

# Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints

## SUPPLEMENTARY MATERIAL

Olga Shurchkov

October 2010

### **Abstract**

This supplementary document contains documentation and additional data analysis. Appendix A provides experimental instructions and questionnaire. Appendix B details the generation of sequences for the math treatments. Appendix C presents additional specifications and robustness checks for the data analysis section of the paper. Finally, Appendix D includes the analysis of data restricting the sample to only the first three rounds of the experiment.

## **Appendix A: Instructions and Questionnaire**

The instructions below were distributed to the subjects and read out loud prior to the start of the experiment in all sessions. Subjects were encouraged to read the information screens between rounds carefully for further instructions. These instructions explained the payment scheme differences from round to round. The questionnaires were distributed electronically at the end of the experiment.

## Instructions for the experiment

Thank you very much for participating in this experiment which involves solving math puzzles. Please read the following instructions very carefully. You will receive all the information you require for participation in the experiment. If you follow these instructions, you will have an opportunity to earn real money that will be paid to you, privately and in cash, at the end. If you do not understand something, please raise your hand and wait for the experimenter to come to your place and answer your question privately.

**Communication between participants is absolutely forbidden during the experiment!** Not obeying this rule will lead to immediate exclusion from the experiment and all payments.

### **Groups:**

Throughout the experiment, you will be a member of a **group of 4 people**. The other three members of your group are sitting in the same row of desks as you.

### **Puzzles:**

You will be solving **Numbers-in-a-Numbers** puzzles. You are NOT allowed use a calculator or any computer program for this game! You are allowed to use scrap paper for calculations by hand.

From a random sequence of digits, you will have to find the combinations of numbers (“solutions”) that add up to a specific target number. Everyone will be working on the same puzzle at the same time.

### **General rules:**

- The goal is to permute the digits in the sequence to come up with number combinations (“solutions”) that add up to a given target number.
- You can use any of the available digits in the sequence for your solution (not just the consecutive digits), but each digit in the sequence can only be used once for each new solution.
- The target number must be obtained through addition; no other arithmetic operation is allowed.
- A valid solution contains sums of 2 or 3 numbers total.
- A valid solution contains sums of 2-digit and 3-digit numbers.
- Each number in a solution must be separated by a “+” sign to signify addition.
- Acceptable characters: digits 0-9; + (no spaces in between characters).

### **Scoring rules:**

- For a valid solution:
  - $\#Points\ Earned = \#Total\ Digits + \#Numbers - 5$
- For invalid solutions:
  - $\#Points\ Lost = \#Total\ Digits + \#Numbers - 5$
- Putting a “,” between numbers makes you lose 1 point for each number
- If the solution has wrong characters, is too long, or is too short, you lose 1 point
- If the solution is a duplicate (reversing the order of numbers counts as a duplicate), you do not lose or earn any points

Example: Sequence 1034582614 and the target number is 117.

<u>Valid solutions:</u>	<u>Score:</u>	<u>Invalid solutions:</u>	<u>Score:</u>
• 103+14	+2	• 103, 14 (need a +, not a comma)	-2
• 54+63	+1	• 5+112 (5 is only a 1-digit number)	-1
• 14+43+60	+4	• 101+16 (1 appears only twice in the sequence)	-2
		• 03+114 (0 cannot be used as a placeholder)	-1

Total score = 7 – 6 = 1 Point.

### **Rounds:**

The entire experiment consists of 7 rounds. Each round asks you to work on a different puzzle. The first round is for practice. The instructions for the following rounds will be given to you on your screens. **Make sure you read the instructions on your screens before every round very carefully!**

### **Exact Procedure**

At the beginning of the experiment, you will receive an ID number. For example, members of the same group might receive ID numbers A1, A2, A3, and A4. Members of another group might receive ID numbers B1, B2, B3, and B4.

On the first screen you see, you will enter your Group Name (A, B, C, etc) and User ID Number (1, 2, 3, etc).



#### **Welcome**

Welcome to NumbersInNumbers. Please enter your user ID number.

Group Name:

User Number:

Submit

Next, you will click “Submit.”

On the next screen, read the directions carefully!



**Start of Round 1 (Round Type: Practice)**

This is a **PRACTICE** round. In this round, you will have 2 minutes to solve one number puzzle for practice. This puzzle will not add or subtract from your total score.

**Proceed**

Then press "Proceed".

On the following screen, you will be presented with the practice puzzle. Remember that you will have 2 minutes to come up with as many valid solutions as you can. After typing in your solutions, you can either press the "Enter" key on the keyboard or click "Submit." Make sure you check the arithmetic carefully, since you will NOT be able to go back and correct your solution once it had been submitted.



**Round 1 (Round Type: Practice)**

Find all combinations of 2 or 3 digit numbers in the sequence **097126525503033** that add up to **149**.

E.g. "25+75"

**Submitted Solutions:**

50+73+26,  
79+70,  
123+26,  
26+123,  
57+92,  
57,92

**Time Left:**

The time will run out automatically. Once it does, you will see your score (the number of points you earned in this practice round). Please, click on "Click to Proceed to Round 2".

# Numbers In Numbers

## Results For Round 1 (Round Type: Practice)

This is the end of the PRACTICE round.

Valid Solutions: 

26+50+73 (+4)  
26+123 (+2)  
57+92 (+1)

Invalid Solutions: 

70+79 (-1)  
26+123 is a duplicate (-0)  
57 (-1)  
92 (-1)

Total Score: 4

[Click To Proceed To Round 2](#)

Please, be patient. You might need to wait a few seconds for other group members to finish each round!

The next screen is an **information screen** for Round 2. At this point, you will get the rules for the up-coming round. Read this information carefully! Your earnings depend on it!

Then, click the “Proceed” button at which point the next round will start. All subsequent rounds follow the same procedure.

After the last round you will be asked to fill out a brief questionnaire. You will receive your payment at the end of the experiment after you complete the questionnaire. Payment calculation is as follows:

	Show-up fee [= \$10]
+	Your income from Round 2
+	Your income from Round 3
+	Your income from Round 4
+	Your income from Round 5
+	Your income from Round 6
+	Your income from Round 7
=	<hr/> Total income from the experiment

## Instructions for the experiment

Thank you very much for participating in this experiment which involves solving word puzzles. Please read the following instructions very carefully. You will receive all the information you require for participation in the experiment. If you follow these instructions, you will have an opportunity to earn real money that will be paid to you, privately and in cash, at the end. If you do not understand something, please raise your hand and wait for the experimenter to come to your place and answer your question privately.

**Communication between participants is absolutely forbidden during the experiment!** Not obeying this rule will lead to immediate exclusion from the experiment and all payments.

### **Groups:**

Throughout the experiment, you will be a member of a **group of 4 people**. The other three members of your group are sitting in the same row of desks as you.

### **Puzzles:**

You will be solving **Word-in-a-Word** puzzles. You will have to find as many smaller sub-words that can be formed out of the letters of the big word as you possibly can. Everyone will be working on the same puzzle at the same time.

### **General rules:**

- The words must be 4-letters long or longer
- Acceptable characters are letters A-Z only. Any other symbol like a number or another symbol will be automatically discarded
- Proper nouns are not allowed
- Plurals are allowed
- Each letter in the word can only be used once

### **Scoring rules:**

- Valid words add  $(N - 3)$  points to your score, where  $N$  is the total number of letters in the word
- Invalid words subtract  $(N - 3)$  points from your score, where  $N$  is the total number of letters in the word
- If the word is a duplicate, no points are subtracted, no points are added
- If the word is too short, 1 point is subtracted from the total score (regardless of whether the word is 1, 2, or 3 letters long)

**Example:** Big word PERSUASIVELY. You enter: persuasive (+7), live (+1), live (0), lyve (-1), lyver (-2), sap (-1), as (-1). The total number of points is 3.

### **Rounds:**

The entire experiment consists of 7 rounds. Each round asks you to work on a different word puzzle. The first round is for practice. The instructions for the following rounds will be given to you on your screens. **Make sure you read the instructions on your screens before every round very carefully!**

## Exact Procedure

At the beginning of the experiment, you will receive an ID number. For example, members of the same group might receive ID numbers A1, A2, A3, and A4. Members of another group might receive ID numbers B1, B2, B3, and B4.

On the first screen you see, you will enter your Group Name (A, B, C, etc) and User ID Number (1, 2, 3, etc).



### Welcome

Welcome to words in a word. Please enter your user ID number.

Group Name:

User Number:

Submit

Next, you will click “Submit.”

On the next screen, read the directions carefully!



### Start of Round 1 (Round Type: Practice)

This is a **PRACTICE** round. In this round, you will have 1 minute to solve one word puzzle for practice. This puzzle will not add or subtract from your total score.

Proceed

Then press “Proceed”.

On the following screen, you will be presented with the practice puzzle. Remember that you will have 1 minute to come up with as many valid words as you can. After typing in your sub-words, you can either press the “Enter” key on the keyboard or click “Submit.” Make sure you check the spelling of your sub-words carefully, since you will NOT be able to go back and edit them once they have been submitted.

# Words In Words

## Round 1 (Round Type: Practice)

Submitted words:

```
infuse,  
fine,  
cats
```

Find all the subwords in the word **INFUSCATE**.

Submit

Time Left:

The time will run out automatically. Once it does, you will see your score (the number of points you earned in this practice round). Please, click on “Click to Proceed to Round 2”.

# Words In Words

## Results For Round 1 (Round Type: Practice)

This is the end of the **PRACTICE** round.

Valid Words: 

infuse (3)  
cats (1)  
fine (1)

Invalid Words: 

Total Score: 5

[Click To Proceed To Round 2](#)

Please, be patient. You might need to wait a few seconds for other group members to finish each round!

The next screen is an **information screen** for Round 2. At this point, you will get the rules for the up-coming round. Please, read this information carefully! **Your earnings depend on it!**

Then, click the “Proceed” at which point the next round will start. All subsequent rounds follow the same procedure.

After the last round you will be asked to fill out a brief questionnaire. You will receive your payment at the end of the experiment after you complete the questionnaire. Payment calculation is as follows:

	Show-up fee [= \$10]
+	Your income from Round 2
+	Your income from Round 3
+	Your income from Round 4
+	Your income from Round 5
+	Your income from Round 6
+	Your income from Round 7
	<hr/>
=	Total income from the experiment

## QUESTIONNAIRE

We ask you now to fill out a brief questionnaire. Please, answer as honestly and completely, as you possibly can, since your answers will help us tremendously!

1. What is your ID number (i.e. A1, B3, etc.)? \_\_\_\_\_
2. What is your gender? \_\_\_\_\_
3. What is your age? \_\_\_\_\_
4. Are you a native English speaker (i.e., is English your first language)? \_\_\_\_\_
5. Are you currently a student? \_\_\_\_\_  
And if yes, what is your field of study? \_\_\_\_\_

6. In your opinion, who would be better in these tasks on average, men or women (circle one)?

Men

Women

7. How did changing the rules from non-competitive to competition change your effort and performance in the puzzles? Why?

8. Circle one answer:

Competition:            a. helped me.                      b. hurt me

I tried more in:            a. non-competitive rounds                      b. competitive rounds

I gave up:                      a. never                                      b. more in non-competitive rounds

                                            c. more in competitive rounds                      d. in both type of rounds

**THANK YOU VERY MUCH FOR YOUR PARTICIPATION IN THIS STUDY!!!**

## Appendix B: Sequence Generation

In selecting the sequences of digits for the math puzzle part of the experiment, I adhere to the following criteria. Note that all sequences are randomly generated.

1. Each sequence contains 15 digits, 0-9.
2. Exactly two digits must be missing from any given sequence (e.g., there are no 2's and no 5's).
3. Within each given sequence, one of the digits must be repeated exactly twice (e.g., there are two 4's); and three digits must be repeated exactly three times (e.g., there are three 1's, three 9's, and three 0's).

The following rules are used to generate a target number for each puzzle.

1. The target number must be a three-digit number
2. The target number must be less than 200
3. The target number must not be a multiple of 10.

In order to preserve consistency across tasks, I picked sequences that are similar along several dimensions. In particular, each sequence has a maximum score in the range of 455-461 points, the total number of solutions in the range of 122-123, and the total number of 2-factor solutions in the range of 14-15. Finally, I picked sequences with target numbers that are as different as I could find.

The resulting sequences I use in the experiment are:

- 436771974115604 with target number 135.
- 647029590696014 with target number 165.
- 903538359266169 with target number 182.
- 497220002195953 with target number 145.
- 845196336864734 with target number 197.
- 674639419829848 with target number 193.

# Appendix C: Additional Tables and Figures

## Descriptive Statistics and Tests

Table C.1. Descriptive Statistics

Variables	Math Sessions			Verbal Sessions		
	Men	Women	p-values	Men	Women	p-values
<u>Score</u>						
Piece-Rate, High Pressure	5.17	5.11	(0.97) [0.43]	12.91	14.19	(0.30) [0.94]
Tournament, High Pressure	6.31**	2.39**	(0.03) [0.05]	9.76*	11.83*	(0.07) [0.09]
Piece-Rate, Low Pressure	29.50	19.94	(0.23) [0.32]	11.74	13.19	(0.34) [0.28]
Tournament, Low Pressure	32.31	23.11	(0.28) [0.51]	17.81***	23.22***	(0.00) [0.00]
<u>Mistake Share</u>						
Piece-Rate, High Pressure	27%	21%	(0.43) [0.33]	8.0%	5.3%	(0.32) [0.68]
Tournament, High Pressure	30%	41%	(0.28) [0.36]	7.8%	4.2%	(0.20) [0.72]
Piece-Rate, Low Pressure	24%	22%	(0.73) [1.00]	13.8%	8.8%	(0.11) [0.39]
Tournament, Low Pressure	31% <sup>‡</sup>	19% <sup>‡</sup>	(0.10) [0.21]	17.1%***	6.5%***	(0.00) [0.00]
<u>% Choosing Tournament</u>						
High Pressure	44%**	19%**	(0.02)	39%	30%	(0.32)
Low Pressure	44%	36%	(0.48)	43%*	57%*	(0.10)
<u>Rank Guess</u>						
Piece-Rate, High Pressure	2.58	2.53	(0.83) [0.80]	2.46	2.50	(0.84) [0.86]
Tournament, High Pressure	2.72	2.53	(0.48) [0.45]	2.46	2.39	(0.73) [0.76]
Piece-Rate, Low Pressure	2.25	2.19	(0.82) [0.89]	2.30	2.48	(0.37) [0.42]
Tournament, Low Pressure	2.14	2.47	(0.23) [0.17]	2.24*	1.88*	(0.07) [0.10]
<u>Questionnaire Statistics</u>						
Mean Age (Years)	20.4	21.3		21.3	22.1	
% Reporting Women Better	31%	31%		61%	86%	
% Studying Science	25%	42%		28%	21%	
% Native English Speakers	92%	92%		89%	87%	
Number of Subjects	36	36		54	54	

Notes: t-test p-values for difference in gender means reported in parentheses; Mann-Whitney U test p-values in square brackets. Significance levels: <sup>‡</sup> 10% (only one of the tests), \*10%, \*\*5%, \*\*\*1%.

## Confidence (Guessed Rank) and Its Determinants

Table C.2. Distribution of Guessed Math Tournament Rank

	High Time Pressure Tournament				Low Time Pressure Tournament			
	Women		Men		Women		Men	
	Guessed	Incorrect	Guessed	Incorrect	Guessed	Incorrect	Guessed	Incorrect
	Rank	Guess	Rank	Guess	Rank	Guess	Rank	Guess
1: Best	9	7	8	3	8	6	16	8
2	8	7	8	4	12	8	7	6
3	10	7	6	5	7	4	5	4
4: Worst	9	6	14	12	9	4	8	3
Total	36	27	36	24	36	22	36	21

Table C.3. Distribution of Guessed Verbal Tournament Rank

	High Time Pressure Tournament				Low Time Pressure Tournament			
	Women		Men		Women		Men	
	Guessed	Incorrect	Guessed	Incorrect	Guessed	Incorrect	Guessed	Incorrect
	Rank	Guess	Rank	Guess	Rank	Guess	Rank	Guess
1: Best	12	6	12	7	20	11	17	13
2	15	11	17	12	25	18	17	17
3	16	13	13	12	4	4	10	9
4: Worst	8	4	12	8	5	3	10	5
Total	51	34	54	39	54	22	54	44

Table C.4. Ordered Probit of Guessed Tournament Rank

Independent Variables:	Dependent Variable: Rank Guess (1, 2, or 3)			
	High Time Pressure		Low Time Pressure	
	Math	Verbal	Math	Verbal
	(1)	(2)	(3)	(4)
Female	-0.30	-0.15	0.49	-0.45*
	(0.31)	(0.55)	(0.15)	(0.10)
Tournament Score	-0.04	-0.09***	-0.02***	-0.03
	(0.21)	(0.01)	(0.01)	(0.19)
Tournament – Piece-Rate Score	0.01	0.002	0.01	0.05**
	(0.81)	(0.93)	(0.13)	(0.04)
No. Observations	72	88	72	89

Notes: Robust standard errors clustered at the group level; p-values in parentheses. Other controls include order of tournament and time, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Significance levels: \* 10%, \*\* 5%, \*\*\* 1%.

## Probit Regressions of Tournament Entry Decision – Additional Specifications

Table C.5 provides probit estimates of tournament entry decision as a function of gender, rank guess, and other controls for the unrestricted sample of rank guesses. Note that the inclusion of all four rank guess categories does not significantly change the main result – the finding that the likelihood of women to self-select into a tournament varies across pressure and task treatments. However, the impact of tournament rank guess on tournament entry is now only significant in the low pressure verbal treatment.

Table C.5. Probit of Tournament Entry Decision, All Sessions

Independent Variables:	Dependent Variable: Choice (Tournament =1)			
	High Pressure	High Pressure	Low Pressure	Low Pressure
	Math	Verbal	Math	Verbal
Female	-0.39*** (0.00)	-0.03 (0.80)	-0.05 (0.70)	0.34*** (0.01)
Tournament Score	0.02*** (0.00)	0.02 (0.25)	0.01*** (0.00)	0.002 (0.84)
Tournament – Piece-Rate Score	-0.02* (0.06)	-0.01 (0.14)	-0.0004 (0.84)	-0.04** (0.02)
Guessed Tournament Rank	-0.02 (0.16)	-0.07 (0.38)	-0.04 (0.57)	-0.14** (0.05)
No. Observations	72	88	72	89

Notes: Standard errors clustered at the group level (p-value in parentheses); marginal effects. Other controls include order of tournament and time, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Significance levels: \*10%, \*\*5%, \*\*\*1%.

## Within-Subject Analysis

The analysis of average performance differences and the distributions of scores presented in the paper does not exploit the individual-level variation I observe in my experiment. Figure C.1 shows the within-subject change in performance from the high to the low time pressure treatments for the math session (left panel) and the verbal sessions (right panel). In particular, I plot the subject-specific difference between the tournament and the piece-rate score in the low time pressure treatments (y-axis) as a function of the difference between the tournament and the piece-rate score in the high time pressure treatments

(x-axis).

In the math task, the majority of men is located in quadrants QI (do better in the tournament than in the piece-rate under both low and high time pressure) and QII (do better in the tournament than in the piece-rate under high time pressure, but worse under low time pressure). Those men that score lower in the tournament in either pressure treatment (QIII) are clustered closer to QI and QII, which means that their tournament-to-piece-rate difference is greater in the low pressure treatments than in the high pressure treatments. Women on the other hand are located mainly in quadrants QIII (do worse in the tournament than in the piece-rate under both low and high time pressure) and QIV (do worse in the tournament than in the piece-rate under high time pressure, but better under low time pressure). Women who score lower in the tournament in either pressure treatment (QIII) are located closer to QIV, which means that their tournament-to-piece-rate difference is greater in the high pressure treatments than in the low pressure treatments. In summary, competition hurts women's performance more than it does men's, and it does so more under high time pressure.

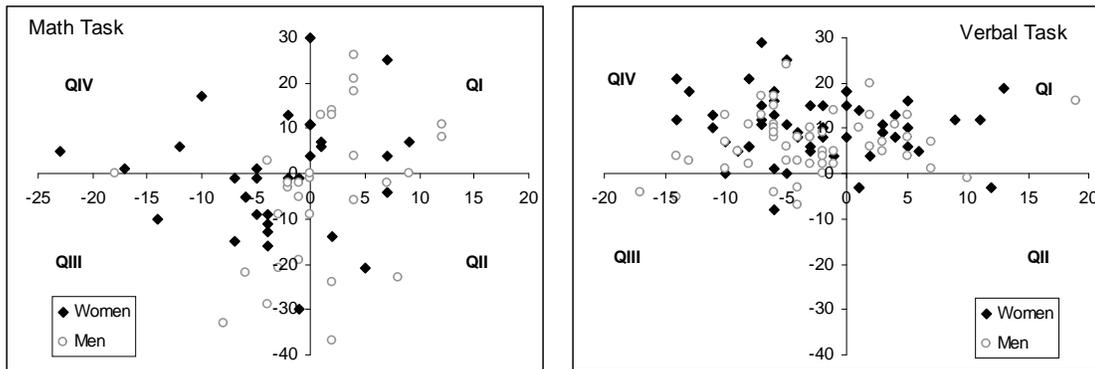


Figure C.1. Tournament – Piece-Rate Score Difference in the Low Time Pressure Treatments as a Function of the Tournament – Piece-Rate Score Difference in the High Time Pressure Treatments by Gender and Type of Task

In the verbal task, most subjects are located in quadrants QI (do better in the tournament than in the piece-rate under both low and high time pressure) and QIV (do worse in the tournament than in the piece-rate under high time pressure, but better under low time pressure). Since it is possible that the task variation from treatment to treatment par-

tially led to this pattern, I focus on the gender differences within each treatment. In QIV, the female observations are on average located above the male observations, which means that, while both genders score higher in the low pressure tournament than piece-rate, women do so to a greater extent than men.

### **Individual-Level Regressions with Controls and Interactions**

Next, I show that the results presented in the paper are robust to the inclusion of various controls and interactions.

### **Performance in Math Task**

Table C.6 gives OLS regression results of performance measures (score and mistake share) as a function of gender, treatment effects, and various controls for the math sessions. In all specifications, baseline ability in the math task (measured by the practice round score) has a significant positive effect on performance (increases the score and reduces the mistake share). Specification 1 shows that men outperform the women in the math game, conditional of time pressure and competition. Note, however, that this is not because men make fewer mistakes in the math tasks (the coefficient on the female dummy is not statistically significant in specification 3). Higher time pressure reduces performance, both in terms of the score (specification 1) and the mistake share (specification 3). Once I control for the interactions between the female dummy, competition, and time pressure, the negative effect of being female on the score disappears (specification 2).<sup>1</sup> This provides some support for the notion that removing time pressure in the mathematical environment helps women perform on par with men even when competition is involved. Specification 4 shows that women do tend to make more mistakes relative to men in the presence of both competition and high time pressure.

---

<sup>1</sup>The omitted category is a man in a low-time-pressure noncompetitive setting.

Table C.6. OLS Regressions of Performance Measures, Math Sessions

Variables	Score		Mistake Share	
	(1)	(2)	(3)	(4)
Score in Practice Round	1.20*** (0.26)	1.20*** (0.26)	-0.01** (0.004)	-0.01** (0.004)
Female Dummy	-6.14** (2.99)	-10.02 (6.46)	-0.01 (0.05)	-0.01 (0.07)
Time Pressure (High = 1)	-21.47*** (2.51)	-24.33*** (6.23)	0.05** (0.02)	0.03 (0.06)
Competition (Tournament = 1)	1.10 (2.11)	2.81 (6.69)	0.07 (0.04)	0.07 (0.06)
Female $\times$ Competition		0.36 (6.81)		-0.10 (0.07)
Female $\times$ Time Pressure		9.5 (7.36)		-0.04 (0.09)
Competition $\times$ Time Pressure		-1.67 (6.67)		-0.04 (0.08)
Female $\times$ Competition $\times$ Time Pressure		-4.22 (6.81)		0.27** (0.11)
Observations	288	288	288	288
$R^2$	0.35	0.36	0.33	0.35
Number of Clusters	18	18	18	18

Notes: Robust standard errors in parentheses (clustered at the group level). Other controls include: order of tournament age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Tournament and piece-rate data only. Significance levels: \* 10%, \*\* 5%, \*\*\* 1%.

## Performance in Verbal Task

Table C.7 gives OLS regression results of performance measures (score and mistake share) as a function of gender, treatment effects, and various controls for the verbal sessions. Once again, the practice round score is significant in all specifications. Specifications 1 and 3 show that women tend to achieve higher scores and make fewer mistakes in the verbal task than men. Higher time pressure seems to reduce performance, both in terms of the score and the mistake share. Finally, competition has a significant positive effect on the score, but shows no effect on the mistake share. Once I control for the interactions between the female dummy, competition, and time pressure, the positive effect of being female disappears (specifications 2 and 4). Finally, I establish that, once again, the presence of both competition and time pressure hurt women relative to men in terms of their score in the verbal task, although competition with reduced time pressure actually

improves their performance (see specification 2, coefficients on the interaction terms Female  $\times$  Competition  $\times$  Time Pressure and Female  $\times$  Competition, respectively)

Table C.7. OLS Regressions of Performance Measures, Verbal Sessions

Variables	Score		Mistake Share	
	(1)	(2)	(3)	(4)
Score in Practice Round	0.70*** (0.12)	0.70*** (0.12)	-0.01** (0.003)	-0.01** (0.003)
Female Dummy	2.57** (1.04)	0.56 (1.30)	-0.05** (0.02)	-0.04 (0.02)
Time Pressure (High = 1)	-4.47*** (0.62)	1.08 (1.24)	-0.05*** (0.01)	-0.06** (0.02)
Competition (Tournament = 1)	2.28*** (0.56)	5.39*** (1.32)	0.01 (0.01)	0.05* (0.03)
Female $\times$ Competition		5.70*** (1.58)		-0.08** (0.04)
Female $\times$ Time Pressure		0.63 (1.51)		0.03 (0.03)
Competition $\times$ Time Pressure		-9.48*** (1.43)		-0.03 (0.03)
Female $\times$ Competition $\times$ Time Pressure		-4.59** (2.23)		-0.03 (0.04)
Observations	362	362	362	362
$R^2$	0.35	0.49	0.23	0.25
Number of Clusters	25	25	25	25

Notes: Robust standard errors in parentheses (clustered at the group level). Other controls include: order of tournament and time, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Tournament and piece-rate data only. Significance levels: \* 10%, \*\* 5%, \*\*\* 1%.

## Preference for Competition

Table C.8 presents the probit estimates of the choice of compensation scheme as a function of gender, past performance, and other controls for the sample pooled across the time pressure conditions.

Table C.8. Probit of Tournament Entry Decision (Pooled Data)

Variables	Math Sessions		Verbal Sessions	
	(1)	(2)	(3)	(4)
Female Dummy	-0.09 (0.13)	-0.07 (0.13)	0.26** (0.12)	0.28** (0.13)
Tournament Score	0.01*** (0.003)	0.01*** (0.003)	0.01 (0.01)	0.004 (0.01)
Tournament – Piece-Rate Score	-0.002 (0.002)	-0.002 (0.002)	-0.02*** (0.01)	-0.02*** (0.01)
Guessed Rank		-0.08*** (0.04)		-0.12*** (0.04)
Time Pressure (High = 1)	0.19* (0.11)	0.21* (0.11)	-0.28*** (0.12)	-0.26** (0.12)
Female × Time Pressure	-0.22** (0.10)	-0.25*** (0.09)	-0.25 (0.11)	-0.27* (0.15)
Observations	288	288	360	357
Number of Clusters	18	18	25	25

Notes: Standard errors clustered at the group level; marginal effects. Other controls include order of tournament and time, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Tournament and piece-rate data only. Significance levels: \* 10%, \*\* 5%, \*\*\* 1%.

## Appendix D: Analysis Using the First Three Rounds of Data (Web Appendix – Not for Publication)

This appendix provides the analysis using the first three rounds of data only in order to check for robustness.

### Average Performance in Competition and Preference for Competition

Figures D.1a and D.1b show the average scores of men and women by treatment in the verbal sessions across all rounds and with the first three rounds of data only, respectively.

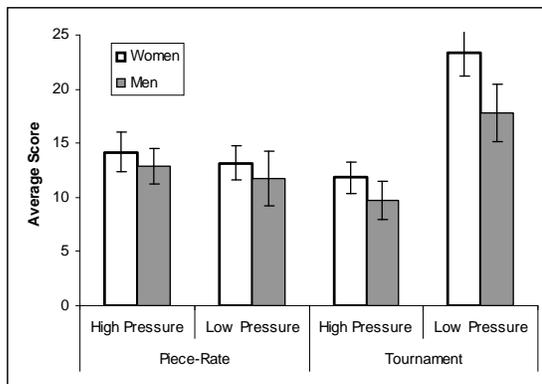


Figure D.1a. Average Verbal Scores by Treatment and Gender (95% C.I.), All Rounds

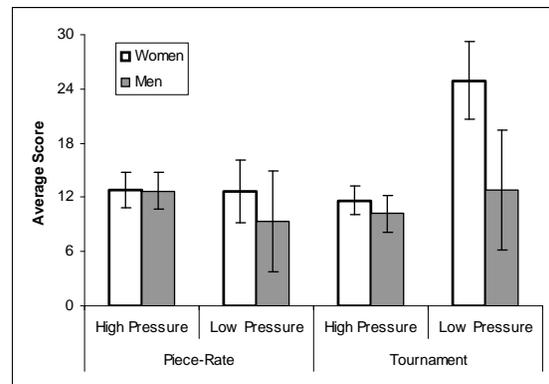


Figure D.1b. Average Verbal Scores by Treatment and Gender (95% C.I.), First 3 Rounds

Note that the figures deliver similar results: women significantly outperform the men in the low time pressure verbal tournament.

Figure D.2 shows the scores in the math sessions by gender and treatment. Even though the 95 percent confidence intervals overlap in the figure, the difference between men and women in the tournament is statistically significant at the 5 percent confidence level. Note also that the figure for the first three rounds only includes high time pressure treatment and is therefore the same whether I use the entire sample or the first three rounds.

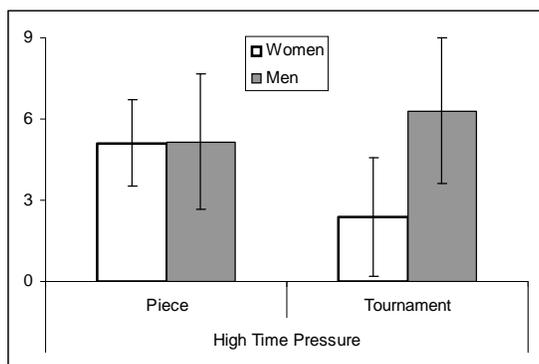


Figure D.2. Average Math Scores by Treatment and Gender

Next, I compare men and women according to their preferences for competition using the first three rounds of data of verbal data. Table D.1 reports the probit estimates of the choice of compensation scheme as a function of gender, past performance, and other controls. Although the samples are now very small, the results are consistent with the full-sample analysis.

Table D.1. Probit of Verbal Tournament Entry Decision

Independent Variables:	Dependent Variable: Choice (Tournament =1)			
	High Time Pressure		Low Time Pressure	
	(1)	(2)	(3)	(4)
Female	-0.14 (0.13)	-0.13 (0.17)	0.58* (0.25)	0.31 (0.37)
Tournament Score	0.05*** (0.02)	0.06*** (0.02)	-0.05*** (0.02)	-0.04 (0.04)
Tournament – Piece-Rate Score	-0.02 (0.14)	-0.06*** (0.01)	0.02 (0.02)	0.01 (0.03)
Guessed Tournament Rank		-0.28*** (0.11)		-0.45** (0.20)
No. Observations	61	51	30	30

Notes: Standard errors clustered at the group level; marginal effects. Other controls include order of tournament and time, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Significance levels: \*10%, \*\*5%, \*\*\*1%. Guesses of 4 are eliminated in Specification (2).

The probit regressions of tournament entry decisions for the first three rounds of the math sessions are not different than the ones reported in the paper.

## Quality and Quantity

Figure D.3b shows that the share of mistakes made by men and women in the first three rounds of the verbal sessions. The figure shows the pattern consistent with the one using the full sample: men make significantly more mistake relative to women in the low time pressure tournament. (Figure D.3a is the same as Figure 7b in the paper. It is provided here for reference and shows the quality-to-quantity ratio using the full verbal sample.)

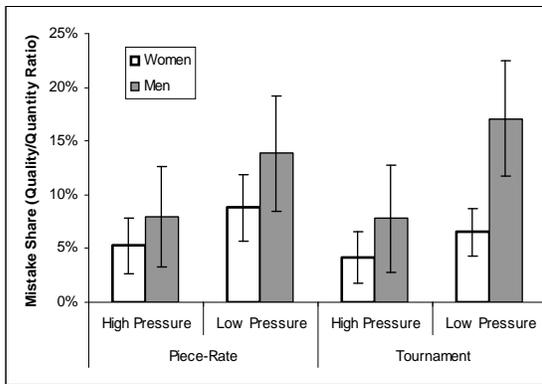


Figure D.3a. Average Mistake Share by Gender for All Verbal Treatments (95% C.I.), All Rounds

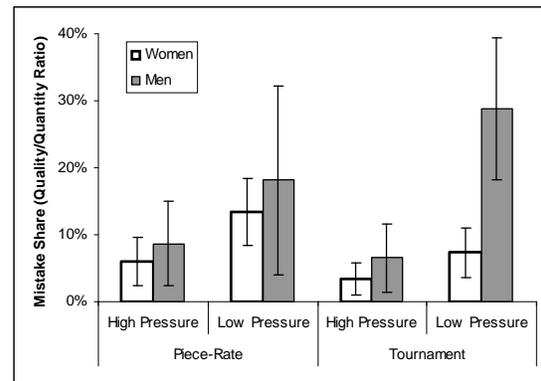


Figure D.3b. Average Mistake Share by Gender for All Verbal Treatments (95% C.I.), First 3 Rounds

## Individual-Level Regressions with Controls and Interactions

Table D.2 reports OLS regressions of my two performance measures as a function of gender, treatment effects and various controls, using only the first three rounds of data for the math sessions. The results are robust to this variation.

Table D.2. OLS Regressions of Performance Measures, Math Sessions

Variables	Score		Mistake Share	
	(1)	(2)	(3)	(4)
Score in Practice Round	0.32 (0.19)	0.32 (0.19)	-0.01 (0.01)	-0.01 (0.01)
Female Dummy	-2.49** (1.16)	-0.56 (1.15)	0.04 (0.06)	0.04 (0.08)
Competition (Tournament = 1)	-0.79 (0.78)	1.14 (1.03)	0.11 (0.07)	0.03 (0.08)
Female $\times$ Competition		-3.86*** (1.01)		0.17* (0.10)
Observations	144	144	144	144
$R^2$	0.26	0.28	0.07	0.08
Number of Clusters	18	18	18	18

Notes: Robust standard errors in parentheses (clustered at the group level). Other controls include: order of tournament, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Tournament and piece-rate data only. Significance levels: \* 10%, \*\* 5%, \*\*\* 1%.

Note that, in the math sessions (Table D.2), the first three rounds only contain the high time pressure treatments.

Table D.3 reports OLS regressions of my two performance measures as a function of gender, treatment effects and various controls, using only the first three rounds of data for the verbal sessions. The results are robust to this variation.

Table D.3. OLS Regressions of Performance Measures, Verbal Sessions

Variables	Score		Mistake Share	
	(1)	(2)	(3)	(4)
Score in Practice Round	0.73*** (0.17)	0.72*** (0.17)	-0.01** (0.003)	-0.01** (0.003)
Female Dummy	3.16** (1.30)	2.84** (1.28)	-0.07** (0.03)	-0.06 (0.06)
Competition (Tournament = 1)	0.83 (1.27)	3.53 (2.15)	-0.004 (0.02)	0.11* (0.05)
Female $\times$ Competition		8.8*** (2.29)		-0.17*** (0.06)
Female $\times$ Time Pressure		-2.68 (1.71)		0.04 (0.07)
Competition $\times$ Time Pressure		-7.25*** (2.37)		-0.10* (0.06)
Female $\times$ Competition $\times$ Time Pressure		-6.56** (2.69)		0.11* (0.06)
Observations	182	182	182	182
$R^2$	0.37	0.52	0.30	0.36
Number of Clusters	25	25	25	25

Notes: Robust standard errors in parentheses (clustered at the group level). Other controls include: order of tournament and time, age, major (1 = science), native language (1 = English), and reported gender stereotype (1 if women perceived to be better). Tournament and piece-rate data only. Significance levels: \* 10%, \*\* 5%, \*\*\* 1%.