Visceral Influences and Gender Difference in Competitiveness

Jingcheng Fu Songfa Zhong*

December 10, 2019

Abstract

In a stylized experimental design, it is commonly observed that, after performing a task first under a piece rate and then a tournament incentive scheme, men are more likely to choose tournament for the next performance than women. We conduct two experiments to investigate whether this difference is influenced by the visceral state as the tournament-entry choice is made right after the first two tasks. Experiment 1 examines the responses of hormones, including progesterone, estradiol, testosterone and cortisol: men have stronger hormonal responses than women, and part of the gender difference in competitiveness is explained by the hormonal responses. Experiment 2 exogenously introduces a resting period before the choice, and shows that it reduces the competitiveness of men but does not affect women. Our results suggest that visceral influences contribute to the widely observed gender difference in competitiveness.

Keywords: tournament, piece rate, competitiveness, visceral influences, sex hormone, gender difference, experiment

JEL classification: C91, D44, D83

*Fu: Department of Economics, National University of Singapore; Email: jingcheng.fu@gmail.com. Zhong: Department of Economics, National University of Singapore; Email: zhongsongfa@gmail.com.
Acknowledgment: To be added.

1 Introduction

Competition is pervasive in our lives: we compete to get into college, to stand out in a job interview, to win the attention of a potential partner. Oftentimes whether to compete is part of a constant stream of decisions and choices that interact with our moods and emotions. How competitive we are is not isolated from emotions experienced at the time of decision making. Our competitiveness may be affected by emotions, drive states and feeling states, referred to as visceral factors (Loewenstein 1996), which are shown to have powerful and systematic influences on choice behavior in many other settings.[1] In this paper, we examine how visceral factors shape competitiveness using laboratory experiments.

The emergence of the experimental literature on competitiveness is motivated by its tight link with the gender gap in labor market outcomes, particularly the underrepresentation of women in high-paying jobs. A substantial part of the gender gap cannot be explained by the difference in human capital accumulation and discrimination (e.g., Altonji and Blank 1999). Research on gender differences in non-cognitive skills, psychological traits and preferences, such as attitude towards competition, turns out to be a major development towards the understanding of the "unexplained" gender gap (Bertrand 2011). In their seminal work, Niederle and Vesterlund (2007) – henceforth, NV – show that men are more competitive than women. It is more likely for a man to choose a winner-takes-all tournament over piece rate, compared to a woman who is equally good at the task. This gender difference is well replicated in the laboratory with few exception (see the review of Niederle and Vesterlund 2011). The competitiveness measured by tournament entry predicts behavior outside the laboratory, such as selecting a more prestigious academic track (Buser et al. 2014, Zhang 2013, Buser, Peter and Wolter 2017), forming higher earnings expectations and staying in a better-paid industry (Reuben et al. 2017, 2015), and taking up additional investment and employment (Berge et al. 2015).

Visceral influences offer a potential channel to give rise to the observed gender difference in competitiveness. More specifically, in the NV design, competitiveness is elicited after a vivid experience of performing the task first under a piece-rate and then under a tournament compensation scheme. The tasks experienced immediately before subjects choose between piece rate and tournament trigger stress and emotional arousal (e.g., see

[1] Among major manifestations of visceral influences is the projection bias, in which decision-makers are biased by the current drive state when they are to predict the preference for the future state (Loewenstein et al. 2003). Empirical evidence of projection bias has been provided by both experimental studies (e.g., Read and van Leeuwen 1998, Ariely and Loewenstein 2006, Badger et al. 2007, Augenblick and Rabin 2019) and studies using naturally-occurring data (e.g., Busse et al. 2015, Chang et al. 2018).

the review of Buckert et al. 2017), which can affect the subsequent willingness to compete. Men and women may undergo different influences if the visceral responses differ by gender, or if the effects on behavior are gender-specific. For instance, men exhibit stronger emotional responses to competition in video games than women (Kivikangas et al. 2014), and temporary hunger changes the behavior of men in the opposite direction of women in terms of decision quality, risk aversion, cognitive performance and generosity (Chen et al. 2018).

We conduct two experiments to investigate whether and how the visceral influences on competitiveness elicited using the NV design differ by gender. Experiment 1 measures individual-level strength of visceral influences using the change of hormonal levels, and tests whether the hormonal responses explain tournament entry. Experiment 2 examines an implication of the visceral influences: whether the insertion of a resting period before decision making affects competitiveness. Both experiments show that visceral influences increase competitiveness, and more importantly, the visceral responses of men are stronger than those of women, which explains part of the gender difference in competitiveness.

This paper builds on and extends the literature which examines the relationship between stress and competitiveness. An indirect evidence of the relationship is that, while men are more willing to compete in a high-pressure mathematical task, the gender gap is reversed for low-pressure verbal tasks (Shurchkov 2012). A growing number of studies directly measure stress, and manipulate either physical or social stress. Apicella et al. (2011) find no significant association between the baseline level of stress hormone, cortisol, and competitiveness in a sample of men. Some studies find no effect of experimentally induced stress (Zhong et al. 2018), others find heterogeneous effects conditional on the level of anxiety (Goette et al. 2015), or specific to one gender (Buser, Dreber and Mollerstrom 2017, Cahlikova et al. 2019, Esopo et al. 2019). More closely related to us are the findings regarding the stress response within the NV design. Again there is no firm consensus, as some studies report a gender-neutral correlation (Zhong et al. 2018) while others report a gender-specific correlation (Buser, Dreber and Mollerstrom 2017, Halko and Sääksvuori 2017) or no correlation (Buckert et al. 2017). Perhaps surprisingly, converging evidence from these studies shows that gender difference in stress response does not account for a meaningful part of the gender gap in tournament entry.

We aim to provide a more comprehensive analysis of visceral factors compared to the existing papers: in addition to the stress hormone of cortisol, we also measure sex hormones, including progesterone, estradiol and testosterone. Sex hormones are major contributors to sexual dimorphism, and may also be relevant to competition and tournament entry. Testosterone has long been shown to play an important role in dominance across species and sex (e.g., Archer 2006). More recently, estradiol is shown to play a role in dominance among women (Ziomkiewicz et al. 2015) and in attachment style in mixed-gender sample (Edelstein et al. 2010). Progesterone is observed to be closely related to stress reduction in both women and men (e.g., Wirth 2011, Childs et al. 2010), potentially through its function in social affiliation (Brown et al. 2009). A small number of studies examine the effects of progesterone and estradiol on competitiveness, with more focus on women.[2] The variation in hormonal levels over the menstrual cycle is found to explain the competitiveness of women (Buser 2012, Wozniak et al. 2014). By contrast, Ranehill et al. (2018) examine the effect of oral contraceptives on the economic behavior of women in a randomized control trial and find no significant effect on the willingness to compete.

This paper investigates the effects of hormonal responses. Based on recent evidence in neuroendocrinology, it has been suggested that the dynamic fluctuations of hormones, rather than the baseline hormonal levels, can better predict competitive and aggressive behaviors (Carré et al. 2011). In Experiment 1, two salivary samples are taken in order to measure the relative change in the levels of hormones before and after the first two tasks of the NV design. We find that the arousal of hormones increases the competitiveness of men but not women, which partially explains the observed gender gap in competitiveness. Among the four hormones, the progesterone response is significantly higher for men than for women, and higher progesterone response predicts a higher rate of tournament entry. The regression analysis shows that after controlling the hormonal responses, the gender difference in tournament entry is reduced by 30%, and becomes insignificant. When we control for other behavioral determinants of tournament entry, progesterone response remains positively correlated with competitiveness.

As the observation in Experiment 1 is correlational, there are several possibilities that visceral reaction can be linked to the gender difference in competitiveness. It is possible that more competition-inclined individuals tend to have higher hormonal responses, which means there is limited scope for applying this result to inform the institutional design to narrow the gender gap. It is also possible that stronger visceral influences induced by the first two tasks drive men to become more competitive. This second possibility implies that the gender difference in competitiveness would be sensitive to the timing of the

[2] Some studies explore the role of the menstrual cycle and intake of hormonal contraceptives in determining the competitiveness of bidding behavior in auctions (Buser 2012, Chen et al. 2013, Pearson and Schipper 2013). Schipper (2015) directly measures the baseline hormonal levels in a mixed-gender sample and finds that progesterone level is positively correlated with bids and negatively correlated with profits, but only among the females who do not use hormonal contraceptives.

tasks, since being time-dependent is a defining feature of visceral influences. Evidence that visceral influences can be cooled off by a delayed reaction has been provided in various contexts, including experimental settings such as gift exchange (Gneezy and List 2006), ultimatum game (Gneezy and List 2006), and risk taking (Bendahan et al. 2017), and field settings such as the issuance of a divorce (Lee 2013), and gun homicide (Luca et al. 2017). If the effects of visceral factors are causal, it means that reshaping the timing structure can help reduce the gender difference in tournament entry.

To test this hypothesis, in Experiment 2, we implement a rest of 40 minutes immediately after the first two tasks and before choosing between piece rate and tournament, and compare this "rest" condition to the control condition that strictly follows the NV design. We choose a 40-minute resting period because this is how long it takes for the hormones to return to baseline (e.g., Dickerson and Kemeny 2004). Experiment 2 provides further support that part of the gender difference in tournament entry is attributed to visceral influences. In the control condition we replicate the finding of a significant gender difference as in NV. By contrast, we observe a much smaller and insignificant difference between men and women in the treatment condition. Specifically, the resting period significantly reduces the competitiveness of men but not women, and therefore narrows the gender difference in competitiveness.

We also examine the effects of visceral factors on confidence and risk preferences, as both are important determinants of competitiveness, discussed by the NV and extensively examined in the related literature (e.g., Croson and Gneezy 2009, Gillen et al. 2019). Putting together the findings from both of our experiments, there is weak evidence of correlations between visceral factors with both confidence and risk attitudes in Experiment 1 and a lack of causal influence of the 40-minute break in Experiment 2. These findings indicate that the visceral influences on tournament entry reflect a change in competitive preference, since they are clearly not fully attributed to changes in confidence or risk attitudes.

We find evidence supporting that the competitiveness elicited using the NV design has two constituent parts: one part is driven by the visceral influences therefore is present in the "hot" state, and the other part which persists after cooling off or in the "cold" state. Importantly, the difference between the hot- and cold-state competitiveness, or the competitiveness attributed to visceral influences, is larger for men than for women. An implication of this finding is that the experimental design should be more closely tailored to the labor market phenomenon that the experiment attempts to address. For instance, the size of the gender difference in choosing a competitive workplace may depend on the extent that the decision-maker is subject to visceral influences.

Our study also sheds light on behavioral spillovers when multiple tasks are performed in the experiment using within-subject design. In the setting of strategic interaction, successful coordination on the Pareto optimal equilibrium in the median game affects behavior in the minimum game when games are played sequentially (Cason et al. 2012). Strategies chosen in one game depend on the other game that the subject plays with different opponents, as the subject may develop heuristics that are applied across games (Bednar et al. 2012). When playing a set of multiple repeated games, subjects use strategies in a new game that reflect both behavioral spillover and cognitive load effects (Liu et al. 2019). In the setting of individual choice, it is observed that subjects become more risk-averse after they make choices in an environment with higher risks (He and Hong 2017). Our study suggests that visceral influences may underpin behavioral spillovers, especially when the earlier tasks induce hormonal responses.

The rest of the paper is organized as follows. Section 2 describes the experimental design and procedure. Section 3 presents the experimental findings. Section 4 offers some concluding remarks.

2 Experimental Design

2.1 The Basic Design

Our experimental design builds on NV. Subjects perform several tasks solving addition problems. Each problem is presented on a computer screen with five randomly generated two-digit numbers, and a blank box for the subject to enter the sum of the numbers. Subjects are allowed to use scratch paper but not calculators. Once the subject submits an answer to a problem, a new problem appears on the screen, with the information on whether the previous answer is correct. Throughout the task, a record of the total number of correct answers is displayed on the right corner of the screen. Subjects have five minutes to solve as many problems as they can, under different monetary incentive schemes as described below. The instructions for each task are only distributed immediately before the task is played. At the end of the task, subjects are informed of their own total number of correct answers, but not the performance of others. There is no feedback on how much they earn until the end of the experiment. Subjects are paid based on their earnings in one randomly-drawn task, to ensure that they have appropriate incentives to perform well in each task and make careful choices, and are unable to hedge across tasks.

In Task 1, subjects are given a piece-rate incentive to perform the task. Subjects earn SGD1.00 per correct answer (approximately USD0.75 at the time of the experiment). This task measures the performance under the piece-rate incentive scheme. After the subjects perform five minutes of additions in Task 1, they proceed to Task 2, in which the incentive scheme for performing the task is a tournament.

In Task 2, subjects are randomly and anonymously divided into groups of four. Within each group, the subject who correctly solves the highest number of problems receives SGD4.00 per correct answer, whereas the other three subjects receive no payment. If a tie occurs, a random draw is used to break the tie. Task 2 measures the performance under tournament incentive. The difference in performances in Task 1 and Task 2 captures the difference in motivations offered by the two schemes, and potentially other effects such as learning, since subjects may get more familiar with the layout of the addition problem. Subjects do not have any choice over the compensation scheme in the first two tasks.

In Task 3, subjects are asked to choose whether they want to be paid according to piece rate or tournament for performing another five minutes of addition. If they select piece rate, they receive SGD1.00 per correct answer in their subsequent performance, as in Task 1. If they select tournament, they receive SGD4.00 per correct answer if their Task 3 performance exceeds the performances of all other group members in the previous tournament, Task 2, regardless of the choices of the other group members; otherwise they receive no payment. The design that subjects compete against the past performance of other subjects simplifies the choice problem. For example, any strategic interaction among the subjects is excluded. The choice provides a measure of the willingness to compete for subsequent performance, that is, competitiveness. This choice is the main dependent variable in the analysis. After subjects make their choices, they perform another five minutes of addition before proceeding to Task 4.

In Task 4, subjects are again asked to choose over the compensation schemes, but this time the choice applies to the past performance in Task 1, so they do not perform the task again. The piece-rate scheme again pays SGD1.00 per correct answer for the Task 1 performance. The tournament scheme pays SGD4.00 per correct answer if the subject's Task 1 performance exceeds the Task 1 performances of the other group members, and otherwise zero. The tournament choice for past performance reflects general factors that may be related to competitiveness, such as overconfidence, risk attitudes, and feedback aversion, therefore provides an important control variable for investigating the preference for entering a competition which requires active performance.

After the four tasks, we elicit the beliefs on the relative performance among the group members: for both Task 1 and Task 2, subjects make a guess on their ranking in the group (1 to 4) and are paid SGD2.00 for each correct guess. This provides an incentive-compatible measure of confidence in the performance under both piece-rate and tournament schemes.

We further include measures of risk attitudes using a price-list design. In the first list, subjects choose between an even-chance lottery that pays either SGD10.00 or SGD0.00 on the left, and 10 levels of sure amounts ranging from SGD2.50 to SGD7.00 on the right, arranged in ascending order. One of the 10 choices is randomly selected to be implemented for payment. A typical subject would select the lottery when the sure amount is low, and switch the choice as the sure amount increases. The switching point provides a proxy for the certainty equivalent of the lottery, which measures the risk attitudes. The same applies to a second list, in which subjects select between a longshot lottery with 10% chance of paying SGD10.00 and 90% chance of paying SGD0.00, and 10 levels of sure amounts ranging from SGD0.40 to SGD3.00.

At the end of the experiment, subjects are paid SGD5.00 as a show-up fee, in addition to their earnings in one task among Tasks 1–4, and the payments from the rank guessing and the risk attitude tasks. Our design is essentially similar to NV, and the only difference is the inclusion of risk attitude tasks at the end of the experiment. The experimental instructions are provided in Appendix B.

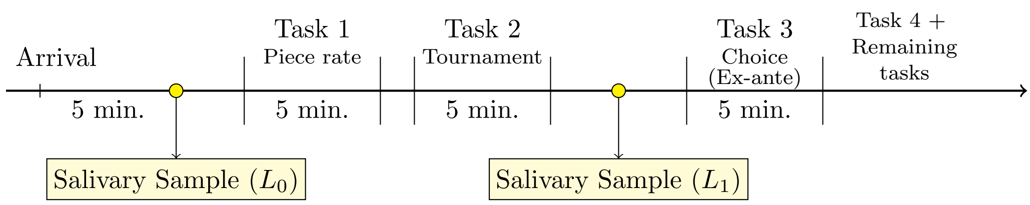
2.2 The Two Experiments

Our two experiments both share the basic design. Experiment 1 examines whether the first two tasks induce changes in visceral factors measured by salivary hormonal levels, and whether those changes predict tournament entry in the third task. We take two salivary samples from the subjects – the timeline is summarized in Figure 1(a). The first sample is collected 5 minutes after subjects arrive at the lab to assess baseline hormone levels (L_0), and the second is taken immediately after Task 2 (L_1). After the second sample, subjects immediately continue with the rest of the tasks until the end of the experiment. Given that the baseline levels tend to vary strongly between individuals, we take the relative change between the two measures ($L_1/L_0 - 1$) as our measure of hormonal response.

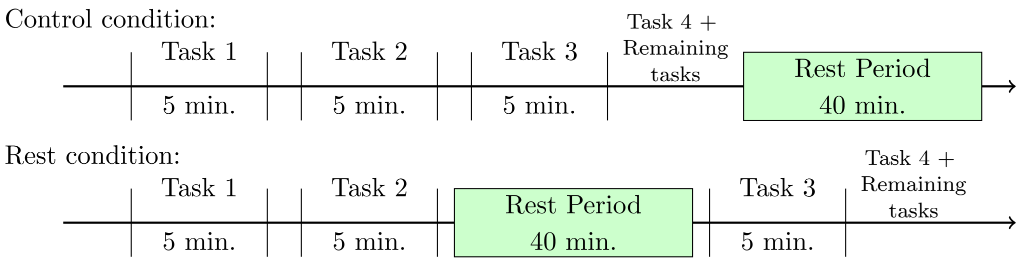
Salivary samples are collected in sterile tubes by passive drool using a short plastic straw for expectorating. Subjects are instructed to rinse the mouth vigorously with

Figure 1: Timelines of the Two Experiments

(a) Experiment 1



(b) Experiment 2



Notes: In Experiment 1, two salivary samples are taken: at the beginning of the experiment and prior to the choice of competitiveness condition. In Experiment 2, a 40-minute rest period is introduced before the choice of competitiveness condition in the treatment condition, and the resting period is placed at the end of the experiment in the control condition.

tap water, wait 2 minutes, and spit the sample into the test tube. When they finish, the experimenters collect the salivary samples from all the subjects in the session. The process takes around 5 minutes. Immediately after collection, samples are refrigerated and stored at −80°C until assayed. Salivary concentrations of cortisol, testosterone, estradiol, and progesterone are measured using a Salimetrics enzyme immunoassay kit. All samples are measured in duplicate, and the mean is obtained as the measure.

Experiment 2 aims to test an implication of the relationship between visceral factors and gender difference in competitiveness. Suppose hormonal responses and competitiveness correlate due to some innate characteristic of an individual, uncovering this relationship is unlikely to inform the institutional design for the purpose of increasing or reducing competitiveness. Instead, if for the same individual competitiveness is affected by the hormonal responses, given that hormonal responses can be cooled off over time, we expect competitiveness to be sensitive to the timing of the tasks.

Experiment 2 consists of a control condition and a treatment condition that differ by the timeline, as summarized in Figure 1(b). The treatment condition is referred to as the "rest" condition: we introduce a 40-minute resting period between Task 2 and Task 3. Subjects are told that there will be a break for 40 minutes before continuing the experiment. During the resting period, subjects sit by themselves quietly without any task to perform. Magazines are provided for light reading, if they choose, but they are not allowed to surf the web, use their mobile phones, or communicate with other subjects in the room. In the control condition, we include a 40-minute rest at the end of the experiment. Our design ensures that the potential learning effects brought by Task 1 and Task 2 are kept constant between the control and the rest conditions: all subjects have performed the task under both the piece-rate and the tournament schemes, before they choose between the two, and would have formed similar expectations regardless of the conditions.

Several design features of the resting period are worth noting. First, as subjects do not know what Task 3 will consist of, it is unlikely that they would be thinking about whether to choose piece rate or tournament during the rest. So if taking a rest affects the choices in the subsequent tasks, it is unlikely to be due to any difference in the time constraint on decision-making. Second, we set the resting period to be 40 minutes for the hormones to return to the baseline level, following the suggestion of the literature (e.g., Dickerson and Kemeny 2004). Third, the overall experimental durations are the same for the control and the rest condition, which allows us to randomize the subjects into the two conditions.

2.3 Implementation

All of the sessions were conducted between 2:00 pm and 3:30 pm on weekdays to avoid any potential time-of-day effect, for instance due to hormones' circadian rhythms. We recruited subjects from the undergraduate student pool of the National University of Singapore, using online advertisements which described the study as a paid decision-making experiment, mentioned the approximate length of study (60 minutes for Experiment 1 and 90 minutes for Experiment 2) but did not provide any detail of the tasks. We sought to test the same number of males and females in each session. The male subjects in our sample were generally two years older than female subjects because of the mandatory military service for men starting at age 18. The sessions were conducted at the laboratory of the Center for Behavioral Economics at the National University of Singapore. The experiment was programmed using z-Tree (Fischbacher 2007). After arriving at the experimental venue, subjects were given the consent forms approved by the Institutional Review Board of the National University of Singapore. Subsequently, general instructions were read aloud to subjects before they began making decisions.

Experiment 1 consists of 7 experimental sessions (the session size varies between 12 and 20). In total, 116 undergraduate students, 58 men and 58 women participated in the experiment. Several subjects had incomplete hormone data due to low saliva volume, resulting in the assay's failure. Three subjects had missing values for cortisol, two for testosterone, five for estradiol and one for progesterone, and the total number of subjects with any of the four hormones missing is 7 (6%).

Experiment 2 consists of 6 experimental sessions (the session size varies between 20 and 28): sessions 1, 3, and 5 are for the control condition and sessions 2, 4, and 6 are for the rest condition. In total, 148 undergraduate students, including 72 men and 76 women were randomly assigned to the control condition (n=76) and the rest condition (n=72). One subject left before the end of the experiment, so we exclude this subject from the analysis. In the risk attitude measure using even-chance lottery price list, two subjects exhibited multiple switching behavior, which is coded as a missing value.

3 Results

3.1 Experiment 1

Table 1 provides the summary statistics, including performances, competitiveness, confidence and risk attitudes. The performance in Task 1 is 2.05 questions higher for men compared to women and the difference is significant ($p = 0.082$).[3] Men also have higher average performance in Task 2 but the gender difference is insignificant (1.53 questions, $p = 0.199$). Between Task 1 and Task 2, the performance increases by 2.90 and 3.41 questions for men and women respectively, and there is no gender difference in the increase ($p = 0.532$). For Task 1, the average guessed rank is 2.46, which is not significantly different from the average of actual ranks, 2.5 ($p = 0.667$), and is similar between men and women (Fisher's exact test, $p = 0.143$). For Task 2, the subjects are significantly overconfident with an average guess of 1.96 ($p < 0.001$), and men are significantly more confident than women (Fisher's exact test, $p = 0.007$). Men are more risk-seeking than women, but the difference is only significant in the even-chance lottery elicitation ($p = 0.012$), not in the longshot lottery elicitation ($p = 0.730$). And there is no gender difference in Task 4 choice for past performance (Fisher's exact test, $p = 1.000$).

We replicate NV's finding that men are more competitive than women. We find that 50% of men and 33% of women choose tournament over piece rate in Task 3, which is a significant difference (Fisher's exact test, $p = 0.089$). Compared to the NV result that 73% of men and 35% of women choose tournament over piece rate (40 men and 40 women,

[3] The p-values are two-sided, and the t-tests are used unless otherwise specified.

Fisher's exact test, $p = 0.002$), we observe a significantly lower tournament entry rate for men (Fisher's exact test, $p = 0.036$). Our results are closer to those in the existing literature using the stylized NV design (Dariel et al. 2017): 56% of the men and 31% women choose tournament over piece rate, averaged across different studies.

Table 1: Summary Statistics for Experiment 1

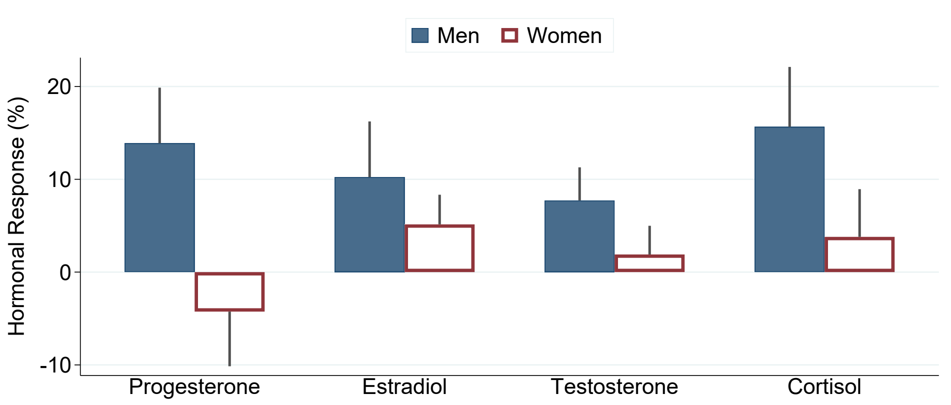
Mean (SD)	Pooled $n = 116$	Men $n = 58$	Women $n = 58$	Gender diff. p-value
Performance: Task 1	12.66	13.69	11.64	0.082
	(6.36)	(6.31)	(6.30)	
Performance: Task 2	15.82	16.59	15.05	0.199
	(6.42)	(6.40)	(6.40)	
Performance: Task 2−Task 1	3.16	2.90	3.41	0.532
	(4.43)	(4.56)	(4.33)	
Competitiveness	0.41	0.50	0.33	0.089
	(0.49)	(0.50)	(0.47)	
Task 4 choice	0.35	0.36	0.34	1.000
	(0.48)	(0.48)	(0.48)	
Guessed rank: Task 1	2.46	2.40	2.52	0.143
	(1.07)	(1.01)	(1.14)	
Guessed rank: Task 2	1.96	1.84	2.07	0.007
	(0.87)	(0.93)	(0.79)	
Risk: even-chance	4.23	4.76	3.71	0.012
	(2.27)	(2.23)	(2.19)	
Risk: longshot	4.00	4.10	3.90	0.730
	(3.21)	(3.28)	(3.16)	
Age	21.56	22.40	20.72	0.000
	(1.72)	(1.54)	(1.47)	

Notes: The p-values are from two-sample t-tests for continuous variables (the performance and risk attitude variables and age) and Fisher's exact tests for binary and discrete variables (competitiveness, guessed ranks and Task 4 choice).

3.1.1 Hormonal responses by gender

Next we examine whether the hormonal responses triggered by the first two tasks differ by gender, and whether such responses are correlated with tournament entry. Compared with women, men have higher baseline levels of testosterone and cortisol ($p < 0.001$ for both), and lower baseline levels of progesterone and estradiol ($p = 0.042$ and 0.029). Figure 2 plots the hormonal responses relative to the baseline level, separately for men and women (detailed statistics provided in Table A1). Significant hormonal responses are observed for men (cortisol, 16%, $p = 0.018$; testosterone, 8%, $p = 0.035$; estradiol, 10%, $p = 0.093$; progesterone, 14%, $p = 0.024$), but not for women (cortisol, 4%, $p = 0.463$; testosterone, 2%, $p = 0.541$; estradiol, 5%, $p = 0.113$; progesterone, −4%, $p = 0.473$). The absence of significant response among women is consistent with the literature on how hormones respond to social stress (Taylor et al. 2000). The progesterone response of men is significantly higher than that of women ($p = 0.033$), and the differences in other hormones are not statistically significant (cortisol, $p = 0.154$; testosterone, $p = 0.220$; estradiol, $p = 0.457$). .

Figure 2: Hormone Response by Gender



Notes: Hormonal response is measured by the change in post-Task 2 measure (L_1) relative to the baseline measure of the hormonal level (L_0), i.e. $(L_1/L_0 - 1) \times 100\%$. The spike indicates the standard error of the mean.

We note that the responses of different hormones tend to be positively correlated, as reported in Table A2. Significant correlations robust to splitting the sample by gender are observed between cortisol and progesterone (Spearman correlation = 0.480, $p < 0.001$), cortisol and testosterone (0.371, $p < 0.001$), and testosterone and progesterone (0.351, $p < 0.001$). The correlation between estradiol and testosterone is also significant (0.232, $p = 0.015$), although it is insignificant if we restrict the sample to men (0.189, $p = 0.167$).

3.1.2 Gender, hormones and competitiveness

To test whether hormonal responses correlate with competitiveness, we report the regression analysis in Table 2. The binary dependent variable of tournament entry is regressed on the gender dummy (which equals one for female subjects) and the hormonal responses,

$$C_i = \beta_0 + \beta_1 Female_i + \beta^P Progesterone_i + \beta^E Estradiol_i + \beta^T Testosterone_i + \beta^C Cortisol_i + \gamma X_i + e_i. \quad (1)$$

Columns (1) and (2) do not include any control, therefore simply summarize how the raw competitiveness differs by gender and changes with the four hormonal responses. Following NV and the subsequent literature, we also investigate the gender difference in the preference for competition, measured by the residual competitiveness after controlling for other factors that may affect tournament entry, reported in columns (3) and (4). The list of control variables (denoted by X_i) includes the absolute performances and the guessed ranks in Tasks 1 and 2, choice in Task 4, risk attitudes, age, year of study, and session fixed effects. Throughout the paper we estimate linear probability models using Ordinary Least Squares (OLS) instead of Probit, as a precaution of potential measurement error in competitiveness (Hausman 2001), and report the robust standard errors. For the ease of interpretation, the hormonal responses are standardized in the pooled sample (i.e., subtracted by the sample mean and divided the sample standard deviation), so that the coefficient indicates the change in the probability of tournament entry associated with one standard deviation difference in the hormonal response. The seven subjects with missing values on any of the hormones are not included in the estimates.

Comparing column (1) and column (2), the raw gender gap ($\hat{\beta}_1$) falls from 15.8% to 12.0% after the hormonal responses are controlled for, and becomes insignificantly different from zero (p =0.093 in column (1) and 0.206 in column(2)). This suggests that hormonal responses can explain around 30% of the gender gap in raw competitiveness. The effect of progesterone response ($\hat{\beta}^P$) is large and significant: for a subject with one standard deviation higher progesterone response, the likelihood of tournament entry increases by 12.1% ($p = 0.020$). Column (3) shows that the behavioral control variables explain a significant proportion of gender difference in our data, as the point estimate of gender gap falls to 7.6%, and becomes insignificant ($p = 0.516$).[4] In column (4) which includes hormonal responses and behavioral controls, the point estimate of gender difference further falls to zero ($p = 0.956$), and the effect of progesterone response remains significant: one standard deviation increase in progesterone response is estimated to increase the likelihood of tournament entry by 9.5% ($p = 0.047$). The control variables mediate the size of $\hat{\beta}^P$, which indicates that progesterone response may also correlate with other behavioral variables such as confidence and risk attitudes, and we examine this in Section 3.1.3.

The result that competitiveness increases with progesterone response is robust to a variety of specifications. Adding the additional controls of the baseline hormonal levels gives $\hat{\beta}^P = 0.102$ ($p = 0.049$), and the baseline levels are all insignificant (joint $F = 1.37$, $p = 0.252$). We also apply the Obviously Related Instrumental Variables (ORIV) method by Gillen et al. (2019) to our duplicate elicitations of risk attitudes. The ORIV estimator gives $\hat{\beta}^P = 0.105$ ($p = 0.021$); the two risk attitude measures have a Spearman correlation of 0.433 ($p < 0.001$) and the instruments have a robust $F = 13.924$ in the first stage. The effect of progesterone is also robust to dropping the demographics ($\hat{\beta}^P = 0.092$, $p = 0.043$). As noted in Section 2.1, the Task 4 choice provides an important control variable

[4] According to a summary of 26 studies using the NV design (van Veldhuizen 2018), while a sizable proportion of the gender gap in tournament entry is attributed to pure competitiveness (on average 69%, close to our finding), the pure competitiveness is not statistically significant in 11 of those studies.

to investigate the competitive preference; some of the behavioral variables captured by Task 4 choice are also measured and directly controlled for in our main specification. The estimated effect of progesterone is very similar if we drop the Task 4 choice ($\hat{\beta}^P = 0.097$, $p = 0.034$), or if we drop the other behavioral variables while keeping the Task 4 choice as the only behavioral control ($\hat{\beta}^P = 0.098$, $p = 0.051$). Following the NV approach, we also check and find that progesterone response has a smaller and insignificant effect on the Task 4 choice itself, controlling the other variables ($\hat{\beta}^P = 0.078$, $p = 0.139$). This suggests that progesterone response has a positive effect on the preference for competition, and not merely the other behavioral determinants of tournament entry.

Table 2: Do Hormonal Responses Explain Gender Difference in Competitiveness

	(1)	(2)	(3)	(4)
Dependent variable: Competitiveness				
β_1 Female	-0.158^{*}	-0.120	-0.076	-0.007
	(0.093)	(0.094)	(0.116)	(0.127)
β^P Progesterone		0.121^{**}		0.095^{**}
		(0.051)		(0.047)
β^E Estradiol		-0.022		-0.047
		(0.046)		(0.045)
β^T Testosterone		0.037		0.018
		(0.039)		(0.036)
β^C Cortisol		-0.018		-0.052
		(0.051)		(0.042)
Controls	No	No	Yes	Yes
Observations	109	109	109	109
R-squared	0.026	0.090	0.408	0.435

Notes: Seven subjects with incomplete data on hormonal responses are excluded from the sample. The dependent variable is 1 if the subject chooses the tournament and 0 otherwise. The control variables include performances and guessed ranks in Tasks 1 and 2, choice in Task 4, two measures of risk attitudes, age, year of study, whether the subject is in the final year, and session fixed effects. Robust standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

The specification of Equation (1) reported in Table 2 assumes that the hormonal responses affect the competitiveness of men and women in the same way. This may not be the case – for instance, Buser, Dreber and Mollerstrom (2017) find that cortisol response does not predict the tournament entry of men, whereas it is significantly positively correlated with the tournament entry of women. And we also evaluate the effect of each hormone separately, to check whether results are affected by the correlations between different hormonal responses. So we independently estimate a model for each hormone, which interacts the hormonal response with the gender dummy,

$$C_i = \beta_0 + \beta_1 Female_i + \beta_2 HR_i + \beta_3 Female_i \times HR_i + \gamma X_i + e_i. \tag{2}$$

The effect of the hormonal response for men is β_2, and that for women is $\beta_2 + \beta_3$, both are plotted in Figure 3. The p-value reported above the two coefficients tests whether β_3 is equal to zero, i.e., whether the effect of the hormone is the same for men and women. Panel (a) does not include any controls, therefore shows how each hormonal response correlates with the raw tournament entry decision. Panel (b) includes all the behavioral and demographic controls aforementioned, to investigate whether each hormone affects the preference for competition.

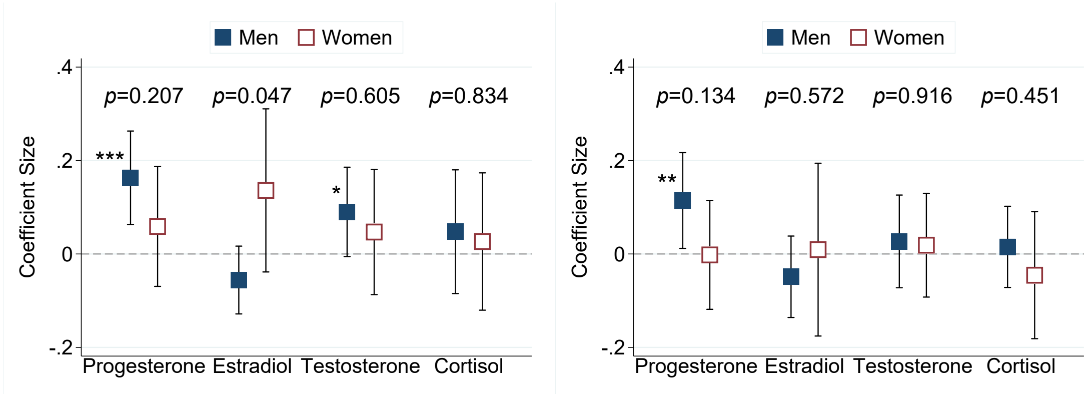
We do not observe much evidence that the hormones affect men and women differently. Panel (a) suggests that a stronger hormonal response tends to increase tournament entry for both men and women. Statistically significant effects are only observed for men: one standard deviation higher progesterone response is estimated to increase tournament entry by 16.3% ($p = 0.002$), one standard deviation higher testosterone response increases tournament entry by 9.0% ($p = 0.065$). The effects on women are positive, smaller and not significantly different from zero ($p = 0.365$ for progesterone and 0.489 for testosterone). Estradiol response has a positive effect on women's tournament entry, despite its negative effect on men; while the difference is significant ($p = 0.047$), the effects on both genders are insignificant ($p = 0.131$ and 0.125). Cortisol response also has no significant effect on men or women ($p = 0.477$ and 0.720).

Panel (b) shows the significant positive correlation between progesterone response and competitive preference reported in column (4) of Table 2 is mainly driven by men: an increase of progesterone response by one standard deviation increases the competitiveness of men by 11.4% ($p = 0.029$). The point estimate for women is zero ($p = 0.972$), although the gender difference is again insignificant ($p = 0.134$). This suggests that either progesterone response is ineffective on women, or perhaps progesterone response needs to reach a threshold to become effective while women's progesterone responses are too small. The effects of the other three hormones are insignificant for both men and women and the point estimates are smaller compared to without the controls (the smallest $p = 0.269$). We also examine the correlation between hormonal responses and the choice in Task 4, and we do not observe a significant correlation for either men or women, as reported in Table A3.

We also analyze the potential effects of hormonal responses on the quality of choice between piece rate and tournament in Task 3. Following NV, for simplicity we assume the subjects are risk-neutral and have fixed effort costs. Tournament entry in Task 3 is optimal, if a subject has a probability of winning which is higher than 0.25 (i.e., the

Figure 3: Do the Effects of Hormone Responses on Competitiveness Differ by Gender

(a) Estimates without controls (b) Estimates with controls



Notes: The dependent variable is tournament entry; the effect of each hormone is estimated separately. Panel (a) estimates the effect on raw competitiveness, and panel (b) on competitive preferences. The marker indicates the estimated effect of one standard deviation increase in hormonal response on competitiveness. The brackets indicate the confidence interval and the stars indicate the significance level of the coefficient. The p-value reported above the coefficients for men and women tests whether they are significantly different from each other.

ratio between the piece rate and tournament winner's rate).[5] We investigate whether hormonal responses correlate with two types of suboptimal choices, namely an over-entry by someone with a low probability of winning, and an under-entry by someone with a high probability of winning. This analysis is potentially subject to power issue: NV only provide summary statistics without drawing any statistical inferences (Table III, p. 1085). The tentative findings from our data are reported in Table A4: progesterone response increases over-entry whereas estradiol and cortisol responses reduce over-entry, and none of the hormones has a significant effect on under-entry.

3.1.3 Correlations with other behavioral variables

Table 2 and Figure 3 suggest that part of the effect of hormonal responses kicks in through changing confidence and risk preference: the estimated effects of hormones on competitiveness reduce when we include the behavioral controls. To investigate the potential effects of hormonal responses on confidence and risk preference, we estimate Equation (2) with the dependent variables of the guessed ranks and risk attitudes in lieu of tournament entry. There is a caveat of this analysis: the hormonal responses are measured before Task 3, whereas confidence and risk attitudes are elicited after Task 3. It

[5]We use the subject's Task 2 performance to estimate the winning probability. Alternatively, the Task 3 performance can be used to calculate an ex-post winning probability. However, subjects tend to have a higher motivation in tournament than in piece rate (as shown in Zhong et al. 2018, controlling for order effects), which means the choice of piece rate in Task 3 can induce a lower subsequent performance. Therefore the ex-post winning probability may underestimate the true winning potential for those who choose piece rate.

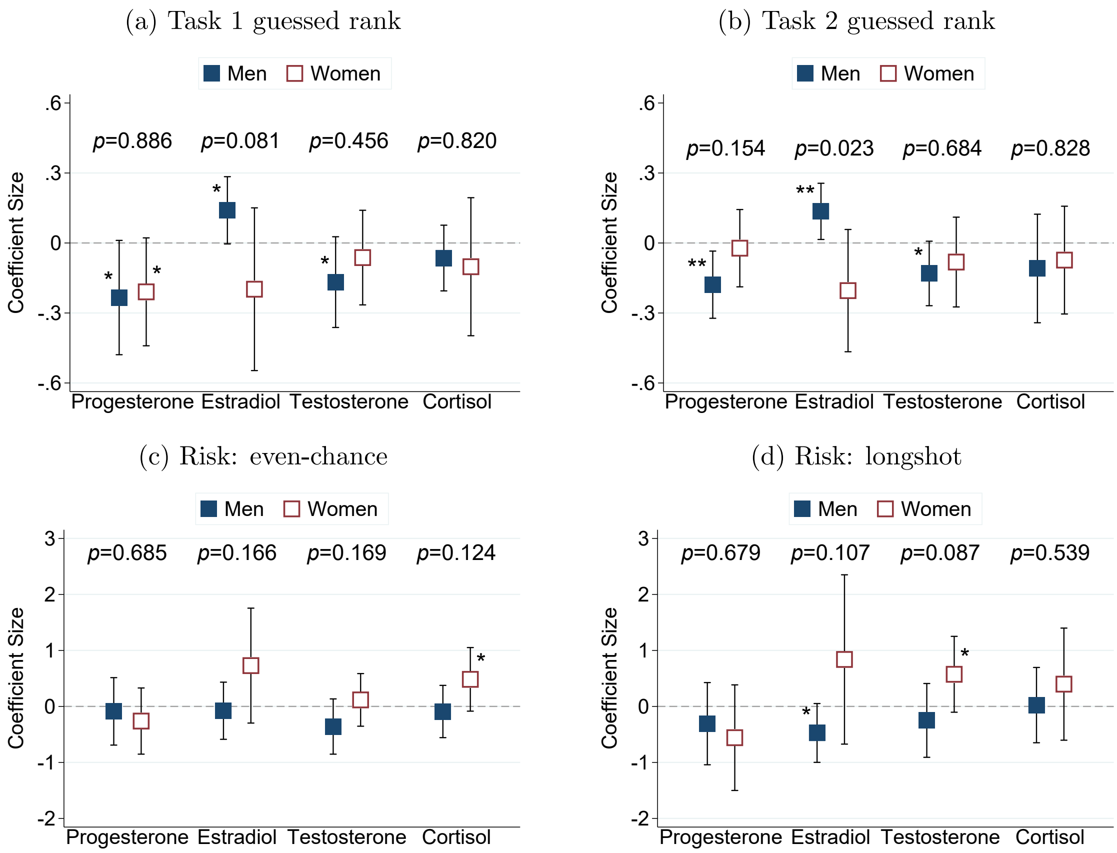
is possible that, since the visceral influences are time-sensitive, by the time that subjects make the decisions which elicit their confidence and risk preference, the visceral state is different from the one before they make the Task 3 choice. Also, the choice and performance in Task 3 can create additional visceral responses that we do not observe.

In analyzing the guessed ranks, we control for the absolute performance, in order to address whether hormones explain over-confidence or under-confidence exhibited in the guesses. The performances in Task 1 and Task 2 are positively correlated with the hormonal responses: cortisol is significantly correlated with the performance in both tasks, and progesterone is significant for Task 2 performance; the correlations are mainly driven by women (full correlations reported in Table A5). The effects of hormones on confidence by task are reported in panel (a) and (b) of Figure 4. Subjects with higher hormonal responses tend to be more confident: the effects of progesterone and testosterone are significant for men in both tasks (progesterone: $p = 0.061$ in Task 1 and 0.016 in Task 2; testosterone: $p = 0.090$ in Task 1 and 0.064 in Task 2), and progesterone is also significant for women in Task 1 ($p = 0.075$). The only exception is that men with higher estradiol responses are less confident ($p = 0.057$ in Task 1 and 0.028 in Task 2).

In terms of risk attitudes, the estimated effects of hormones reported in panel (c) and (d) of Figure 4 are small. Although the results are qualitatively similar between our two measures, significant effects are never observed for both measures. For women, cortisol response is positively correlated with risk-seeking choices in the even-chance lottery ($p = 0.094$), testosterone response is positively correlated with risk-seeking in the longshot lottery ($p = 0.096$); for men, estradiol is negatively correlated with risk-seeking in the longshot lottery ($p = 0.077$). The findings that estradiol response mildly reduces confidence and risk-seeking choices of men may help explain its negative but insignificant effect on their tournament entry.

To summarize, Experiment 1 investigates the potential visceral influences embedded in the NV design by investigating both stress and sex hormones. We observe that visceral responses after the first two tasks are stronger for men: for each of the four hormones measured, the hormonal level significantly increases among men but not among women. We find that the progesterone response positively correlates with competitiveness, and can explain a part of the gender gap in competitiveness – the result is present after controlling other behavioral variables, suggesting that visceral influences may have a direct effect on competitive preference.

Figure 4: Correlations between Hormonal Responses and Confidence / Risk Attitudes



Notes: In panels (a) and (b), the dependent variables are the guessed ranks for performance in Task 1 and Task 2 (taking the value of 1–4, a lower value indicates higher confidence), and we control for the actual performance. In panels (c) and (d) the dependent variables are the risk attitudes measured with even-chance and longshot lotteries (taking the value of 0–10, a higher value indicates more risk-seeking). The marker indicates the estimated effect of one standard deviation increase in hormonal response. The brackets indicate the 95% confidence interval and the stars indicate the significance level of the coefficient. The p-value reported above the coefficients for men and women tests whether the hormone has a different impact by gender.

3.2 Experiment 2

Table 3 provides summary statistics by gender, for both control and rest conditions in Experiment 2. There are three sets of p-values, the first two compare men and women within the same condition, and the third one (in the leftmost column) compares between the control and the rest condition, pooling the genders. The differences between the two conditions are small for all characteristics and performance variables, and is insignificant apart from the difference in age ($p = 0.075$). These suggest that the randomization between the two conditions is largely valid.

We move on to investigate the gender difference within each condition. We observe that in the control, men and women have very similar performance ($p = 0.561$ for Task 1 and $p = 0.766$ for Task 2) whereas in the rest condition, men happen to have higher performances than women in both tasks ($p = 0.017$ and $p = 0.002$, respectively). We note that these differences in relative performance can introduce noise when we compare the gender gap in competitiveness between the control and the rest condition, so that the performances need to be controlled for in the analysis. The behavioral variables elicited after the rest are potentially subject to the treatment effects. In the control, men are more confident than women in their guessed ranks for Task 2 (Fisher's exact test, $p = 0.022$), and are more risk-seeking according to their longshot lottery choices ($p = 0.046$). In the rest condition, there is a significantly higher proportion of men choosing tournament in Task 4 than women (Fisher's exact test, $p = 0.035$), perhaps reflecting the difference in performance, and men are more risk-seeking in the even-chance lottery choices ($p = 0.003$).

Table 3: Summary Statistics for Experiment 2

	Control Condition				"Rest" Condition				
Mean (SD)	Pooled $n = 76$	Men $n = 36$	Women $n = 40$	p-value (1)	Pooled $n = 71$	Men $n = 35$	Women $n = 36$	p-value (2)	p-value (3)
Performance: Task 1	13.11	13.53	12.72	0.561	14.37	16.69	12.11	0.017	0.283
	(5.95)	(7.67)	(3.88)		(8.14)	(10.12)	(4.74)		
Performance: Task 2	16.41	16.64	16.20	0.766	17.31	20.60	14.11	0.002	0.485
	(6.36)	(7.43)	(5.30)		(9.09)	(11.42)	(4.17)		
Performance: T2-T1	3.30	3.11	3.48	0.677	2.94	3.91	2.00	0.072	0.599
	(3.76)	(4.37)	(3.16)		(4.49)	(5.11)	(3.63)		
Competitiveness	0.39	0.56	0.25	0.010	0.25	0.31	0.19	0.285	0.080
	(0.49)	(0.50)	(0.44)		(0.44)	(0.47)	(0.40)		
Task 4 choice	0.24	0.31	0.17	0.280	0.18	0.29	0.08	0.035	0.544
	(0.43)	(0.47)	(0.38)		(0.39)	(0.46)	(0.28)		
Guessed rank: Task 1	2.57	2.47	2.65	0.166	2.42	2.20	2.64	0.295	0.543
	(1.01)	(1.13)	(0.89)		(1.09)	(1.08)	(1.07)		
Guessed rank: Task 2	2.13	1.97	2.27	0.022	2.07	1.86	2.28	0.262	0.984
	(0.97)	(1.08)	(0.85)		(0.98)	(0.94)	(0.97)		
Risk: even-chance[a]	3.89	3.78	4.00	0.689	4.06	4.82	3.31	0.003	0.667
	(2.39)	(2.09)	(2.66)		(2.15)	(1.71)	(2.29)		
Risk: longshot	3.74	4.47	3.08	0.046	4.15	4.14	4.17	0.974	0.411
	(3.06)	(3.31)	(2.68)		(3.08)	(3.18)	(3.03)		
Age	22.33	22.75	21.95	0.054	21.82	22.40	21.25	0.002	0.075
	(1.81)	(1.87)	(1.69)		(1.63)	(1.46)	(1.61)		

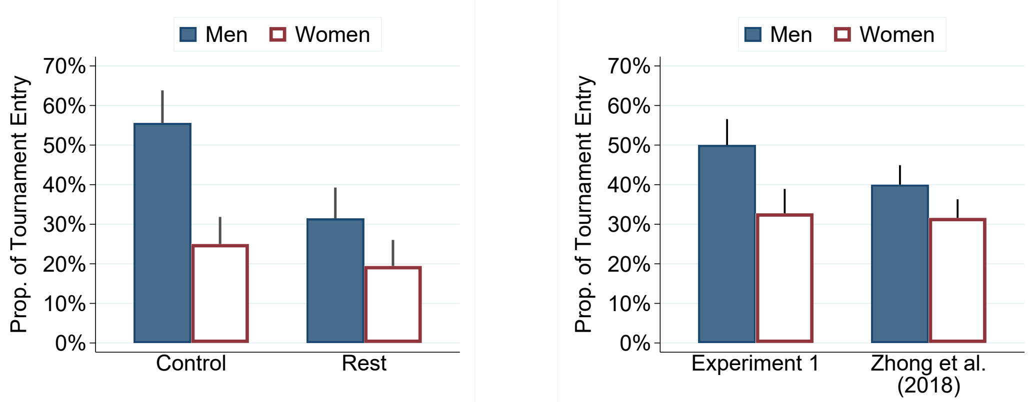
Notes: [a]One man and one woman, both in the rest condition, have missing values due to multiple switching behavior. The p-values in column (1) and (2) test for gender difference within the control and the rest conditions; column (3) provide p-values testing whether the variable differs by experimental condition. The p-values are from two-sample t-tests for continuous variables (the performance and risk attitude variables and age) and Fisher's exact tests for binary and discrete variables (competitiveness, guessed ranks and Task 4 choice).

3.2.1 Effects of taking a rest on competitiveness

Figure 5(a) displays the proportion of tournament choice separately by gender and condition. In the control condition, we find that more than twice as many men as women choose tournament over piece rate (56% versus 25%), which is a significant difference (Fisher's exact test, $p = 0.010$). In the rest condition, men are still slightly more likely to choose tournament than women (31% versus 19%), but the difference is insignificant (Fisher's exact test, $p = 0.285$). Note that the higher relative performance of men in the rest condition should predict a higher rate of tournament entry among men, thus undermines the effect of the rest condition if it reduces the competitiveness of men. Comparing within gender, we find that men are less competitive in the treatment condition than in the control condition (Fisher's exact test, $p = 0.056$), while women do not differ significantly between the two conditions (Fisher's exact test, $p = 0.594$). These results suggest that the resting period cools off the gender difference in competitiveness, by reducing men's willingness to compete.

Figure 5: Competitiveness by Gender and Condition

(a) Experiment 2: Control vs. Rest (b) Experiment 1 vs. Zhong et al. (2018)



Notes: The spike indicates the standard error of the proportion. The control condition strictly follows the NV design, while the rest condition has a 40-minute break. Experiment 1 takes around 5 minutes to collect a salivary sample after Task 2, which is the only interruption. In Zhong et al. (2018) there is a 40-minute break after both Task 1 and Task 2.

We can also examine the cooling-off effect by comparing the results of Experiment 1 and those in Zhong et al. (2018), plotted in Figure 5(b). Both studies departure from the design of NV by taking salivary samples to measure hormonal responses, and both studies use the same subject pool of the undergraduate students from the National University of Singapore. Zhong et al. (2018) introduce a 40-minute break after Task 1 and a 40-minute break after Task 2 to measure cortisol response induced by each of the first two tasks. We do not explicitly introduce any breaks in Experiment 1, except that it takes about 5 minutes to collect the salivary samples from all subjects in each session. Naturally, this comparison should be interpreted with caution, since the subjects are not randomly assigned into the two experiments.[6] There is no significant gender difference in Zhong

[6]The duration of Zhong et al. (2018) is about an hour longer than Experiment 1 and this is compensated by higher payments. Also, the two experiments are conducted at different points of time.

et al. (2018): 40% of men and 32% of women choose tournament over piece rate (99 men and 98 women, Fisher's exact test, $p = 0.236$). By contrast, in Experiment 1 we observe a tournament-entry rate of 50% for men and 33% for women ($p = 0.089$). This supports the observation in Experiment 2 that the resting period leads to a reduction of the gender difference in competitiveness.

Table 4 reports regression analyses to examine whether a significant gender difference is observed in tournament entry in the control and the rest conditions. Again we first present the gender difference in raw competitiveness, then the gender difference controlling for the aforementioned behavioral factors. The two subjects with missing values for the risk attitude measure are excluded. For the control condition, column (1) shows that there is a significant gender difference in raw competitiveness: women are 30.6% less likely to compete than men ($p = 0.006$). After controlling for the behavioral factors in column (2), the coefficient remains significant but drops to 20.2% ($p = 0.043$). By contrast, column (3) shows that in the rest condition, gender is no longer significantly correlated with the raw competitiveness and the point estimate drops to 12.4% ($p = 0.250$). With the inclusion of the control variables in column (4), the gender difference changes direction and remains small and insignificant (3.9%, $p = 0.738$).

Table 4: Treatment Effect on Competitiveness by Gender

	Control		Treatment		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Competitiveness						
δ_1 Female	-0.306^{***}	-0.202^{**}	-0.124	0.039	-0.306^{***}	-0.213^{**}
	(0.109)	(0.098)	(0.106)	(0.115)	(0.109)	(0.095)
δ_2 Rest					-0.232^{**}	-0.262^{*}
					(0.117)	(0.153)
δ_3 Female × Rest					0.182	0.287^{**}
					(0.152)	(0.143)
Controls	No	Yes	No	Yes	No	Yes
Observations	76	76	69	69	145	145
R-squared	0.097	0.505	0.020	0.354	0.083	0.413

Notes: Two subjects have missing values for the even-chance risk attitude task. The control variables include performances and guessed ranks in Tasks 1 and 2, choice in Task 4, two measures of risk attitudes, age, and session fixed effects (the only difference from Experiment 1 is that we did not collect information on year of study in Experiment 2). Robust standard errors are reported in parentheses.

Subsequently, we compare the two conditions in columns (5) and (6) using pooled estimates, regressing competitiveness on gender, the rest condition, and the interaction between gender and the rest condition:

$$C_i = \delta_0 + \delta_1 Female_i + \delta_2 Rest_i + \delta_3 Female_i \times Rest_i + \gamma X_i + \epsilon_i. \quad (3)$$

The coefficient of the interaction, δ_3, is the key parameter which tells us whether the making decision after the resting period significantly reduces the gender gap, by affecting men and women differently. Without including any controls, column (5) shows that women in the control condition ($\hat{\delta}_1$) are 30.6% less likely to choose tournament than men ($p = 0.006$). The resting period is estimated to reduce the willingness to compete of men by 23.2% ($\hat{\delta}_2 = -0.232$, $p = 0.049$) whereas its effect on women is much smaller: resting only reduces women's competitiveness by 5.0% and the effect is insignificant ($\hat{\delta}_2 + \hat{\delta}_3 = -0.050$, $p = 0.609$). The gender difference in response to the resting period, although sizable, is not statistically significant ($\hat{\delta}_3 = 0.182$, $p = 0.234$). With all the control variables in column (6), the coefficients remain significant for both the gender dummy ($\hat{\delta}_1 = -0.213$, $p = 0.027$) and the rest condition dummy ($\hat{\delta}_2 = -0.262$, $p = 0.090$) , and notably, the interaction term is estimated to be larger, and becomes significantly positive ($\hat{\delta}_3 = 0.287$, $p = 0.046$). Overall, these analyses suggest that the resting period significantly reduces the competitive preference of men but does not affect women, making the decisions of men more similar to those of women.

For robustness we also estimate Equation (3) using the ORIV method, which gives $\hat{\delta}_3 = 0.262$ ($p = 0.087$). However, the method may not be appropriate here: the Spearman correlation between our two risk attitude measures is only 0.252 ($p = 0.002$), much smaller than in Experiment 1; consequently in the first-stage estimation the risk attitude measures only serve as very weak instruments for each other (robust $F = 0.732$). We also obtain results similar to the main specification when we drop age ($\hat{\delta}_3 = 0.283$, $p = 0.044$), drop the Task 4 choice ($\hat{\delta}_3 = 0.268$, $p = 0.090$), or keep the Task 4 choice as the only behavioral control ($\hat{\delta}_3 = 0.232$, $p = 0.072$). Therefore the result that the gender difference in competitiveness is reduced by the resting period is qualitatively robust.

Table A6 reports regression analyses to examine whether a significant gender difference is observed in tournament choice in Task 4. We find that neither the gender coefficient nor the treatment coefficient is significant. The interaction term between gender and treatment has a coefficient of -0.078 without control ($p = 0.566$) and -0.035 with control ($p = 0.807$). The lack of significance of the interaction term on Task 4 choice confirms that the resting period has some direct effect on the competitive preference.

We also analyze the welfare effects of the reduced tournament entry, and find some evidence that the rest condition reduces the rate of over-entry among low-performing men, while there is no evidence that it affects the under-entry of the high-performing men, or the over-entry and under-entry of women, as reported in Table A7.

3.2.2 Effects on other behavioral variables

The changes in risk and confidence variables are unlikely to explain a large part of the treatment effect on competitiveness in Experiment 2, which is estimated to be stronger when the behavioral control variables are included (column (6) compared to column (5) in Table 4). Since Experiment 1 provides some evidence that confidence and risk preference are correlated with the hormonal responses, we also test whether these variables are affected by the resting period. We estimate regressions analogous to Equation (3), taking guessed ranks and risk preferences as the dependent variables, and report the results in Table 5. Columns (1) and (2) suggest that, controlling for absolute performance, the guessed ranks for both task 1 and task 2 are not significantly affected by the rest for both men and women (smallest $p = 0.440$). In terms of risk attitudes, the gender-specific treatment effects are in opposite directions for the two measures. According to the even-chance lottery measure analyzed in column (3), men and women make similar choices in the control ($p = 0.685$), but women are significantly less risk-seeking than men in the rest condition ($p = 0.002$). According to the longshot lottery measure analyzed in column (4), women are significantly less risk-seeking than men in the control ($p = 0.047$), but women are similar to men in the rest condition ($p = 0.822$).

The strong similarity in the guessed ranks in the control and the rest conditions suggests that the relationship between hormonal responses and confidence documented in Experiment 1 (panel (a) and (b) of Figure 4) more likely reflects some innate characteristic of the subjects. The treatment effect on risk attitudes is far from clear-cut: although the correlations observed in Experiment 1 (panel (c) and (d) of Figure 4) are qualitative similar between the two measures, the treatment effects observed in Experiment 2 are sensitive to the measure. As noted in Section 3.1.3, since confidence and risk preference measures are elicited after Task 3, these variables are potentially subject to the visceral response induced by Task 3, which brings a confound to this analysis.

In summary, Experiment 2 shows that a resting period reduces the tournament entry of men, and therefore effectively alleviates the gender difference, making it insignificant. Visceral influences can be present in real-life economic contexts, suggesting that a delay of reaction which cools off the visceral influences can be a cost-effective intervention for reducing the gender difference in competitiveness.

Table 5: Treatment Effects on Confidence and Risk Attitudes by Gender

Dependent variable:	(1) Task 1 guessed rank	(2) Task 2 guessed rank	(3) Risk attitude: even-chance	(4) Risk attitude: longshot
Female	0.111	0.275	0.222	-1.397**
	(0.193)	(0.192)	(0.546)	(0.696)
Rest	-0.019	0.161	1.046**	-0.413
	(0.225)	(0.208)	(0.455)	(0.776)
Female × Rest	0.006	-0.275	-1.731**	1.567
	(0.291)	(0.280)	(0.730)	(1.024)
Performance: Task 1	-0.083***			
	(0.015)			
Performance: Task 2		-0.062***		
		(0.013)		
Observations	145	145	145	145
R-squared	0.332	0.263	0.055	0.032

Notes: Two subjects have missing values for the even-chance risk attitude task therefore are dropped from the regression analysis.

4 Concluding Remarks

This study investigates the visceral influences on competitiveness. Our two-experiment design is related to the two-stage assessment of the role of hormones: eliciting the correlation between endogenous hormone release and behavior in the first stage, and testing the effect of exogenous manipulation of the hormone on the behavior in the second stage. Yet our approach is different, to some extent more exploratory, as we are mainly interested in assessing the overall visceral influences rather than the effect of specific hormones. In Experiment 1, we demonstrate that hormonal responses account for at least 30% of the gender difference in competitiveness. Since the visceral factors at work are not restricted to the hormones we measure, it is hard to manipulate them as a whole. As such we test the cooling-off effect on behavior by manipulating the timing structure of the tasks in Experiment 2. We find that the introduction of a resting period reduces the gender difference in competitiveness through reducing the competitiveness of men. Taken together, the results consistently suggest that the first two tasks of piece rate and tournament in the NV design induce changes in visceral factors, which drive the immediate attitude towards competition especially for men.

It is possible that visceral influences are moderated or changed by factors underpinning the gender difference in competitiveness. A large-scale field experiment by Flory et al. (2015) shows that, consistent with the experimental literature, male job-seekers are more likely to apply for jobs with a competitive compensation scheme than females.

Importantly, this gender gap is significantly attenuated by mentioning teamwork and avoiding masculine connotations in the advertisement, and is much smaller among older job-seekers. It has been suggested that the sharp changes in hormone levels at adolescence and menopause explain the changes in competitiveness among women (Flory et al. 2018). Experimental evidence also suggests that culture (Gneezy et al. 2009, Zhang 2018, Booth et al. 2019), gender composition of the environment (Booth and Nolen 2012), team competition (Dargnies 2012, Healy and Pate 2011), task stereotype (Dreber et al. 2014, Shurchkov 2012) and the beneficiary of the reward (Cassar et al. 2016) significantly affect the gender difference in competitiveness. It has not escaped our notice that these factors may transform the underlying visceral influences. For example, working in a team can generate a sense of social support for the decision-makers to buffer pressure and stress in a competitive situation, and math tasks may create a stronger stress response compared to verbal tasks. It will be useful to systematically examine whether visceral influences indeed foster these phenomena.

Our study also contributes to the literature on institutional design aiming to promote gender equality in competitive settings. For instance, affirmative action policies such as quotas and a head start for women (Balafoutas and Sutter 2012, Niederle et al. 2013), and providing feedback on relative performance (Wozniak et al. 2014) are shown to reduce the gender difference in competitiveness, mainly by increasing the tournament entry of high-performing women. Sponsorship in the form of having someone who provides a vote of confidence and shares the success, on the other hand, does not reduce the gender difference, because it increases the competitiveness of men but has little effect on women (Baldiga and Coffman 2018). We show that cooling off the visceral influences using a resting period could be a simple method to narrow the gender difference in competitiveness, with the limitation that it does not encourage the tournament entry of women but rather reduces the tournament entry of men. Given the potential negative impacts of excessive competitive pressure on well-being, we believe that lowering the aggregate level of competitiveness to achieve a gender-balanced outcome may be socially desirable in some settings.

References

Altonji, J. G. and Blank, R. M. (1999), 'Race and gender in the labor market', *Handbook of Labor Economics* **3**, 3143–3259.

Apicella, C. L., Dreber, A., Gray, P. B., Hoffman, M., Little, A. C. and Campbell, B. C. (2011), 'Androgens and competitiveness in men', *Journal of Neuroscience, Psychology, and Economics* **4**(1), 54–62.

Archer, J. (2006), 'Testosterone and human aggression: An evaluation of the challenge hypothesis', *Neuroscience and Biobehavioral Reviews* **30**(3), 319–345.

Ariely, D. and Loewenstein, G. (2006), 'The heat of the moment: The effect of sexual arousal on sexual decision making', *Journal of Behavioral Decision Making* **19**(2), 87–98.

Augenblick, N. and Rabin, M. (2019), 'An Experiment on Time Preference and Misprediction in Unpleasant Tasks', *Review of Economic Studies* **86**(3), 941–975.

Badger, G. J., Bickel, W. K., Giordano, L. A., Jacobs, E. A., Loewenstein, G. and Marsch, L. (2007), 'Altered states: The impact of immediate craving on the valuation of current and future opioids', *Journal of Health Economics* **26**, 865–876.

Balafoutas, L. and Sutter, M. (2012), 'Affirmative action policies promote women and do not harm efficiency in the laboratory', *Science* **355**(6068), 579–582.

Baldiga, N. R. and Coffman, K. B. (2018), 'Laboratory evidence on the effects of sponsorship on the competitive preferences of men and women', *Management Science* **64**(2), 888–901.

Bednar, J., Chen, Y., Liu, T. X. and Page, S. (2012), 'Behavioral spillovers and cognitive load in multiple games: An experimental study', *Games and Economic Behavior* **74**(1), 12–31.

Bendahan, S., Goette, L., Thoresen, J., Loued-Khenissi, L., Hollis, F. and Sandi, C. (2017), 'Acute stress alters individual risk taking in a time-dependent manner and leads to anti-social risk', *European Journal of Neuroscience* **45**(7), 877–885.

Berge, L. I. O., Bjorvatn, K., Pires, A. J. G. and Tungodden, B. (2015), 'Competitive in the lab, successful in the field?', *Journal of Economic Behavior & Organization* **118**, 303–317.

Bertrand, M. (2011), *New perspectives on gender*, Handbook of Labor Economics, volume 4b, chapter 17, pp. 1543–1590.

Booth, A., Fan, E., Meng, X. and Zhang, D. (2019), 'Gender differences in willingness to compete: The role of culture and institutions', *Economic Journal* **129**(618), 734–764.

Booth, A. and Nolen, P. (2012), 'Choosing to compete: How different are girls and boys?', *Journal of Economic Behavior and Organization* **81**(2), 542–555.

Brown, S. L., Fredrickson, B. L., Wirth, M. M., Poulin, M. J., Meier, E. A., Heaphy, E. D., Cohen, M. D. and Schultheiss, O. C. (2009), 'Social closeness increases salivary progesterone in humans', *Hormones and Behavior* **56**(1), 108–111.

Buckert, M., Schwieren, C., Kudielka, B. M. and Fiebach, C. J. (2017), 'How stressful are economic competitions in the lab? An investigation with physiological measures', *Journal of Economic Psychology* **62**, 231–245.

Buser, T. (2012), 'The impact of the menstrual cycle and hormonal contraceptives on competitiveness', *Journal of Economic Behavior and Organization* **83**(1), 1–10.

Buser, T., Dreber, A. and Mollerstrom, J. (2017), 'The impact of stress on tournament entry', *Experimental Economics* **20**, 506–530.

Buser, T., Niederle, M. and Oosterbeek, H. (2014), 'Gender, competition and career choices', *The Quarterly Journal of Economics* **129**(3), 1409–1447.

Buser, T., Peter, N. and Wolter, S. C. (2017), 'Gender, competitiveness, and study choices in high school: Evidence from Switzerland', *American Economic Review* **107**(5), 125–130.

Busse, M. R., Pope, D. G., Pope, J. C. and Silva-Risso, J. (2015), 'The psychological effect of weather on car purchases', *Quarterly Journal of Economics* **130**(1), 371–414.

Cahlikova, J., Cingl, L. and Levely, I. V. (2019), 'How stress affects performance and competitiveness across gender', *Management Science* pp. 1–16.

Carré, J. M., McCormick, C. M. and Hariri, A. R. (2011), 'The social neuroendocrinology of human aggression', *Psychoneuroendocrinology* **36**(7), 935–944.

Cason, T. N., Savikhin, A. C. and Sheremeta, R. M. (2012), 'Behavioral spillovers in coordination games', *European Economic Review* **56**(2), 233–245.

Cassar, A., Wordofa, F. and Zhang, Y. J. (2016), 'Competing for the benefit of offspring eliminates the gender gap in competitiveness', *Proceedings of the National Academy of Sciences* **113**(19), 5201–5205.

Chang, T. Y., Huang, W. and Wang, Y. (2018), 'Something in the air: Pollution and the demand for health insurance', *Review of Economic Studies* **85**, 1609–1634.

Chen, Y., Jiang, M. and Krupka, E. L. (2018), 'Hunger and the gender gap', *Experimental Economics* **22**(4), 885–917.

Chen, Y., Katuščák, P. and Ozdenoren, E. (2013), 'Why can't a woman bid more like a man?', *Games and Economic Behavior* **77**(1), 181–213.

Childs, E., Dlugos, A. and De Wit, H. (2010), 'Cardiovascular, hormonal, and emotional responses to the tsst in relation to sex and menstrual cycle phase', *Psychophysiology* **47**(3), 550–559.

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

Dargnies, M.-P. (2012), 'Men too sometimes shy away from competition: The case of team competition', *Management Science* **58**(11), 1982–2000.

Dariel, A., Kephart, C., Nikiforakis, N. and Zenker, C. (2017), 'Emirati women do not shy away from competition: evidence from a patriarchal society in transition', *Journal of the Economic Science Association* **3**(2), 121–136.

Dickerson, S. S. and Kemeny, M. E. (2004), 'Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research', *Psychological Bulletin* **130**(3), 355–391.

Dreber, A., von Essen, E. and Ranehill, E. (2014), 'Gender and competition in adolescence: Task matters', *Experimental Economics* **17**(1), 154–172.

Edelstein, R. S., Stanton, S. J., Henderson, M. M. and Sanders, M. R. (2010), 'Endogenous estradiol levels are associated with attachment avoidance and implicit intimacy motivation', *Hormones and Behavior* **57**(2), 230–236.

Esopo, K., Haushofer, J. and Kleppin, L. (2019), Acute stress decreases competitiveness among men. Working paper.

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

Flory, J. A., Gneezy, U., Leonard, K. L. and List, J. A. (2018), 'Gender, age, and competition: A disappearing gap?', *Journal of Economic Behavior and Organization* **150**, 256–276.

Flory, J. A., Leibbrandt, A. and List, J. A. (2015), 'Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions', *Review of Economic Studies* **82**(1), 122–155.

Gillen, B., Snowberg, E. and Yariv, L. (2019), 'Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study', *Journal of Political Economy* **127**(4), 000–000.

Gneezy, U., Leonard, K. L. and List, J. A. (2009), 'Gender differences in competition: Evidence from a matrilineal and a patriarchal society', *Econometrica* **77**(5), 1637–1664.

Gneezy, U. and List, J. A. (2006), 'Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments', *Econometrica* **74**(5), 1365–1384.

Goette, L., Bendahan, S., Thoresen, J., Hollis, F. and Sandi, C. (2015), 'Stress pulls us apart: Anxiety leads to differences in competitive confidence under stress', *Psychoneuroendocrinology* **54**, 115–123.

Halko, M.-L. and Sääksvuori, L. (2017), 'Competitive behavior, stress and gender', *Journal of Economic Behavior and Organization* **141**, 96–109.

Hausman, J. (2001), ‘Mismeasured variables in econometric analysis: problems from the right and problems from the left’, *Journal of Economic perspectives* **15**(4), 57–67.

He, T.-S. and Hong, F. (2017), ‘Risk breeds risk aversion’, *Experimental Economics* **21**.

Healy, A. and Pate, J. (2011), ‘Can teams help to close the gender competition gap?’, *Economic Journal* **121**(555), 1192–1204.

Kivikangas, J. M., Kätsyri, J., Järvelä, S. and Ravaja, N. (2014), ‘Gender differences in emotional responses to cooperative and competitive game play’, *PLoS ONE* **9**(7).

Lee, J. (2013), ‘The impact of a mandatory cooling-off period on divorce’, *The Journal of Law and Economics* **56**(1), 227–243.

Liu, T. X., Bednar, J., Chen, Y. and Page, S. (2019), ‘Directional behavioral spillover and cognitive load effects in multiple repeated games’, *Experimental Economics* **22**(3), 705–734.

Loewenstein, G. (1996), ‘Out of control: Visceral influences on behavior’, *Organizational Behavior and Human Decision Processes* **65**(3), 272–292.

Loewenstein, G., O’Donoghue, T. and Rabin, M. (2003), ‘Projection bias in predicting future utility’, *the Quarterly Journal of Economics* **118**(4), 1209–1248.

Luca, M., Malhotra, D. and Poliquin, C. (2017), ‘Handgun waiting periods reduce gun deaths’, *Proceedings of the National Academy of Sciences* **114**(46), 12162–12165.

Niederle, M., Segal, C. and Vesterlund, L. (2013), ‘How costly is diversity? Affirmative action in light of gender differences in competitiveness’, *Management Science* **59**(1), 1–16.

Niederle, M. and Vesterlund, L. (2007), ‘Do women shy away from competition? Do men compete too much?’, *Quarterly Journal of Economics* **122**(3), 1067–1101.

Niederle, M. and Vesterlund, L. (2011), ‘Gender and competition’, *Annual Review of Economics* **3**(1), 601–630.

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

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.

Read, D. and van Leeuwen, B. (1998), ‘Predicting hunger The effects of appetite and delay on choice-annotated.pdf’, *Organizational Behavior and Human Decision Processes* **76**(2), 189–205.

Reuben, E., Sapienza, P. and Zingales, L. (2015), Taste for competition and the gender gap among young business professionals. NBER Working Paper No. 21695.

Reuben, E., Wiswall, M. and Zafar, B. (2017), 'Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender', *Economic Journal* **127**(604), 2153–2186.

Schipper, B. C. (2015), 'Sex hormones and competitive bidding', *Management Science* **61**(2), 249–266.

Shurchkov, O. (2012), 'Under pressure: Gender differences in output quality and quantity under competition and time constraints', *Journal of the European Economic Association* **10**(5), 1189–1213.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. and Updegraff, J. A. (2000), 'Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight', *Psychological Review* **107**(3), 411–429.

van Veldhuizen, R. (2018), Gender differences in tournament choices: Risk preferences, overconfidence or competitiveness? WZB Berlin Social Science Center Discussion Paper.

Wirth, M. (2011), 'Beyond the HPA axis: progesterone-derived neuroactive steroids in human stress and emotion', *Frontiers in Endocrinology* **2**, 19.

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.

Zhang, Y. J. (2013), 'Can experimental economics explain competitive behavior outside the lab?', *SSRN Electronic Journal* .

Zhang, Y. J. (2018), 'Culture, institutions and the gender gap in competitive inclination: Evidence from the communist experiment in china', *Economic Journal* **129**(671), 509–552.

Zhong, S., Shalev, I., Koh, D., Ebstein, R. P. and Chew, S. H. (2018), 'Competitiveness and stress', *International Economic Review* **59**(3), 1263–1281.

Ziomkiewicz, A., Wichary, S., Gomula, A. and Pawlowski, B. (2015), 'Trait anxiety moderates the association between estradiol and dominance in women', *Physiology & Behavior* **143**, 97–103.

Appendix A Supplementary Tables

Table A1: Summary Statistics for Hormonal Measures in Experiment 1

	Men			Women			Difference	
	Obs.	Mean	SD	Obs.	Mean	SD	(M−W)	p-value
Baseline (L_0)								
Cortisol	58	6.24	(3.77)	57	4.19	(1.85)	2.05	0.000
Testosterone	57	132.44	(38.44)	57	57.89	(23.50)	74.55	0.000
Estradiol	56	5.41	(2.02)	55	6.33	(2.31)	−0.91	0.029
Progesterone	57	74.21	(51.10)	58	100.19	(80.92)	−25.98	0.042
After Task 2 (L_1)								
Cortisol	57	7.09	(4.85)	57	4.17	(1.97)	2.92	0.000
Testosterone	57	142.27	(54.05)	58	57.57	(21.62)	84.70	0.000
Estradiol	56	5.67	(2.10)	57	6.35	(2.17)	−0.68	0.092
Progesterone	57	79.34	(50.02)	58	88.79	(72.08)	−9.45	0.417
Relative ($L_1/L_0 - 1$)								
Cortisol	57	0.16	(0.49)	56	0.04	(0.38)	0.12	0.154
Testosterone	57	0.08	(0.27)	57	0.02	(0.23)	0.06	0.220
Estradiol	56	0.10	(0.45)	55	0.05	(0.24)	0.05	0.457
Progesterone	57	0.14	(0.45)	58	−.04	(0.45)	0.18	0.033

Notes: The units of measure are nmol/L for cortisol, pg/mL for testosterone and progesterone, and pmol/L for estradiol. 116 subjects, 58 men and 58 women participated in Experiment 1. The assay failed for a small number of subjects due to low saliva volumes, which results in a reduced number of observations. Complete measures of all four hormonal responses are obtained for 109 subjects, 55 men and 54 women.

Table A2: Correlations between Hormonal Responses

(a) Full Sample

	Cortisol	Testosterone	Estradiol
Testosterone	0.371***		
	[0.000]		
Estradiol	0.009	0.232**	
	[0.927]	[0.015]	
Progesterone	0.480***	0.351***	0.104
	[0.000]	[0.000]	[0.282]

(b) Men

	Cortisol	Testosterone	Estradiol
Testosterone	0.293**		
	[0.030]		
Estradiol	−0.126	0.189	
	[0.360]	[0.167]	
Progesterone	0.519***	0.347***	0.118
	[0.000]	[0.009]	[0.391]

(c) Women

	Cortisol	Testosterone	Estradiol
Testosterone	0.412***		
	[0.002]		
Estradiol	0.165	0.260*	
	[0.234]	[0.057]	
Progesterone	0.380**	0.252*	0.067
	[0.005]	[0.066]	[0.633]

Note: The table presents the Spearman correlations, with p-values in square brackets. The full sample consists of 109 subjects with complete data on hormonal measures, 55 men and 54 women.

Table A3: Hormonal Responses and Task 4 Choice

	(1) Progesterone	(2) Estradiol	(3) Testosterone	(4) Cortisol
Dependent variable: Task 4 choice				
Female	0.106	0.136	0.129	0.066
	(0.140)	(0.139)	(0.141)	(0.137)
Hormonal Response	0.144	0.016	0.351	−0.034
	(0.149)	(0.140)	(0.221)	(0.112)
Female × Hormonal Response	−0.052	−0.338	−0.290	0.200
	(0.164)	(0.325)	(0.294)	(0.160)
Controls	Yes	Yes	Yes	Yes
Observations	115	111	114	113
R-squared	0.352	0.368	0.371	0.369

Notes: Subjects with incomplete data on the hormonal response are excluded from the sample. Dependent variable is 1 if the subject chooses the tournament and 0 otherwise. We examine the hormonal responses of cortisol, testosterone, estradiol, and progesterone and their joint effects. All the behavioral controls, excluding Task 4 choice itself which is the dependent variable, are included in all columns. Robust standard errors are reported in parentheses.

Table A4: Do Hormonal Responses Explain Over- and Under-entry of Tournament

Dependent variable:	Over-entry (among those who should not enter)		Under-entry (among those who should enter)	
	(1)	(2)	(3)	(4)
Female	-0.079	0.117	0.118	-0.001
	(0.118)	(0.191)	(0.169)	(0.192)
Progesterone	0.135*	0.146**	-0.030	-0.008
	(0.073)	(0.066)	(0.082)	(0.069)
Estradiol	-0.033	-0.092**	-0.055	0.036
	(0.039)	(0.045)	(0.081)	(0.075)
Testosterone	-0.050	-0.039	-0.082	-0.064
	(0.059)	(0.067)	(0.052)	(0.066)
Cortisol	-0.081	-0.128**	-0.009	0.103
	(0.052)	(0.052)	(0.075)	(0.083)
Controls	No	Yes	No	Yes
Observations	70	70	39	39
R-squared	0.090	0.483	0.111	0.686

Notes: Seven subjects with incomplete hormonal response measure are excluded from the sample. Columns (1) and (2) consider those for whom the expected earnings are maximized by choosing piece rate; the dependent variable is equal to 1 if the subject chooses tournament, and 0 otherwise. Columns (3) and (4) consider those for whom the expected earnings are maximized by choosing tournament; the dependent variable is equal to 1 if the subject chooses piece rate, and 0 otherwise.

Table A5: Correlations between Hormonal Responses and Performances in Task 1 and 2

	(a) Full sample		(b) Men		(c) Women	
	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Cortisol	0.189**	0.171*	0.099	0.090	0.275**	0.197
	[0.049]	[0.075]	[0.472]	[0.516]	[0.044]	[0.154]
Testosterone	0.117	0.065	0.054	0.058	0.154	0.042
	[0.226]	[0.503]	[0.693]	[0.675]	[0.267]	[0.765]
Estradiol	0.018	0.069	−0.200	−0.161	0.263*	0.317**
	[0.856]	[0.477]	[0.144]	[0.241]	[0.054]	[0.019]
Progesterone	0.138	0.163*	0.028	−0.030	0.214	0.358***
	[0.153]	[0.090]	[0.839]	[0.830]	[0.120]	[0.008]

Note: The table presents the Spearman correlations, with p-values in square brackets. The full sample consists of 109 subjects with complete data on hormonal measures, 55 men and 54 women.

Table A6: Treatment Effect on Task 4 Choice by Gender

	Control		Treatment		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Task 4 choice						
Female	−0.131	−0.158	−0.208**	−0.116	−0.131	−0.118
	(0.099)	(0.113)	(0.093)	(0.117)	(0.099)	(0.105)
After-break					−0.011	0.188
					(0.111)	(0.153)
Female × After-break					−0.078	−0.035
					(0.135)	(0.143)
Controls	No	Yes	No	Yes	No	Yes
Observations	76	76	69	69	145	145
R-squared	0.024	0.261	0.071	0.198	0.047	0.186

Note: Two subjects have missing values for the even-chance risk attitude task therefore are dropped from the regression analysis.

Table A7: Treatment Effects on Over- and Under-entry of Tournament by Gender

Dependent variable:	Over-entry (among those who should not enter)		Under-entry (among those who should enter)	
	(1)	(2)	(3)	(4)
Female	-0.224	-0.202*	0.429**	0.319*
	(0.138)	(0.110)	(0.178)	(0.161)
Rest	-0.219	-0.362**	0.303*	0.355
	(0.151)	(0.179)	(0.177)	(0.270)
Female × Rest	0.096	0.255	-0.588**	-0.488
	(0.183)	(0.167)	(0.291)	(0.357)
Controls	No	Yes	No	Yes
Observations	93	93	52	52
R-squared	0.087	0.426	0.111	0.502

Notes: Two subjects have missing values for the even-chance risk attitude task therefore are dropped from the regression analysis.

Appendix B Experimental Instructions

Welcome to our study on decision-making. The descriptions of the study contained in this experimental instrument will be implemented fully and faithfully. Each participant will receive on average $20 for the study. The overall compensation includes a $5 show-up fee in addition to earnings based on how you make decisions.

All information provided will be kept CONFIDENTIAL. Information in the study will be used for research purposes only. You are not to discuss with anyone any aspect of the specific tasks during or after the study.

1. The set of decision-making tasks and the instructions for each task are the same for all participants.
2. It is important to read the instructions CAREFULLY so that you understand the tasks and make better decisions.
3. PLEASE DO NOT communicate with others during the experiment. Cell phones and other electronic devices are not allowed.

If you have any questions, please raise your hand to ask our experimenters at ANY TIME.

In the experiment today you will be asked to complete four different tasks involving calculation. None of these will take more than 5 minutes. At the end of the 4 tasks, there is a decision making task involving 20 choices.

The method we use to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined. Your total earnings from the experiment are the sum of your payment for the randomly selected task from task 1 to task 4, a randomly selected choice from the 10 choices, and a $5 show-up fee. At the end of the experiment you will be asked to come to the side, where you will be paid in private.

1. Task 1

For Task 1 you will be asked to calculate the sum of five randomly chosen two-digit numbers. You will be given 5 minutes to calculate the correct sum of a series of these problems. You cannot use a calculator to determine this sum; however, you are welcome to write the numbers down and make use of the provided scratch paper. You submit an answer by clicking the submit button with your mouse. When you enter an answer the computer will immediately tell you whether your answer is correct or not. Your answers to the problems are anonymous.

If Task 1 is the one randomly selected for payment, then you get $1 per problem you solve correctly in the 5 minutes. Your payment does not decrease if you provide an incorrect answer to a problem. We refer to this payment as the piece-rate payment. Please do not talk with one another for the duration of the experiment. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

2. Task 2

As in Task 1, you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However, for this task your payment depends on your performance relative to that of a group of other participants. Each group consists of four people randomly chosen in the room. If Task 2 is the one randomly selected for payment, then your earnings depend on the number of problems you solve compared to the three other people in your group. The individual who correctly solves the largest number of problems will receive $4 per correct problem, while the other participants receive no payment. We refer to this as the tournament payment. You will not be informed of how you did in the tournament until all four tasks have been completed. If there are ties, the winner will be randomly determined. Please do not talk with one another. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

[Second salivary sample was taken at this point in Experiment 1. A rest period is introduced here in the treatment condition in Experiment 2.]

3. Task 3

As in the previous two tasks, you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However, you will now get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task. If Task 3 is the one randomly selected for payment, then your earnings for this task are determined as follows. If you choose the piece rate you receive $1 per problem you solve correctly. If you choose the tournament your performance will be evaluated relative to the performance of the other three participants of your group in the Task 2-tournament. The Task 2-tournament is the one you just completed. If you correctly solve more problems than they did in Task 2, then you receive four times the payment from the piece rate, which is $4 per correct problem. You will receive no earnings for this task if you choose the tournament and do not solve more problems correctly now than the others in your group did in the Task-2 tournament. You will not be informed of how you did in the tournament until all four tasks have been completed. If there are ties, the winner will be randomly determined.

The next computer screen will ask you to choose whether you want the piece rate or the tournament applied to your performance. You will then be given 5 minutes to calculate the correct sum of a series of five randomly chosen two-digit numbers. Please do not talk with one another. If you have any questions, please raise your hand.

4. Task 4

You do not have to add any numbers for the fourth and final task of the experiment. Instead, you may be paid one more time for the number of problems you solved in Task

1–Piece Rate. However, you now have to choose which payment scheme you want applied to the number of problems you solved. You can either choose to be paid according to the piece rate, or according to the tournament.

If the fourth task is the one selected for payment, then your earnings for this task are determined as follows. If you choose the piece rate you receive $1 per problem you solved in Task 1. If you choose the tournament your performance will be evaluated relative to the performance of the other three participants of your group in Task 1 – piece rate. If you correctly solved more problems in Task 1 than they did, then you receive four times the earnings of the piece rate, which is equivalent to $4 per correct problem. You will receive no earnings for this task if you choose the tournament and did not solve more problems correctly in Task 1 than the other members of your group.

This is the last stage of the calculation tasks. You will now need to choose which payment scheme (piece rate or tournament) to determine your payoff in this task. You don't have to perform calculation in this task, as we will use the number of correct answers in your piece-rate round (the first task) to calculate your payoff.

The next computer screen will tell you how many problems you correctly solved in Task 1, and will ask you to choose whether you want the piece rate or the tournament applied to your performance. Please do not talk with one another. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

Reminder: In the piece-rate task, the number of your correct answers is <ScoreTask1>

5. Rank Guess

We would like you to guess which was your rank compared to your group members in Tasks 1 and 2. In case of ties in the actual ranks, we counted every answer that could be correct as correct. For example, if the performance in the group was 10, 10, 11, 11, then the correct guess for score 10 is either 3rd or 4th, and the correct guess for score 11 is either 1st or 2nd.

Please enter a number between 1 (meaning that you were the best in your group of four) to 4 (meaning that you were the 4th in your group of four) for each task.

For each right guess $2 will be added to your payment from this part of the experiment.

Please enter your guess for your rank in Task 1:

Please enter your guess for your rank in Task 2:

6. Risk Tasks

This set comprises 1 decision sheet of the form illustrated in the table below.

	Option A	Option B	Decision
1	A	B1	A ☐ B ☐
2	A	B2	A ☐ B ☐
3	A	B3	A ☐ B ☐
4	A	B4	A ☐ B ☐
5	A	B5	A ☐ B ☐
6	A	B6	A ☐ B ☐
7	A	B7	A ☐ B ☐
8	A	B8	A ☐ B ☐
9	A	B9	A ☐ B ☐
10	A	B10	A ☐ B ☐

Each such table lists 10 choices to be made between a fixed Option A and 10 different Option B's arranged in an ascending manner in terms of value of the amount of money. For each row, you are asked to indicate your choice in the final "Decision" column – A or B – with a tick (✓).

Selection of decision sheet to be implemented: One of your 10 choices will be further selected randomly and implemented. The amount of money will be added to your earnings in the previous tasks. At any time during the study, should you have questions, please raise your hand. An experimenter will come to you and answer your questions individually.

DECISION LIST 1

Option A: 50% chance of receiving $10 and 50% chance of receiving $0.

The Option B column lists 10 amounts (displayed in an ascending manner) each corresponding to what you will receive for sure if you choose this option.

DECISION: For each of the 10 rows, please indicate your decision in the final column with a tick (✓).

	Option A	Option B	Decision
1	50% of $10, 50% of $0	$2.5	A ☐ B ☐
2	50% of $10, 50% of $0	$3	A ☐ B ☐
3	50% of $10, 50% of $0	$3.5	A ☐ B ☐
4	50% of $10, 50% of $0	$4	A ☐ B ☐
5	50% of $10, 50% of $0	$4.5	A ☐ B ☐
6	50% of $10, 50% of $0	$5	A ☐ B ☐
7	50% of $10, 50% of $0	$5.5	A ☐ B ☐
8	50% of $10, 50% of $0	$6	A ☐ B ☐
9	50% of $10, 50% of $0	$6.5	A ☐ B ☐
10	50% of $10, 50% of $0	$7	A ☐ B ☐

DECISION LIST 2

Option A: 10% chance of receiving $10 and 90% chance of receiving $0.

The Option B column lists 10 amounts (displayed in an ascending manner) each corresponding to what you will receive for sure if you choose this option.

DECISION: For each of the 10 rows, please indicate your decision in the final column with a tick (✓).

	Option A	Option B	Decision
1	10% of $10, 90% of $0	$0.4	A ☐ B ☐
2	10% of $10, 90% of $0	$0.6	A ☐ B ☐
3	10% of $10, 90% of $0	$1	A ☐ B ☐
4	10% of $10, 90% of $0	$1.2	A ☐ B ☐
5	10% of $10, 90% of $0	$1.4	A ☐ B ☐
6	10% of $10, 90% of $0	$1.6	A ☐ B ☐
7	10% of $10, 90% of $0	$1.8	A ☐ B ☐
8	10% of $10, 90% of $0	$2	A ☐ B ☐
9	10% of $10, 90% of $0	$2.5	A ☐ B ☐
10	10% of $10, 90% of $0	$3	A ☐ B ☐

7. Questionnaire on gender and age.