Working Hard or Working Smart? [1]

Anat Bracha[2] and Chaim Fershtman[3]

Wednesday, May 16, 2018

Abstract: Almost all jobs require a combination of cognitive effort and labor effort. The focus of the paper is on the effect of different incentive schemes on the chosen combination of these types of efforts. We use an experimental approach to show that tournament incentives may induce agents to work harder but not necessarily smarter. This effect was stronger for women. We then ran a "managerial bonus" experiment in which a preassigned manager receives a bonus whenever the overall performance of his/her group is above a given threshold level. Although the bonus does not affect the participants' direct incentives, it induces participants to lower their cognitive effort and increase their labor effort.

[1] This work was supported by ISF Grant No. 0610114192. We thank Lynn Conell-Price for her excellent research assistance.

[2] Research Center for Behavioral Economics, Research Department, The Federal Reserve Bank of Boston, 600 Atlantic Avenue, Boston, MA 02210 USA. Tel: 617-973-3096; Fax 617-973-4218; Email: anat.bracha@bos.frb.org

[3] The Eitan Berglas School of Economics, Tel Aviv University, Tel Aviv, Israel 69978, CEPR and the Tinbergen Institute; Email: fersht@post.tau.ac.il.

1. Introduction

Almost all jobs require a combination of cognitive effort and labor effort. Research and development, running a company, building a house, teaching a class or even working on a factory floor requires both cognitive and labor efforts. The tradeoff between these types of efforts exists whenever agents need to think about how to perform a task or to choose a method of solving a problem before they actually implement it. Take for instance the task of trying to find the highest value of a function, people may try to analyze the function (cognitive effort), they may try to check it for many parameters value (labor effort) or they may try a combination of the two methods. It is the combination of cognitive and labor effort that determines whether people work hard or smart.

The general intuition in economic literature is that providing performance-based (or competitive) incentives motivates individuals to exert more effort.[4] But once we distinguish between cognitive and labor efforts, an additional question emerges: what is the effect of different incentive schemes on the combination of these two types of efforts? This is the focus of our paper.

More specifically, we examine and compare how people work under three types of incentives: (i) A simple pay for performance incentive scheme (hereinafter PFP) in which agents are paid according to their own performance. (ii) Tournament incentives in which pairs of participants compete for a prize and (iii) "managerial bonuses" in which groups are randomly formed and one of the participants is randomly selected as the group manger. Members of the group receive the same incentives as in the PFP treatment. However if the overall performance of the group exceeds a given threshold, the group's manager gets a managerial bonus. Taking the PFP treatment as our benchmark case our focus is on the effect of tournament incentives and managerial bonuses on the combination of cognitive and labor efforts chosen by participants, where we expect competitive incentives and managerial bonus to lead people to choose to work harder using more of labor effort rather than smarter using cognitive effort. We expect competition to lead people to work harder but not smarter due to competitive pressure,

[4] For a survey see Lazear (2000).

while managerial bonus to have the same effect due to emotional response such as envy, disappointment and fairness concerns.

To test our claims, we designed a simple lab experiment with two computerized tasks, the "sequences" task where participants were asked to solve numeric sequences that required cognitive effort, and a "filing" task which is a simple number categorizing task that requires mainly manual dexterity. Participants in this study could engage in either task and were free to switch between the two during the entire duration of the study. We use a between subject design where different groups face different incentives (PFP, tournament, or managerial bonus) and then compared the average performance under these different incentives. In particular we examine the allocation of time between the sequence task and the filing task and the players' success rate in solving sequences[5], since these variables capture different aspects of cognitive effort.

Our main results are that in the tournament and the managerial bonus treatments participants devote less time to the sequence task and have a lower success rate compared with PFP incentives. In other words, under tournament incentive and the managerial bonuses participants work harder but not smarter.

Our results are however gender sensitive. Under the PFP incentives the performance of women is lower than the performance of men, they attempt to solve less sequences and devote more of the allotted time to the simpler task of categorizing numbers. This is despite the fact that men and women have the same success rate in solving sequences under the PFP incentive scheme. Analyzing the effect of tournament incentives by gender, we find that relative to the PFP treatment, tournament incentives induce both men and women to spend less time on the sequence task and more on the routine filing task. However, the negative effect of the tournament on the success rate is entirely a female effect.

Managerial bonuses also induce participants to exert less cognitive effort relative to the PFP treatment. However under such bonuses men spend less time on sequence without any effect on their success rate, while women choose to spend similar amount of time on sequences compared with PFP but have a lower success rate. Note that managerial bonuses in our experiment do not affect the participants' direct incentives –

[5] Success rate is the percentage of sequences solved correctly over the number of sequences attempted.

we examine only those participants who are not assigned to be managers. That is, these participants do not have the opportunity to get a bonus, and their overall performance simply determines whether *another* individual will get a bonus. Therefore, managerial bonuses may affect behavior only by affecting participants' emotions. Providing managerial bonuses may lead participants to be envious, angry, disappointed or having fairness concerns which may affect participants' choices, efforts and performances.

This paper adds to the recent literature on the psychological foundation of incentives, that provides an important critical view of the traditional incentive theory (for a survey see Fehr and Falk, 2002). The main claim in this literature is that considering monetary incentives alone is too narrow, empirically questionable, and limits our understanding of incentives. Nonpecuniary motives such as reciprocity, the desire for social status and fairness concerns are powerful drives of human motivation. The paper extends this literature by focusing on the combination of cognitive and labor effort which may be affected by nonpecuniary motives. For example, fairness concerns may reduce the player's overall effort but it may also trigger a switch from a cognitive effort to a labor effort.

The paper is also related to the psychological literature that identified several mechanisms resulting in "choking under pressure". This literature suggests that pressure may lower performance (see Baumeister (1984) and Baumeister and Showers (1986)). In a recent paper Ariely, Gneezy, Loewenstein and Mazar (2009) demonstrated the "choking under pressure" effect and showed that excessively high rewards have a detrimental effect on performance. Competitive pressure that is associated with tournaments may indeed affect the players' choices and performance.[6] Our intuition suggests, however, that managerial bonuses give rise to other types of emotions which may affect performance in a different way than competitive pressure.

Finally, Economics derives part of its strength from its ability to analyze the effect of incentives on people's behavior. The effect of incentives is relevant for analyzing virtually most economic problems. An important part of the economic framework is the assumption that individual abilities are exogenously given and are not affected by incentives. This assumption has been very instrumental in building a workable paradigm

[6] Our focus however is not on the effect of such a pressure on performance but on its effect on the combination of cognitive and labor effort.

for economic analysis. One of the findings of our experiment is that incentives affected agents' success rate. Clearly success rate is a performance measure and as such it is determined by both ability and effort. We followed the conservative approach and interpret the reduction in the success rate as an outcome of a lower cognitive effort. But an alternative interpretation would be that ability is not exogenously given and can be affected by incentives.[7]

2. Experimental design

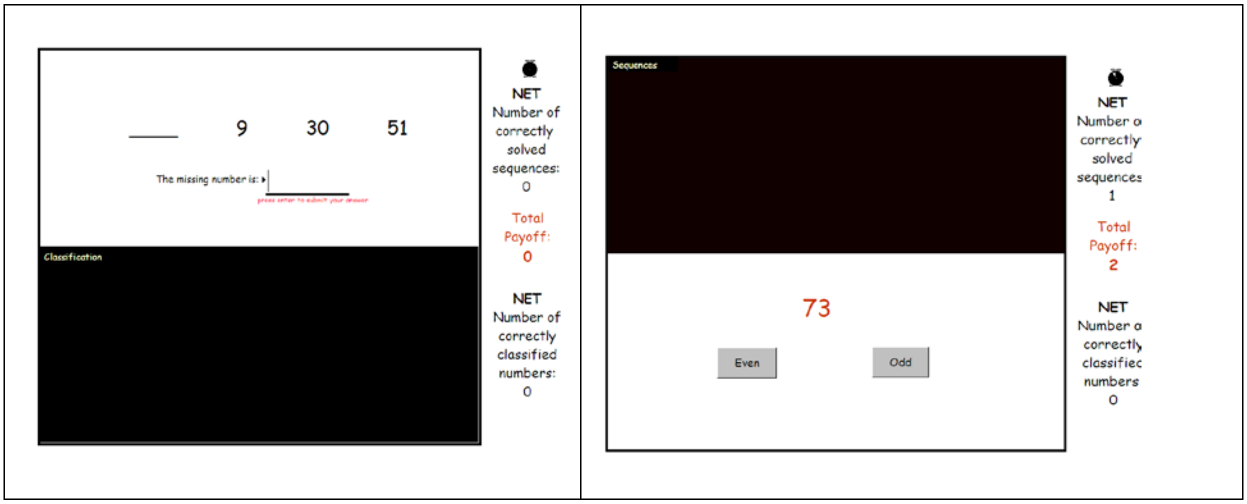
To test the effect of incentives on people choice of cognitive and labor effort combination, we need to design a task that captures and distinguishes between cognitive and labor effort, and to allow incentives to vary across treatments. Below is a detailed description of the task we used and the incentive treatments.

2.1 The Tasks

To capture different types of efforts we introduce two tasks. Subjects could engage in either solving sequences ("sequences" task)—finding a missing number in a sequence of 4 numbers—or classifying a random number into an "odd" or "even" category ("filing" task) by pressing an appropriate button on the screen. The sequences task requires cognitive effort in the form of abstract thinking, while filing numbers mainly requires labor effort. Both tasks were available during the study; engaging in any of the two tasks was done simply by clicking on the section of the screen with the desired task (see Figure 1). In our design the goal is to earn money, and the complementary between the tasks is achieved using the compensation scheme, as described below.

[7] In the tournament treatment success rate (of women) went down sharply at the end of the tournament. It seems unlikely that lower effort levels explain these findings.

Figure 1: Sequences Task and Filing Task



2.2 Treatments

To analyze the effect of incentives we use a between subject design with three treatments: PFP, head-to-head tournament, and managerial bonus. The exact compensation is described below.

Pay for Performance (PFP): subjects were paid $2 per *net* correctly solved sequences, 3 cents per *net* correctly filed numbers and a 1 cent extra reward for the product net sequences and net filed numbers.[8] Net number of correctly solved sequences is the number of correctly solved sequences minus half the number of incorrectly solved sequences. Penalizing incorrectly solved sequences was designed to prevent guessing. Net filed numbers equals correctly filed numbers minus incorrectly filed numbers. Penalizing incorrectly filed numbers was designed to prevent random clicking. The extra reward introduces a complementary term as greater number of net correctly solved sequences (filed numbers) increases the marginal return to successful filing (sequence).

Tournament: in this treatment, subjects were randomly paired using a randomly generated subject ID number. The pairs were announced before the beginning of the task and by subject ID, such that the identity of one's opponent was not revealed. The winner was determined according to the accumulated number of points for each of the opponents in a pair. The point schedule was exactly as under the PFP compensation scheme—2

[8] This compensation is different from a piece rate since there is a multiplicative term in their incentives.

points per net sequence, 0.03 of a point per net number filed and an extra 0.01 of a point for the product of net sequences times the net numbers filed. The winner's prize was $60, and the loser received the minimum guarantee of $10, such that the expected earning was $35 similar to the average earning under PFP. At the end of the study, after completing the time devoted to the task, the accumulated number of points for each participant was announced (by the randomly generated subject ID), and the earnings were determined and announced.

Managerial Bonus: We had two "managerial bonus" treatments. In both of these treatments subjects faced the PFP compensation scheme. However, before the subjects began working on the task, one participant in each group was selected at random by drawing a note from an envelope. The randomly selected participant engaged in the same task under the same compensation schedule as the other members of the group, but unlike the others he or she was entitled to $100 instead if the overall earning of the other group members exceeded a certain threshold level. If the other members (excluding the randomly selected participant) did not reach the earnings target, the selected person would receive his or her PFP earning on the task. This was explained both in the instructions on the screen and then it was repeated again by the experimenter. In the first managerial bonus treatment the group size was two players and a bonus was given if the threshold of $40 was reached. In the second managerial bonus treatment the group size was six and the threshold was $200.

2.3 Procedure

The sessions were conducted at the Harvard Decision Science Laboratory at the Harvard Kennedy School. 268 Harvard students participated in the study, 74 participated in the PFP treatment, 60 in the tournament treatment, and 134 in the managerial bonus treatment—55 in groups of six and 79 in groups of two. In each session, participants sat at an individual station and read the instructions on their individual screen. Once all were done with the instructions, they were given a code to proceed such that all started working on the task at the same time. They were given, under all conditions, 10 minutes to work on the two tasks. In the tournament treatment, once all were done with the instructions, and before giving the code to proceed, the experimenter announced the pairs by subject ID. In the bonus treatment, once all were done with the instructions, and before giving the code to proceed, the experimenter randomly determined the person with the opportunity to earn the $100. In the "group of six" treatment, the experimenter announced that each row of participants is a group (each row in the lab has six terminals) and then went one by one to draw a note from an envelope. The envelope had six notes, five of which were marked “000” and one was marked “100”. The first person in a row who drew the “100” note was the one selected for the possibility of earning the managerial bonus of $100. When we had a group of two, the experimenter announced the groups and approached one person in each pair. That person was asked to draw a note from an envelope with three notes marked “000” and three marked “100.” If the first person in a pair drew a “100” note he or she were announced as the selected person to have the possibility of earning $100; otherwise the second person in the pair was announced to be the selected one.

3. The Effect of Competitive Incentives

Competitive incentives typically induce agents to exert more effort. Our focus is on the effect of such incentives on the combination of efforts that agents choose.

3.1 The effect of a tournament on cognitive effort.

Our hypothesis is that tournament incentives induce agents to reduce their cognitive effort. To test this hypothesis we examine two variables that are affected by such a shift: time allocation between the two tasks and the success rate in solving sequences. We analyze these effects using the appropriate t-test. [9]

Comparing the data from the PFP and the tournament treatment yields that the overall performance in the tournament treatment was slightly (but not significantly) higher than in the PFP treatment. However, under PFP, the average number of attempted sequences was 10.7 (with 8.5 solved correctly) while under tournament the average number of sequences attempted was 9.2 (with 7.2 solved correctly).[10] Under PFP

[9] In calculating the various averages, we first calculate the particular measure (such as success rate) for each individual and then average across individuals. When comparing across gender with no a-priori hypothesis we use a two sided t-test and when we examine whether tournament incentives reduce success rates and the time devoted to sequences, as hypothesized, we use one-sided t-test.

[10] Under PFP, the average number of sequences attempted was 11 with a standard deviation of 6.69; under tournament, the average number of sequences attempted was 9.6 with a standard deviation of 6.78. There

incentives participants devoted on average 381 seconds to solving sequences, while in the tournament treatment they spent only 330 seconds on solving sequences (p=0.025; see Figure 2A).

Observation 1 (Tournament and cognitive effort): (i) Under the tournament incentives participants devoted less time to the cognitive task than under the PFP incentives. (ii) Success rate in solving sequences is lower under tournament incentives—78.6% under PFP while only 72% under tournament incentives (p=0.047).

Clearly the two parts of Observations 1 may be interdependent: if participants are aware of the fact that under tournament they have a lower success rate then their rational reaction would be to reduce the time they spend on solving sequences. For our purposes, whether the two are interdependent or not, they both represent measures of cognitive effort.

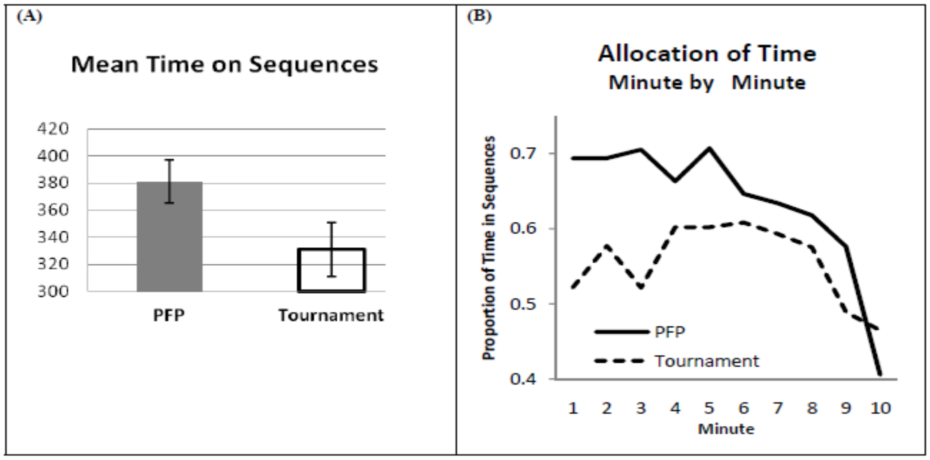
The effect of tournament on success rate may be due to an effect on effort, an effect on ability, or both. We take the conservative economic interpretation of an effect on effort and treat ability as exogenous. However, under pressure it is possible that even if one tries as much as she could she would still not be able to perform as well as when she was not under pressure. It is therefore possible that incentive scheme affect ability and that such an effect should be taken into account whenever incentives are designed.

The effect of tournament incentives on time allocation is best seen in Figure 2B which compares the minute by minute percentage of time devoted to solving sequences in the PFP and the tournament treatments. Figure 2B strikingly illustrates that the above effect is neither due to a single episode nor due to a particular stage of the task. The effect of tournament incentives on time allocation stems from different time allocation throughout the entire 10 minutes of the study.

---

were two outliers, one in each condition, who attempted over 30 sequences in 10 minutes (32 sequences under PFP and 33 sequences under Tournament.) Examining the average and standard deviation by gender, men have a higher number of attempts on average (14.84 under PFP; 11.76 under tournament) and greater standard deviation (7.37 under PFP; 7.90 under tournament). Yet, over 30 attempts are more than two standard deviations from the mean. Therefore, in our analysis and the numbers presented above we exclude these two outliers.

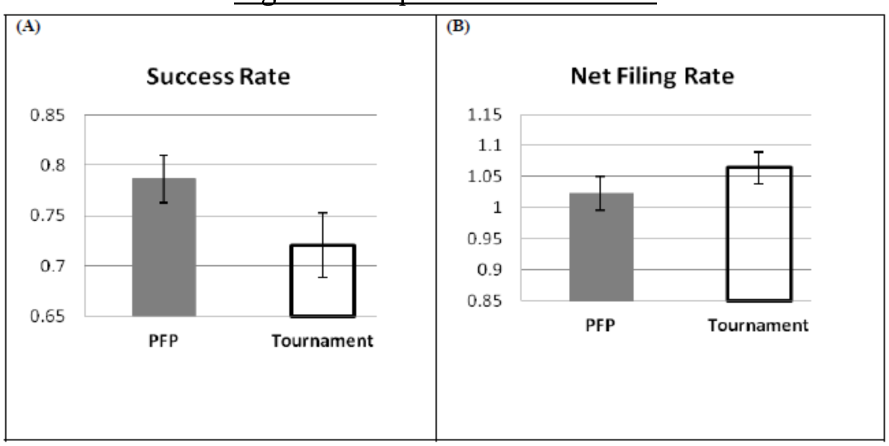
Figure 2: Allocation of Time



N=73 under PFP, N=59 under Tournament. Bars in panel A represent SEM

The second type of cognitive effort is captured by the players' success rate in solving sequences. We calculate this measure for each participant and present in Figure 3 the average success rate across all individuals. In Figure 3B we present the average net filing rate which is the average speed of net filing across participants. Having a significantly lower success rate together with higher net filing rate (although the latter is not statistically significant) indeed demonstrate that competitive incentives induce participants to work harder but not smarter.

Figure 3: Sequence Success Rate



Panel A: N= 73 under PFP, N=57 under Tournament; Panel B: N= 69 under PFP, N=56 under Tournament. Bars represent SEM

The effect of competitive incentives is not necessarily uniform. Some people may perform better in a competitive environment while others may suffer from a competitive pressure. We therefore distinguish between winners and losers in the tournament treatment and then compare the performance of each group to the performance of the appropriate comparison group in the PFP treatment (see Table A1 in the Appendix). We find that winners and losers spend statistically the same amount of time on sequences—winners spend 347 seconds on average while losers spend 313 second on average, but this difference is not significant (p=0.19). Nevertheless, the average score of winners in the tournament treatment is 44.38 points, they solve on average 11.3 sequences and their success rate is 84% while the losers' average score is 21.2 points, they solve only 7.20 sequences and their success rate was 58%.[11]

Observation 2: The tournament winners are the participants with the higher success rate. But there was no difference between winners and losers with respect to the time they spent on sequence solving.

Next we compare the top performing individuals in the PFP treatments with the winners in the tournament treatment. We split the PFP participants into two groups—above and below median performers. We then compare the tournament winners to the above-median performers under PFP, and the tournament losers to the below-median PFP performers. Note that in the tournament treatment we have a random matching of pairs. Thus in the tournament treatment the losers are not necessarily all of low ability, as it is possible that two strong participants were competing against each other. Therefore, if incentives do not affect performance we would expect that winners' success rate in the tournament treatment to be below the success rate of the high performance individuals in the PFP treatment and that the losers in the tournament treatment would have a higher success rate, on average, than the below-median group in the PFP treatment. Our findings were different: we found that the success rate of tournament winners and the above-median PFP performers was similar (85.29% under PFP and 84.64% for tournament winners). However, the tournament losers' success rate was 58.96% and it was significantly lower than the 73.42% success rate of the below-median PFP performers ($p<0.01$). Interestingly, although both winners and losers seem to reduce the time spent on sequence solving compared to above and below PFP performers (respectively), the

[11] These differences between winners and losers (score, success rate, and number of sequences solved) are all significant at the 1% level.

decline is significant only for the winners (399 seconds under PFP vs. 347 seconds under tournament; p=0.074.)

If lower success rate in the tournament treatment is the outcome of competitive pressure that (some) participants were facing, we would expect this effect to be stronger at the end of the tournament when the competitive pressure is possibly at its highest level. We therefore divide the ten minute experiment into two parts: the first seven minutes and the last three minutes. This division is arbitrary but we expect that in the tournament treatment participants would be more "pressured" in the last three minutes than in the first seven minutes. We compare the participants' success rate at the beginning and at the end of the treatment. In the tournament treatment success rate was 77.5% in the first seven minutes and only 57.9% in the last three minutes. This decline is highly significant (p<0.01.) In the PFP treatment success rate in the first seven minutes was 76.4% and 77.2% in the last three minutes. It is of course possible that in the tournament treatment some low ability individuals chose to solve sequences only in the last 3 minutes. To exclude this possibility we compare the success rate only for those individuals who solved sequences both during the first 7 minutes and the last 3 minutes. We find the same pattern: no effect under PFP (76.12% success in the first 7 minutes, and 77% success in the last 3 minutes) and a highly significant decline under competition (76.3% success rate in the first 7 minutes compared with 57.5% success rate in the last 3 minutes; p<0.01.)

Observation 3 (Competitive Time Pressure): In the last three minutes of the tournament participants' success rate was significantly lower than in the first seven minutes. However, time pressure in and by itself has no such effect, as in the PFP treatment there was no reduction of participants' success rate at the last three minutes.

One may wonder whether reducing cognitive effort is an optimal reaction to competitive incentives since exerting labor effort is “safer”, yielding a sure benefit to effort. However, one can make the case for the exact opposite claim: to increase winning probability a better strategy is to engage in the “risky”, high variance task. Furthermore, to argue that reducing cognitive effort is an optimal reaction to incentives demands that participants both form expectations on their opponent’s strategy and solve a dynamic tournament, two very complicated tasks which participants in our study were unlikely to solve. And, of course, lower cognitive effort in the form of reduced success rate can never be an optimal reaction to competition. Finally, calculating the optimal time allocation empirically is problematic since the number of sequences one solves is an endogenous choice variable. Nevertheless, whether the observed lower cognitive effort under tournament is optimal or not, this paper is the first to show that given the opportunity to trade off cognitive and labor effort people choose different mix under different incentive schemes, and in particular they exert more labor effort and less cognitive effort under tournament.

3.2. The Gender Effect

Recent studies (see Gneezy, Niederle and Rustichini, 2003; Croson and Gneezy, 2009) indicate that men and women respond differently to competitive incentives. Much to our surprise, not only that we find gender differences in response to competition, we also find gender differences in the benchmark PFP treatment.

Observation 4 (PFP: Gender Effect):

(i) Women devoted 360 seconds to solving sequences while men devoted 419 seconds to solving sequences ($p=0.055$).

(ii) On average women solved 8.91 sequences correctly while men solved on average 14.16 sequences correctly ($p<0.01$.)

(iii) Men and women had a similar success rate; 76.2% for men and 79.8% for women.[12]

Women's choice to devote less time to sequence solving cannot be the outcome of a lower success rate since they had similar success in solving sequences as men; in fact, although not statistically significant, women's success rate in this treatment is slightly higher. It is possible that the observed time allocation choice is the outcome of the gender difference in risk aversion and self-confidence (see Croson and Gneezy, (2009)) as solving sequences is a riskier task than filing numbers.

Given the above gender differences, the reader may wonder whether the observed effect of tournament incentives is an artifact of having a different gender mix across treatments. That is, if there were relatively more women in the tournament treatment than

[12] There was also no gender difference in the speed of the filing task—women's net filing rate was 1.00 while for men it was 1.04.

in the PFP treatment, then the gender mix could explain the decline in the time devoted to solving sequences in the tournament treatment. However in our experiments there were relatively more women in the PFP treatment than in the tournament treatment so we would expect the opposite effect[13]. Nevertheless, next we will examine the effect of tournament incentives on each gender separately and at the end we will present a simple OLS regression that takes into account these relevant variables.

The overall performance, as measured by the average achieved points, was not significantly different in the tournament treatment and the PFP treatment. The average number of points for women under tournament was 29.12 compared with 30.98 under PFP. The average number of points for men under tournament was 38.64 compared with 39.64 under PFP. Differences for both women and men are insignificant.

The Effect of tournament incentives on women: Women’s success rate declined from 79.87% under PFP to 67.18% under tournament (p<0.01), a sharp and strong decline of over 15 percent of the success rate. This decline was evident in the last 3 minutes of the experiment—a decline from 77.11% under PFP to 49.74% under tournament (p<0.01)—but not in the first 7 minutes. Under PFP women spent on sequence solving an average of 360 seconds, while under tournament incentives they spent only 308 second on overage (p=0.066). [14] See figure 2A (in the Appendix) for a minute by minute time allocation in the PFP and the tournament treatments for both men and women.

The Effect of tournament incentives on men: under tournament incentives men reduced the amount of time they devoted to sequence solving from 419 seconds to 363 seconds (p=0.059). The average number of sequences they attempt to solve decreased from 14.16 under PFP to 10.87 under tournament (p=0.045.). However, the tournament incentives did not affect men's success rate that was 76.23% in the PFP treatment and 78.69% in the tournament treatment.

Observation 5 (Gender and Tournament): (i) Both men and women reduce the time they spent on sequence solving when facing tournament incentives. (ii) The effect of

[13] In the PFP treatment, 48 females and 25 males participated; in the tournament treatment, 35 females and 24 males participated.
[14] This led to a significantly higher net filing among women (339.52 under tournament vs. 258.71 under PFP; p=0.024).

tournament on success rate was entirely a female effect. This effect is mainly due to pressure at the end of the tournament.

Interestingly, despite the different reaction to tournament incentives women in our settings won the tournament at a similar proportion as did men. Specifically, 16 out of the 35 women who participated in the tournament treatment won while 14 out 24 men who participated in the tournament treatment won (Fisher exact test; p=0.43). Furthermore, matching participants randomly in the tournament condition, we find that out of 100 such matching only in eight instances there were significant differences in winning proportions across genders. The result is that we reject the hypothesis that the average z statistics across all 100 random matching is equal or greater than 1.96, and therefore the result of no difference in the winning proportion across gender is not an artifact of the actual matching we used in the study[15].

These results are also reflected in a regression analysis which controls for gender and age: we used an OLS regression of success rate and time devoted to sequences on a treatment dummy variable (takes a value of 1 for a tournament treatment), gender (takes a value of 1 for females), and age. The results are presented in Table 1.

Table 1: OLS Regressions

	Success	Time Allocated to Sequences	Success	Time Allocated to Sequences
Treatment (=1 for tournament)	-.075 (-1.90)	-52.381 (-2.10)	.017 (0.28)	-59.322 (-1.45)
Gender (=1 for Females)	-.030 (-0.77)	-53.030 (-2.06)	.038 (0.71)	-58.177 (-1.65)
Treatment x Gender			-.150 (-1.89)	11.157 (0.22)
Age	-.018 (-1.19)	-7.209 (-1.90)	-.017 (-1.19)	-7.268 (-1.90)
Const	1.179 (3.74)	564.671 (7.00)	1.130 (3.61)	569.255 (6.80)
N	130	132	130	132
R^2	0.037	0.089	0.064	0.089

t-statistics are in parenthesis; the number of observations is lower when analyzing the success rate compared with time allocation. This is due to participants who did not solve a single sequence.

[15] In generating the random matching we tried both (1) using all participants, including those who attempted over 30 sequences, and (2) excluding those who attempted over 30 sequences. In the latter case we simply dropped the person left with no competitor, and we found similar results.

As Table 1 shows, we find that under tournament success rate is lower by 7.5 percent, which is approximately 9.5 percent of the average success rate under PFP (p=0.06). The time allocated to sequences is lower under tournament incentives by 52.38 seconds which is about 13.7 percent of the average time devoted to sequences under PFP (p=0.038.) Adding an interaction term to examine whether competition has a differential gender effect, confirms that the decline in the success rate under tournament is a female effect (p=0.06) while the decline in time allocation is similar for both men and women (interaction term is not significant).

4. The Effect of Managerial Bonus

Next we turn to examine the effect of managerial bonus on the allocation of cognitive and labor efforts. Our experimental design was such that we randomly assigned participants into groups, and then randomly selected one participant from each group to be a "manager"; that is, the selected person from each group had the chance to earn a bonus if the overall earning of her group (excluding herself) was above a certain threshold. We had two treatments: in the first one the group size was two, and in the second treatment the group size was six. Unlike others in the literature, we examine the effect of this “managerial bonus” on the performance of participants who were *not* assigned to be managers. The incentives for those participants were exactly the same as the incentives in the PFP treatment. The only difference was that there was another player who could receive a large bonus if the overall performance of the group would reach a certain (known) target.

While the above managerial bonus does not affect the direct incentives, it may lead participants to be envious, disappointed, or have fairness concerns. The expectation is that these emotions, if they are present, would lead participants to devote less effort, in particular cognitive effort, out of spite or anger. These emotions may depend on the size of the group. We expect that this elevated emotional state will be highest in the "group of two" treatment, as the chance of being selected to be the manager is higher and there is only one person in the group to be compared to.

4.1 Managerial Bonus (group of 2)

Table 3 provides the details of the players' time allocation and success rate in the managerial bonus treatment.

Table 3: Managerial Bonus (groups of 2)

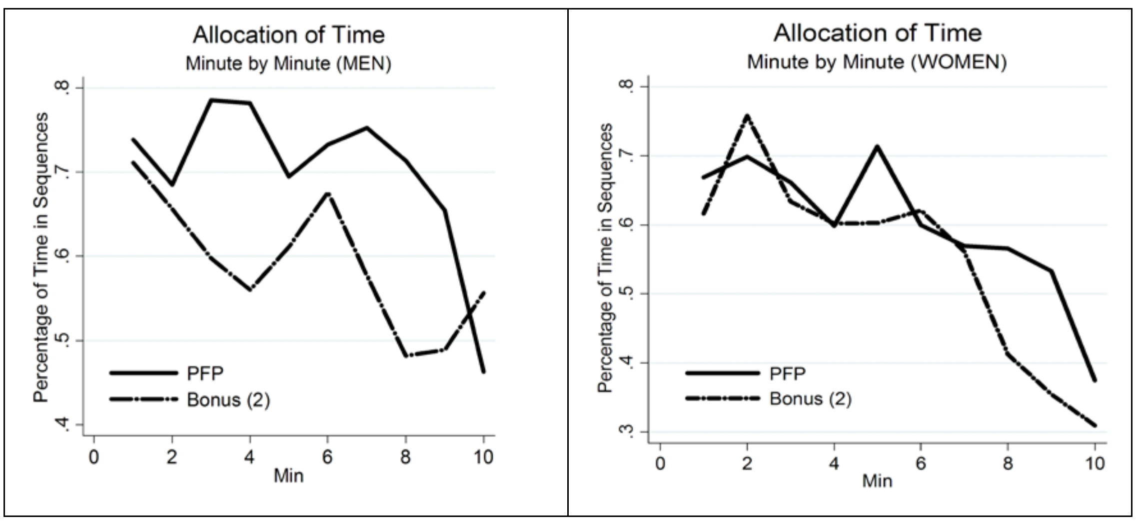
Men	**PFP**	**Bonus (2)**	**Ttest**
Success Rate	76.23%	76.61%	p=0.52
Time in sequences	419.82	350.62	p=0.029

Women	**PFP**	**Bonus (2)**	**Ttest**
Success Rate	79.87%	72.63%	p=0.04
Time in sequences	360.87	331.57	p=0.17

Observation 6 (Bonus group of 2): Managerial Bonuses induce participants to reduce their cognitive effort relative to the PFP treatment. In the bonus treatment (group of 2) men spent less time on solving sequence and indeed solved significantly less sequences than in the PFP treatment while women spent a similar amount of time on sequences as in the PFP treatment, but with a lower success rate.

To illustrate the players' choice of time allocation we provide in Figure 4 below the minute by minute allocation of time between sequences and filing.

Figure 4: Allocation of Time, Minute by Minute (PFP vs. Bonus 2)



4.2 Managerial Bonus (group of 6).

We now turn to examine the effect of managerial bonus when the group size is six. The intuition was that when the group was larger the effect of managerial bonus would be lower. In Table 4 below we compare time spent on sequences and success rates under PFP and managerial bonus, for the "group of six" treatment.

Table 4: Managerial Bonus (group of 6)

Men	**PFP**	**Bonus (6)**	**Ttest**
Success Rate	76.23%	76.58%	p=0.52
Time in sequences	419.82	355.20	p=0.027

Women	**PFP**	**Bonus (6)**	**Ttest**
Success Rate	79.87%	80.86%	p=0.57
Time in sequences	360.87	349.06	p=0.36

Observation 7 (bonus group of 6): Group size affects men and women in a different manner. Men were not affected by the group size and solved fewer sequences and devoted less time to sequence solving relative to the PFP treatment even when the group size was six. Women, on the other hand, reduce their cognitive effort only when managerial bonuses were provided in a group of two. When the group became larger (group of six) they exert the same cognitive effort as in the PFP treatment.

Our intuition is that the decline of cognitive effort is due to an emotional reaction to managerial bonuses. We expect however that envy, anger or disappointed would decline with group size. Our findings that cognitive effort exerted by women is sensitive to the group size may indicate that these emotions may play an important role in explaining their reaction to managerial bonuses. On the other hand we expect that fairness concern is less sensitive to group size. It seems indeed unfair that one of the participants will receive a much larger payment independent of his/her performance. The fact that in our experiment men exert less cognitive effort whether the group size was two or six, may imply that effort choices by men in this treatment may be affected by fairness concerns.

5. Concluding Comment

Economic theory typically argues that workers, students or researchers should be given strong and competitive incentives in order to induce them to exert more effort. There are however several studies showing that this intuitive effect does not always hold. For example, Gneezy and Rustichini (2000) and Frey and Jegen (2001) demonstrated the crowding out effect, where strong explicit incentives may crowd out different types of social motivation and may result in less effort.[16] The main result of this paper focuses on yet another shortcoming of strong competitive incentives – it may induce agents to work harder but not necessarily smarter.

[16] Another example is Fershtman and Gneezy (2011) who showed that strong competitive incentives in dynamic tournaments may induce a large extent of quitting.

References

1. Ariely D., U. Gneezy, G. Loewenstein, and N. Mazar, (2009), "Large stakes and big mistakes", *Rev. Econ. Studies*. 76: 451-469.

2. Baumeister, R.F. (1984), "Choking under pressure: self conscious and paradoxical effect of incentives on skillful performance", *J. Pers. Soc. Psychol.* 46: 610-620.

3. Baumeister, R.F. and C.J. Showers, (1986), "A review of paradoxical performance effects: choking under pressure in sports and mental tests", *Eur. J. Soc. Psychol.* 16: 361-383.

4. Croson, R. and U. Gneezy, (2009), "Gender differences in preferences", *J. Econ. Lit.* 47: 448-474.

5. Fershtman, C. and U. Gneezy, (2011), "The Tradeoff between Performance and Quitting in High Power Tournament", *Journal of the European Economic Association*, Vol. 9(2) pp. 318-336.

6. Frey, B.S. and R. Jergen (2001), "Motivation crowding theory", *J. Econ. Surveys* 15: 589-611.

7. Fehr, E., and A. Falk, (2002), "Psychological foundation of incentives", *Eur. Econ. Rev.* 46: 687-724.

8. Gneezy, U. and A. Rustichini (2000), "Pay enough or don't pay at all" *Q. J. Econ.* 115: 791-810.

9. Gneezy, U. M. Niederle, and A. Rustichini (2003), "Performance in competitive environments: gender differences" *Q. J. Econ*. 118: 1049-1074.

10. Lazear, E. (2000), "The future of personnel economics", *Econ. J.* 110: F611-F639.

Appendix

Table A1: Winners and Losers in the tournament treatment

	Winners N=30 Men=14, Women=16	**Losers** N=30 Men=11, Women=19	Two-sided Ttest
Total Performance	44.38	21.21	p=0.00
Women	39.75	20.17	p=0.00
Men	49.67	23.18	p=0.00
Success rate	0.84	0.58	p=0.00
Women	0.80	0.56	p=0.00
Men	0.89	0.64	p=0.02
Time on Sequences	347.83	313.35	p=0.39
Women	302.97	313.37	p=0.85
Men	399.11	313.30	p=0.14
Net File Rate	1.10	1.02	p=0.13
Women	1.12	1.00	p=0.057
Men	1.07	1.06	p=0.88

Figure A1: Allocation of Time, Minute by Minute