

The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices¹

David Wozniak²

William T. Harbaugh³

Ulrich Mayr⁴

October 28, 2010

¹We are grateful for support and direction from Chris Minson and Paul Kaplan of the University of Oregon Human Physiology Department. Seminar participants at the (June) 2009 ESA and (July) 2009 IAREP/SABE meetings provided useful ideas and comments. Correspondence should be addressed to David Wozniak (Eastern Michigan University) at dwozniak@gmail.com. We would also like to thank Jack Wozniak for his programming assistance and Cathleen Leue, the director of the Social Science Instructional Labs and her staff at the University of Oregon. A final thanks to Sarah Wunderlich for all her help scheduling and running experiments.

²Eastern Michigan University

³University of Oregon

⁴University of Oregon

The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices.

October 28, 2010

Abstract

Economic experiments have shown that in mixed gender groups women are more reluctant than men to choose tournaments when given the choice between piece rate and winner-take-all tournament style compensation. These gender difference experiments have all relied on a framework where subjects were not informed of their abilities relative to potential competitors. We replicate these findings with math and word tasks, and then show that feedback about relative performance moves high ability females towards more competitive compensation schemes, moves low ability men towards less competitive schemes such as piece rate and group pay, and removes the average gender difference in compensation choices. We also examine between and within-subjects differences in choices for females across the menstrual cycle. We find women's relative reluctance to choose tournaments comes mostly from women in the low hormone phase of their menstrual cycle. Women in the high hormone phase are substantially more willing to compete than women in the low phase, though still somewhat less willing to compete than men. There are no significant differences between the choices of any of these groups after they receive relative performance feedback.

Introduction

Economic experiments have shown when given the choice between piece rate and winner-take-all tournament style compensation, women are more reluctant than men to choose tournaments (e.g. Niederle and Vesterlund 2007). These gender difference experiments have all relied on a framework where subjects were not informed of their abilities relative to potential competitors. We use a within-subjects design and replicate these previous findings for a math task, and show they also exist for a word task. We then show that feedback about relative performance moves high ability females towards more competitive compensation schemes, moves low ability men towards less competitive schemes such as piece rate and group pay, and removes the average gender difference in compensation choices. We also examine between and within-subjects differences in choices for females across the menstrual cycle. We find that the relative reluctance to choose tournaments on the part of women comes mostly from women in the low hormone phase of their menstrual cycle. Women in the high hormone phase are substantially more willing to compete than women in the low phase, though still somewhat less willing to compete than men. There are no significant differences between the choices of any of these groups after they receive relative performance feedback.

In low information settings, the effects of gender and menstrual phase are large. A female has a 0.14 lower probability of choosing a tournament than a male, even when controlling for performance and confidence. For a female to be as likely to choose a tournament as an average male, she must believe she is 40% better than average in performance. The within gender menstrual phase effect is larger than the across gender effect. Females in the low hormone phase of their cycle have a 0.16 lower probability of choosing a tournament than females in higher hormone phase. A low phase female must believe she has 50% better performance to be as likely to compete as a female in the high hormone phase.

Without feedback, high ability females and males are both more reluctant to enter tournaments than expected value maximization would require. This effect is larger for high ability females. Too many low ability types enter competitive environments, and this effect is larger for males. Relative performance feedback moves all these groups toward more optimal choices. This result suggests that gender differences in choices, and the choice differences we identify across the menstrual cycle, are not driven by stable preference differences.

The corporate ladder can be considered a tournament where a number of individuals compete

for promotion based on the results of individual performance. Females make up a small portion of top-level executive positions. Bertrand and Hallock (2001) found that in 1997 the fraction of females in top level management positions was 3% and only 15% of firms had at least one female in a top level executive position. This underrepresentation of females in executive positions may be partially explained by the roles females have in the traditional family with the raising of children affecting their career choices and human capital investments (Polachek 1981). Part of this underrepresentation may be caused by a preference by females against tournaments or for piece rate compensation. Jirjahn and Stephan (2004) find that the attractiveness of piece rate schemes for females is likely caused by less wage discrimination in such a setting when performance can easily be measured. It could be for this reason that firms with a higher proportion of females are more likely to offer piece rate compensation (Brown 1990).

Niederle and Vesterlund (2007) find that when given a choice between a piece rate or a winner-take-all tournament compensation scheme, females overwhelmingly choose the piece rate while males choose the tournament. In their experiment, subjects were given information only about their own absolute performance and were not informed of the quality of competitors. By having participants choose a compensation scheme for past performance from another treatment, they find that the gender differences for competition still remain even when they control for confidence and risk aversion. They find that the gender difference remains large while controlling for overconfidence as 14% of the variation in competitive choices is driven by gender, suggesting that females prefer not to compete when compared to males.

Datta Gupta, Poulsen and Villeval (2005) find similar results in an experiment examining the effects of gender in tournaments. In another experiment, Dohmen and Falk (2007) find that females are less likely to choose variable pay schemes such as tournaments and piece rates when given the choice of a fixed rate for their time. All these experiments used a similar protocol where subjects were given their absolute performance, but were never informed of their relative standing within the group. These economic experiments have been interpreted as meaning that gender differences in self selection are derived from a preference difference between the genders where females seem to have a greater distaste for competition than males.

Gneezy, Niederle, and Rustichini (2003) find that females see lower performance gains from participating in competitive environments. In observing children's performance in running races, Gneezy and Rustichini (2004) also find that competition increases performance of boys, but not girls.

In an experimental setting, Gneezy, Niederle, and Rustichini (2003) find that in a mixed-gender competitive environments, males have significant performance increases when an environment is made more competitive and females do not. However, when females compete only against other females, their performance increases as the environment becomes more competitive. Gupta, Poulsen and Villeval (2005) find that females are more competitive when given the opportunity to choose the gender of a potential competitor. Specifically, females are more likely to choose to enter a tournament if they first choose to be paired against another female before making the competitive environment decision. Similar results have been found by Grosse and Reiner (2010) where females seem to be more competitive when competing against other females. These results suggest that performance in competitive environments is different for males and females and that the gender composition of groups may play a role in performance gains from competition, as well as in the selection into competitive environments. Both performance differences under competition and gender composition effects may help explain the underrepresentation of females in the corporate business world.

Grossman and Eckel (2003) provide a review of gender differences in risk preferences and find that results seem more supportive of females being more risk averse than man, but these results are not entirely robust. The experimental studies examining gender differences for competition argue that gender differences in competitive choices are not driven by gender differences in risk preferences. Ambiguity aversion might be an explanation, since in the competition experiments subjects had little information about the probability distribution of payoffs. But ambiguity aversion has not been found to vary systematically across gender. Moore and Eckel (2003) find that females are more ambiguity averse for specific contexts and domains, while Borghans et al. (2009) find that males are initially more ambiguity averse than females, but as ambiguity increases, males and females behave similarly.

None of the results cited so far show that these gender differences are biologically determined. In fact, Gneezy et al. (2009) report results from experiments in a matrilineal society in India where women are more likely to compete than men. Cross cultural studies of this sort are one way to show the importance, or lack of importance, of biological factors in gender differences. In this paper, we use natural variations in the levels of hormones across the menstrual cycle to examine this same question. We find that competitive choices do vary across the cycle. This shows that while biology may not be the only driving factor, it does seem to play a role. Interestingly, the effect is such that

during the low hormone phase the behavioral differences are quite large, while in the high hormone phase females choices are very similar to those of males.

Hormones have been found to affect behaviors in humans. The hormone oxytocin has been found to increase trusting behavior of individuals (Baumgartner et al. 2008). Fehr (2009) suggests that due to such results, preferences towards trust are affected by biological mechanisms. Other hormones have also been found to matter for economic behaviors and outcomes. For males, testosterone levels of financial traders in the morning can predict profitability through the rest of the day. Cortisol levels in these same traders were found to rise with increased volatility in their market returns (Coates and Herbert 2008). Testosterone levels are correlated with behaviors in economic experiments such as offers and acceptances in ultimatum games (Burnham 2007). Financial risk taking has also been linked to circulating levels of testosterone in men (Apicella et al. 2008).

Males and females have very different levels of a number of hormones, including estrogen, luteinizing hormone (LH), follicle stimulating hormone (FSH), progesterone, and testosterone (Speroff and Fritz 2005). However, it is not necessarily the case that men and women experience the same effects from these hormones. In fact, Shepard et al. (2009) point out that there are thousands of citations in medical journals on the effect of hormones on sex differences and brain organization. This review article concludes that the expression of hormonal effects within a gender may be very dependent on the social environment.

Women exhibit large and predictable hormonal variations across the menstrual cycle (Speroff and Fritz 2005). For females, estrogen and progesterone have received most of the attention in studies examining neuroendocrinological, psychological and behavioral effects. Estrogen has been found to increase power motivation— a reported preference for impact or dominance, suggesting that it may affect competitive preferences (Stanton and Schultheiss 2007). Such a hormonal link with appetites for competition may be very dependent on the social context. Gneezy et al. (2003), Datta Gupta et al. (2005), Delfgaauw et al. (2009), and Grosse and Reiner (2010) have all shown that females are less responsive to competition and also less likely to compete against males than in all-female settings. Using all-female groups and cross-sectional data, Buser (2010) finds that when females have the opportunity to compete with other females they behave less competitively during phases of the cycle typically associated with higher levels of progesterone.

Neuroendocrinology has demonstrated the existence of hormonal effects on brain activity. Results

show that major depression may be linked to reduced density of hydroxytryptamine (5-HT), also known as serotonin, binding sites (Malison et al. 1998). By injecting estradiol in rats, Fink et al. (1996) find that estrogen stimulates an increase in the density of 5-HT binding sites in certain areas of the brain. Injections of estradiol significantly increase the density of binding sites in the anterior cingulate cortex (24%), anterior frontal cortex (41%) and the nucleus accumbens (12%). These areas of the brain have been variously linked with mood, memory, and the anticipation and receipt of monetary rewards (Fink et al. 1996, McEwen 2002, Bethea et al. 2002, Platt and Huettel 2008). Progesterone has been shown to inhibit neurotransmission, and as a result it may decrease anxiety and increase sedation (Vliet 2001). Other research suggests that progesterone may decrease the degradation of 5-HT (Bethea et al. 2002).

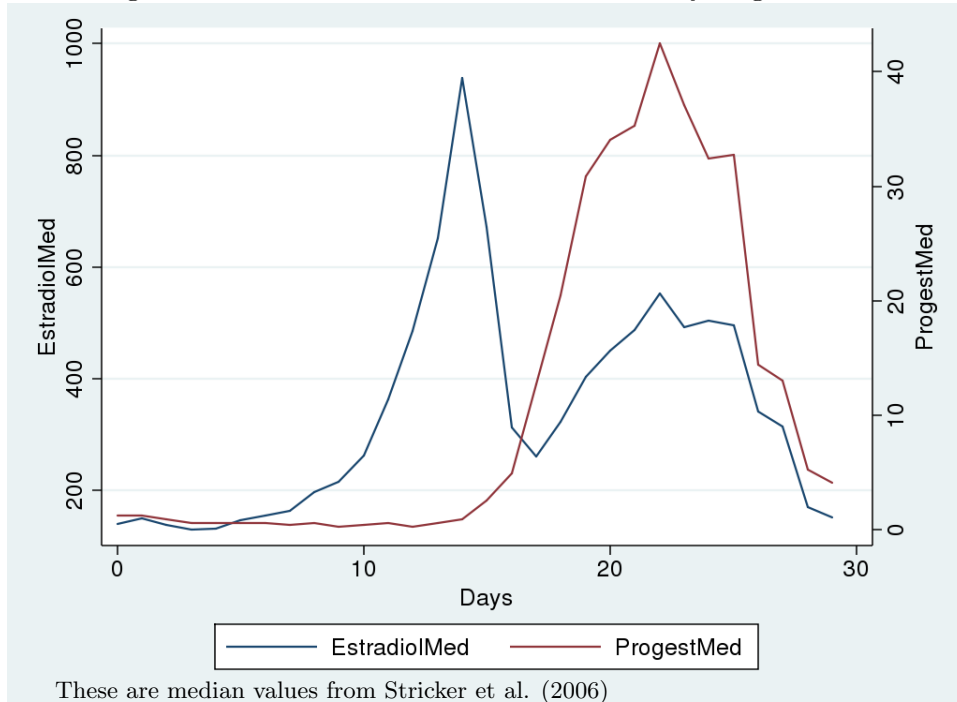
Sex differences in the brain develop during perinatal development where both females and male brains are organized differently from different exposure to steroid hormones (Gagnidze and Pfaff 2009). For female brains, estrogen masculinizes aromatase-expressing neural pathways and also masculinizes territorial behavior in female rodents (Wu et al. 2009).¹ For females, estrogen is required and received by the estrogen receptor to express male-type aggressive and territorial behavior in mice (Gagnidze and Pfaff 2009). Thus, estrogen for females may lead to similar behavior for females as that induced by testosterone in males. In pre-menopausal women, it has also been found that brain activity and activations change across different phases of the menstrual cycle (Dreher et al. 2007).

In this study, we focus on two stable and predictable phases of estrogen and progesterone levels for females, an elevated hormonal phase and a low hormonal phase. As shown in Figure 1, both progesterone and estrogen remain low during the early part of the menstrual cycle. This first week of the cycle is when normal cycling females menstruate and can be considered a low-hormone phase. Estrogen rises quickly and spikes just prior to ovulation, this is referred to as the follicular spike as it occurs during the follicular phase, which is just prior to ovulation. After ovulation (approximately day 14 in the graph), progesterone levels spike in the latter half of the female menstrual cycle, which is called the luteal phase. During the luteal phase, females who ovulate experience heightened levels in both progesterone and estrogen. This second spike in both hormones may be referred to as the luteal spike or high-hormone phase (Speroff and Fritz 2005, Stricker et al. 2006).

Females using a hormonal contraceptive experience suppression of endogenous hormone production

¹Aromatase is an enzyme that converts testosterone to estradiol. In the adult male brain testosterone acting through androgen receptors is necessary to complement male type behavior (Gagnidze and Pfaff 2009).

Figure 1: Hormonal Fluctuations in Normal Cycling Females



when in the active phase of their contraceptive regimen (Speroff and Fritz 2005). Both progesterone and estrogen levels remain fairly constant as the body receives a dose of hormones exogenously (Aden et al. 1998). During the placebo or non-active phase of the contraceptive regimen there are no exogenous hormones being provided to the body; this lowering of both progesterone and estrogen through exogenous means leads to withdrawal bleeding caused by the withdrawal of exogenous hormones (Speroff and Fritz 2005). There is no biological or medical necessity to induce this withdrawal bleeding (Anderson and Hait 2003). Since menstrual bleeding is caused by low hormones, it allows for easy identification of a low hormone phase for research purposes.

Monthly bleeding, periodic bleeding, or no bleeding—this is an individual woman’s choice (Speroff and Fritz 2005, 908).

It is suspected that keeping withdrawal bleeding as part of the hormonal contraceptive regimen was a marketing ploy to make the birth control pill seem more socially acceptable (Coutinho and Segal 1999). Thus, females could entirely avoid hormonal fluctuations by sustained contraceptive use. Indeed, some forms of contraception do ensure that this occurs (Anderson and Hait 2003). If the decrease in hormones affects females in an economically costly manner then one could expect that there is some demand to prolong contraceptive use among the female population which has

become evident by recent prolonged-use contraceptives such as *Seasonale*.

Other economic studies have examined the effects of the menstrual cycle and therefore implicitly or explicitly hormonal fluctuations for females. Hormonal fluctuations occur across the female menstrual cycle and contribute to premenstrual syndrome effects, which may have significant economic consequences. Ichino and Moretti (2009) use detailed employee attendance data from a large Italian bank and find that absences for females below the age of 45 tend to occur according to a 28-day cycle. These 28-day cycle absences explain about one-third of the gender gap in employment absences at the firm. The female menstrual cycle is approximately 28 days and they focus on females below the age of 45, who are more likely to be pre-menopausal. Ichino and Moretti's result provides significant support for the hormonal fluctuations of the menstrual cycle having significant economic impacts.

In an experimental study, Chen et al. (2010) explore hormonal differences according to menstrual cycle phase for differences between males and females auctions. They find bidding differences in first-price auctions between females with higher bids for females in the follicular phase versus females in the luteal phase though most of this variation is found to be driven by contraceptive users. In contrast, Pearson and Schipper (2009) find that women bid higher than men and earn lower profits only during the menstrual and premenstrual phases of the cycle when estrogen and progesterone levels are lower. Another between subject experimental study, Buser (2010), finds that in all-female groups, females participating during high levels of progesterone tend to be less competitive.

Not all economic studies have found support for hormonal effects being significant in economic decision making. Zethraeus et al. (2009) examine 200 post-menopausal women in a double-blind study. Participants were given either estradiol (2 mg), testosterone (40 mg) or a placebo daily for a four week period. Then they participated in an economic experiment session that included a variety of different tasks looking at risk aversion, altruism, reciprocal fairness, trust and trustworthiness. No significant differences were found when comparing the behaviors between the three different treatment groups of females. One should note that the lack of effects in post-menopausal women may be caused by neural receptors no longer being as sensitive to hormonal changes due to the effects of aging (Chakraborty et al. 2003).

Though a double-blind study using exogenous delivered hormones would be ideal to examine the effects of hormones for females it would not be feasible since low hormone levels cause bleeding

in pre-menopausal females (Speroff and Fritz 2005). Instead of a double-blind study, we include a within-subject design in this experiment focusing on a comparison between the midluteal peak and the low hormone phase that occurs during menses. This type of scheduling has been included in experiment designs used in neuroscience to study menstrual-cycle dependent neural plasticity (Fernandez et al. 2003). These behavioral and neurological studies lay the foundation for our exploration of whether females behave more competitively during a high hormone phase within the menstrual cycle.

1 Experiment Design

We use an experimental design such that females are scheduled to participate in two sessions, during both a low and high hormonal phase of the menstrual cycle. Since females may alter their hormonal fluctuations by taking a hormonal contraceptive, we include females taking hormonal contraceptives. For females using hormonal contraception, during the placebo phase of the hormonal contraception regimen, their hormone levels are the same as normal cycling females in the low-hormone phase of their cycle. This design allows for the comparison of behaviors between females in the low phase to those not in the low phase, while also comparing the behavior of male participants. In the following section we explain the experimental design used to examine both the effects of relative performance feedback and the role of hormones for competitive environment selection.

Our experiment used a within-subjects design with two sessions. For normally cycling females these sessions were scheduled during a low hormone phase (days 2 to 7 in Figure 1) and a high hormone phase (days 18 to 25 in Figure 1) of the menstrual cycle. We intentionally avoided the estradiol spike around day 14, because of its short duration and variability within and across females. For females on a monthly hormonal contraceptive with a placebo week (i.e. “the pill”) hormone levels in the placebo week are the same as normal cycling females in the low-hormone phase (Speroff and Fritz 2005). They were scheduled accordingly. During the active phase of the contraceptive regimen females using hormonal contraceptives get the higher hormonal levels exogenously from the pill.

An online pre-screening survey was used to recruit and schedule subjects for experiment sessions. We limited the sample of females to those using a monophasic hormonal contraceptive or not using

hormonal contraceptives at all.² Using this data females were first scheduled during a predicted high or predicted low hormone phase. This created a random assignment of hormonal phases for experiment sessions. In comparison to simply asking females that show up for a session as to the date of their last or expected next menstruation, we focused on scheduling across two specific hormonal phases to minimize identification error as the menstrual cycle has a large degree of variability, so much so, that subjects may have trouble accurately recalling and predicting menses (Crenin et al. 2004).

To help minimize identification errors, we combined the pre-screening survey with an exit survey. The low hormone phase is easily identified by the presence of withdrawal bleeding (in both normal cycling females and those in the placebo phase of contraceptives). The high hormone phase is more difficult to pinpoint, particularly for subjects not on contraceptives, because of variability in the cycle. The combination of pre-screening, scheduling, and the exit survey were designed to address this.

Previous studies on differences in competitive choices use a strictly between-subjects design (Niederle and Vesterlund 2007, Gneezy et al. 2009). In our within-subjects design each subject participated in one session of math tasks, and another of word tasks. It is generally believed that females may view themselves as having worse math skills than males (Niederle and Vesterlund 2010). For this reason females may be less likely to compete in math tasks than in word tasks. This design is the first to examine whether there are stable differences between subjects and within subjects for math and word based tasks.

Subjects were randomly assigned to start with a math or a word based session. In each session tasks were performed for five different treatments, one of which was randomly chosen for payment at the end of the experiment. Each treatment lasted 4 minutes. In the first treatment participants performed the task under a non-competitive piece rate compensation scheme, where pay was entirely dependent on the individual's own performance. In the second treatment, participants were randomly assigned to a winner-take-all tournament with a size of two, four, or six competitors. This second treatment provided participants with experience in a situation where their pay depended on their own performance as well as the performance of others. In the third treatment, participants performed the task with a group pay (revenue sharing) form of compensation. This treatment

²Monophasic hormonal contraceptives, release the same level of exogenous hormones each day for the entire non-placebo phase of the hormonal contraceptive regimen. We excluded users of biphasic and triphasic pills, with varying daily hormone doses.

randomly paired participants and payment for the group's total production was split evenly. This third treatment can be considered the least competitive, in that pay was identical for the two players. This type of revenue sharing allows for an individual to free-ride and not compete at all if they so choose. As long as an individual is paired with an individual that exerts effort then he or she can still be paid without exerting any effort at all.

In the final two treatments subjects were able to choose between piece rate, group pay, or a two, four, or six person tournament. Before the fourth treatment, subjects were told their own absolute performance, but were not told anything about the performance of others. Just before the fifth treatment, participants were shown how all individuals in the session had performed in the first treatment. They then chose their compensation method and performed the task again.

The math task was similar to that in Niederle and Vesterlund (2007). Participants were asked to add four randomized two-digit numbers and complete as many of these summations as possible in 4 minutes. Equations were presented to participants on a computer screen and they typed in their answer and pressed the *Enter* key or clicked a *Submit* button on the screen. After each submission participants were promptly shown the next equation to solve. On the screen, the equations looked like the following:

$$12 + 57 + 48 + 52 =$$

A sheet of paper and a pencil were provided for this task, but no other form of assistance was available.

The word task was similar to that used by Günther et al. (2008). In this task participants are shown a letter on a computer screen and have four minutes to form as many unique words as possible that begin with that specific letter. The letter remains on the screen for the entire four minutes and participants enter in their word submissions in a text box below the letter. The attempted word formations are then listed below the text box to help subjects minimize duplicate answers, since these are counted as incorrect. Common place names (cities, countries) are acceptable, but proper names are counted as incorrect. Plural and tense changes to root words are counted as separate and correct answers as long as these words still begin with the appropriate letter. In the experiment, participants were informed of the rules before beginning the task. All participants were informed that everybody in the same session and same treatment received the same letter; thus, providing them with a task of equivalent difficulty for all participants in each treatment.

The word list used for grading words is a common English word list used by open source word processors.³ We used a restricted group of letters for this study to limit the variation of difficulty between treatments and sessions (e,f,g,h,i,l,n,o). Between 2.7% to 3.8% of all words in the word list began with these letters.

For the piece rate compensation, the payoff an individual receives is equal to the piece rate multiplied by the production of the individual for that particular treatment. Payoffs for both the math and verbal tasks were calculated in a similar manner though the base rate, in the form of the piece rate (w), was different for word formation tasks ($w_w = \$0.25$) and math addition tasks ($w_m = \$0.50$), to adjust for generally higher performance in the word task. In a tournament, if an individual has the best performance in his group then he receives the piece rate multiplied by the size of the tournament, multiplied by his individual performance. If an individual does not have the best performance in his tournament then he receives nothing. In the event of a tie, the individual receives a fraction of the tournament winnings based on the number of individuals he tied with. Subjects were not informed about whether they won or lost a tournament until all five treatments were complete. After each treatment, and before seeing their score, subjects were asked how well they thought they did and how well they thought the average person in the session did, and they were paid for having accurate predictions.

Subjects were told that they could be randomly grouped with people that did not necessarily choose the same compensation option and were therefore playing under different rules than their potential competitors or group members. This rule strengthens the incentive for high ability types to choose a more competitive tournament as there is a positive probability that they may compete against lower ability individuals. This rule also creates an incentive for low ability individuals to choose group pay as they may be matched with high ability individuals, which would increase their expected payoffs.⁴

2 Results

Experiment sessions took place in a computer lab at a large public university. The majority of the 219 participants were university students, characteristics are in Table 1. The average size of the 26 sessions was 14.5 participants (with a standard deviation of 4.15). The gender composition of

³Spell Checking Oriented Word Lists (SCOWL), Revision 6, August 10, 2004 by Kevin Atkinson.

⁴The text for experiment instructions is available in Appendix D.

two session participants consisted of 62 females and 64 males. Using the pre and post surveys we identified 45 females as participating in a session during a low-hormone phase of their menstrual cycle, and 34 during both a low and a high hormone phase. The word task was used in 12 of the sessions and the math task was used in 14 sessions. Of the 345 individual subject sessions, 165 involved the use of the word task and 180 used the math task.⁵

Table 1: Summary statistics of session attendees

Participant Characteristics					
Variable	Mean	Std. Dev.	Min.	Max.	N
Age	20.52	2.81	18	33	218
Years PS	2.18	1.48	0	6	217
GPA	3.29	0.47	2	4.1	218
Live Independently	0.82	0.39	0	1	219
Female	0.5	0.5	0	1	219
Psych meds	0.09	0.28	0	1	219
Characteristics by Sessions*					
Low Phase	0.14	0.34	0	1	345
Word task	0.48	0.5	0	1	345
Session Size	14.54	4.15	7	21	345
Second session	0.37	0.48	0	1	345

*126 individuals attended a second session.

Table 2 shows that men and women were similar in terms of age, GPA, and were assigned to sessions with similar characteristics, except that on average females were in slightly larger sessions. The session female to male ratio ranged from 0.3 to 2.3 and averaged 1.01. Thus, all sessions had some degree of gender mix and on average this mix was about one-to-one.

Table 2: Mean values of individual and session characteristics by gender.

Sex	Age	Years PS	GPA	Indep.	Meds	Word	Size	Sess. 2
Male	20.70	2.14	3.25	0.83	0.09	0.47	13.79	0.37
Female	20.35	2.21	3.33	0.80	0.08	0.48	15.30	0.36
Total	20.52	2.18	3.29	0.82	0.09	0.48	14.54	0.37

Sessions took place three to four times a week and were held in the morning. Each session took slightly less than an hour, including approximately 10 minutes at the beginning of the session during which participants waited together in a foyer. This allowed participants to see that sessions included both males and females. Once participants entered the lab, partitions were used so that

⁵One individual (female) was removed from the data due to inability to follow the instructions.

participants could not see each other’s computer screens or facial responses from the feedback received. Competition and group memberships were also anonymous.

Payouts were based on one randomly chosen treatment, excluding the flat rate show-up payment, payouts averaged \$7.38 for the math session and \$15.01 for the word sessions. Participants who attended two sessions were later asked to perform a risk aversion task similar to that used in Holt and Laury (2002). The risk aversion tasks were performed a few days after the second session to avoid endogeneity with competition task earnings. A total of 112 participants (56 male and 56 females) participated in the risk aversion task. The average payout for the risk aversion task was \$6.57.

2.1 Task Performance

Each individual participated in five different treatments in each session. For the first three treatments the compensation schemes were as follows:

Treatment 1: Piece rate (\$0.50 per sum and \$0.25 per word).

Treatment 2: Random sized tournament of 2, 4, or 6 individuals (the winner earned the piece rate multiplied by the size of tournament).

Treatment 3: Group pay: an individual was paired with a randomly chosen partner and the total production of the 2 individuals was multiplied by the piece rate and then split evenly.

Table 3: Performance Across Treatments and Gender

Math	T1	T2	T3	T4	T5	Word	T1	T2	T3	T4	T5
Male	10.9	12.1	12.3	12.7	12.8	Male	38.2	39.4	43.0	45.3	47.0
Female	9.9	11.4	11.8	12.3	12.1	Female	41.0	41.1	45.0	48.4	47.3
Both	10.4	11.8	12.0	12.5	12.5	Both	39.6	40.3	44.0	46.9	47.1

Table 3 shows mean performance by gender over treatments and tasks. The increasing mean values over the first three treatments in both the math and the word tasks suggest that subjects are learning to do the task better during the session. The across gender performance differences in the math task are not statistically significant at 5% using two-tailed t-tests. In the word task, the differences in performance are not statistically significant for any of the treatments. Thus, there are no noticeable performance differences between males and females after the first treatment.⁶ This

⁶An analysis of the performance effects that occur from learning and different competitive settings are available on request.

lack of a performance difference by gender, for either task, removes one obvious potential reason for gender differences in choices.

2.2 Gender Differences in Compensation Choice

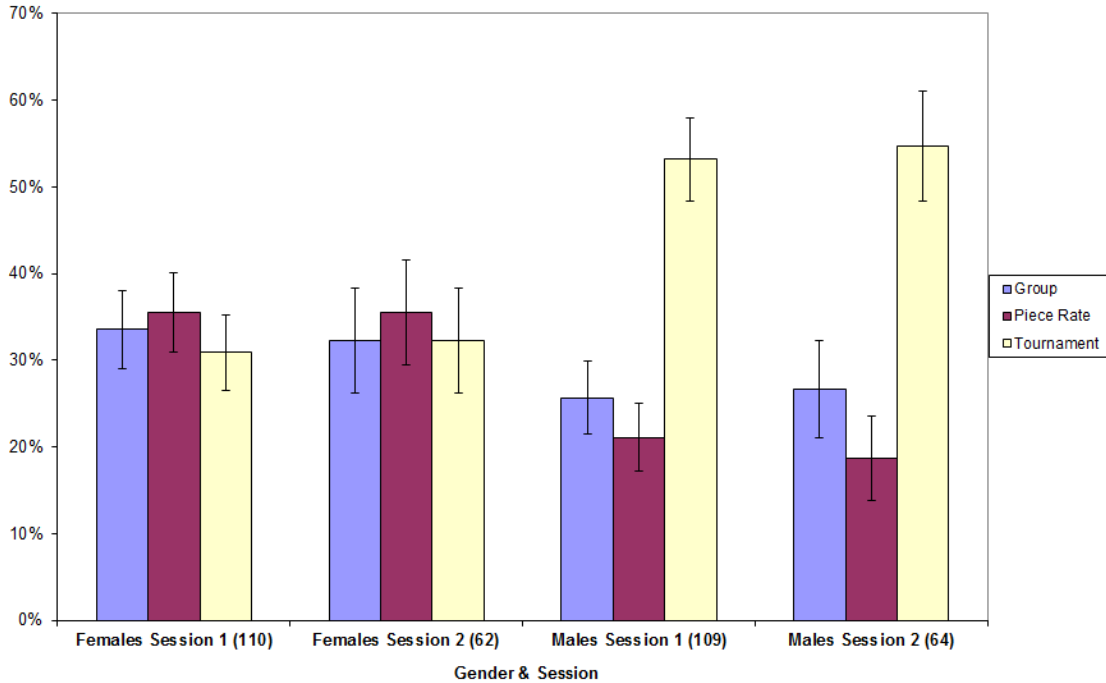
Niederle and Vesterlund (2007) and Datta Gupta et al. (2005) find that females are less likely than males to enter tournaments when given the option between a tournament and piece rate. To test whether these results can be replicated with our protocol, we first focus on choices made in Treatment 4. In this treatment, as in previous studies, individuals were not given information about their relative performance. In Treatment 5 we consider the effect of relative performance information on competitive choices.

In both Treatment 4 and Treatment 5, individuals had the opportunity to choose between group pay (revenue sharing), piece rate, or a tournament size of two, four or six individuals. These choices are ordered by increasing competitiveness, with sharing less competitive than not sharing and larger tournaments being more competitive. In the figures and empirical analysis we lump the tournaments together regardless of size, the results are robust when using an ordered scale for tournament size.

Figure 2 shows the distribution of choices made by males and females in the first and second sessions for Treatment 4. The gender differences are large: pooling over sessions we find that only 31% of females chose to compete in tournaments while 54% of males chose the tournaments. The gender difference persists for the piece rate as 36% of females chose the piece rate compared to only 20% of males. These differences are all significant at the 2% level or better with chi-square tests. On average males and females chose consistently across the two repeated sessions, which are often separated by a week or longer. This consistency suggests that the gender differences for competition we measure are stable with repetition.

Figure 3 shows the effect of the relative performance feedback that occurs between Treatments 4 and 5 on choices. The two groups of bars on the left side of the figure suggest that females' choices are barely affected by this information. Comparing females and males before feedback, we see that females select both group pay and piece rate much more frequently than males. This replicates earlier findings, and shows that gender differences for competitive choices are robust to the addition of a group pay option and different sized tournaments.

Figure 2: Choice Differences Before Feedback, by Gender and Session



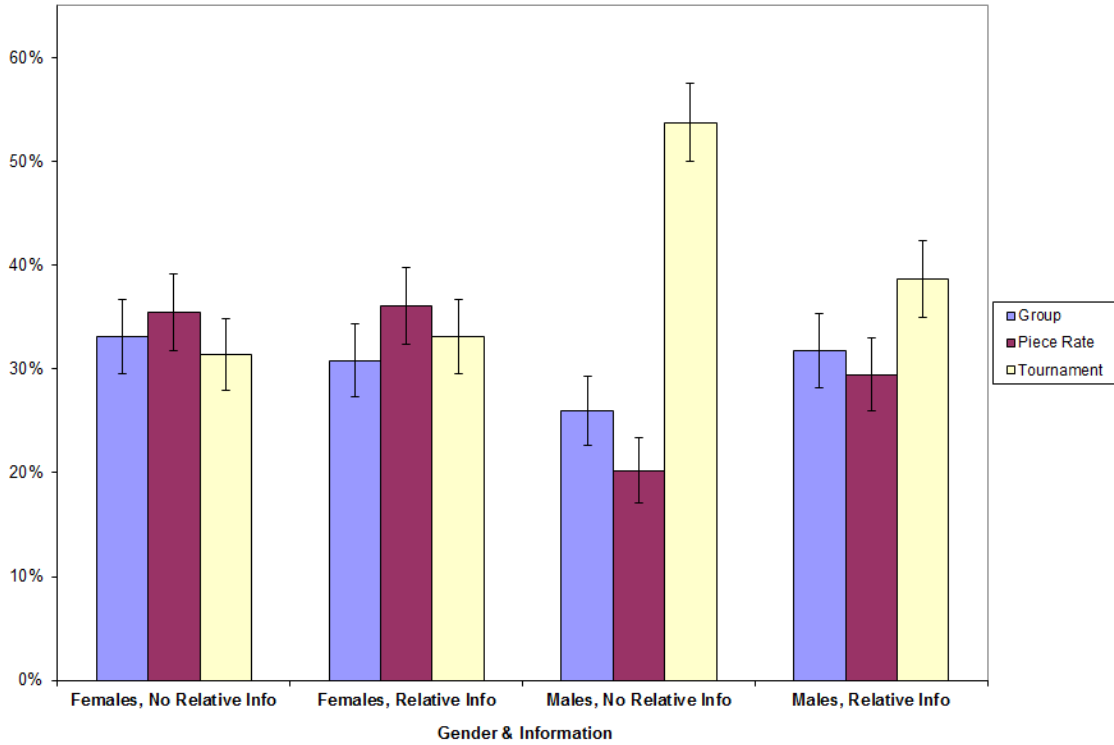
Sample size in parentheses.

The right side of Figure 3 shows that male’s choices are dramatically affected by information about the performance of potential competitors. There is a significant increase in the proportion of males choosing piece rate (5% significance level) and group pay (10% significance level), and a significant decrease in the proportion choosing tournaments (5% significance level). After receiving the relative performance information (Relative Info), the choices of men and women become more similar. Comparing the distributions of men’s and women’s choices in Treatment 4 gives a Pearson chi-square statistic of 18.79 (p-value: 0.000). With relative performance feedback, in Treatment 5, the chi-square statistic across genders is 1.91 (p-value: 0.385) and the proportions of males and females choosing tournaments are not statistically different from each other.

With information, on average men become less competitive while women’s choices seem to barely change. However, there is significant movement within both genders based on ability. These within-subject responses to feedback, by gender and ability, are discussed in a later section.

While there are no significant differences in performance between males and females, other factors such as age and GPA might affect compensation choices. We use an ordered probit

Figure 3: Selection Differences by Gender and Information Treatment (Treatments 4 and 5)



Females (172), Males (173). Sample size in parentheses.

estimation model to test whether the gender differences in the probability of selections for Treatment 4 remain after controlling for these other potentially relevant factors. We show that the gender differences persist with these controls, along with the addition of control variables for confidence, performance, and improvement in the repetition of tasks in a tournament.

Table 4 shows the results from ordered probit estimations of choices in Treatment 4. Columns 1 to 3 use CompScale as the dependent ordinal variable, where group pay compensation is less competitive than piece rate which is less competitive than a tournament of any size.⁷

Table 4 replicates the results of Niederle and Vesterlund (2007) with Treatment 4, before relative performance feedback. Females are less likely than males to enter tournaments, even when controlling for individual confidence (Confidence (T1)) and relative rank of performance within the session (%-tile Rank (T1)) from the first treatment. The %-tile Rank (T1) variable gives the rank of an individual based on her or his performance in Treatment 1 in the session. Using rank allows us to

⁷Our results are consistent with those from multinomial logit and from using ordered probit with rankings that treat larger tournaments as more competitive.

Table 4: Ordered Probit Estimates: Choices for No Relative Information Treatment

VARIABLES	(1)	(2)	(3)
	Pooled CompScale	RE CompScale	RE Risk CompScale
Female	-0.36 (0.13) (***)	-0.40 (0.15) (***)	-0.49 (0.19) (**)
Confidence (T1)	0.86 (0.25) (***)	0.98 (0.29) (***)	0.99 (0.34) (***)
Improve (T2)	0.61 (0.20) (***)	0.72 (0.23) (***)	0.73 (0.32) (**)
%-tile Rank (T1)	1.05 (0.23) (***)	1.08 (0.26) (***)	0.85 (0.32) (***)
Characteristic Controls	Yes	Yes	Yes
Risk Controls	No	No	Yes
Observations	343	343	224
ll	-336.6	-335.6	-212.3
chi2	66.91	61.00	48.81

Pooled means pooled cross section. RE means that random effects were used.

Standard errors in parentheses. $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

have the same measure for both math and word tasks.⁸ Confidence is measured by an individual’s predicted performance at the end of Treatment 1 (prior to finding out their actual performance) divided by that individual’s prediction of the average performance of all session participants.⁹ To control for performance, we use the relative rank from Treatment 1, but the results are unchanged when using absolute performance along with an interaction term for word based tasks.

As expected, both confidence and the percentile rank from the first treatment are positively correlated with the selection of more competitive environments. Improvement in task performance between the first and second task (Improve (T2)) also has a significant positive effect. These regressions include controls for individual specific characteristics, including the number of years of college, psychoactive medication, GPA, and age.¹⁰ The results are similar when using a random effects ordered probit, in column 2. Column 3 includes a measure of risk aversion for individuals

⁸Using a variable that measures actual performance with an interaction term for the type of task, gives the same results as are presented here.

⁹Females tend to be less confident than males in the math task, but females and males have no significant differences in confidence in the word task and selection differences still remain.

¹⁰Details on these controls are in the appendix.

that participated in a task similar to the one used by Holt and Laury (2003). We find that this measure of risk aversion is not significantly correlated with competitive choices in Treatment 4.

The marginal effects (calculated from column 1) show that a female has a 0.14 lower probability of choosing a tournament than a male, even when controlling for performance and confidence. For a female to be as likely to choose a tournament as an average male, we would have to increase her belief about her performance relative to the average by 40%, which is a significant increase in overconfidence. A ten-percentile improvement in actual relative performance would increase the probability of entering a tournament by 0.04. A female would have to improve her percentile rank by 34% to be as likely to enter a tournament as a male. Thus, these gender differences are not just significant, but they are also large.

After each treatment, before receiving any feedback, subjects were asked how many correct answers they believed they submitted. Subjects were paid (\$0.25) for each correct answer to encourage accurate answers. We create a measure of confidence by dividing an individual's prediction of how well he or she did divided by his or her prediction of the session average for that treatment. Since the average individual should believe they did not perform any better than the session average, this confidence measure should have a mean of one— in the absence of overconfidence.¹¹ We could have asked for rank estimates instead of performance estimates, but rank is a poor measure of the degree of over or under confidence. Consider two individuals that think they are ranked first in their respective group. One may think that he is 10% better than the average while the other may think she is 50% better. Both these individuals would be treated as having the same level of confidence with the rank measure, but one individual is actually much more confident. We use the measure of confidence from the first treatment because every subject performed the task for this treatment under the same piece rate form of compensation. This confidence measure also provides the earliest measure of overconfidence before experiencing any feedback or differing experimental manipulations.

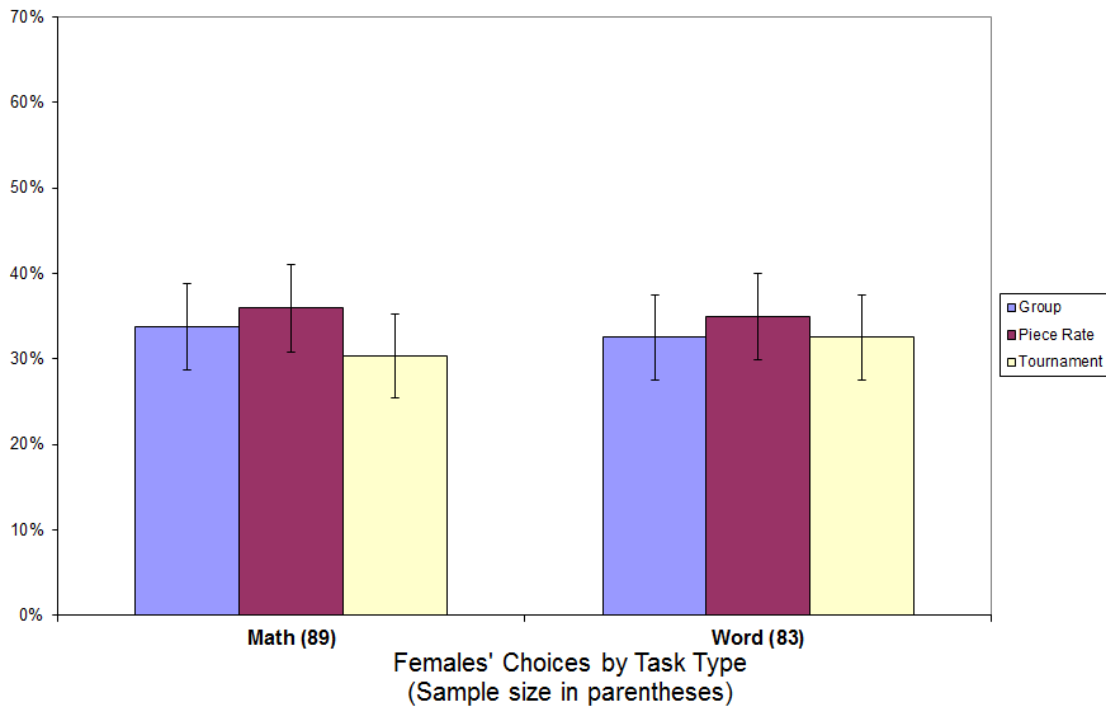
Of course, changes in performance as the experiment proceeds could change confidence. The variable Improve (T2), measured as the ratio of the individual's performance from Treatment 2 divided by the performance in Treatment, captures the effect of individual improvement between

¹¹We also asked how many correct answers they believed were submitted by the most productive person, the least productive person, as well as the average number of correct answers, for each session and treatment. We use the average instead of the prediction of the best or worst individual in the session because it provides a clean measure of overconfidence. In other estimations, not included here, these measures were separated and variations of using both the best performance and the worst performance as the denominator were used with little difference in our results.

Treatment 1 (piece rate) and Treatment 2 (tournament). There are two possible reasons that this variable should matter: First, individuals may feel that they improve more than the average individual or that they were unlucky in Treatment 1 compared to how others would have performed. Second, it may be the case that individuals become more motivated to put in greater effort because of the competitive nature of the tournament in Treatment 2. Individuals that improve a lot from competing in such settings would be more likely to choose to compete than individuals whose performances are not positively affected by competitive settings.

Niederle and Vesterlund (2007) found that part of the difference between male and female willingness to compete was driven by males being more overconfident than females. In their study, independent of confidence, females had a 0.16 lower probability of entering a tournament than males. Using our measure of confidence we find that the gender difference is nearly the same, 0.14. Our

Figure 4: Selection Differences for Females by Task Type for No Information Treatment



within-subjects design includes one session of math treatments and one of word treatments. Günther et al. (2010) found that in a maze task, men increased performance in reaction to competitive pressures by more than women did. In a word task the improvements were the same. They attribute

this to a "stereotype threat" arising from beliefs that women are not good at the maze task. This could logically lead to different choices by women to compete, with different tasks. Figure 4 shows that in our data there is little difference in the selection of competitive environments by females regardless of the type of task used. We also find little difference in choices by males as more than 50% of males chose to compete in tournaments in both math and word tasks.

There still may be gender differences in confidence across tasks that do not show up in choice differences. Comparing females and males, in Table 5, we find both genders are overconfident on average, but the degree of this overconfidence varies by task. Males are significantly more overconfident in their math abilities than females. But there is no significant difference in confidence between males and females in the word task. There is also no significant difference among females between the math and word tasks, though males seem significantly more confident in their math abilities than in their abilities in performing the word task.

Table 5: Confidence Differences by Gender and Task Type

Task Type	Gender	Confidence (T1)	S.E.
Math	Female	1.081	0.030
Math	Male	1.209	0.037
Math	Both	1.145	0.024
Word	Female	1.046	0.029
Word	Male	1.039	0.030
Word	Both	1.043	0.021
Both	Female	1.064	0.021
Both	Male	1.128	0.025
Both	Both	1.096	0.016

When comparing aggregated confidence for Both, removing an outlier makes the gender difference insignificant at 5%.

On average, males are slightly more confident in their abilities than females, but this is partly driven by a few high ability and overconfident males, who are correct in believing they are better than the average, but overestimate the degree. For example, the highest level of confidence for a male is 3.375 times his prediction of the average. His actual performance is 2.29 times the actual session average. Overall, males and females are fairly consistent in their choices to compete in both types of task: males choose to compete more than females in both math and word tasks even though male overconfidence is higher in the math task. The type of task was not significant in regressions for choices, with or without confidence controls.

2.3 Does Relative Performance Feedback Eliminate Gender Differences in Choices to Compete?

If the gender differences in competitive choices that have been identified are driven by preference differences for competition, then providing information about the quality of possible competitors might reduce mistakes in competitive choices, but there is no obvious reason feedback should reduce the gender difference if it is driven by preferences. We test the effect of performance feedback on choices by providing subjects with a list of the performance of all the participants in their session from Treatment 1, before they choose their Treatment 5 compensation scheme. This provides information about the quality of the competitors that participants may end up competing against, if they choose to enter a tournament.

Table 6 shows the results from three different types of ordered probits for Treatment 5 choices, using the same CompScale dependent variable definition as in the Treatment 4 analysis. Columns 1 through 3 show, that once performance feedback is provided, there are no significant differences between males and females in choices. Instead, we find that choices are very dependent on the relative performance information and on the individual's improvement from Treatment 1 to Treatment 2. Risk aversion control variables are not significantly correlated with compensation choices on average; though risk aversion measures were significant when only examining high ability individuals' choices in Treatment 5. The one variable that consistently affects individual choices in Treatment 5 is an individual's percentile rank from Treatment 1, which is a summary statistic of the feedback information provided to individuals prior to making a competitive choice.

The overall conclusion from Table 6 is that there are no significant differences between females' and males' selections of competitive environments when they are fully informed of their relative performance compared to potential competitors. In the next section we consider the costs of the selection differences between men and women when they lack information about the quality of competitors and whether there are gender differences according to ability levels.

2.4 The Cost of Gender Differences in Compensation Choice.

To give some sense of the costs of gender differences in choices, we simplify and assume people maximize expected payoffs, keep effort constant across compensation choices, and take the choices and performance of others as given. Table 7 shows the average expected value losses for the

Table 6: Ordered Probit Estimates: Choices for Relative Information Treatment

	(1)	(2)	(3)
	Pooled	RE	RE Risk
VARIABLES	CompScale	CompScale	CompScale
Female	0.00	-0.02	0.13
	(0.13)	(0.18)	(0.21)
	()	()	()
Confidence (T1)	0.34	0.44	0.65
	(0.24)	(0.30)	(0.35)
	()	()	(*)
Improve (T2)	0.81	1.01	0.65
	(0.20)	(0.26)	(0.32)
	(***)	(***)	(**)
%-tile Rank (T1)	2.17	2.59	2.31
	(0.25)	(0.34)	(0.37)
	(***)	(***)	(***)
Risk Controls	No	No	Yes
Characteristic Controls	Yes	Yes	Yes
Observations	343	343	224
ll	-320.6	-316.7	-194.5
chi2	110.9	98.51	79.67

Pooled means pooled cross section. RE means that random effects were used

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

suboptimal selections by males and females in Treatment 4 and Treatment 5.¹² Each column represents the optimal choice that should have been made. The numbers represent the average expected value cost for choosing something other than that optimal choice. For example, in the first row under column 6 (for the 6 person sized tournament), the 27.27 represents the average loss to females whose optimal choice would have been a tournament of six, but who instead chose a different form of compensation. The Avg Loss column provides the average loss by gender and treatment. The average loss of 6.78 in the first row means that females lost an average of \$6.78 from their suboptimal choices in Treatment 4.

In Treatment 4, the average expected value loss from selection mistakes was \$4.91 for males and \$6.78 for females, a statistically insignificant difference with a t-test. These loss differences are mostly driven by high ability females choosing not to compete, and to a lesser extent by low ability males choosing to compete. Column 6 shows that many high ability females (those who should select a tournament size of 6) are instead selecting smaller tournaments or group pay or

¹²The method used to calculate expected values is provided in the Appendix.

Table 7: Selection Losses

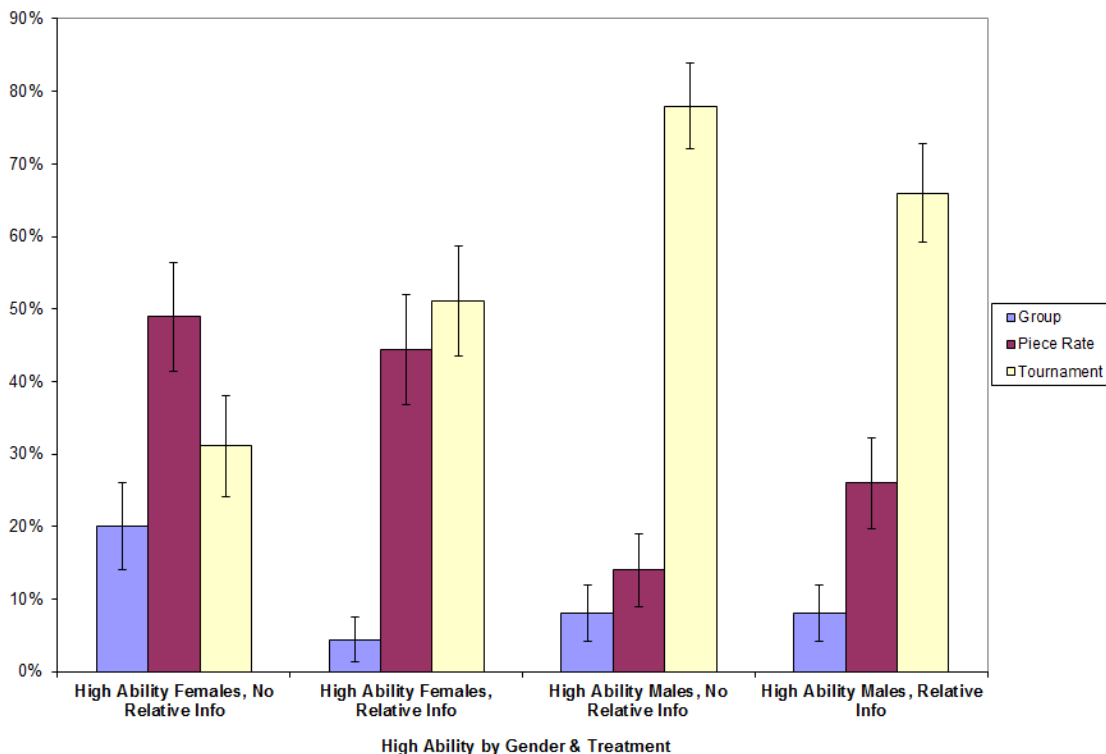
Average Loss from Suboptimal Decisions							
Treatment	Gender	Optimal Choice					Avg Loss
		Grp	PR	2	4	6	
4	Female Avg Loss	1.58	2.28	2.91	6.80	27.27	6.78
4	Male Avg Loss	2.42	2.97	2.31	3.29	12.60	4.91
5	Female Avg Loss	0.88	1.88	2.21	5.93	18.70	4.80
5	Male Avg Loss	1.39	1.49	2.02	4.79	10.98	3.95

piece rate, at a large cost. The top females lose \$27.27 in expected value compared to \$12.60 for the top males. In contrast, low ability males make only slightly more costly decisions than low ability females, averaging \$2.42 versus \$1.58 for the lowest types of each gender. We find that high ability females and high ability males are not entering competitive environments enough. But the high ability females overwhelmingly select the noncompetitive environments of piece rate and group pay, which are very costly decisions. In contrast, too many low ability males are entering competitive environments, but these mistakes are not particularly costly, on average, because low ability males would not perform well in the piece rate either.

Table 7 also shows that relative performance feedback decreases the average expected value losses for both males and females and shrinks the gender gap as well. The decreases in expected value losses are greatest for high ability females, whose average expected loss fell from \$27.27 in Treatment 4 to \$18.70 in Treatment 5, while losses for high ability males fell from \$12.60 to \$10.98. Low ability females and males tend to move towards group pay as they get performance feedback. While a gender difference remains, with low ability males making more expensive mistakes than women, the cost differences are small.

In Figure 5, we turn to the question of how relative choice information affects the choices of high ability females and males. A high ability individual is defined as an individual who should enter a four person tournament or larger to maximize expected returns from competition. Figure 5 shows that the relative performance information leads to a large increase in the proportion of high ability females entering tournaments. Over 50% of high ability females enter tournaments when given relative performance feedback, which is significantly more than the 31% that choose tournaments before receiving the performance feedback. In testing for distributional changes, we find that there is a significant difference in choices for females between Treatment 4 and Treatment 5; using a Pearson chi-square test the level of significance is $p = 0.034$.

Figure 5: Information Effects for Decisions by High Ability Types

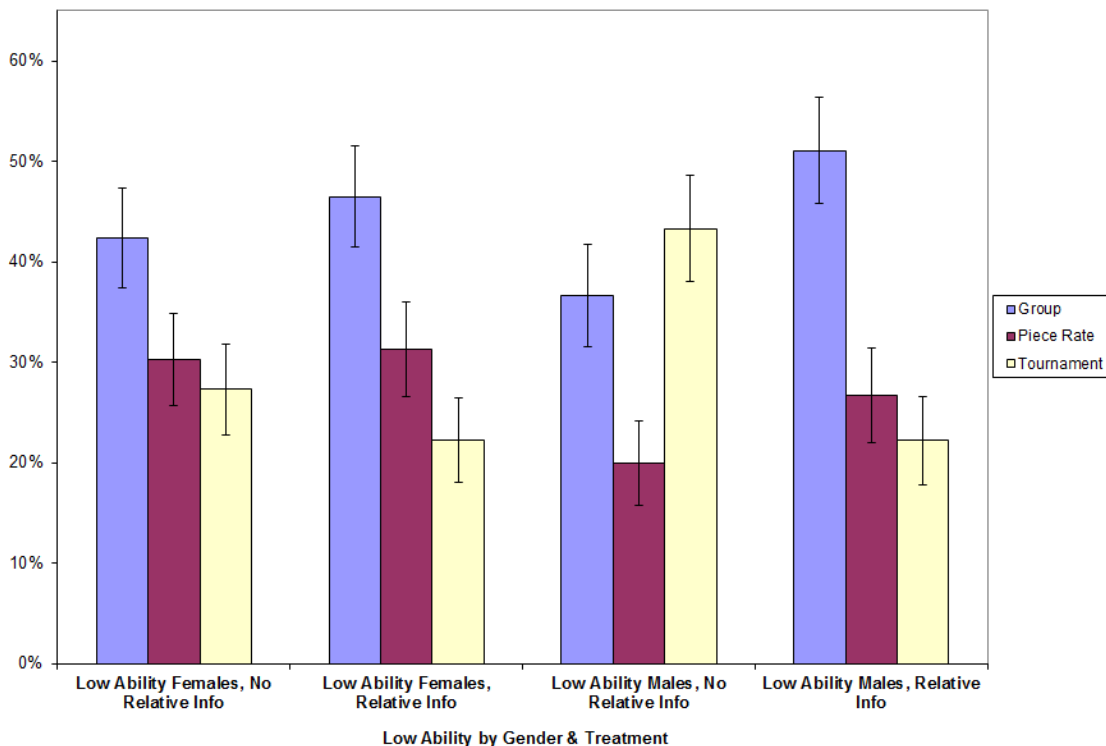


Females (45), Males (50). Sample size in parentheses.

With information, fewer high ability males enter tournaments (12% less), but this change in tournament selection is not statistically significant (at a 5% significance level). The distributional difference of choices for high ability males coming from information feedback is not significant as a chi-square test comparing high ability males between treatments produces a level of significance of $p = 0.317$. Without feedback in Treatment 4, there is a significant difference in the distributions of competitive choices between males and females ($p = 0.000$). After receiving feedback as the level of significance using a chi^2 test is $p = 0.158$. Thus, relative performance feedback seems to eliminate most of the differences in choices between the high ability females and high ability males.

Figure 6 shows the effect of relative performance information on choices by low ability types, where low ability is defined as those individuals with performance below the median in their respective session from Treatment 1. The largest effects are for males. Information drops the percentage of low ability males choosing tournaments from 43% to 22%, and increases the percentage of low ability males choosing group pay from 37% to 51%. For low ability males, the difference in the distribution of competitive choices between Treatment 4 and Treatment 5 is significant at a

Figure 6: Information Effects for Decisions by Low Ability Types



Females (99), Males (90). Sample size in parentheses.

$p = 0.010$ with a chi-square test. No such significant difference occurs for low ability females. The distributions of choices are significantly different for low ability females and males in Treatment 4 as chi-square test lead to a $p = 0.054$. But in Treatment 5 there are no significant differences between distributions for low ability females and males.

Information about relative performance moves high ability females towards more competitive choices and low ability males away from tournaments towards less competitive types of pay. Low ability females show only a small movement away from group pay towards piece rate. Overall, providing relative performance feedback information leads to more efficient sorting by both genders.

2.5 Competitiveness of Females Differs Between Low and High Hormone Phases

Normal cycling women experience large changes in hormone levels across the menstrual cycle. Hormonal contraceptives deliver hormones exogenously, suppressing the endogenous production of hormones that occurs across the natural cycle. Both groups of females experience a low hormone phase, during menses for normal cycling females and during the placebo (non-use) phase for females

using hormonal contraceptives. We look for differences in competitive behavior across these phases. This section starts with an explanation of how we determined the hormonal phase for the females subjects. We then show the differences in choices across phases.

We used a screening survey with retrospective questions on menses to schedule participants for a session during both a low-hormone phase and a high-hormone phase. The high-hormone phase of interest corresponds to the mid-luteal peak for normal cycling females (Figure 1). Females using hormonal contraceptives experience elevated exogenous hormones about three days into contraceptive use.

Table 8: Menstrual Cycle Regularity

Regularity of Period	Percent	Count
Identical	14.3%	55
Within 1-2 days	42.3%	163
Within 3-7 days	34.3%	132
Very Irregular (7+)	9.1%	35
Total		385
Missed Period in Last 3 Months	Percent	Count
Yes	14.7%	57
No	85.3%	330
Total		387

Numbers may not add up due to item non-response in screening survey.

Table 8 summarizes screening survey responses of females. Of the females who completed the screening survey almost 15% missed a menstrual period during the previous 3 months. Over 43% of these females experienced menstrual cycle irregularity of 3 days or more, suggesting that predicted menstrual periods may have significant measurement error. Due to the potential inaccuracies introduced by this prospective survey, we used an exit survey with both retrospective and prospective questions on the cycle, to help with the identification of hormonal phases.¹³

The screening survey also provided information on the proportion of females that use hormonal contraceptives. Over 54% of females in our sample used some form of hormonal contraceptive in the form of the pill or ring. This makes for easier predictability of low and high phases for these females, since hormonal fluctuations are exogenously determined by hormonal contraceptive use. To help identify hormonal phases for females using a hormonal contraceptive, we asked all female participants for the start day of their hormonal contraceptive regimen.

¹³Missed periods are a problem for identification purposes in normal cycling females as they imply that a female may not have ovulated during that month, and thus did not experience a mid-luteal peak in hormones. Furthermore, without a recent menstrual period it is difficult to determine the phase in the hormonal cycle.

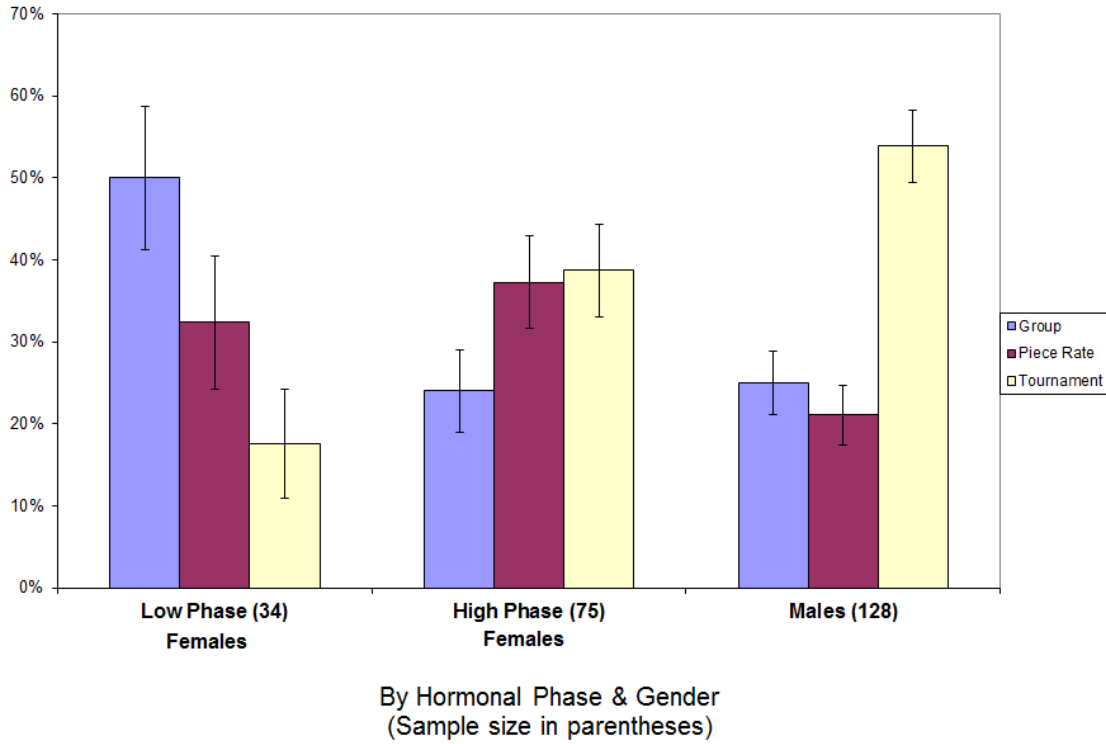
Of the females that participated in experiment sessions, 62.7% of those attending a first session were following a hormonal contraceptive regimen, as were 62.9% of those at second sessions. The American College Health Association found that about 72% of sexually active females were using some form of hormonal contraceptive in 2008 . In examining contraceptive use by females in the United States, it was found that for women between the ages of 15 to 44, over 82% had at one time taken oral hormonal contraceptives (Mosher et al. 2004), suggesting that our sample is not unusual in terms of contraceptive use.

Females during the placebo phase of the hormonal contraceptive regimen experience lower levels of estrogen and progesterone than during the active phase of their hormonal contraceptive regimen; this is due to the lack of exogenous hormones being provided (Speroff and Fritz 2005). Females with a naturally occurring menstrual cycle experience a drop in hormones during menses (Aden et al. 1998). We hypothesize that the low hormone phase, whether induced through endogenous or exogenous means, is associated with similar behavioral changes for both hormonal contraceptive users and normal cycling females. We tested this by including a control variable for hormonal contraceptive use and found no systematic significant difference in behavior between hormonal contraceptive users and normal cycling females. We therefore pool both groups of females and focus on differences across phases.

Figure 7 shows the distribution of competitive choices of females by phase, along with choices by males, before participants had relative performance feedback. Female behavior is very different in the two phases. They are more than twice as likely to choose group pay when they are in the low phase, and twice as likely to choose tournament when they are in the high phase, though still not as likely as men. When we include controls in regressions, this last difference will become insignificant. The data for the histogram includes all females and males that attended two sessions and all females who could be identified as being in the low or the high phase. Due to the difficulty of predicting the low phase, some females were identified by the exit survey as being in the same phase for both word and math tasks. As well, some phases could not be accurately identified and those subjects are not included in the analyses.

These differences in competitive environment choices across hormonal phases may result from differences in expected performance changes across the menstrual cycle or from different preferences for competition. We find that for the most part, there are no significant performance differences

Figure 7: Competitive Choice by Gender and Hormonal Phase
Subjects Attending Two Sessions

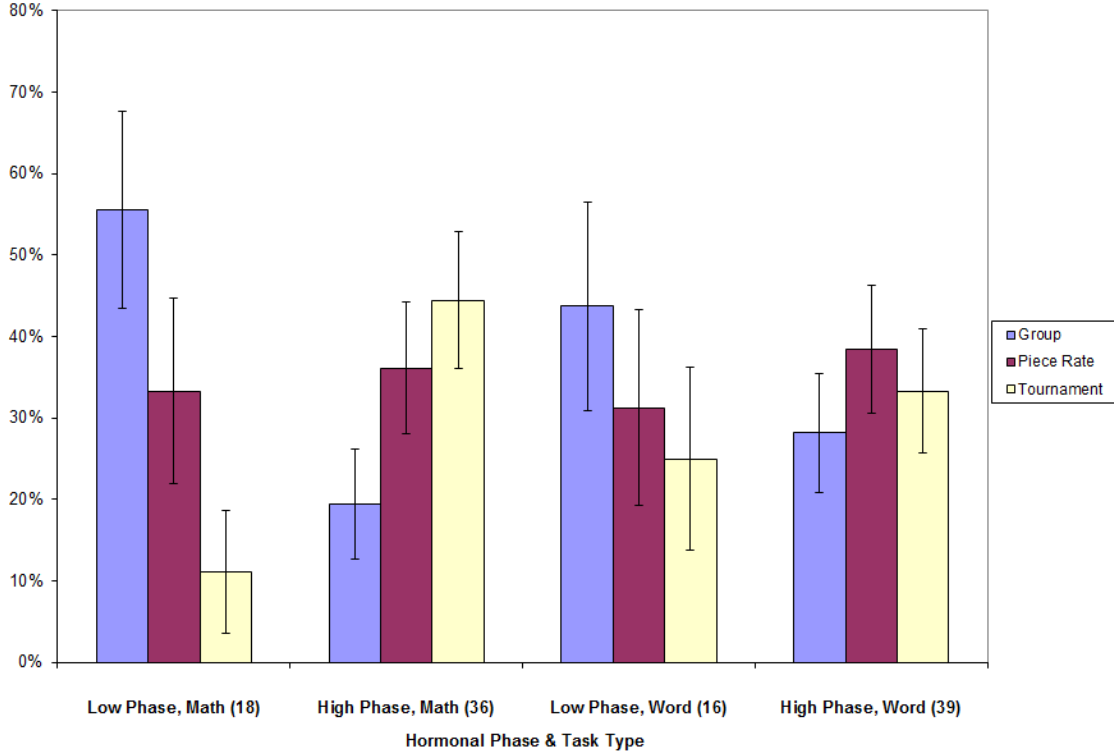


between females in the low phase and those that are not in the low phase.¹⁴ It is also possible that females in a specific hormonal phase might experience greater aversion to certain types of tasks; therefore, we separate out these results by math and word tasks. Figure 8 shows female compensation choices before feedback by hormonal phase and task type. Recall that we alternated the type of task used in sessions; females that participated in a math or word task during the low phase would then participate in the other type of task when in a higher hormonal phase (High Phase), and vice-versa, therefore these are averages across subjects. The figure shows that the general correlation between competitive choice and menstrual phase holds: high phase females are less likely to choose group pay and more likely to choose tournaments in both word and math tasks.

Next we use ordered probits to examine the statistical significance of gender and menstrual phase before feedback, while including control variables. Table 9 uses the CompScale variable, an ordered categorical variable with choices ranked from group pay, piece rate, to tournament. The first column provides pooled cross-sectional results including all subjects, the second to fourth

¹⁴See Appendix B.1.

Figure 8: Compensation Choice by Hormonal Phase and Task Type for Females.



We are comparing females that attended two sessions.
 Sample size in parentheses.

columns provide estimates using random effects ordered probit. The second column includes all males and females, the third column consists of a female only sample and the fourth column takes into account only males and females for which risk aversion measures were available.¹⁵

We find that females in the low phase select less competitive compensation plans than females in the high hormone phase. Without relative performance information, much of the difference between genders is driven by choices of the low phase females. This result holds even when controlling for confidence levels of individuals. It is worth noting that there are no significant differences in confidence levels between low hormonal phase and high hormonal phase females and yet females in the low phase avoid the competitive environments of tournaments and are more likely to choose the least competitive setting possible, group pay.

These differences could potentially result from discomfort during the low hormone phase of menstruation. But females in the low hormone phase do not behave differently from any other group once they receive relative performance feedback. Thus, physical discomfort is an unlikely

¹⁵All regressions include controls for session ordering, GPA, age, education, and psychoactive medications.

Table 9: Ordered Probit: Hormone Effects for No Relative Information (Treatment 4)

	(1)	(2)	(3)	(4)
Sample	All	All	Females Only	Risk
VARIABLES	Pooled	RE	RE	RE
Female	-0.26 (0.14) (*)	-0.29 (0.16) (*)		-0.26 (0.21) ()
Low Phase	-0.44 (0.21) (**)	-0.46 (0.22) (**)	-0.53 (0.26) (**)	-0.76 (0.27) (***)
Confidence (T1)	0.81 (0.26) (***)	0.91 (0.30) (***)	1.08 (0.54) (**)	0.90 (0.35) (**)
Improve (T2)	0.60 (0.20) (***)	0.69 (0.23) (***)	0.79 (0.38) (**)	0.72 (0.32) (**)
%-tile Rank (T1)	0.97 (0.23) (***)	0.99 (0.26) (***)	0.52 (0.43) ()	0.72 (0.32) (**)
Risk Controls	No	No	No	Yes
Characteristic Controls	Yes	Yes	Yes	Yes
Observations	328	328	155	211
ll	-322.3	-321.7	-156.0	-197.4
chi2	64.32	58.60	19.76	51.31

Dependent variable is CompScale where -1 is group pay, 0 is piece rate, and 1 is a tournament.

The total low phase females that could be identified for data analysis is 45.

Pooled means pooled cross section. RE means that random effects were used.

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

explanation for these systematic differences in low information settings.

The magnitudes of the marginal effects (calculated using the pooled cross sectional estimates) of being in the low hormone phase are substantial and are larger than the average gender effects. For group pay, females on average have a 0.08 higher probability of choosing group pay than males. Females in the low phase have an additional 0.16 higher probability of choosing group pay. For tournaments, females have a 0.10 lower probability of choosing a tournament when compared to males, and females in the low hormone phase have an additional 0.16 decrease in the probability of choosing a tournament.

These hormonal effects can also be compared to confidence and performance. For a female in the low phase to have the same probability of entering a tournament as a female in the high phase, we would have to increase her belief about her performance relative to the average by 50%.

In terms of an equivalent performance effect, a female in the low hormone phase would have to improve her percentile rank by 42% to be as likely to enter a tournament as a female in the high hormone phase. These differences across the menstrual cycle are not just significant, they are also large relative to other effects.

Relative performance feedback makes these cycle specific effects disappear. Table 10 provides the results from ordered probit estimations for Treatment 5, where subjects were provided with relative performance information from Treatment 1 prior to making their competitive environment selections. Table 10 shows that when participants are informed of their relative performance compared to other potential competitors, then there is little difference in selection between genders or across the menstrual cycle.

As with the gender differences, we find that after participants are informed of the quality of potential competitors choice differences across the menstrual cycle become insignificant. We find that choices with feedback mainly depend on the relative performance information provided prior to making the decision and, to a certain extent, on an individual's improvement from Treatment 1 to Treatment 2. Though females' choices to select less competitive environments are most frequent in the low hormone phase, these results suggest that this effect seems to be linked with the information available about the quality of potential competitors.

As discussed previously, there is a cost associated with high ability individuals avoiding competitive settings, and with low ability individuals choosing tournaments. We find that females in the low hormone phase make more costly mistakes than high phase females, and males.¹⁶ The average expected value losses for males, low phase and high phase females in Treatment 4 are shown in Table 11. Low phase females sacrifice the greatest amount of expected value from making suboptimal choices, \$8.50. The expected value losses for high phase females and males are \$6.52 and \$4.91. The differences between low and high phase females and between males and high phase females are not statistically significant, but low phase females make more costly choices than males at the 5% significance level.

These results imply that hormones may matter in the selection of competitive environments, but only if the strength of the competition or the probability of winning is relatively unknown. If there is little information, then females in the low-hormone phase make more costly decisions than males and non-low phase females. But there are no significant differences in expected value losses

¹⁶These expected values are calculated in the same way as discussed in the previous section.

Table 10: Ordered Probit: Hormone Effects after Feedback (Treatment 5)

	(1)	(2)	(3)	(4)
Sample	All	All	Females Only	Risk
VARIABLES	Pooled	RE	RE	RE
Female	-0.07 (0.15) ()	-0.13 (0.20) ()		0.10 (0.24) ()
Low Phase	0.03 (0.21) ()	0.13 (0.25) ()	0.12 (0.27) ()	-0.12 (0.29) ()
Confidence (T1)	0.23 (0.25) ()	0.32 (0.31) ()	0.21 (0.53) ()	0.54 (0.36) ()
Improve (T2)	0.76 (0.20) (***)	0.92 (0.26) (***)	1.06 (0.40) (***)	0.49 (0.33) ()
%-tile Rank (T1)	2.18 (0.25) (***)	2.61 (0.35) (***)	2.63 (0.55) (***)	2.33 (0.38) (***)
Risk Controls	No	No	No	Yes
Characteristic Controls	Yes	Yes	Yes	Yes
Observations	328	328	155	211
ll	-307.8	-303.9	-143.3	-183.4
chi2	104.7	93.82	45.89	75.79

We identified 45 females as low phase for this analysis.

Pooled is pooled cross section, RE is random effects.

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

between genders or between different hormonal phases for females if good relative performance information is available.

2.6 Systematic Absenteeism, Cancellations, and Tardiness

Absenteeism, cancellations, and tardiness are frequent in experiments, but their effect on sample composition and results are poorly understood. In this study, because of the screening and exit surveys, we know some of the characteristics of those who missed a scheduled session, canceled at the last moment, or showed up late. Our recruiting procedures were designed to ensure we would observe females during both the high phase and the shorter low hormone phase. Due to the variation of the menstrual cycle, it is difficult to predict the low hormonal phase for females. Using the screening survey, female subjects were scheduled for their sessions according to their

Table 11: Expected Value Loss in Treatment 4

	Mean	Std. Error
Male	4.91	0.72
Female Non-Low	6.52	1.30
Female Low	8.50	2.57

predicted cycle day, calculated using self reported data about the start of previous menstrual periods. Whenever possible these data were combined with self reported data concerning females' hormonal contraceptive regimens. Once the cycle day could be predicted, then a set of possible session days were provided to potential participants and they chose and confirmed a day with a research assistant. For individuals to be considered as scheduled, they had to confirm that they would attend a specific session.

Table 12: Session Attendance After Confirmation

Gender		
	Proportion	N
Male	0.79	217
Female	0.68	243
Total	0.73	460
Females by Predicted Phase		
	Proportion	N
High	0.72	141
Low	0.62	102
Total	0.68	246

Table 12 shows the proportion of participants that attended the experiment sessions as scheduled, meaning they were present and punctual. Females are noticeably less likely to show up than males: 79% of the males showed up as scheduled compared to 68% of females. Based on predicted phases, only 62% of low phase females attended as scheduled, while high phase females attended 72% of the time. These differences between attendance rates of low phase and non-low phase females were significant at the 5% level. High phase females were less likely to attend compared to males, but this was significant only at the 10% level. Since this study was partly focused on hormonal fluctuations we made attempts to incentivize more low phase females to attend. Part way through the study, the participation payment of \$5 was raised to \$10 for low phase females.¹⁷ Even with this increase in participation payments a hormonal phase related difference in attendance remained. This result is consistent with Ichino and Moretti (2009), who found that female worker absenteeism was highly

¹⁷These females were not told why they were receiving a higher participation payment.

correlated with the female menstrual cycle.

These attendance results suggest there may be a systematic bias in the hormonal phase of females who show up to experiments. Coupled with our evidence of behavioral differences, this bias should be considered in interpreting the results of experiments that report gender differences. Additionally, if females who do not show up to sessions are the ones who have worse symptoms during the low phase, or are more likely to behave differently, then a selection bias may add a downward bias in the hormonal effects found in this study. Though one should still note that any such differences are removed with relative performance feedback.

2.7 Confidence and Robustness

It is possible that during the low hormone phase females are not as confident in their abilities as during elevated hormonal phases. Here we show that confidence differences do not explain the choice differences across phases. Table 13 shows that there are no significant differences in confidence for females across phases. There are also no significant differences between females using or not using hormonal contraceptives. Thus, differences in confidence do not seem to explain why females in the low hormonal phase choose not to compete without relative performance feedback.

Table 13: Confidence Levels of Females by Hormonal Phase

Confidence (T1)				
Phase	Task	Mean	Std. Error	N
High	Math	1.068	0.032	58
Low	Math	1.041	0.078	23
Both	Math	1.060	0.031	81
High	Word	1.021	0.031	54
Low	Word	1.042	0.057	22
Both	Word	1.027	0.028	76

The sample consists of all females whose phase could be accurately identified.

2.8 Robustness Checks

We conducted a variety of robustness checks. First, we checked to see whether the results for gender differences in competition and differences by hormonal phase remain consistent when clustering

standard errors on experiment sessions and then also separately on individuals. We found that the results of both the gender and hormonal effects are robust to such error clustering.¹⁸

Individuals in this study were asked to attend two sessions. Out of 219 total subjects, 126 attended two sessions and 93 attended only a single session. Session composition was random, sessions contained individuals that were attending sessions for either a first or a second time. The type of task individuals were using to compete was different for the first and second session, and the ordering was random. Neither order, the type of task, or group composition in terms of these measures had significant performance or behavioral effects.

The possibility exists that attendance in the second session was affected by how much subjects earned in the first session. Individuals that attended only a single session earned \$9.36 for their performance, while individuals that returned for a second session were paid \$11.73 for their first session performance. This difference in payouts between these groups of individuals is not statistically significant.

Another possibility is that the value of payment in the first session affects confidence in the second session. We exploit the random element in the payment scheme and estimate that, controlling for performance, a person who earns \$10 more than average in the first session increases their prediction of their performance relative to the session average by just 4%. Looking just at the second session females, we find that their confidence in the second session is more sensitive to how much they got paid in the first session than is that of males. But we find no significant difference between females' and males' confidence when attending a second session, or across menstrual cycle phase. Any such difference should in fact make females more competitive in second sessions, but we find no systematic behavior in this direction.

It is still possible that there is an interaction between payments and gender driving choice differences. Table 14 provides ordered probit estimations to examine whether females are still less likely to compete when controlling for payments in the first session. Though the first session payout variable is significant at the 10% level, its inclusion does not affect the significance of the gender difference of females to avoid competition.

Results seem consistent for both the sample where risk aversion measures were available and for the sample of participants that did not participate in the risk aversion task. The third column in Table 14 provides the coefficients from the ordered probit estimation for the full information

¹⁸Results are available upon request.

Table 14: Robustness Check for Gender: Controlling for First Session Payment

VARIABLES	(1)	(2)	(3)
	RE CompScale	RE & Risk CompScale	RE & Full Info (T5) CompScale
Female	-0.40 (0.156) ***	-0.50 (0.201) **	-0.02 (0.178)
Confidence (T1)	1.00 (0.301) ***	1.01 (0.361) ***	0.53 (0.309) *
Improve (T2)	0.78 (0.245) ***	0.81 (0.333) **	1.10 (0.276) ***
%-tile Rank (T1)	1.08 (0.264) ***	0.82 (0.333) **	2.56 (0.345) ***
First Session Payout	0.02 (0.009) *	0.02 (0.009) *	0.00 (0.008)
Risk Controls	No	Yes	No
Characteristic Controls	Yes	Yes	Yes
Observations	343	224	343
chi2	65.63	52.89	100.7
ll	-333.3	-210.2	-315.6

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

treatment (Treatment 5). Again, with feedback we find that gender does not matter in this setting and neither does the payout from the first session. Thus it seems that the gender difference for competition in low information settings is not affected by the payment received in a previous session.

We now use similar estimations to test whether the hormone specific results are sensitive to the inclusion of a payment variable. We find no significant differences in confidence or payouts in the first session between females who could be identified as low phase and the females that were in a second session and were in the high hormone phase. The first column of 15 provides ordered probit coefficients for the entire sample of individuals, including both those that attended a single session and those that attended two sessions. In this sample, we find that the first session payout variable is significant at the 10% level, but the effect is quite small. The effect of females being in the low hormonal phase remains large and significant at the 5% level.

In the second column in Table 15 we limit the sample to those individuals who participated in the risk aversion task and therefore attended two sessions. In this sample, while controlling for risk

Table 15: Robustness Check for Hormones: Controlling for First Session Payment

VARIABLES	(1)	(2)
	RE CompScale	RE & Risk CompScale
Female	-0.29 (0.167) *	-0.27 (0.212)
Low Phase	-0.44 (0.224) **	-0.74 (0.275) ***
Confidence (T1)	0.95 (0.306) ***	0.94 (0.366) **
Improve (T2)	0.74 (0.245) ***	0.77 (0.328) **
%-tile Rank (T1)	0.98 (0.261) ***	0.68 (0.329) **
First Session Payout	0.02 (0.009) *	0.01 (0.009)
Risk Controls	No	Yes
Characteristic Controls	Yes	Yes
Observations	328	211
chi2	62.07	53.99
ll	-320.0	-196.0

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

aversion, we find that the effect of the low hormonal phase is large and significant, but that the level of the payout from the first session is insignificant. As well, in this group of individuals who did the risk aversion task, we find that the gender effect is insignificant, but that the effect of the low phase for females remains significant.

Thus, these checks of the effects of first session payouts on second session attendance and behavior show that some of the gender difference in competitive choices may be entirely driven by females in the low hormone phase of their cycle. They further confirm that the gender and menstrual phase differences are only relevant in low information settings. In the following section we discuss the results and implications of our findings.

3 Discussion

We show that the gender differences in compensation choices that have been reported in other studies are robust to a variety of protocol changes, including different tasks and variations in the degree of competitiveness of the available choices. These gender differences are stable, persisting in a second session of the experiment with a different task, performed days or weeks after the first session. We also find that female choices in a mixed gender setting vary across the menstrual cycle. In the low hormone phase, females are less likely to enter tournaments than during a high hormone phase. All these differences for competitive environment selections become noticeably smaller and statistically insignificant after subjects are provided with feedback on their relative performance.

Previous studies have shown that gender differences in compensation choices remain even after controlling for confidence and performance. The general conclusion has been that an underlying gender difference in preferences for competition must be driving the result. But if the choice differences originate from different preferences for competitive environments, then selection differences should remain even after relative performance feedback is provided, in our experiment they do not.

Further, we find gender differences in competitive choices in both math and word tasks. In the word tasks, confidence levels of females and males are the same and a gender difference in choices persists. In the math task the gender difference persists even after controlling for confidence. The information about relative abilities may help subjects form more accurate beliefs about the possibilities of succeeding in competitive situations, but since females hold similar overconfident beliefs as males in word tasks, we would expect females to choose in a similar fashion as males.

Instead, females choose not to compete in low information settings, irrespective of confidence or the type of task.

Another possible explanation for the difference in gender behavior with and without feedback would be a gender difference in risk aversion or beliefs. However, we control for risk aversion and beliefs about relative ability and find that the gender difference still remains, in the uninformed treatment. Since beliefs are updated and become more accurate in the full information treatment, we should find that risk aversion matters in such a setting, we do not find that result. Another explanation that would be consistent with our results would be a gender difference in ambiguity aversion rather than risk aversion. There is little support in the literature for a stable gender difference in ambiguity.

Our finding of a variation in an important economic behavior across the menstrual cycle is consistent with a large endocrinological literature showing the importance of hormonal variations to behavioral differences within females. There are neurological reasons to expect such an effect, for example higher levels of hormones such as estrogen have been found to increase activation areas of the brain that are related to anticipation of uncertain rewards. One other paper, Buser (2010) looks at competition and the menstrual cycle. Using between subject differences in females competing against other females, he finds that in the low hormone phase females compete more. Other studies have found that female behavior in single gender groups is very different from that in mixed gender groups (Datta Gupta et al. 2005, Grosse and Reiner 2010)– females are generally much more willing to compete when the group is all female.

We find that, in mixed gender groups, both the gender and hormone specific effects that both effects are removed with information. In the mixed gender setting, relative performance feedback leads more high ability females to compete and more low ability males to less competitive settings. Thus, the gender composition of the particular labor market being examined may play an important role in regards to the effects of relative performance feedback and decisions to enter more competitive settings by both males and females. Our within-subject design has obvious advantages to an across subjects approach. The potential disadvantage would be that experiences in the first session might affect choices in the second. We use two different tasks to minimize this problem, and by showing the results are robust to order and these tasks, we provide additional evidence of the strength of the effect. Of course, as the gender composition effects show, competitive entry is very much context dependent, a result that is consistent with other findings in the neuroendocrinological

literature (Anders and Watson 2006).

The existence of a regular and predictable within-subjects variation in behavior across the menstrual cycle is a novel finding that raises many interesting questions for optimal choice. Women who know that they are subject to this variation might well take steps to increase or decrease their competitive choices by changing when they make decisions involving competition, such as job choices and college applications. The fact that hormonal contraceptives allow for control of the timing of the low hormone phase provides another potential control mechanism.

Interestingly, there is no medical reason for the low hormone (placebo) phase of the hormonal contraceptive regimen (Anderson and Hait 2003). As a leading textbook on clinical gynecological endocrinology states: *Monthly bleeding, periodic bleeding, or no bleeding—this is an individual woman’s choice (Speroff and Fritz 2005, pg 908)*. It has been suggested that the placebo and withdrawal bleeding was a marketing effort to make the birth control pill seem less novel and more acceptable (Coutinho and Segal 1999). In fact, females can entirely avoid hormonal fluctuations and the low hormone phase by sustained contraceptive use. Indeed, some forms of contraception, such as the oral contraceptive *Seasonale*, now eliminate the placebo and ensure that low hormone phase does not occur (Anderson and Hait 2003). If the decrease in hormones affects females in a costly manner then there is an incentive for avoiding this phase. Women who are unaware of the behavioral effects of hormonal variations, or who are unable to commit to these timing or contraceptive changes, might find themselves experiencing costly regret. Of course, our results also show that such behaviors would only matter when there is poor information about ability relative to potential competitors.

Firms, governments, and schools sometimes implement affirmative action policies to encourage females to apply for competitive jobs and scholarships. Affirmative action policies typically focus efforts on recruiting females or changing the acceptance or promotion process to favor females. Niederle et al. (2009) report on experiments showing affirmative action can encourage competition by females at low efficiency cost. The performance feedback condition result from our study suggests an alternative to affirmative action— providing better information about relative abilities.

We find that information settings play the greatest role in gender differences to compete. In low information settings females, and more specifically, females in the low hormone phase chose to be less competitive. This behavior was consistent in second iterations of the experiment sessions by the same subjects. This was also consistent regardless of whether individuals were choosing

to compete in a math or word task. In all these variations of the experimental setting, the effect of relative performance feedback removed any differences between the groups. This information did more than just remove differences in confidence and the effects of those differences. The fact that both relative performance feedback and menstrual cycle variations affect competitive choices suggests that both information and biological factors contribute to gender differences in competitive choice. Further investigation of these factors holds the promise for a better understanding of the mechanisms behind these gender differences.

Prior to this study, three distinct methods were used to explain gender differences in competitive choices. Gneezy et al. (2009) showed that gender effect actually reversed in matrilineal societies. Females have also been found to perform very differently in all female competitions (Gneezy et al. 2003); suggesting that restricting the gender of potential competitors may also remove gender differences in choices to compete. Affirmative action policies have been shown to help remove gender differences to compete with little efficiency cost (Niederle et al. 2009). We have shown that the hormonal cycle of females partly explains gender differences for competition. But, if individuals are given relative performance feedback prior to making competitive decisions, the gender difference shrinks and becomes insignificant, and differences across the hormonal cycle for women are also eliminated. Thus, this study has shown the importance of biological factors for the gender difference in competitive choices, and has also shown that good information helps remove the gender difference. These results are found using a within-subjects design and they hold over multiple sessions and with both math and word tasks.

References

- Aden, U., Jung-Hoffmann, C. and Kuhl, H.: 1998, A randomized cross-over study on various hormonal parameters of two triphasic oral contraceptives, *Contraception* **58**(2), 75–81.
- American College Health Association: 2009, ACHA National College Health Assessment II: Reference Group Data Report, *Technical report*, ACHA.
- Anders, S. and Watson, N.: 2006, Social Neuroendocrinology: Effects of Social Contexts and Behaviors on Sex Steroids in Humans, *Human Nature* **17**(2), 212–237.
- Anderson, F. and Hait, H.: 2003, A multicenter, randomized study of an extended cycle oral contraceptive, *Contraception* **68**(2), 89–96.
- Apicella, C., Dreber, A., Campbell, B., Gray, P., Hoffman, M. and Little, A.: 2008, Testosterone and financial risk preferences, *Evolution and Human Behavior* **29**(6), 384–390.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U. and Fehr, E.: 2008, Oxytocin shapes the neural circuitry of trust and trust adaptation in humans, *Neuron* **58**(4), 639–650.
- Bertrand, M. and Hallock, K.: 2001, The gender gap in top corporate jobs, *Industrial and Labor Relations Review* pp. 3–21.
- Bethea, C., Lu, N., Gundlach, C. and Streicher, J.: 2002, Diverse Actions of Ovarian Steroids in the Serotonin Neural System, *Frontiers in Neuroendocrinology* **23**(1), 41–100.
- Borghans, L., Golsteyn, B., Heckman, J. and Meijers, H.: 2009, Gender Differences in Risk Aversion and Ambiguity Aversion, *Journal of the European Economic Association* **7**(2-3), 649–658.
- Brown, C.: 1990, Firms’ choice of method of pay, *Industrial and Labor Relations Review* pp. 165–182.
- Burnham, T.: 2007, High-testosterone men reject low ultimatum game offers, *Proceedings of the Royal Society B: Biological Sciences* **274**(1623), 2327.
- Buser, T.: 2010, The impact of female sex hormones on competitiveness, *Unpublished manuscript* .
- Chakraborty, T., Ng, L. and Gore, A.: 2003, Age-Related Changes in Estrogen Receptor {beta} in Rat Hypothalamus: A Quantitative Analysis, *Endocrinology* **144**(9), 4164.
- Chen, Y., Katuscak, P. and Ozdenoren, E.: 2009, Why Can’t a Woman Bid More Like a Man?, *Working Paper*, *Ann Arbor* **1001**, 48109–1220.
- Coates, J. and Herbert, J.: 2008, Endogenous steroids and financial risk taking on a London trading floor, *Proceedings of the National Academy of Sciences* **105**(16), 6167.
- Coutinho, E. and Segal, S.: 1999, *Is menstruation obsolete?*, Oxford University Press, USA.
- Creinin, M., Keverline, S. and Meyn, L.: 2004, How regular is regular? An analysis of menstrual cycle regularity, *Contraception* **70**(4), 289–292.
- Delfgaauw, J., Dur, R., Sol, J. and Verbeke, W.: 2009, Tournament Incentives in the Field: Gender Differences in the Workplace, *IZA Discussion Papers* .
- Dohmen, T. and Falk, A.: 2007, Performance pay and multi-dimensional sorting: productivity, preferences and gender, *Institute for the Study of Labor (IZA); CESifo (Center for Economic Studies and Ifo Institute for Economic Research)* .

- Dreher, J., Schmidt, P., Kohn, P., Furman, D., Rubinow, D. and Berman, K.: 2007, Menstrual cycle phase modulates reward-related neural function in women, *Proceedings of the National Academy of Sciences* **104**(7), 2465.
- Fehr, E.: 2009, On the economics and biology of trust, *Journal of the European Economic Association* **7**(2-3), 235–266.
- Fernandez, G., Weis, S., Stoffel-Wagner, B., Tendolkar, I., Reuber, M., Beyenburg, S., Klaver, P., Fell, J., de Greiff, A., Ruhlmann, J. et al.: 2003, Menstrual cycle-dependent neural plasticity in the adult human brain is hormone, task, and region specific, *Journal of Neuroscience* **23**(9), 3790.
- Fink, G., Sumner, B., Rosie, R., Grace, O. and Quinn, J.: 1996, Estrogen control of central neurotransmission: effect on mood, mental state, and memory., *Cell Mol Neurobiol* **16**(3), 325–44.
- Gagnidze, K. and Pfaff, D.: 2009, Sex on the Brain, *Cell* **139**(1), 19–21.
- Gneezy, U., Leonard, K. and List, J.: 2009, Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society, *Econometrica* **77**(5), 1637–1664.
- Gneezy, U., Niederle, M. and Rustichini, A.: 2003, Performance in Competitive Environments: Gender Differences, *Quarterly Journal of Economics* **118**(3), 1049–1074.
- Gneezy, U. and Rustichini, A.: 2004, Gender and Competition at a Young Age, *American Economic Review* **94**(2), 377–381.
- Grosse, N. and Riener, G.: 2010, Explaining Gender Differences in Competitiveness: Gender-Task Stereotypes, *Jena Economic Research Papers* **2010**, 017.
- Grossman, P. and Eckel, C.: 2003, Men, women and risk aversion: Experimental evidence, *Unpublished manuscript* .
- Gunther, C., Ekinci, N., Schwier, C. and Strobel, M.: 2008, Women can't jump?—An experiment on competitive attitudes and stereotype threat, *Technical report*, Mimeo.
- Gupta, N., Poulsen, A. and Villeval, M.: 2005, Do (wo) men prefer (non-) competitive jobs, *Technical report*, GATE Working Paper.
- Gupta, N., Poulsen, A., Villeval, M. and Gade, H.: 2005, Male and Female Competitive Behavior- Experimental Evidence, *IZA Working Paper* .
- Holt, C. and Laury, S.: 2002, Risk Aversion and Incentive Effects, *American Economic Review* **92**(5), 1644–1655.
- Ichino, A. and Moretti, E.: 2009, Biological Gender Differences, Absenteeism and the Earnings Gap, *American Economic Journal: Applied Economics* **1**(1), 183–218.
- Jirjahn, U. and Stephan, G.: 2004, Gender, piece rates and wages: evidence from matched employer-employee data, *Cambridge Journal of Economics* **28**(5), 683–704.
- Malison, R., Price, L., Berman, R., van Dyck, C., Pelton, G., Carpenter, L., Sanacora, G., Owens, M., Nemeroff, C., Rajeevan, N. et al.: 1998, Reduced brain serotonin transporter availability in major depression as measured by [123I]-2 β -carbomethoxy-3 β -(4-iodophenyl) tropane and single photon emission computed tomography, *Biological Psychiatry* **44**(11), 1090–1098.
- McEwen, B.: 2002, Estrogen Actions Throughout the Brain, *Recent Progress in Hormone Research*

- 57(1), 357–384.
- Moore, E. and Eckel, C.: 2003, Measuring Ambiguity Aversion. Virginia Tech.
- Mosher, W., Martinez, G., Chandra, A., Abma, J. and Willson, S.: 2004, Use of contraception and use of family planning services in the United States: 1982–2002, *Advance data from vital and health statistics* **350**, 1.
- Niederle, M., Segal, C. and Vesterlund, L.: 2009, How costly is diversity? Affirmative action in light of gender differences in competitiveness, *NBER Working Paper* .
- 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.: 2010, Explaining the Gender Gap in Math Test Scores: The Role of Competition, *The Journal of Economic Perspectives* **24**(2), 129–144.
- Pearson, M. and Schipper, B.: 2009, Menstrual cycle and competitive bidding, *MPRA Paper* .
- Platt, M. and Huettel, S.: 2008, Risky business: the neuroeconomics of decision making under uncertainty., *Nat Neurosci* **11**(4), 398–403.
- Polachek, S.: 1981, Occupational self-selection: A human capital approach to sex differences in occupational structure, *The Review of Economics and Statistics* pp. 60–69.
- Shepard, K., Michopoulos, V., Toufexis, D. and Wilson, M.: 2009, Genetic, epigenetic and environmental impact on sex differences in social behavior, *Physiology & behavior* **97**(2), 157–170.
- Speroff, L. and Fritz, M.: 2005, *Clinical Gynecologic Endocrinology and Infertility*, Lippincott Williams & Wilkins.
- Stanton, S. and Schultheiss, O.: 2007, Basal and dynamic relationships between implicit power motivation and estradiol in women, *Hormones and Behavior* **52**(5), 571–580.
- Stricker, R., Eberhart, R., Chevallier, M., Quinn, F., Bischof, P. and Stricker, R.: 2006, Establishment of detailed reference values for luteinizing hormone, follicle stimulating hormone, estradiol, and progesterone during different phases of the menstrual cycle on the Abbott ARCHITECT analyzer., *Clinical Chemistry and Laboratory Medicine* **44**(7), 883.
- Vliet, E.: 2001, *Screaming to be heard: hormone connections women suspect and doctors still ignore*, M. Evans and Co., New York, NY.
- Wu, M., Manoli, D., Fraser, E., Coats, J., Tollkuhn, J., Honda, S., Harada, N. and Shah, N.: 2009, Estrogen masculinizes neural pathways and sex-specific behaviors, *Cell* **139**(1), 61–72.
- Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Hirschberg, A. and Johannesson, M.: 2009, A randomized trial of the effect of estrogen and testosterone on economic behavior, *Proceedings of the National Academy of Sciences* **106**(16), 6535.

A (Not For Publication) Appendix I: About controls

A number of control variables were used in this analysis including whether a subject was attending a first or second session, or a word based session, neither were ever significant. Also, other individual characteristic control variables such as the number of years of post secondary schooling, age and GPA were included, but were never significant with the exception of number of years of post secondary schooling, which was negative and significant. In estimations not included here categorical dummy variables were included for the majors of students, but none were significant and were removed due to their irrelevancy.

The other set of controls that are not listed are measures controlling for risk aversion. After participating in both a math and word session, subjects were invited to participate in a Holt and Laury (2002) task a few days after the competition experiment sessions. These risk aversion sessions were done separately so as not to affect behavior in the competition task. Consequently, the sample of these individuals is slightly different from the other specifications, but the results remain fairly consistent. The risk aversion coefficients were not found to be significant in the treatment with no relative information feedback.¹⁹ The only specification where these measures are significant is in the full information treatment for a limited sample of high ability individuals. The fact this measure does not significantly predict behavior in the no information treatment is intuitive as individuals have little information regarding their probabilities of success; thus, beliefs may matter more in such situations than risk preferences. Since risk aversion is significant for high ability individuals in the full information treatment; this result suggests that risk appetites matter only to individuals that have higher returns from making an optimal choice. For those individuals of low ability, their expected value is higher by avoiding the tournament regardless of risk preferences.²⁰

Three variables that are used and presented in the tables include, Confidence (T1), Improve (T2), %-tile Rank (T1). These variables deserve further explanation as they are measuring some very relevant factors in the individual decision making process. In other estimations, variations of these variables were used, but these variables provide the cleanest and most intuitive interpretation possible and results remain consistent with other possible measures of the same factors.

¹⁹For brevity these results are not included in this discussion, but are available upon request.

²⁰There were no significant gender differences for risk aversion in the sample.

B (Not For Publication) Appendix II: Performance

The main focus of this paper is on the compensation choices that individuals made in the experiment; however, these choices may have been affected by performance differences. This section focuses on task performance for the different treatments. To consider how individuals are affected by the different incentives of each type of compensation, we focus on the performance of individuals in the first three treatments. In these first three treatments, individuals had no choice over the type of compensation they received for their efforts; thus, the performance effects from different compensation environments are exogenously determined.

According to the theory of piece rates and tournaments, one would expect greater effort for a higher piece rate. Similarly, an individual of higher ability and higher probability of winning should increase effort in a tournament. As the tournament gets larger and more competitive, one would expect that individuals would increase effort or set their effort levels to zero. Before considering the effects of tournament size on effort, we first focus on possible order effects and gender differences between treatments.

The regressions in Table 16 are used to consider gender differences in performance, learning effects and the incentive effects of increasing tournament size. The performance in the word task, but not the math task, is highly correlated with the GPA of participants. Regression estimates for both word and math show an order effect suggesting that subjects are learning in the first three treatments. Regression 1 in Table 16 shows that the tournament size has a statistically significant positive effect on individual performance in the math task. Increasing the competitiveness of a compensation environment from the piece rate to a tournament size of 6 should increase performance of an individual by 0.65 problems in a four-minute task. This is an increase of 5.7% for the average individual. In columns 2, 4, and 5, categorical variables are used to investigate whether tournament size is actually leading to the increase or whether just competing against someone in a tournament of any size leads to performance increases.

In the second column in Table 16, categorical variables were used for each of the possible competitive environments, for group pay and for tournaments. To test whether the tournament size matters a separate dummy variable, *Tourney* ($ts > 1$) was used to identify if an individual had to compete against someone else. This categorization created a separate baseline for tournaments consisting of six individuals. Once controlling for tournament competition it was found that the

size of the tournament does not matter and that group pay performance is not significantly different from the piece rate environment in the math task. There is also a positive effect for age as the average effect of a year of life leads to an increase in performance in the math task of 0.5 problems, though this may be offset by further post-secondary schooling. It is worth noting that overall the competitive environment and individual characteristics explain very little of the variation in performance for the math task in terms of goodness of fit measures such as R^2 .

The results from the word task (Regressions 3 and 4) in Table 16 suggest that neither the tournament nor tournament size increase performance. There is a significant amount of learning that occurs with each treatment. GPA has a significant positive effect in terms of performance. This likely occurs because an individual's vocabulary expands with age and individuals with a higher GPA probably have richer vocabularies than individuals with lower GPAs. More of the variation in performance can be explained in regressions using the word task than the math task; this mainly stems from the inclusion of control variables for the random letters used for each task.

The math task results suggest that being in a tournament does increase performance, but the size of the tournament is irrelevant. One might expect that only high ability individuals would increase performance from the incentive effects from being in a tournament, but we find the opposite. In splitting the sample for high and low ability individuals, according to their performance in the first task and whether they are above or below the median, we find that the low ability individuals increase performance in response to being in a tournament (significant at 1%) in the math task. We find no significant effects from tournament size for high ability individuals in the math task.²¹

The competitiveness of the environment has a significant impact on performance only in math tasks and once an individual is participating in a tournament, then the number of competitors does not lead to further performance benefits. Competitive environments (tournament size) had no influence on the performance of individuals in the word task. Thus, depending on the type of task, competition between individuals may increase performance. Therefore, due to mixed results, one cannot conclude that tournament size increases performance or effort of agents.

Overall, we find that a more competitive work environment may not lead to performance increases as the incentive effects of competitiveness are not robust across different types of tasks. Another important result shown by these regressions is that there are no significant performance

²¹The estimation results for high ability individuals are not shown here, but none of the competitive environment variables were significant in these estimations.

differences between males and females. In terms of performance effects, some learning occurs across the different treatments and only low ability individuals tend to increase performance when they are put in a tournament of any size— it is enough to be competing against someone.

B.1 Performance differences according to hormonal phase.

To examine performance differences across the menstrual cycle, we consider the word and math tasks separately and estimate effects using linear specifications similar to the ones used to examine exogenous performance effects of tournaments (Appendix B). Table 17 provides the random effects OLS estimates for a number of factors that may explain performance differences in both word and math tasks for all treatments where the participants could not choose their competitive environments. The estimations for math tasks are in columns 1 to 3 and the estimations for word tasks are in columns 4 to 6.

There seems to be no correlation with the low-hormone phase and performance in the word task (columns 4 to 6) for females. The Low Phase coefficient is insignificant for all the different samples in the word task. Focusing on performance in the math task (columns 1 to 3), there seems to be no effect from the low phase in the sample of both females and males (column 1), and only females (column 2). There seems to be some effect for low phase females when including controls for risk aversion for the portion of the participants for which such measures were available. On average, performance in the math task decreases by about 2.2 correct answers for low phase females when controlling for risk aversion, which is a 20% decrease for the average female. Though the low phase effect is only significant at the 10% level when taking into account individuals for which measures of risk aversion can be used as controls, it still suggests that anticipated performance differences may play some role in differences in selection exhibited by females in the low phase. Therefore, if performance or confidence differences are driving selection differences, then controlling for performance and confidence in a discrete choice model should help isolate the effect of the low phase on selection choices.

Table 16: RE Performance Regression of No Choice Treatments

VARIABLES	(1) Math	(2) Math	(3) Word	(4) Word	(5) Low Math
Task Order	0.88 (0.09) (***)	0.82 (0.09) (***)	3.49 (0.39) (***)	3.50 (0.38) (***)	1.10 (0.11) (***)
Tourney Size	0.13 (0.04) (***)		0.04 (0.15) ()		
Tourney(ts>1)		0.63 (0.27) (**)		-0.16 (1.02) ()	1.18 (0.31) (***)
Tourney Size=2		-0.07 (0.38) ()		2.08 (1.49) ()	-0.24 (0.42) ()
Tourney Size=4		-0.21 (0.37) ()		0.22 (1.40) ()	-0.73 (0.45) ()
Female	-0.45 (0.59) ()	-0.46 (0.59) ()	2.31 (1.87) ()	2.29 (1.88) ()	0.44 (0.39) ()
Years PS	-0.50 (0.29) (*)	-0.50 (0.30) (*)	0.16 (0.87) ()	0.15 (0.87) ()	0.17 (0.23) ()
Age	0.40 (0.15) (***)	0.40 (0.15) (***)	0.33 (0.49) ()	0.34 (0.49) ()	-0.07 (0.14) ()
GPA	-0.26 (0.63) ()	-0.26 (0.64) ()	6.60 (1.94) (***)	6.57 (1.95) (***)	0.32 (0.41) ()
Constant	3.22 (3.42) ()	3.39 (3.43) ()	4.11 (11.57) ()	0.00 (0.00) ()	6.43 (2.58) (**)
Letter Controls	No	No	Yes	Yes	No
Observations	534	534	492	492	303
Number of id	178	178	164	164	101
R-sq	0.0745	0.0721	0.367	0.368	0.196
chi2	102.5	104.5	565.4	570.9	136.0

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 17: Hormonal Effects for Performance ($t < 4$)

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	Female	Female	All	Female	Female
Task	Math	Math	Math	Word	Word	Word
VARIABLES	RE	RE	RE Risk	RE	RE	RE Risk
Task Order	0.87 (0.09) (***)	0.94 (0.13) (***)	1.01 (0.18) (***)	3.46 (0.39) (***)	3.92 (0.57) (***)	3.09 (0.70) (***)
Tourney Size	0.14 (0.04) (***)	0.12 (0.06) (**)	0.09 (0.08) ()	0.05 (0.16) ()	-0.03 (0.23) ()	-0.20 (0.29) ()
Low Phase	-1.07 (0.98) ()	-1.09 (0.71) ()	-2.21 (0.89) (**)	2.22 (3.00) ()	2.09 (2.98) ()	4.57 (3.58) ()
Female	-0.32 (0.67) ()			1.72 (2.11) ()		
Years PS	-0.41 (0.30) ()	-0.70 (0.40) (*)	-1.19 (0.52) (**)	0.31 (0.87) ()	-0.87 (1.39) ()	-3.41 (2.10) ()
Age	0.38 (0.15) (**)	0.55 (0.23) (**)	0.89 (0.27) (***)	0.31 (0.49) ()	1.48 (0.90) ()	3.52 (1.22) (***)
GPA	-0.12 (0.65) ()	-0.43 (0.78) ()	-0.46 (1.03) ()	5.73 (2.00) (***)	7.56 (3.19) (**)	6.22 (4.13) ()
Psych meds	-1.48 (1.10) ()	-0.88 (1.19) ()	-1.17 (1.65) ()	3.37 (3.23) ()	4.80 (4.95) ()	3.56 (6.78) ()
Constant	3.26 (3.47) ()	0.78 (4.18) ()	-3.87 (5.28) ()	6.92 (11.70) ()	-18.13 (19.47) ()	-43.30 (26.05) (*)
Letter Controls	No	No	No	Yes	Yes	Yes
Risk Controls	No	No	Yes	No	No	Yes
Observations	510	237	147	471	225	147
Number of id	170	79	49	157	75	49
R-sq w	0.206	0.242	0.248	0.624	0.647	0.665
R-sq b	0.0740	0.147	0.365	0.291	0.340	0.430
R-sq o	0.0909	0.168	0.339	0.369	0.410	0.496
chi2	100.5	62.28	54.45	563.4	289.1	209.3
df_m	9	8	10	16	15	17

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C (Not For Publication) Appendix III: Expected value from competitive choice

To maximize expected value from choosing to compete in a tournament, or to accept the piece rate, or to select the group pay scheme, an individual cares about his probability of winning a tournament and his expected output. Expected values are based on the performance from Treatment 1 and the improvement of the individual from repeating the task and being in a tournament in Treatment 2.²² Assume an exogenous probability of winning a two person tournament in the form the percentile rank from Treatment 1 for the individual. If the individual is the best in the session then he receives a probability of winning a tournament of 0.99 and if he is the worst performing individual in the group his probability of winning is set to 0.01. Let this rank be equal to p . An individual's probability of winning an n person sized tournament is then p^{n-1} . Let b be the base piece rate and y be the expected output of the individual, then the expected value of an n person tournament is:

$$EV_n = p^{n-1} y b n \quad (\text{C.1})$$

This expected value form includes the piece rate which is the equivalent of a tournament size of 1. For group pay selection, assume that an individual accurately predicts the mean performance of the group. If an individual expects his output to be lower than the group average and if his probability is less than $\frac{1}{2}$ then he should choose the group pay compensation scheme. Otherwise the rank ordering for the different sized tournaments in terms of expected values is as follows:

$$\left\{ \begin{array}{ll} EV_6 \geq EV_4 & \text{if } y > \bar{y} \text{ and } p \geq \sqrt{\frac{2}{3}} \\ EV_4 \geq EV_6 & \text{if } y > \bar{y} \text{ and } p \leq \sqrt{\frac{2}{3}} \\ EV_2 \geq EV_4 & \text{if } y > \bar{y} \text{ and } p \leq \sqrt{\frac{1}{2}} \\ EV_1 \geq EV_2 & \text{if } y > \bar{y} \text{ and } p \leq \frac{1}{2} \\ EV_{Grp} \geq EV_n & \text{if } y \leq \bar{y}, \forall n \in \{1, 2, 4, 6\} \end{array} \right\} \quad (\text{C.2})$$

²²For simplicity, the incentive effects of effort from entering a tournament are ignored in calculating these expected values.

D (Not For Publication) Appendix IV: Experiment Instructions

D.1 Math Task Instructions

General Instructions

In this experiment you will be performing a task five different times. The task will consist of having you solve 2-digit 4-number addition problems in a 4 minute period.

The addition problems will look similar to the following equation:

$$12 + 57 + 48 + 52 =$$

In some cases, you will be asked to make decisions about how you will potentially be paid for your performance.

Only one of the five tasks will determine your payout for the experiment and it will be randomly chosen at the end.

To answer a problem, you will simply type the numbers on the keyboard, then press enter and another problem will appear. You can choose not to answer a question by pressing the Enter button or clicking on submit. The answer will then be recorded as being incorrect and you will be moved to the next problem.

To help with time management, there will be a clock counting down the seconds for the 4 minute duration.

Task 1

For Task 1, you will be paid \$0.50 for each correctly answered addition problem during the 4 minute time limit.

You will not know how many problems you answered correctly until the end of the 4 minutes.

At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen showing your payout method and confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 2

For this task, you will be randomly placed in a tournament.

The tournament will have a size of *2 or 4 or 6* people, including yourself.

If you win the tournament you will be paid \$0.50 multiplied by the number of people in the tournament for each correctly answered problem. For example, if you are in a 4 person tournament then you will be paid \$2.00 for each correct answer so long as you win the tournament.

If you do not win the tournament then you will receive nothing.

In the event of a tie for first place, you will split your tournament winnings evenly with the number of people that tied for first place in the tournament.

You will not know who you are competing against.

You will not know how many problems you answered correctly until the end of the 4 minutes.

To help with time management, there will be a clock counting down the seconds for the 4 minute duration. At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen showing your tournament size and confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 3

For this task, you will be randomly put in a group with one other person.

The group has a total size of 2 people (including yourself) and group members will be paid an equal amount.

You will be paid according to the productivity of your group. The number of correct answers by you and your partner will be added together and as a group you will be paid \$0.50 per correct answer. Your group earnings will be divided evenly between each partner ensuring you each get equal pay.

You will not know who you are a partner with.

At the end of the 4 minutes you will see a screen asking about your performance.

You will not know how many problems you answered correctly until the end of the 4 minutes.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen showing your payout method and confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 4

You will now have the opportunity to choose the individual pay rate of \$0.50 per correct answer, or to enter in a tournament, or to join a partner and receive the group pay rate.

Your possible choices are: *Group Pay, Individual Pay, 2 Person Tournament, 4 Person Tournament, 6 Person Tournament*

Whenever possible, you will be paired with other people that chose the same pay format as yourself. If there are not enough people for your chosen format, then individuals who have selected another pay format will be randomly selected to meet the participation requirements for your chosen format.

Tournament:

If you win the tournament, then you will be paid \$0.50 multiplied by the number of people in the tournament for each correctly answered problem. In the event of a tie for first place, you will split your tournament winnings evenly with the number of people that tied for first in the tournament.

If you do not win the tournament then you will receive nothing.

You will not know who you are competing against.

Individual Pay:

You will be paid \$0.50 for each correctly answered problem.

Group Pay:

For the Group Pay format, you will be joining one other person in a group.

You will be paid according to the productivity of your group. The number of correct answers by you and your partner will be added together and as a group you will be paid \$0.50 per correct answer. Your group earnings will be divided evenly between each partner ensuring you each get equal pay.

You will not know who you are a partner with.

You will not know how many problems you answered correctly until the end of the 4 minutes.

At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below, you will be taken to a screen to make your payout method choices and confirm that you understand.

The experiment will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 5

Every task that you have done so far has involved no feedback about other's performance. Now we will provide you with feedback regarding all session participants' performance from Task 1.

Your performance: *Individual performance.*

Everyone's performance: *List of everyone's performance (high to low).*

You will now have the opportunity to choose the individual pay rate of \$0.50 per correct answer, or to enter in a tournament, or to join a partner and receive the group pay rate.

Your possible choices are: *Group Pay, Individual Pay, 2 Person Tournament, 4 Person Tournament, 6 Person Tournament*

Whenever possible, you will be paired with other people that chose the same pay format as yourself.

If there are not enough people for your chosen format, then individuals who have selected another pay format will be randomly selected to meet the participation requirements for your chosen format.

Tournament:

If you win the tournament, you will be paid the rate of \$0.50 times the number of people in the tournament for each correctly answered problem.

In the event of a tie for first place, you will split your tournament winnings evenly with the number of people that tied for first in the tournament.

If you do not win the tournament then you will receive nothing.

You will not know who you are competing against.

Individual Pay:

You will be paid \$0.50 for each correctly answered problem.

Group Pay:

For the Group Pay format, you will be joining one other person in a group.

You will be paid according to the productivity of your group. The number of correct answers by you and your partner will be added together and as a group you will be paid \$0.50 per correct answer. Your group earnings will be divided evenly between each partner ensuring you each get

equal pay.

You will not know who you are a partner with.

You will not know how many problems you answered correctly until the end of the 4 minutes.

At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below, you will be taken to a screen to make your payout method choices and confirm that you understand.

The experiment will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

D.2 Word Task Instructions

General Instructions

In this experiment you will be performing a task five different times. The task will consist of giving you an alphabetical letter and having you type as many words as possible that begin with that letter in a 4 minute period.

To enter a word, you will simply type the word using the keyboard, then press enter or submit. You will then be able to enter another word. Using an English dictionary, the computer program will verify that the words you have spelled are correct. Misspelled words will be counted as incorrect and will not be included in your task total. During the task, you will see the words you have typed, but you will not know if you have spelled them correctly.

- Capitalization will not affect spelling.
- Duplicate words will be counted as incorrect.
- Proper names will be counted as incorrect.
- Common place names (cities, countries) will be counted as correct.
- Plurals and tense changes to root words will count as separate correct answers.

Everybody in the session will be given the same letter.

In some cases you will be asked to make decisions about how you will potentially be paid for your performance.

Only one of the five tasks will determine your payout for the experiment and it will be randomly

chosen at the end.

To help with time management, there will be a clock counting down the seconds for the 4 minute duration.

Task 1

For Task 1, you will be paid \$0.25 for each correctly spelled word during the 4 minute time limit.

You will be informed of how many words you spelled correctly at the end of the task.

At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen showing your payout method and confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 2

For this task, you will be randomly placed in a tournament.

The tournament will have a size of *2 or 4 or 6* people, including yourself.

If you win the tournament, you will be paid \$0.25 multiplied by the number of people in the tournament for each correctly spelled word you enter beginning with the designated letter.

For example, if you are in a 4 person tournament, then you will be paid \$1.00 for each correctly spelled word so long as you win the tournament.

If you do not win the tournament, then you will receive nothing.

In the event of a tie for first place, you will split your tournament winnings evenly with the number of people that tied for first place in the tournament.

You will not know who you are competing against. Everybody in the session will be given the same letter.

You will be informed of how many words you spelled correctly at the end of the task.

At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen showing your payout method and

confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 3

For this task, you will be randomly put in a group with one other person.

The group has a total size of 2 people (including yourself) and group members will be paid an equal amount.

You will be paid according to the productivity of your group. The number of correctly spelled words by you and your partner will be added together and as a group you will be paid \$0.25 per word. Your group earnings will be divided evenly between each partner ensuring you each get equal pay.

You will not know who you are a partner with. Everybody in the session will be given the same letter.

You will be informed of how many words you spelled correctly at the end of the task.

At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen showing your payout method and confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 4

You will now have the opportunity to choose the individual pay rate of \$0.25 per correctly spelled word, or to enter in a tournament, or to join a partner and receive the group pay rate.

Your possible choices are: *Group Pay, Individual Pay, 2 Person Tournament, 4 Person Tournament, 6 Person Tournament*

Whenever possible, you will be paired with other people that chose the same pay format as yourself.

If there are not enough people for your chosen format, then individuals who have selected another

pay format will be randomly selected to meet the participation requirements for your chosen format.
Tournament:

If you win the tournament, then you will be paid \$0.25 multiplied by the number of people in the tournament for each correctly spelled word. In the event of a tie for first place, you will split your tournament winnings evenly with the number of people that tied for first in the tournament.

If you do not win the tournament then you will receive nothing.

You will not know who you are competing against.

Individual Pay:

You will be paid \$0.25 for each correctly spelled word.

Group Pay:

For the Group Pay format, you will be joining one other person in a group.

You will be paid according to the productivity of your group. The number of correctly spelled words by you and your partner will be added together and as a group you will be paid \$0.25 per correctly spelled word. Your group earnings will be divided evenly between each partner ensuring you each get equal pay.

You will not know who you are a partner with. Everybody in the session will be given the same letter. You will be informed of how many words you spelled correctly at the end of the task. At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen to make your payout method choices and confirming that you understand.

The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.

Task 5

Every task that you have done so far has involved no feedback about other's performance. Now we will provide you with feedback regarding all session participants' performance from Task 1.

Your performance: *Individual performance.*

Everyone's performance: *List of everyone's performance (high to low).*

You will have the opportunity to choose the individual pay rate of \$0.25 per correctly spelled word, or to enter in a tournament, or to join a partner and receive the group pay rate.

Your possible choices are: *Group Pay, Individual Pay, 2 Person Tournament, 4 Person Tournament, 6 Person Tournament*

Whenever possible, you will be paired with other people that chose the same pay format as yourself. If there are not enough people for your chosen format, then individuals who have selected another pay format will be randomly selected to meet the participation requirements for your chosen format.

Tournament:

If you win the tournament, you will be paid the rate of \$0.25 multiplied by the number of people in the tournament for each correctly spelled word.

In the event of a tie for first place, you will split your tournament winnings evenly with the number of people that tied for first in the tournament.

If you do not win the tournament then you will receive nothing.

You will not know who you are competing against.

Individual Pay:

You will be paid \$0.25 for each correctly spelled word you submit.

Group Pay:

For the Group Pay format, you will be joining one other person in a group.

You will be paid according to the productivity of your group. The number of correctly spelled words by you and your partner will be added together and as a group you will be paid \$0.25 per word. Your group earnings will be divided evenly between each partner ensuring you each get equal pay.

You will not know who you are a partner with.

Everybody in the session will be given the same letter. You will be informed of how many words you spelled correctly at the end of the task. At the end of the 4 minutes you will see a screen asking about your performance.

Do not discuss your performance with anyone else at anytime.

Once you press the button below you will be taken to a screen to make your payout method choices and confirming that you understand. The task will begin once everyone in this session is ready to begin.

If you feel like you understand the instructions and are ready to begin then press the "I understand" button below.