THE PINHAS SAPIR CENTER FOR DEVELOPMENT
TEL AVIV UNIVERSITY

"Tournament and the Quality of Work"

Anat Bracha

Discussion Paper No. 5-12

January 2012

I thank The Pinhas Sapir Center for Development, Tel Aviv University for their financial support.

Abstract

Almost all jobs require a combination of cognitive effort and labor effort. This paper focuses on the effect that competitive incentive schemes have on the chosen combination of these two types of efforts. We use an experimental approach to show that competitive incentives may induce agents to work harder but not necessarily smarter. This effect was stronger for women.

1. Introduction

Competitive incentives are commonly used by firms and modern organizations to motivate their workers. While competitive incentives were proven to be effective in inducing more effort the focus of this paper is whether such incentives induce workers to work smarter, or just work harder.

Almost all jobs require a combination of cognitive effort and labor effort. Research and development (R&D), running a company, building a house, teaching a class, or working on a factory floor requires both cognitive and labor efforts. The tradeoff between these two types of effort 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. For instance, consider 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 parameter values (labor effort), or they may attempt a combination of the two methods. It is the combination of cognitive and labor effort that determines whether people work *harder or smarter.*

The general intuition is that providing competitive incentives motivates individuals to exert more effort.[2] But once we distinguish between cognitive and labor efforts an additional question arises: what is the effect of different incentive schemes on the *combination* of the two types of efforts? If different incentive schemes do indeed influence workers' chosen combination of effort this could be important for firms, as different firms may wish to elicit different combinations of effort. Firms that specialize in R&D and innovation, for instance, may want workers to exert more cognitive effort—to work smarter—while some manufacturing firms may wish to motivate workers' labor effort—to work harder. The main application of our study is that the optimal incentive scheme also depends on the type of effort firms wish to elicit.

To answer the question of whether incentive schemes affect the chosen combination of effort, we consider and examine players' effort allocation under two types of incentives: (i) a simple pay-for-performance incentive scheme (hereinafter PFP), in which agents are paid according to their own performance, and (ii) tournament incentives

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

in which pairs of participants compete for a prize. Our claim is that, due to competitive pressure, competitive incentives will lead participants to exert less cognitive effort but more labor effort, relative to PFP. That is, we expect that under a competitive incentive scheme participants will work harder but not smarter.

This conjecture is related to the psychological literature identifying several mechanisms that result in "choking under pressure." This literature suggests that pressure in various forms such as large stakes, performing in front of an audience, and competition may lower performance in various types of tasks (see for example Baumeister1984; Baumeister and Showers 1986; Beilock, Kulp, Holt, and Carr 2004). In economics, Ariely, Gneezy, Loewenstein, and Mazar (2009) recently demonstrated the choking under pressure effect and showed that excessively high rewards have a detrimental effect on performance. However, choking under pressure does not necessarily occur in all circumstances: for instance, Beilock, Kulp, Holt, and Carr (2004) show that pressure adversely affect performance in solving novel but not heavily practiced math problems[3]. Where solving novel math problems required high working memory—cognitive effort—while the practiced questions did not.

To test our claim that under competition people will work harder, not smarter, we designed a simple lab experiment with two computerized tasks, a “sequences” task where participants were asked to solve numeric sequences that require cognitive effort and a “filing” task that is a simple number categorizing task mainly requiring 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 experiment. In examining performance across the two incentive schemes we focus on whether the competitive incentives affected time allocation between the sequence task and the filing task as well as the players' success rate in solving sequences, compared with PFP.[4]

Our main results are that under competitive incentives participants devote less time to the sequence task and have a lower success rate than when they are provided with PFP incentives. In other words, the tournament incentive did indeed lead participants to work harder but not smarter.

[3] To correctly answer a practiced problem one only needed to retrieve the solution from memory.
[4] Success rate is the percentage of sequences solved correctly over the number of sequences attempted.

However, our results are gender sensitive. Under the PFP incentives the performance of women is lower than the performance of men, as women attempted to solve less sequences and devoted more of the allotted time to the simpler task of categorizing numbers. This is despite the fact that in the PFP treatment men and women had the same success rate in solving sequences. Analyzing the effect of competitive incentives by gender, we find that relative to the PFP treatment, competitive 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 competitive incentives on the success rate is entirely a female effect.

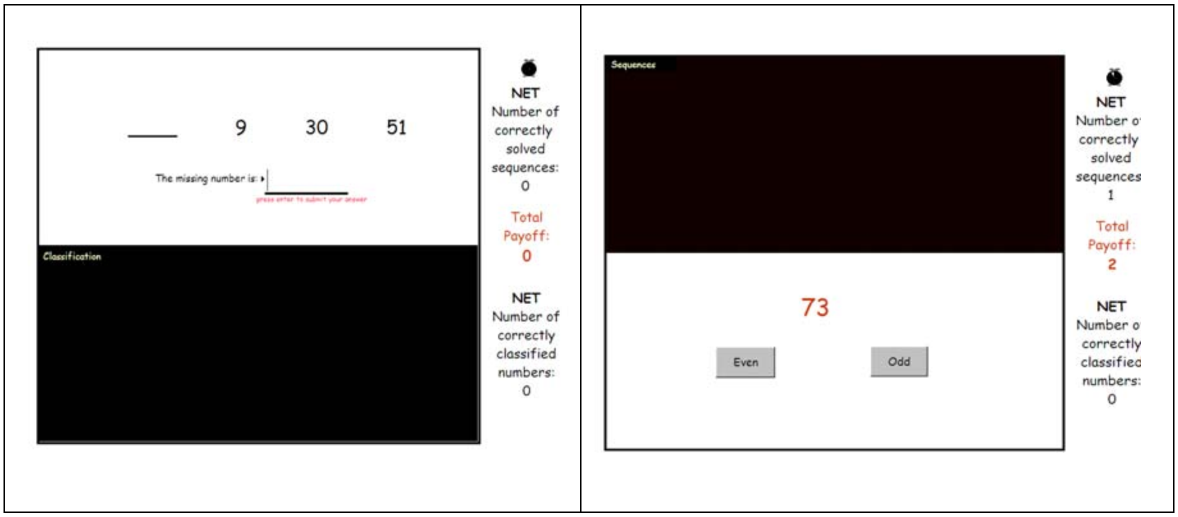
This paper adds to the experimental investigation of tournament relative to pay-for-performance incentives (see, for example, Bull, Schotter and Weigelt 1987) and the recent literature on gender differences in response to competition (for example, Gneezy, Niederle, and Rustichini 2003; for a survey see Croson and Gneezy 2009 and Bertrand 2011). However, previous studies either use non real-effort settings where effort is a number chosen by a participant in the study, or a real-effort task which requires one type of effort. Hence, this paper adds to the literature by considering a new aspect of competitive incentives—its effect on effort allocation—and points at the differential effect of competition on effort allocation across gender.

This paper also adds to the recent literature on the psychological foundation of incentives, which 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. Our results extend this literature by focusing on the combination of cognitive and labor effort.

2. Experimental Design

To capture the two different types of effort we introduced two tasks, each of which emphasized either cognitive or labor effort. Subjects could freely choose to engage in either solving sequences ("sequences" task)—finding a missing number in a sequence of four numbers—or classifying a random number into an “odd” or “even” category ("filing" task) by pressing the appropriate button on the computer 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 experiment, and engaging in each of the two tasks was done by simply clicking on the section of the screen with the desired task (see figure 1).

Figure 1: Sequences Task and Filing Task



2.1 Treatments

To analyze the effect of incentives we use a between subject design with two treatments: Pay for Performance and head-to-head tournament. In our experimental design the goal was to earn money, and the complementarity between the tasks was achieved using the compensation scheme 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 obtained by multiplying the net sequences by the net filed numbers.[5] The 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. The net filed numbers equals the correctly filed numbers minus the incorrectly filed numbers. Penalizing incorrectly filed numbers was designed to prevent random clicking. The extra reward introduces a complementary term as a greater number

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

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 identification (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 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.

2.3 Procedure

The sessions were conducted at the Harvard Decision Science Laboratory at the Harvard Kennedy School. A total of 134 Harvard students participated in the study, 74 in the PFP treatment and 60 in the tournament treatment. In each session, participants sat at individual computer stations and read the instructions on their individual screens. 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. Under both conditions, they were given 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.

3. Competitive Incentives

Economic intuition suggests that competitive incentives may induce individuals to exert more effort.[6] In contrast, our focus is on the effect of competitive incentives on the chosen combination of cognitive and labor effort. Our hypothesis is that competitive incentives have two effects. (i) These incentives induce individuals to choose a different combination of cognitive/labor effort reducing their *cognitive* effort while increasing their labor. (ii) Competitive incentives reduce the effectiveness of cognitive effort. To test these hypotheses we examine two variables: time allocation between the two tasks and the success rate in solving sequences. We analyze these effects using the appropriate t-test. [7]

3.1 The Effect of Competitive Incentives

Comparing the data from the PFP and the tournament treatment, we find that the overall performance in the tournament treatment was slightly (but not significantly) higher than in the PFP treatment. However, under the PFP incentive, the average number of attempted sequences was 10.7 (with 8.5 solved correctly) while under the tournament treatment the average number of sequences attempted was 9.2 (with 7.2 solved correctly).[8] Under PFP incentives participants devoted on average 381 seconds (approximately 6.33 minutes) to solving sequences, while in the tournament treatment

[6] See for example Lazear and Rosen (1981) for a theoretical argument. Empirically, Ehrenberg and Bognanno (1990) demonstrated the positive effect of tournament incentives on effort in golf tournaments, and Kremer et al (2009) find positive effect of scholarship competition for girls in Kenya on their school achievements. Experimentally, results of studies such as Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007), although aimed at gender differences, show that competitive incentives increase performance among men compared with piece rate, and even among women, performance is at least as good under tournament as under piece rate.

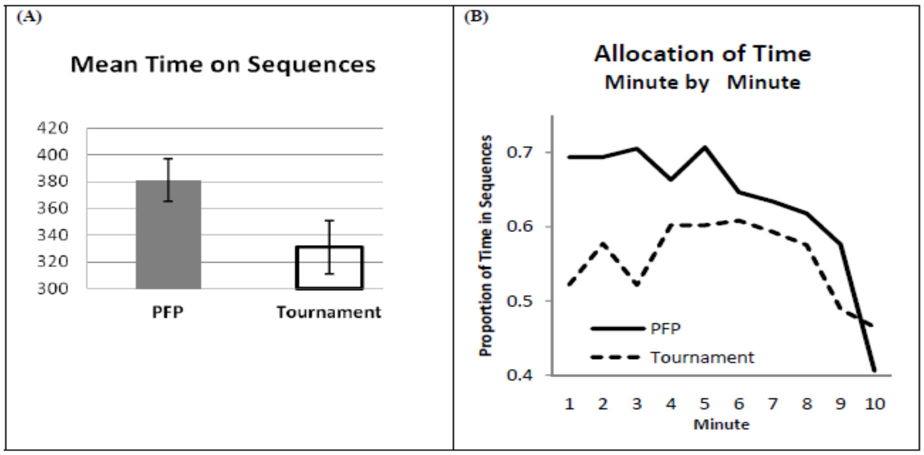
[7] In calculating the various averages, we first calculate the particular measure (such as success rate) for each individual and then average across individuals. When comparing across gender with no a priori hypothesis we use a two sided t-test and when we examine whether tournament incentives reduce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

[8] Under PFP incentives, 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 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.

they spent only 330 seconds (5.5 minutes) on solving sequences (p=0.025; see Figure 2A).

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 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 decisions made throughout the entire 10 minutes of the study.

Figure 2: Allocation of Time

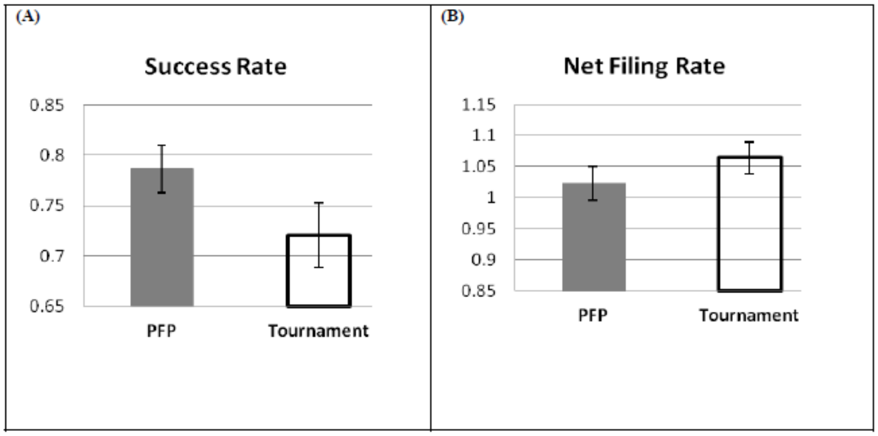


N=73 under PFP, N=59 under Tournament. The bars in panel A represent the standard error of the mean.

Regarding our second hypothesis we calculate success rate in solving sequences for each participant and then average these results across all individuals (see figure 3A). We find that the success rate is lower under competitive incentives (78.6 percent under PFP while only 72 percent under tournament). Labor effort can be captured by examining the average net filing rate which is the average speed of net filing across participants (see figure 3B).

Having a significantly lower success rate together with higher net filing rate (although the latter is not statistically significant) indeed demonstrates 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 the standard error of the mean.

Observation 1 (Tournament and cognitive effort): (i) Under competitive incentives participants devoted less time to the cognitive task than under the PFP incentives (330 second under competition versus 381 under PFP; p=0.025). (ii) The success rate in solving sequences is lower under tournament incentives—78.6 percent under PFP while only 72 percent under tournament incentives (p=0.047).

Clearly the two parts of Observation 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. The effect of competitive incentives on the success rate may be due to an effect on the effort expended, an effect on ability, or both. We may take the conservative economic interpretation of an effect on effort and treat ability as exogenous. However, under competitive 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 acting under pressure. Therefore it is possible that incentive schemes affect ability and such an effect should be taken into account whenever incentives are designed.

3.2 Time Pressure

If the 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 manifested more strongly at the end of the tournament when the competitive pressure is likely at its highest level. We therefore divide the 10 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 percent during the first seven minutes and only 57.9 percent during the last three minutes. This decline is highly significant ($p<0.01$). In the PFP treatment the success rate in the first seven minutes was 76.4 percent and 77.2 percent in the last three minutes. Of course it is possible that in the tournament treatment some low-ability individuals chose to solve sequences only during the last three minutes. To exclude this possibility we compared the success rate only for those individuals who solved sequences during both the first seven minutes and the last three minutes. We find the same pattern: there was no effect under PFP (76.12 percent success rate during the first seven minutes, and 77 percent success rate during the last three minutes) and there was a highly significant decline under competition (76.3 percent success rate during the first seven minutes compared with 57.5 percent success rate during the last three minutes; $p<0.01$.)

Observation 2 (Competitive Time Pressure): During the last three minutes of the tournament, the participants' success rate was significantly lower than in the first seven minutes. However, time pressure in and of itself has no such effect, as in the PFP treatment there was no reduction in participants' success rate during the last three minutes.

3.3 Winners and Losers.

Competitive incentives do not necessarily have a uniform effect on individuals: some people may be encouraged by and do better under competition, while others may get discouraged and perform worse due to competitive pressure. We therefore examine the effect of competitive incentive by subgroups. We first distinguish between winners and losers in the tournament treatment and compare their performance to the performance of the appropriate comparison group in the PFP treatment. We then split our sample by the overall performance quartile and examine the effect of competitive incentives on each quartile separately.

Table 1 presents the performance of the winners and the losers separately. We find that winners and losers spend statistically the same amount of time on solving sequences—the winners spend 347 seconds (approximately 5.78 minutes) on average while losers spend 313 second (approximately 5.22 minutes) on average, but this difference is not significant (p=0.39). Nevertheless, the winners' average score is 44.38 points, they solve on average 11.3 sequences, and their success rate is 84 percent while the losers' average score is 21.2 points, they solve only 7.20 sequences, and their success rate was 58 percent.[9]

Table 1: Winners and Losers in the Tournament Treatment

	Winners N=30	**Losers** N=29	Ttest
Total Performance	44.38	21.21	p=0.00
Success rate	0.84	0.58	p=0.00
Time on Sequences	347.83	313.35	p=0.39
Net File Rate	1.10	1.02	p=0.13

Observation 3: The winners in the tournament treatment are the participants with the higher success rate. Nevertheless, there was no difference between the tournament winners and losers with respect to the time they spent on solving sequences.

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 had 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 the winners' success rate in the tournament treatment to be lower than the success rate of the high-performing individuals in the PFP treatment, and for the losers in the tournament treatment to have a higher success rate compared with the below-median group in the PFP treatment. Contrary to our expectations, our findings were different: we found that the success rate of tournament winners and the above-median PFP performers was similar (85.29 percent under PFP and 84.64 percent for tournament winners). However, the tournament losers'

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

success rate was 58.96 percent and it was significantly lower than the 73.42 percent success rate of the below-median PFP performers ($p<0.01$).[10]

Since the above analysis indicates that competitive pressure may have a different effect on players, we now examine the performance of each quartile separately.[11] Splitting our population into quartiles will provide a better understanding of our findings but at the same time due to this division the results for the different quartiles are not significant. Table 2 presents the success rate, the net filing rate, and individuals' time allocation for each quartile separately.

Table 2: Performance by Quartile

Tournament	PFP		Quartile
.4380952	.6956078	Success Rate	1
1.023022	.8869749	Net Filing Rate	
310.9984	384.209	Time in Sequences	
.7162018	.7520596	Success Rate	2
1.008487	.9973465	Net Filing Rate	
309.9976	347.9667	Time in Sequences	
.8464196	.8409308	Success Rate	3
1.087253	1.076832	Net Filing Rate	
360.1371	411.3077	Time in Sequences	
.8622351	.8584484	Success Rate	4
1.137562	1.129284	Net Filing Rate	
339.7352	382.5957	Time in Sequences	

Our main observation that competitive incentives reduce the time devoted to the cognitive task applies to all quartiles. Moreover we can see that competitive incentives do induce individuals to devote more effort to their labor task as the net filing rate is higher under competitive incentives in each quartile (although this effect is significant only for the first quartile). The main difference between the groups is with respect to their success rate. The success rate for the first quartile is 69.5 percent under PFP incentives and went down by 37 percent to 43.8 percent under competitive incentives. For the second quartile success rate was 75.2 percent under PFP and decreased by 5 percent (to 71.6 percent)

[10] Interestingly, although both the tournament winners and losers seem to reduce the time they spent on solving sequences compared to above- and below-median PFP performers (respectively), the decline is significant only for the winners (399 seconds, or approximately 6.66 minutes under PFP versus 347 seconds, or approximately 5.66 minutes, under tournament incentives; p=0.074.)

[11] We define quartiles separately for the PFP and the tournament treatments.

under competitive incentives. At the same time success rate is slightly (but not significantly) higher under the competitive incentives for the third and fourth quartiles.

Note however that we compare different quartiles and not specific individuals. It is still possible that some individuals choke under pressure and move, for example, from the fourth quartile under PFP to the first quartile under competition while others thrive under competitive incentives and improve their performance in all parameters.

Observation 4: The decline in the time spent on sequences under competitive incentives and the higher net filing rate are evident across all quartiles. However, the decline in the success rate characterizes only the first two quartiles.

4. The Gender Effect

Recent studies have indicated that men and women respond differently to competitive incentives.[12] Much to our surprise, not only did we find gender differences in response to competitive incentives, we also found gender differences in the benchmark PFP treatment.

Observation 5 (PFP: Gender Effect):

(i) Women devoted on average 360 seconds (6 minutes) to sequences solving while men devoted 419 seconds (approximately 7 minutes) to solving sequences ($p=0.055$).
(ii) On average women solved 8.91 sequences correctly while on average men solved 14.16 sequences correctly ($p<0.01$).
(iii) Men and women had similar average success rates; 76.2 percent for men and 79.8 percent for women.[13]

Women's choice to devote less time to solving sequences cannot be the outcome of a lower success rate since in solving sequences they had success rates similar to men; in fact, although not statistically significant, women's average success rate in the PFP

[12] See Gneezy, Niederle, and Rustichini (2003) for gender differences in response to competition among college students, and Gneezy and Rutchini (2004) for gender differences in response to competition among children. See Niederle and Vesterlund (2007), Sutter and Rutzler (2010), as well as Datta Gupta, Poulsen, and Villeval (2011) for gender differences in selecting into competitive environment. Gneezy, Leonard, and List (2009) suggest that these differences may be due to nurture rather than nature; see also Booth and Nolen (2011). For recent reviews see Croson and Gneezy (2009) and Bertrand (2011). Interestingly, two forthcoming papers find no such gender differences in competitiveness (either performance or preference) among Swedish and Colombian children (Dreber, Von Essen, and Ranehill 2011; Cárdenas, Dreber, Von Essen, and Ranehill 2011).

[13] There was 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.

treatment is slightly higher than the men's. It is possible that the observed time allocation choice is the outcome of the gender difference in risk aversion and self-confidence (see for example Eckel and Grossman 2008a; for a review see Eckel and Grossman 2008b, as well as Croson and Gneezy 2009 and Bertrand 2011), 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. In other words, if there were relatively more women participating in the tournament treatment than in the PFP treatment, then the different gender mix could explain the decline in the overall time devoted to solving sequences in the tournament treatment. However in our experiments there were relatively more women participating in the PFP treatment than in the tournament treatment, so we would expect the opposite effect.[14] Nevertheless, below we 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 was not different in the tournament treatment and the PFP treatment (measured by the average score in points in the tournament treatment, and average payoff in 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. The differences for women and men are insignificant.

The Effect of Competitive Incentives on Women: Women's success rate in solving sequences declined from 79.87 percent under PFP to 67.18 percent under tournament ($p<0.01$), a sharp and strong decline of over 15 percent in the success rate. This decline was evident in the last three minutes of the experiment—a decline from 77.11 percent under PFP to 49.74 percent under tournament ($p<0.01$)—but was not evident during the first seven minutes. Under the PFP incentives women spent an average of 360 seconds (6 minutes) on solving numerical sequences, while under tournament incentives they spent

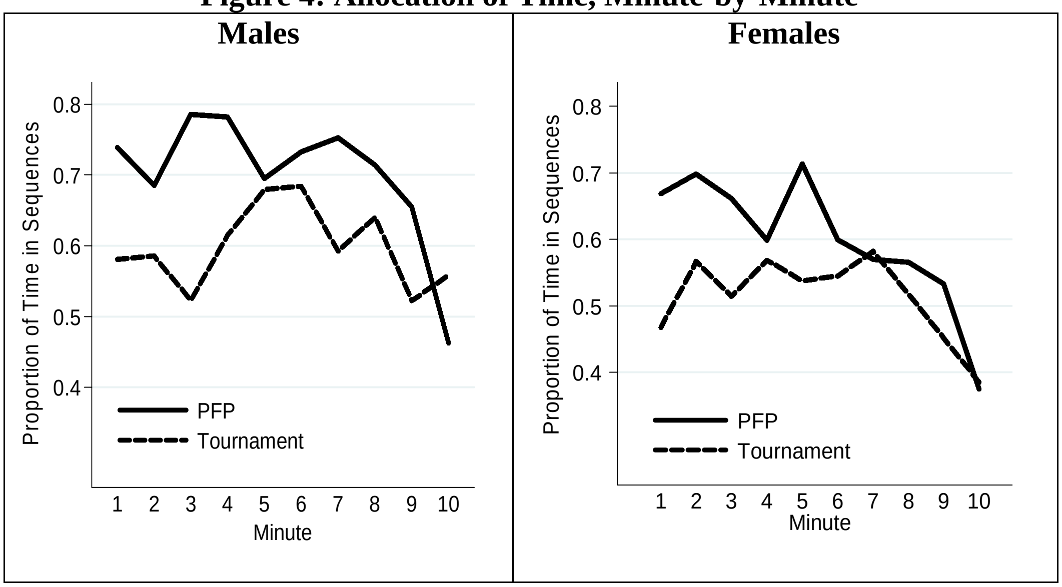
[14] In the PFP treatment, 48 females and 25 males participated; in the tournament treatment, 35 females and 24 males participated.

on average only 308 seconds (about 5.13 minutes; p=0.066).[15] See figure 4 below for the minute-by-minute time allocation in the PFP and the tournament treatments.

The Effect of Competitive Incentives on Men: In the tournament treatment men reduced the amount of time they devoted to solving sequences from 419 seconds (almost 7 minutes) to 363 seconds (a little over 6 minutes) (p=0.059). The average number of sequences men attempted 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, which was 76.23 percent in the PFP treatment and 78.69 percent in the tournament treatment.

Observation 6 (Gender and Tournament): (i) Both men and women reduce the time they spent on sequence solving when facing tournament incentives. (ii) Tournament incentives only affected the success rate of women. This effect is mainly due to pressure at the end of the tournament.

Figure 4: Allocation of Time, Minute-by-Minute



These results are also reflected in a regression analysis which controls for gender and age: we used an OLS regression of the success rate and the time devoted to solving sequences on a treatment dummy variable (that takes a value of 1 for a tournament

[15] This led to a significantly higher net filing among women (339.52 under tournament versus 258.71 under PFP; p=0.024).

treatment), gender (that takes a value of 1 for females), and age. The results are presented in table 3.

Table 3: OLS Regressions

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

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.

As table 3 shows, we find that under competitive incentives success rate is lower by 7.5 percent, which is approximately 9.5 percent of the average success rate under PFP (p=0.06). Under tournament incentives the time allocated to solving sequences is lower 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 solely a female effect (p=0.06), while the decline in the time allocated to solving sequences is similar for both men and women (the interaction term is not significant).

Interestingly, in our settings women won the tournament at a similar proportion as did men in contrast to previous findings (see table A1; Gneezy, Niederle, and Rustichini 2003, and Gneezy and Rustchini 2004). 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, by randomly matching participants in the tournament condition, and repeating this test, we find that out of 100 random matchings there were only eight instances with significant differences in the winning proportions across genders. The result is that we reject the hypothesis that the average z statistics across all 100 random matches is equal to or greater than 1.96. Therefore, this lack of difference in the winning proportions across gender is not an artifact of the actual matching we used in the study.[16]

5. Concluding Comment

Modern organizations typically provide workers, managers, students, or researchers strong and competitive incentives in order to induce them to exert more effort. However, there are several studies showing that this intuitive effect of competitive incentives 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 (for a recent survey see Gneezy, Meier and Rey-Biel 2011).

The main result of this paper focuses on yet another shortcoming of strong competitive incentives— these may induce agents to work harder but not necessarily smarter. The distinction between these two types of efforts should be taken into account by organizations when they design their incentives schemes. There are many situations in which the labor and the cognitive efforts cannot be directly observed, are not verifiable, and cannot be incentivized separately. In these cases organizations need to evaluate the relative importance and implications of different combinations of cognitive and labor efforts. If firms have strong preferences for motivating their employees to work smarter but cannot directly incentivize cognitive effort, the implication of our finding is that these organizations should refrain from competitive incentives.

[16] 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.

References

1. Ariely Dan, Uri Gneezy, George Loewenstein, and Nina Mazar. 2009., "Large Stakes and Big Mistakes." *Review of Economic Studies*76 (2): 451–469.

2. Baumeister, Roy F. 1984.) "Choking under Pressure: Self-Consciousness and Paradoxical Efects of Incentives on Skillful Performance."*Journal of Personality and Social. Psychology* 46(3): 610–620.

3. Baumeister, Roy F. and Carolin J. Showers. 1986. "A Review of Paradoxical Performance Effects: Choking under Pressure in Sports and Mental Tests." *European Journal of Social Psychology* 16(4): 361–383.

4. Beilock Sian L., Catherine A. Kulp, Lauren E. Holt, and Thomas H. Carr. 2004. “More on the Fragility of Performance: Choking Under Pressure in Mathematical Problem Solving.” *Journal of Experimental Psychology: General*133(4):584–600.

5. Bertrand Marianne. 2011. “New Perspectives on Gender.” In *Handbook of Labor Economics*, Volume 4b, ed. Orly Ashenfelter and David Card 1543–1590. Amsterdam: North Holland.

6. Booth Alison, and Patrick Nolen. Forthcoming. “Choosing to Compete: How Different are Girls and Boys?” *Journal of Economic Behavior & Organization.*

7. Bull Clive, Andrew Schotter, and Keith Weigelt. 1987. “Tournaments and Piece Rates: An Experimental Study.” *Journal of Political Economy*95(1): 1–33.

8. Cárdenas Juan-Camilo, Anna Dreber, Emma Von Essen, and Eva Ranehill. . Forthcoming. “Gender Differences in Competitiveness and Risk Taking: Comparing Children in Colombia and Sweden.” *Journal of Economic Behavior & Organization*.

9. Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2): 448–474.

10. Datta Gupta, Nabanita, Anders Poulsen, and Marie Claire. Villeval. Forthcoming. “Gender Matching and Competitiveness: Experimental Evidence.” *Economic Inquiry.*

11. Dreber, Anna, Emma Von Essen, and Eva Ranehill. Forthcoming. "Outrunning the Gender Gap – Boys and Girls Compete Equally." *Experimental Economics*.

12. Eckel Catherine C. and Philip J. Grossman 2008a. “Forecasting Risk Attitudes: An Experimental Study using Actual and Forecast Gamble Choices.” *Journal of Economic Behavior &Organization* 68 (1): 1–17.

13. Eckel Catherine .C. and Philip J. Grossman. 2008b. “Men, Women and Risk Aversion: Experimental Evidence.” In *Handbook of Experimental Economics Results,* ed.Charles R. Plott and Vernon L. Smith, 1061–1073. Amsterdam: Elsevier.

14. Ehrenberg Ronald G., and Michael L. Bognanno. 1990. “Do Tournaments Have Incentive Effects?” *Journal of Political Economy* 98(6): 1307–1324.

15. Frey, Bruno. S., and Reto Jergen. 2001. "Motivation Crowding Theory." *Journal of Economic Surveys* 15(5): 589–611.

16. Fehr, Ernst, and Armin Falk. 2002. "Psychological Foundation of Incentives." *European Economic Review* 46(4–5): 687–724.

17. Gneezy, Uri, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics* 115(3): 791–810.

18. Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics* 118(3): 1049–1074.

19. Gneezy, Uri, and Aldo Rustichini. 2004. “Gender and Competition at a Young Age.” *American Economic Review Papers and Proceedings* 94(2): 377–381.

20. Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. “Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society.” *Econometrica* 77(5): 1637–1664.

21. Gneezy, Uri, S. Meier and Pedro Rey-Biel. 2011. "When and Why incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives* 25(4): 191–210.

22. Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. “Incentives to Learn.” *Review of Economics and Statistics* 91(3): 437–456.

23. Lazear Edward P., and Sherwin Rosen. 1981., “Rank-Order Tournaments as Optimum Labor Contracts.” *Journal of Political Economy* 89(5): 841–864.

24. Lazear, Edward P. 2000. "The Future of Personnel Economics." *Economic Journal* 110(467): F611–F639.

25. Niederle, Muriel, and Lise Vesterlund. 2007. “Do Women Shy Away from Competition? Do Men Compete Too Much?” *Quarterly Journal of Economics* 122(3): 1067–1101.

26. Sutter Matthias, and Daniela Rützler. 2010., "Gender Differences in Competition Emerge Early in Life." *IZA Discussion Paper Number 5015.* Available at http://ftp.iza.org/dp5015.pdf.

Appendix

Table A1: Winners and Losers in the Tournament Treatment

	Winners N=30 Men=14, Women=16	**Losers** N=29 Men=10, Women=19	Ttest
Total Performance Women Men	44.38 39.75 49.67	21.21 20.17 23.18	p=0.00 p=0.00 p=0.00
Success rate Women Men	0.84 0.80 0.89	0.58 0.56 0.64	p=0.00 p=0.00 p=0.00
Time on Sequences Women Men	347.83 302.97 399.11	313.35 313.37 313.30	p=0.39 p=0.85 p=0.14
Net File Rate Women Men	1.10 1.12 1.07	1.02 1.00 1.06	p=0.13 p=0.057 p=0.88