.

Miller, G., Tybur, J. M., and Jordan, B. D. (2007). Ovulatory cycle effects on tip earnings by lap dancers: economic evidence for human estrus? *Evolution and Human Behavior*, 28(6):375–381.

Muehlheusser, G., Roider, A., and Wallmeier, N. (2015). Gender differences in honesty: Groups versus individuals. *Economics Letters*, 128:25–29.

Muggleton, N. K., Tarran, S. R., and Fincher, C. L. (2019). Who punishes promiscuous women? Both women and men are prejudiced towards sexually-accessible women, but only women inflict costly punishment. *Evolution and Human Behavior*, 40(3):259–268.

Muthukrishna, M. and Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, 3(3):221–229.

Nadler, A., Wibral, M., Dohmen, T., Falk, A., Previtero, A., Weber, B., Camerer, C., Dreber, A., and Nave, G. (2021). Does testosterone increase willingness to compete, confidence, and risk-taking in men? Evidence from two randomized placebo-controlled experiments. *PsyArXiv*, (1942917):1–47.

Nelson, R. J. (2011). *An introduction to behavioral endocrinology.* Sinauer Associates.

Pearson, M. and Schipper, B. C. (2013). Menstrual cycle and competitive bidding. *Games and Economic Behavior*, 78:1–20.

Pillsworth, E. G. and Haselton, M. G. (2006). Women's sexual strategies: The evolution of long-term bonds and extrapair sex. *Annual Review of Sex Research*, 17(June 2014):59–100.

Pletzer, B., Harris, T.-a., Scheuringer, A., and Hidalgo-Lopez, E. (2019). The cycling brain: menstrual cycle related fluctuations in hippocampal and fronto-striatal activation and connectivity during cognitive tasks. *Neuropsychopharmacology*, 44(11):1867–1875.

Quinn, D. M. and Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, 57(1):55–71.

Ranehill, E., Zethraeus, N., Blomberg, L., Von Schoultz, B., Hirschberg, A. L., Johannesson, M., and Dreber, A. (2018). Hormonal contraceptives do not impact economic preferences: Evidence from a randomized trial. *Management Science*, 64(10):4515–4532.

Richardson, J. T. (1992). The menstrual cycle, cognition, and paramenstrual symptomatology. In *Cognition and the menstrual cycle*, pages 1–38. Springer.

Roney, J. R., Mei, M., Grillot, R. L., and Emery, M. (2023). No effects of exposure to women's fertile window body scents on men's hormonal and psychological responses. *Evolution and Human Behavior*, (March).

25

Roney, J. R. and Simmons, Z. L. (2013). Hormonal predictors of sexual motivation in natural menstrual cycles. *Hormones and Behavior*, 63(4):636–645.

Schipper, B. C. (2014). Sex hormones and competitive bidding. *Management Science*, 61(2):249–266.

Schipper, B. C. (2023). Sex hormones and choice under risk. *Journal of Economic Psychology*, 96:102607.

Schleifenbaum, L., Driebe, J. C., Gerlach, T. M., Penke, L., and Arslan, R. C. (2021). Women feel more attractive before ovulation: Evidence from a large-scale online diary study. *Evolutionary Human Sciences*, 3.

Schleifenbaum, L., Stern, J., Driebe, J. C., Wieczorek, L. L., Gerlach, T. M., Arslan, R. C., and Penke, L. (2022). Men are not aware of and do not respond to their female partner ' s fertility status : Evidence from a dyadic diary study of 384 couples. *Hormones and Behavior*, 143(November 2021):105202.

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

Smith, T. G. (2023). Endocrine state is the physical manifestation of subjective beliefs. *Journal of Economic Psychology*, 96:102620.

Spielberger, C., Gorsuch, R. L., Lushene, R., Vagg, P., and Jacobs, G. (1983). *Manual for the State-trait anxiety inventory (form Y) ("self-evaluation questionnaire")*. Consulting Psychologists Press.

Stalder, T., Steudte-Schmiedgen, S., Alexander, N., Klucken, T., Vater, A., Wichmann, S., Kirschbaum, C., and Miller, R. (2017). Stress-related and basic determinants of hair cortisol in humans: A meta-analysis. *Psychoneuroendocrinology*, 77:261–274.

Stallmann, R. R. and Froehlich, J. W. (2000). Primate sexual swellings as coevolved signal systems. *Primates*, 41(1):1–16.

Stern, J., Gerlach, T. M., and Penke, L. (2020). Probing Ovulatory-Cycle Shifts in Women's Preferences for Men's Behaviors. *Psychological Science*, 31(4):424–436.

Stern, J., Kordsmeyer, T. L., and Penke, L. (2021). A longitudinal evaluation of ovulatory cycle shifts in women's mate attraction and preferences. *Hormones and Behavior*, 128(February).

Thomas, A. G., Armstrong, S. L., Stewart-Williams, S., and Jones, B. C. (2021). Current Fertility Status Does Not Predict Sociosexual Attitudes and Desires in Normally Ovulating Women. *Evolutionary Psychology*, 19(1):147470492097631.

Uban, K. A., Rummel, J., Floresco, S. B., and Galea, L. A. (2012). Estradiol modulates effort-based decision making in female rats. *Neuropsychopharmacology*, 37(2):390–401.

Welling, L. L. M. and Burriss, R. P. (2019). Investigating the Ovulatory Cycle. *The Oxford Handbook of Evolutionary Psychology and Behavioral Endocrinology*, (May):107–123.

West, B. T., Welch, K. B., and Galecki, A. T. (2006). *Linear mixed models: a practical guide using statistical software*. Chapman and Hall/CRC.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, 32(1):161–198.

Zak, P. J., Kurzban, R., and Matzner, W. T. (2004). The neurobiology of trust. *Annals of the New York Academy of Sciences*, 1032:224–227.

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., Von Schoultz, B., Hirschberg, A. L., and Johannesson, M. (2009). A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16):6535–6538.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13:75–98.

Zsido, R., Barth, C., Kurth, L., and Witte, A. V. (2022). Menstrual cycle brain plasticity : Ultra-high field 7T MRI reveals changes in human medial temporal lobe volume in female adults. *preprint*, i:1–21.

Title: Ovulatory Shift Hypothesis, Hormonal Changes, and No Effects on Incentivized Decision-Making

Highlights

- We investigate the effects of ovulatory shift on salient behavioral outcomes.

- Within-subject study with free cycling women during their ovulation and menstruation.

- We measured changes in their a risk-taking, rule-violating, and exploring behavior.

- Results reveal no systematic differences between ovulation and menstruation.