**Title:** No Gender Difference in Willingness to Compete When Competing against Self

**Authors:**

Coren L. Apicella

Affiliation: University of Pennsylvania

Mailing address: 3720 Walnut Street, Philadelphia, PA 19104, USA

Email address: capicella@psych.upenn.edu

Elif E. Demiral

Affiliation: Interdisciplinary Center for Economics Science, George Mason University

Mailing address: 4400 University Dr, 22030 Fairfax, VA, USA

Email address: edemiral@gmu.edu

Johanna Mollerstrom (corresponding author)

Affiliations: Humboldt University, German Institute for Economic Research (DIW), Research Institute for Industrial Economics (IFN).

Mailing address: DIW, Mohrenstrasse 58, Berlin, Germany

Email address: jmollerstrom@diw.de

**Abstract**

We report on two experiments investigating whether there is a gender difference in the willingness to compete against oneself (self-competition), similar to what is found when competing against others (other-competition). In one laboratory and one online market experiment, involving a total of 1,200 participants, we replicate the gender-gap in willingness to other-compete but find no evidence of a gender difference in the willingness to self-compete. We explore the roles of risk and confidence and suggest that these factors can account for the different findings. Finally, we document that self-competition does no worse than other-competition in terms of performance boosting.

In *The Descent of Man,* Darwin described men as "rivals of other men" and as the sex that "delight[s] in competition" (Darwin, 1888). Nearly 150 years later, economists have substantiated this narrative: men are more likely to enter competitive fields, pursue competitive promotions and select into competitive payment schemes over piece rate schemes (for review, Niederle and Vesterlund, 2011). The gender difference in willingness to compete has been documented in diverse societies, including isolated hunter-gatherers (Apicella and Dreber, 2015) suggesting that differences in competitiveness between men and women is a relatively robust characteristic of humanity– though this does not imply that the difference is necessarily large, or present for all types of tasks. Additionally, it has been suggested that this difference in the willingness to compete may help to explain the persistent labor and economic disparities that exist between the sexes, such as the gender gap in earnings. Indeed, the predictive power of laboratory measures of competitiveness on career choices and labor market outcomes has been shown to be substantial (e.g. Buser et al., 2014, Reuben et al., 2015, Buser et al., 2017).

Following the work of Niederle and Vesterlund (2007), economists have shifted their level of analysis from the descriptive to explanatory: *Why* is there a gender difference in willingness to compete and how can it be eliminated? Whereas earlier work claimed that there is a specific "preference for competitions" that is distinct from risk preferences and overconfidence, more recent work suggests that gender differences in these factors may actually explain all or most of the competitiveness gap (e.g. Gillen et al, 2015, van Veldhuizen, 2016). Other work has focused on implementing institutional changes to increase the number of women entering competitive environments. Such changes include: providing feedback about individual performance, instituting affirmative action policies and assembling gender-specific competitions (for review, Niederle and Vesterlund, 2011). Many of these policies however, are unfeasible and impractical

from a firm's perspective. Moreover, restructuring firms to be "competition-free" can be suboptimal, as competition often enhances performance and productivity (e.g. Gneezy et al., 2003, see also Shurchkov and van Geen, 2016, on gender differences in choice of incentives for others).

To our knowledge, research on gender differences in willingness to compete has exclusively focused on competitions against others (*other-competition*). We consider a different type of competition: competition against one's own previous performance. This type of competition embodies notions of self-improvement, progress and mastery and the idea of such *self-competition* has previously been discussed in relation to sports performance and business related goal-setting (e.g. Locke, 1968; Howe, 2008; Brown et al, 1998). We examine whether there is a gender difference in the willingness to self-compete, similar to the gender difference found when competing against others. We document that this is not so, and investigate why.

## I The Two Experiments[1]

### *A. Laboratory Experiment*

The laboratory experiment, which follows the original Niederle and Vesterlund (2007) design closely had two treatments: *Other* and *Self*. In the *Other* treatment, subjects performed a series of simple arithmetic problems in three rounds (each round lasting five minutes, with no feedback given between rounds and the incentives for each round outlined just before the start of that particular round). In the first round, subjects were paid a fixed amount for every correctly solved problem (piece rate). In the second round, subjects were anonymously matched in pairs and the subject with the highest score in the pair was paid double the piece rate for every correctly solved problem, whereas the other subject received nothing (tournament rate). Subjects were

[1] Further details about the experiments, including all instructions, are available in the Online Appendix.

then given a choice about which payment scheme to apply in round three. If a subject chose the piece rate, she was paid the fixed amount for every problem she correctly solved in the third round. If she chose the tournament rate, her third round performance was compared to the second round performance of the matched participant, such that if she performed better than they did in the second round, she would earn double the piece rate for every correctly solved problem; otherwise she earned nothing. The *Self* treatment was identical with the following exceptions: 1) the subjects were not matched to another player; instead their scores in the second round tournament were compared to their own scores in the first round, 2) when subjects chose whether to apply the piece rate or the tournament rate in the third round, a choice of tournament meant that their score in the third round would be compared to their own score in round two.

After the three rounds, all subjects filled out a questionnaire. Basic demographics, and self-reported risk aversion using a ten-point likert scale, were collected. Subjects were also incentivized to correctly rank their own performance across the rounds, and to guess whether they outperformed their opponent in round two. These questions provided measures of confidence.[2] Subjects were paid in private for a randomly selected round before leaving the laboratory.

The experiment was programmed in Z-Tree (Fischbacher, 2007) and conducted at the ICES laboratory at George Mason University in October 2016. The 204 subjects (50.5% female) earned an average of $17.42 for their participation. Sessions lasted approximately 40 minutes.

[2] Our main confidence variable is a dummy equal to 1 for subjects who believe that they improved their performance between rounds 2 and 3 ("Self") or that they performed better than the person they were matched to in the second round (the three "Other"-treatments). The results are robust to using other specifications.

*B. Online Experiment*

We used the labor market Amazon Mturk to ensure that the results from the laboratory replicate. The online experiment also entailed additional treatments to further investigate the mechanisms underlying the findings. Specifically, we implemented two additional versions of the *Other* treatment. In *Other, Same Gender* we matched participants of the same gender in the competition, and in *Other, Same Ability* participants who did the same number of tasks correctly in the first round were matched. Ahead of the second round, subjects were informed about these aspects of the matching. We use these treatments in order to mirror certain features of self-competition –the fact that the person knows their own gender and also has additional information about her own ability to perform in the task, respectively – and to investigate whether these alone cause the gender difference in competitiveness to diminish. The online experimental design also differed from the laboratory design in two other ways: 1) math tasks were replaced with a Captcha-style counting task to prevent cheating and 2) the rounds were shortened to 90 seconds.

994 subjects (49.9% female) took part in the online experiment conducted in November 2016. On average, participants earned $1.20 for an approximately twelve-minute-long session.

## II Results[3]

*A. Laboratory Results: No Gender Difference in the Willingness to Self-Compete*

Panel A of Table 1 reports the percentage of subjects choosing the tournament rate by gender and treatment. We replicate the oft-documented finding that women are less willing than men to other-compete: In the *Other* treatment 57.7 percent of men chose to compete in the third round, compared to merely 37.5 percent of women. This results in a gender gap of 20 percentage points

[3] Summary statistics for both experiments as well as additional analysis and robustness checks are available in the Online Appendix.

(p=0.044[4]). When the competition is against the participant's own previous score the size of the gender gap is reduced by a third and is no longer statistically significant (p=0.176). The difference in difference is, however, not significant (p=0.612) in the laboratory data alone.[5]

Panel A of Table 2 outlines the regression analysis for the lab experiment. A comparison of (1) and (2) in Panel A indicates that risk preferences and confidence explain parts of the gender difference in the willingness to other-compete. Risk preferences are related also to the willingness to self-compete, but here the coefficient on Female is not significant ((3) and (4)).

*B. Online Results: Replication and Investigating Mechanisms*

Panel B of Table 1 shows that the gender gap in the willingness to compete is 12 percentage points in the *Other*-treatment in the online experiment (p=0.045). In the *Self*-treatment the sign of the gap reverses and it is no longer significant (p=0.446). Further, a difference-in-difference estimation reveals that the gender gaps in the two treatments differ marginally significantly from each other (p=0.052). The analysis of the *Self* and *Other* treatments in the online experiment (specifications 5-8 in Table 2, Panel B) also gives similar results to the regression analysis for the laboratory experiment.

Considering the two additional versions of the *Other*-treatment, we note that there is still a significant gender difference in competitiveness in the *Other, Same Gender* but not in the *Other, Same Ability* treatment (Panel B of Table 1 and Panel C of Table 2). The latter result indicates that receiving a signal that the matched opponent has an ability akin to one's own (similar to what happens in self-competition), is enough to eliminate the gender difference.

[4] All tests, unless otherwise noted, are two-sided t-tests of proportions. Our results are robust to using the non-parametric Fisher Exact test instead, see Online Appendix.

[5] In addition to the analysis of the online data below, we also perform the analysis on the pooled data. Here the difference in difference is significant, see Online Appendix.

That women are more risk averse (e.g. Croson and Gneezy, 2009) and less overconfident (Niederle and Vesterlund, 2007) than men has been documented previously. We largely replicated these findings (see Online Appendix). As these factors have been shown to be important mechanisms underlying the gender difference in the willingness to other-compete, we investigate if the roles of risk and confidence are different in self-competition, which could be one reason that we find no gender gap in that type of competition. We confirm this by regressing a dummy indicating the choice to compete in the third round on either risk or confidence, a dummy that takes the value 1 if the treatment is *Other* and 0 if it is *Self*, and the interaction between the two (controlling for ability with round 1 score as in all regressions). The results, which are outlined in more detail in the Online Appendix, indicate that especially confidence (p=0.014), and to some extent risk aversion (p=0.095), have a larger impact on the choice of whether or not to compete in the *Other*-treatment than in the *Self*-treatment.[6] Additionally, we document in the Online Appendix that whereas men report to be less risk averse than women overall, the gender difference in confidence is significant for other-competition but not for self-competition.

*C. Self- Competition is No Worse than Other-Competition for Performance Boosting*

Previous literature (e.g. Gneezy et al., 2003) has documented that competitions can boost performance. With our design we cannot distinguish competition boosting effects from learning effects, as the order of the piece rate and the tournament rate is not randomized. We do however investigate if the performance improvement between rounds 1 and 2 is significantly different in the two treatment and find that it is not. In our laboratory experiment we document an average score improvement of 23.9 percent in the *Other*-treatment and 18.2 percent in the *Self*-treatment,

[6] Using the laboratory data the same analysis yields p=0.068 for confidence and p=0.469 for risk. See Online Appendix.

(p=0.444 for t-test of difference). In the online experiment the improvement between the first and the second round is 18.0 percent in the *Self*-treatment and 22.2 percent in the three *Other*-treatments (p=0.456).

*Table 1 Percentage Choosing Tournament Rate, by Treatment and Gender*

Panel A: Laboratory Experiment

| Treatment: | **Women** | **Men** | **Total** |
|---|---|---|---|
| **Other** | 37.5 (7.1) | 57.7 (6.9) | 48.0 (5.0) |
| **Self** | 41.8 (6.7) | 55.1 (7.2) | 48.1 (4.9) |
| **Total** | 39.8 (4.8) | 56.4 (5.0) | 48.0 (3.5) |

Panel B: Online Experiment

| Treatment: | **Women** | **Men** | **Total** |
|---|---|---|---|
| **Other** | 27.8 (4.2) | 40.0 (4.3) | 34.3 (3.0) |
| **Other, Same Gender** | 21.9 (3.7) | 34.1 (4.2) | 28.0 (2.8) |
| **Other, Same Ability** | 30.6 (4.2) | 33.3 (4.3) | 32.0 (3.0) |
| **Self** | 35.7 (4.2) | 31.1 (4.3) | 33.5 (3.0) |
| **Total** | 29.0 (2.0) | 34.7 (2.1) | 31.9 (1.5) |

*Notes: Standard errors in parentheses*

*Table 2 Regression Analysis*

Panel A: Lab Experiment

| | (1) (Other) | (2) (Other) | (3) (Self) | (4) (Self) |
|---|---|---|---|---|
| Female | -0.195** | -0.114 | -0.132 | -0.029 |
| | (0.10) | (0.10) | (0.10) | (0.10) |
| Confidence | | 0.246** | | -0.013 |
| | | (0.11) | | (0.10) |
| Risk | | 0.039* | | 0.091*** |
| | | (0.02) | | (0.02) |
| Constant | 0.177 | -0.212 | 0.503*** | -0.008 |
| | (0.14) | (0.22) | (0.16) | (0.20) |
| N | 100 | 100 | 104 | 104 |
| R-square | 0.116 | 0.180 | 0.019 | 0.140 |

Panel B: Online Experiment

| | (5) (Other) | (6) (Other) | (7) (Self) | (8) (Self) |
|---|---|---|---|---|
| Female | -0.126** | -0.090 | 0.052 | 0.083 |
| | (0.06) | (0.06) | (0.06) | (0.06) |
| Confidence | | 0.246*** | | 0.128** |
| | | (0.06) | | (0.06) |
| Risk | | 0.045*** | | 0.032** |
| | | (0.01) | | (0.01) |
| Constant | 0.297*** | -0.114 | 0.371*** | 0.120 |
| | (0.07) | (0.10) | (0.08) | (0.12) |
| N | 245 | 245 | 248 | 248 |
| R-square | 0.028 | 0.172 | 0.006 | 0.042 |

Panel C: Online Experiment, ctd

| | (9) (Other, Same Gender) | (10) (Other, Same Gender) | (11) (Other, Same Ability) | (12) (Other, Same Ability) |
|---|---|---|---|---|
| Female | -0.122** | -0.094* | -0.028 | 0.030 |
| | (0.06) | (0.05) | (0.06) | (0.06) |
| Confidence | | 0.269*** | | 0.287*** |
| | | (0.06) | | (0.05) |
| Risk | | 0.027** | | 0.042*** |
| | | (0.01) | | (0.01) |
| Constant | 0.349*** | 0.063 | 0.307*** | -0.117 |
| | (0.07) | (0.09) | (0.07) | (0.11) |
| N | 257 | 257 | 244 | 244 |
| R-square | 0.019 | 0.158 | 0.002 | 0.158 |

*Notes: Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments). All results are robust to using probit instead of OLS. Significance: ***p<0.01 **p<0.05 *p<0.1*

## III Conclusions

While women are less willing than equally able men to compete against other people, we find no gender difference in the willingness to compete against one's own, previous score. For those worried about the inequalities resulting from women shying away from competitive settings, our results provide an alternative to simply removing competitive features from the environment – features which are partly employed to boost performance. Instead, we suggest that a restructuring of institutions, such that competitive pressure primarily concerning comparisons with oneself is enhanced, could be tried to reduce gender disparities in economic and labor market outcomes. This will be especially appropriate when ratchet-effects are not a concern, and when the competition is mainly used for motivation and remuneration rather than for selection. The fact that self-

competition leads to a boost in performance similar to other-competition, suggests that firms would not have to sacrifice these properties of the competitive environment.

**References**

**Apicella, Coren L., and Anna Dreber.** 2015. "Sex Differences in Competitiveness: Hunter-Gatherer Women and Girls Compete Less in Gender-Neutral and Male-Centric Tasks." *Adaptive Human Behavior and Physiology* 1 (3): 247-269.

**Brown, Steven P., William L. Cron, and John W. Slocum Jr.** 1998. "Effects of Trait Competitiveness and Perceived Intraorganizational Competition on Salesperson Goal Setting and Performance." *The Journal of Marketing* 62 (4): 88-98.

**Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek.** 2014. "Gender, Competitiveness and Career Choices." *The Quarterly Journal of Economics* 129 (3): 1409-1447.

**Buser, Thomas, Noemi Peter, and Stefan Wolter.** 2017. "Gender, Competitiveness and Study Choices – Evidence from Switzerland" American Economic Review P&P (forthcoming)

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

**Darwin, Charles.** 1888. *The Descent of Man, and Selection in Relation to Sex.* London: Murray.

**Fischbacher, Urs.** 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10 (2): 171–78.

**Gillen, Ben, Erik Snowberg, and Leeat Yariv.** 2015. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." National Bureau of Economic Research Working Paper 21517.

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

**Howe, Leslie A.** 2008. "On Competing Against Oneself, Or 'I need to get a different voice in my head'." *Sports Ethics and Philosophy* 2 (3): 353-366.

**Locke, Edwin A.** 1968. "Toward a Theory of Task Motivation and Incentives." *Organizational Behavior and Human Performance* 3 (2): 157-189.

**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.

**Niederle, Muriel and Lise Vesterlund.** 2011. "Gender and Competition." *Annual Review of Economics* 3 (1): 601-630.

**Reuben, Ernesto, Paola Sapienza, and Luigi Zingales.** 2015. "Taste for Competition and the Gender Gap Among Young Business Professionals." National Bureau of Economic Research Working Paper 21695.

**Shurchkov, Olga, and Alexandra van Geen. 2016.** "The determinants of incentive choice: Do female decision-makers shy away from promoting competition?" Working Paper

**van Veldhuizen, Roel.** 2016. **"**Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness?" WZB Berlin Social Science Center Discussion Paper SP II 2016–207.

We gratefully acknowledge excellent comments on an earlier draft from Pio Baake, Thomas Buser, Katie B. Coffman, Anna Dreber, Christine Exley, Roel van Veldhuizen and participants in the session "Understanding Gender Differences in Labor Market Outcomes - Experimental Evidence" at the 2017 ASSA meetings.

**No Gender Difference in Willingness to Compete When Competing against Self: Online Appendix**

Coren L. Apicella

Elif E. Demiral

Johanna Mollerstrom

## Appendix A: Instructions for the Lab Experiment

### Welcome

Hi and welcome! In this experiment you will be asked to complete different tasks.

Please press OK to get started with the experiment.

### General Instructions

In this experiment you will be asked to complete three different tasks. None of these will take more than 5 minutes.

At the end of the experiment we will randomly select one of the tasks. This is the task that will be relevant for your profit. Once you have completed the three tasks we determine which task counts for your profit by randomly drawing a number between 1 and 3.

The method we use to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined.

### Rules for Task 1

For Task 1 you will be asked to calculate the sum of five randomly chosen two-digit numbers.

You will be given 5 minutes to do a series of these problems. You are not allowed to use a calculator to determine the sum. However you are welcome to write the numbers down and make use of the provided scratch paper. You submit an answer by clicking the OK button with your mouse.

If Task 1 is the one randomly selected for your profit, then you get 1 dollar per problem you solve correctly in the 5 minutes. Your profit does not decrease if you provide an incorrect answer to a problem. We refer to this task as the Piece Rate task.

If you have any questions before we begin, please raise your hand.

**Rules for Task 2 – *Other treatment***

As in Task 1 you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers.

However for Task 2 your payment depends on your performance relative to that of another participant who is here right now, and who has been put in a group together with you. Each group consists of two randomly grouped people.

If Task 2 is the one randomly selected for payment, then your profit depends on the number of problems you solve compared to the other person in your group. The individual who solves the most problems correctly will receive 2 dollars for every problem he or she solved correctly, while the other participant receives no profit. If there is a tie the payment will be split between the two of you.

We refer to this as the Tournament Task. You will not be informed of how you did in the tournament until the end of the experiment.

If you have any questions before we begin, please raise your hand.

**Rules for Task 2 – *Self treatment***

As in Task 1 you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers.

However for Task 2 your payment depends on your performance relative to your own performance in Task 1.

If Task 2 is the one randomly selected for payment, then your profit depends on the number of problems you solve in Task 2 compared to the number of problems you solved in Task 1. If you solve more problems correctly than you did in Task 1, you will receive 2 dollars for every correct answer you give in Task 2. Otherwise, you will receive no profit from Task 2. If there is a tie with your previous Task 1 score, then you will receive 1 dollar for every correct answer in Task 2.

We refer to this as the Tournament Task. You will not be informed of how you did in the tournament until the end of the experiment.

If you have any questions before we begin, please raise your hand.

**Rules for Task 3 – *Other treatment***

As in the previous two tasks you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However you will now get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task.

If Task 3 is the one randomly selected for profit, then your earnings for this task are determined as follows. If you choose the Piece Rate, you receive 1 dollar per problem you solve correctly. If you choose the Tournament Rate, your performance will be evaluated relative to the performance of the other participant in your group in the Task 2-tournament. If you correctly solve more problems than s/he did during Task 2, then you receive two times the profit from the piece rate, which means you will get 2 dollars per problem you solve correctly. You will receive no earnings for this task if you choose the

tournament and do not solve more problems correctly now, than the other person in your group did during Task 2.

The next screen will ask you to choose whether you want the piece rate or the tournament rate applied to your performance in Task 3. You will then be given 5 minutes to calculate the correct sum of a series of five randomly chosen two-digit numbers in the same way as before.

If you have any questions before we begin, please raise your hand.

**Rules for Task 3 – *Self treatment***

As in the previous two tasks you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However you will now get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task.

If Task 3 is the one randomly selected for profit, then your earnings for this task are determined as follows. If you choose the Piece Rate, you receive 1 dollar per problem you solve correctly. If you choose the Tournament Rate, your performance will be evaluated relative to your own performance in the Task 2-tournament. If you correctly solve more problems than you did during Task 2, then you receive two times the profit from the piece rate, which means you will get 2 dollars per problem you solve correctly. You will receive no earnings for this task if you choose the tournament and do not solve more problems correctly now, than you did during Task 2.

The next screen will ask you to choose whether you want the piece rate or the tournament rate applied to your performance in Task 3. You will then be given 5 minutes to calculate the correct sum of a series of five randomly chosen two-digit numbers in the same way as before.

If you have any questions before we begin, please raise your hand.

**Task 3 Payment Scheme Choice:**

Which compensation scheme do you prefer for Task 3?

**Rank Guess**

In this part we want you to make some guesses.

For each correct guess 1 dollar will be added to your profit from the experiment.

First we would like you to guess in which round your own performance was the best.

- I did more tasks correctly in Task 2 than I did in Task 1.
- I did more tasks correctly in Task 3 than I did in Task 2.

Now we would like you to guess how you performed compared to the other person in your group.

(***Self treatment***: Now we would like you to guess how you performed compared to a randomly chosen person in this room.)

- In Task 2 I did more tasks correctly than what the other person in my group did.

**End of Experiment Survey**

- How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

Not at all willing Very willing

1 2 3 4 5 6 7 8 9 10

- Do you consider yourself a "competitive" person? Please rate on a scale of 1 to 10

Not competitive at all Extremely competitive

1 2 3 4 5 6 7 8 9 10

- Do you think men or women would do better in this addition task?
- Are you an undergraduate or graduate student?
- Have you seen math task in an ICES experiment before?
- Please indicate your gender.
- What is your age?
- What is your major?
- Please specify your ethnicity:

  White

  Hispanic/Latino

  Black/African American

  Asian/Pacific Islander

  Other

  Prefer not to say

- Was there any part of the experiment that confused you? Please explain:
- Do you have any comments or other suggestions on today's experiment? Please explain:

## Appendix B: Instructions for the Online Experiment

Thank you for participating in our study. We estimate that this study will take about 5-10 minutes to complete. After you have finished, you will receive a completion code. Please return to the HIT on MTurk and enter the completion code in the space provided, in order to receive your credit. You will receive $0.25 for completing the HIT. In addition to that, you can earn a bonus of up to $3.50 based on your, and others', performance. The additional money will be paid to you as a bonus through Amazon Mturk in the next few business days. We will now go through the instructions. Please read them carefully. You are only eligible for bonus payment if you adhere to the instructions. As established researchers and long-term Requesters on Amazon MTurk, we promise that the information in this survey is truthful and accurate. We never use deception: the decisions you make are real, any groups that you participate in is real and we always send you the money that you earn in your interactions with others in this HIT. If you have any questions about this research, please feel free to email us at mturk.surveys.research@gmail.com.

Before we move on, please answer the following demographic questions:

What is your age (in years)?

Ethnicity Ethnicity:

- ❍ Asian (1)
- ❍ Black (2)
- ❍ Hispanic-Latino (3)
- ❍ Native American (4)
- ❍ White (5)
- ❍ Other (6)

Gender What is your gender?

- ❍ Male (1)
- ❍ Female (2)

Residence What is your country of residence?

- ❍ United States (1)
- ❍ Canada (2)
- ❍ Other (3)

Student Are you currently a student?

- ❍ Yes (1)
- ❍ No (2)

You will participate in an experiment. This experiment has many other participants in addition to you. Your payoffs will be paid to you as a bonus on Mturk and will depend on your performance and/or on the performance of others. In this experiment you will be asked to complete three tasks that will each take 90 seconds. At the end of the experiment we will randomly select one of the tasks. This is the task that will be relevant for your profit. We determine which task counts for your profit by randomly drawing a number between 1 and 3. The method used to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined.

**Rules for task 1**

For task 1, you will be asked to solve a series of problems by counting the number of zeros (0) in tables consisting of zeros (0) and ones (1). You will be given 90 seconds to count the zeros in as many tables as possible. After the 90 seconds are up you will automatically continue to the next page. That means that you do not need to keep time yourself, but can concentrate on solving the tables. If you solve all available tables before the time is up, please just wait for the survey to continue automatically. In task 1, you get 15 cents per table you solve correctly in the 90 seconds. Your profit does not decrease if you provide an incorrect answer to a table. We refer to this task as the Piece Rate task. Now click to continue to get started with task 1

This is task 1. Please count the number of 0s in each table below and provide the answer.

**Rules for task 2 – *Self treatment***

As in task 1, you will be given 90 seconds to count the zeros in a series of tables with ones and zeros. However, for task 2 your payment depends on your performance relative to that of your own performance in task 1. In task 2, your profit depends on the number of tables you solve in task 2 compared to the number of tables you solved in task 1. If you solve more tables correctly now than you did in task 1, you will receive 30 cents for every correct answer you give in task 2. Otherwise you will receive no profit from task 2. If there is a tie with your previous task 1 score, you will receive 15 cents for every correct answer in task 2. We refer to this as the Tournament Task. You will not be informed of how you did in the tournament until you receive your bonus payment. Now click to continue to get started with task 2.

**Rules for task 2 – *Other treatment***

As in task 1, you will be given 90 seconds to count the zeros in a series of tables with ones and zeros. However for task 2 your payment depends on your performance relative to that of another participant who is doing the same experiment with the same tables as you, and who has been put in a group together with you. Each group consists of two randomly grouped people. You will not be given any information about the other person in your group, and that person will not be given any information about you. In task 2, your profit depends on the number of tables you solve compared to the other person in your group. The individual who solves the most tables correctly will receive 30 cents for every table s/he solved correctly, while the other participant receives no profit. If there is a tie the payment will be split between the two of you. We refer to this as the Tournament Task. You will not be informed of how you did in the tournament until you receive your bonus payment. Now click to continue to get started with task 2.

**Rules for task 2 – *Other Same Ability treatment***

As in task 1, you will be given 90 seconds to count the zeros in a series of tables with ones and zeros. However for task 2 your payment depends on your performance relative to that of another participant who is doing the same experiment with the same tables as you, and who has been put in a group together with you. Each group consists of two randomly grouped people. The only information that you will be given about the other person in your group is that your performance in task 1 was the same, that is you solved the same number of tables in task 1. This information is also given to the other person. In task 2, your profit depends on the number of tables you solve compared to the other person in your group. The individual who solves the most tables correctly will receive 30 cents for every table s/he solved correctly, while the other participant receives no profit. If there is a tie the payment will be split between the two of you. We refer to this as the Tournament Task. You will not be informed of how you did in the tournament until you receive your bonus payment. Now click to continue to get started with task 2.

**Rules for task 2 – *Other Same Gender treatment*** (the information women/man and she/he is varied after the person's own gender)

As in task 1, you will be given 90 seconds to count the zeros in a series of tables with ones and zeros. However for task 2 your payment depends on your performance relative to that of another participant who is doing the same experiment with the same tables as you, and who has been put in a group together with you. Each group consists of two randomly grouped people. The only information that you will be given about the other person in your group is that she (he) is a woman (man). She (He) will get the same information about you. In task 2, your profit depends on the number of tables you solve compared to the other person in your group. The individual who solves the most tables correctly will receive 30 cents for every table s/he solved correctly, while the other participant receives no profit. If there is a tie the payment will be split between the two of you. We refer to this as the Tournament Task. You will not be informed of how you did in the tournament until you receive your bonus payment. Now click to continue to get started with task 2.

**Rules for task 3 – *Self treatment***

As in the previous two tasks you will be given 90 seconds to count the zeros in a series of tables with ones and zeros. However, for task 3 you will get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task. In task 3 your earnings are determined as follows: If you choose the Piece Rate, you receive 15 cents per table you solve correctly. If you choose the Tournament Rate, your performance will be evaluated relative to your own performance in task 2. If you correctly solve more tables now than you did during task 2, then you receive double the profit from the piece rate. That means that you will get 30 cents per table you solve correctly. You will receive no earnings for this task if you choose the tournament and do not solve more tables correctly now, than you did during Task 2. The next screen will ask you to choose whether you want the Piece Rate or the Tournament Rate applied to your performance in task 3. You will then be given 90 seconds to count the number of zeros in a series of tables with ones and zeroes, in the same way as before. Now click to continue.

**Rules for task 3 – *Other treatments***

As in the previous two tasks you will be given 90 seconds to count the zeros in a series of tables with ones and zeros. However, for task 3 you will get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task. In task 3 your earnings are determined as follows: If you choose the Piece Rate, you will receive 15 cents per table you solve correctly. If you choose the Tournament Rate, your performance will be evaluated relative to the performance of the other participant in your group in the Task 2-tournament. If you correctly solve more tables than s/he did during Task 2, you will receive double the profit from the piece rate. That means you will get 30 cents per table you solve correctly. You will receive no earnings for this task if you choose the tournament rate and do not solve more tables correctly now, than the other person in your group did during Task 2. The next screen will ask you to choose whether you want the Piece Rate or the Tournament Rate applied to your performance in task 3. You will then be given 90 seconds to count the number of zeros in a series of tables with ones and zeroes, in the same way as before. Now click to continue.

**Ranks Guess**

In this part we want you to make some guesses. For each correct guess 0.1 dollar will be added to your profit from the experiment.

I did more tasks correctly in Task 2 than I did in Task 1:

- ❍ Yes (1)
- ❍ No (2)

I did more tasks correctly in Task 3 than I did in Task 2:

- ❍ Yes (1)
- ❍ No (2)

If my performance is compared to that of the person I was matched to (*for Self treatment:* that of a randomly chosen person who also participated in this experiment), I think that I did more tasks correctly in Task 2 than s/he did:

- ❍ Yes (1)
- ❍ No (2)

How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

- ❍ 1(Not at all willing to take risks) (1)
- ❍ 2 (2)
- ❍ 3 (3)
- ❍ 4 (4)
- ❍ 5 (5)
- ❍ 6 (6)

- 7 (7)
- 8 (8)
- 9 (9)
- 10 (very willing to take risks) (10)

Do you think men or women generally do better in the "counting zeros"-task that you just did?

- 1(Women do a lot better) (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 (6)
- 7 (7)
- 8 (8)
- 9 (9)
- 10(Men do a lot better) (10)

The experiment is now finished. Please answer the following questions. You will then see the completion code on the screen.

What was your total income last year? Take into account all your sources of income, including scholarships, health benefits, fringe benefits, and others. Please note that this is your personal income, not the income of your household.

- Less than $10,000 (1)
- $10,000 to $20,000 (2)
- $20,000 to $30,000 (3)
- $30,000 to $40,000 (4)
- $40,000 to $50,000 (5)
- $50,000 to $60,000 (6)
- $60,000 to $70,000 (7)
- $70,000 to $80,000 (8)
- $80,000 to $90,000 (9)
- $90,000 to $100,000 (10)
- Over $100,000 (11)

What is your highest level of education completed?

- Less than a high school degree (1)
- High School Diploma (2)
- Vocational Training (3)
- Some College (4)
- Bachelor's Degree (5)
- Graduate Degree (6)

What was/is your major in college/graduate school?

- Economics (1)

- ❍ Psychology (2)
- ❍ Sciences / Math (3)
- ❍ Humanities / Arts (4)
- ❍ Business / MBA (5)
- ❍ Medical (6)
- ❍ Law (7)
- ❍ Others (8)
- ❍ Not applicable (9)

Was anything unclear in the instructions or survey questions? (Optional)

## Appendix C: Summary Statistics and Additional Analysis for the Lab Experiment

**Table C1: Summary Statistics**

| | | Women | Men | Total |
|---|---|---|---|---|
| **Other** | N | N=48 | N=52 | N=100 |
| | Age | 21.02 (0.31) | 20.50 (0.23) | 20.75 (0.19) |
| | Task 1 Score | 8.25 (0.36) | 8.40 (0.45) | 8.33 (0.29) |
| | Task 2 Score | 9.71 (0.38) | 9.42 (0.52) | 9.56 (0.32) |
| | Task 3 Score | 9.98 (0.36) | 10.40 (0.51) | 10.20 (0.31) |
| | Confidence 1 | 0.69 (0.07) | 0.73 (0.06) | 0.71 (0.05) |
| | Confidence 2 | 0.56 (0.07) | 0.67 (0.07) | 0.62 (0.05) |
| | Confidence 3 | 0.71 (0.07) | 0.85 (0.05) | 0.78 (0.04) |
| | Risk Attitudes | 6.00 (0.29) | 7.25 (0.24) | 6.65 (0.20) |
| **Self** | N | N=55 | N=49 | N=104 |
| | Age | 21.09 (0.25) | 21.33 (0.26) | 21.20 (0.18) |
| | Task 1 Score | 8.29 (0.35) | 8.49 (0.48) | 8.38 (0.29) |
| | Task 2 Score | 9.49 (0.42) | 9.47 (0.46) | 9.48 (0.31) |
| | Task 3 Score | 9.53 (0.43) | 9.76 (0.52) | 9.63 (0.33) |
| | Confidence 1 | 0.64 (0.07) | 0.73 (0.06) | 0.68 (0.05) |
| | Confidence 2 | 0.56 (0.07) | 0.61 (0.07) | 0.59 (0.05) |
| | Confidence 3 | 0.71 (0.06) | 0.76 (0.06) | 0.73 (0.04) |
| | Risk Attitudes | 6.16 (0.29) | 7.35 (0.27) | 6.72 (0.21) |
| **Total** | N | N=103 | N=101 | N=204 |
| | Age | 21.06 (0.20) | 20.90 (0.18) | 20.98 (0.13) |
| | Task 1 Score | 8.27 (0.25) | 8.45 (0.33) | 8.36 (0.20) |
| | Task 2 Score | 9.59 (0.28) | 9.45 (0.34) | 9.52 (0.22) |

| | | | | |
|---|---|---|---|---|
| | Task 3 Score | 9.74 (0.28) | 10.09 (0.36) | 9.91 (0.23) |
| | Confidence 1 | 0.66 (0.05) | 0.73 (0.04) | 0.70 (0.03) |
| | Confidence 2 | 0.56 (0.05) | 0.64 (0.05) | 0.60 (0.03) |
| | Confidence 3 | 0.71 (0.04) | 0.80 (0.04) | 0.75 (0.03) |
| | Risk Attitudes | 6.09 (0.21) | 7.30 (0.18) | 6.69 (0.14) |

*Notes: Standard errors in parentheses. Confidence 1 is the incentivized belief question on scoring better in Task 2 than in Task 1 (1 if Yes, 0 if No). Confidence 2 is the incentivized belief question on scoring better in Task 3 than in Task 2 (1 if Yes, 0 if No). Confidence 3 is the incentivized belief question on scoring better than the other person in Task 2 (1 if Yes, 0 if No). Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk".*

**Table C2: Proportion (in percentages) Who Chooses Tournament Pay in Task 3**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test of Proportions | Fisher's Exact test |
| **Other** | 37.5 (7.1) | 57.7 (6.9) | 48.0 (5.0) | p=0.044 | p=0.048 |
| **Self** | 41.8 (6.7) | 55.1 (7.2) | 48.1 (4.9) | p=0.176 | p=0.238 |
| **Total** | 39.8 (4.8) | 56.4 (5.0) | 48.0 (3.5) | p=0.018 | p=0.025 |

*Notes: Standard errors in parentheses.*

**Table C3: Risk Preferences**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test | Wilcoxon test |
| **Other** | 6.00 (0.29) | 7.25 (0.24) | 6.65 (0.20) | p=0.001 | p=0.002 |
| **Self** | 6.16 (0.29) | 7.35 (0.27) | 6.72 (0.21) | p=0.004 | p=0.005 |
| **Total** | 6.09 (0.21) | 7.30 (0.18) | 6.69 (0.14) | p=0.000 | p=0.000 |

*Notes: Standard errors in parentheses. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk".*

**Table C4: Confidence**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test for Proportions | Fisher's Exact test |
| **Other** | 0.71 (0.07) | 0.85 (0.05) | 0.78 (0.04) | p=0.097 | p=0.146 |
| **Self** | 0.56 (0.07) | 0.61 (0.07) | 0.59 (0.05) | p=0.615 | p=0.692 |
| **Total** | 0.63 (0.04) | 0.73 (0.05) | 0.68 (0.03) | p=0.119 | p=0.134 |

*Notes: Standard errors in parentheses. Confidence is a binary variable that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments).*

**Table C5: Improvement in Task Scores (in percentages) Between Task 1 and Task 2, by Treatment and Gender**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test | Wilcoxon test |
| **Other** | 32.8 (13.0) | 15.5 (4.4) | 23.9 (6.7) | p=0.196 | p=0.641 |
| **Self** | 18.2 (4.4) | 18.2 (5.4) | 18.2 (3.4) | p=0.992 | p=0.628 |
| **Total** | 25.0 (6.5) | 16.8 (3.5) | 21.0 (3.7) | p=0.271 | p=0.568 |

*Notes: Improvement in score between Task 1 and Task 2. Standard errors in parentheses. Score improvement is calculated by ([Task2Score/Task1Score] – 1)*100.*

**Table C6: Probit Regressions for the Tournament Entry Decision (marginal effects from probit):**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | (Other) | (Other) | (Self) | (Self) |
| Female | -0.193** | -0.113 | -0.130 | -0.026 |
| | (0.09) | (0.09) | (0.10) | (0.09) |
| Confidence | | 0.242** | | -0.019 |
| | | (0.12) | | (0.10) |
| Risk | | 0.039* | | 0.089*** |
| | | (0.02) | | (0.02) |
| N | 100 | 100 | 104 | 104 |
| Pseudo R-square | 0.091 | 0.144 | 0.014 | 0.108 |

*Notes: This table is the probit analysis version of Table 2 – Panel A in the main paper. Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the "Other"-treatment). Significance: ***p<0.01 **p<0.05 *p<0.1*

**Table C7: OLS Regressions for the Difference in Difference Estimates in Tournament Entry Decision**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Female | -0.162** | -0.128 | | |
| | (0.07) | (0.10) | | |
| Treatment (Other) | | 0.028 | 0.155 | -0.223* |
| | | (0.10) | (0.21) | (0.12) |
| Female*Treatment | | -0.070 | | |
| | | (0.14) | | |
| Risk | | | 0.081*** | |
| | | | (0.02) | |
| Risk*Treatment | | | -0.022 | |
| | | | (0.03) | |
| Confidence | | | | 0.070 |
| | | | | (0.10) |
| Confidence*Treatment | | | | 0.268* |
| | | | | (0.15) |
| Constant | 0.347*** | 0.333*** | -0.209 | 0.280** |
| | (0.11) | (0.12) | (0.16) | (0.13) |
| N | 204 | 204 | 204 | 204 |
| R-square | 0.050 | 0.051 | 0.105 | 0.063 |

*Notes: Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the "Other"-treatment). Significance: ***p<0.01 **p<0.05 *p<0.1*

## Appendix D: Summary Statistics and Additional Analysis for the Online Experiment

**Table D1: Summary Statistics**

| | | Women | Men | Total |
|---|---|---|---|---|
| **Other** | N | N=115 | N=130 | N=245 |
| | Age | 36.66 (1.10) | 36.68 (0.95) | 35.61 (0.72) |
| | Task 1 Score | 2.97 (0.14) | 2.87 (0.12) | 2.92 (0.09) |
| | Task 2 Score | 3.27 (0.15) | 3.33 (0.12) | 3.30 (0.10) |
| | Task 3 Score | 3.57 (0.15) | 3.75 (0.13) | 3.67 (0.10) |
| | Confidence 1 | 0.43 (0.05) | 0.62 (0.04) | 0.53 (0.03) |
| | Confidence 2 | 0.62 (0.05) | 0.65 (0.04) | 0.63 (0.03) |
| | Confidence 3 | 0.53 (0.05) | 0.61 (0.04) | 0.57 (0.03) |
| | Risk Attitudes | 5.41 (0.22) | 5.78 (0.20) | 5.60 (0.15) |
| | Income ($) | 41,957 (3,068) | 40,000 (2,626) | 40,918 (2,001) |
| **Other, Same Gender** | N | N=128 | N=129 | N=257 |
| | Age | 37.34 (1.03) | 33.44 (0.96) | 35.39 (0.71) |
| | Task 1 Score | 3.10 (0.12) | 2.92 (0.13) | 3.01 (0.09) |
| | Task 2 Score | 3.23 (0.14) | 3.02 (0.14) | 3.12 (0.10) |
| | Task 3 Score | 3.50 (0.14) | 3.21 (0.13) | 3.35 (0.09) |
| | Confidence 1 | 0.42 (0.04) | 0.49 (0.04) | 0.46 (0.03) |
| | Confidence 2 | 0.64 (0.04) | 0.55 (0.04) | 0.60 (0.03) |
| | Confidence 3 | 0.48 (0.04) | 0.50 (0.04) | 0.49 (0.03) |
| | Risk Attitudes | 5.00 (0.21) | 5.84 (0.20) | 5.42 (0.15) |
| | Income ($) | 33,984 (2,533) | 39,496 (2,576) | 36,751 (1,811) |
| | N | N=124 | N=120 | N=244 |
| | Age | 35.37 (0.97) | 33.88 (1.01) | 34.64 (0.70) |

| | | | | |
|---|---|---|---|---|
| **Other, Same Ability** | Task 1 Score | 3.02 (0.15) | 2.92 (0.14) | 2.97 (0.10) |
| | Task 2 Score | 3.49 (0.14) | 3.13 (0.15) | 3.31 (0.10) |
| | Task 3 Score | 3.78 (0.14) | 3.29 (0.16) | 3.54 (0.11) |
| | Confidence 1 | 0.57 (0.04) | 0.53 (0.05) | 0.55 (0.03) |
| | Confidence 2 | 0.62 (0.04) | 0.54 (0.05) | 0.58 (0.03) |
| | Confidence 3 | 0.53 (0.04) | 0.63 (0.04) | 0.58 (0.03) |
| | Risk Attitudes | 5.02 (0.22) | 5.76 (0.21) | 5.39 (0.15) |
| | Income ($) | 33,951 (2,431) | 38,417 (2,744) | 36,148 (1,831) |
| **Self** | N | N=129 | N=119 | N=248 |
| | Age | 36.93 (1.06) | 32.74 (0.82) | 34.92 (0.69) |
| | Task 1 Score | 3.14 (0.12) | 2.86 (0.13) | 3.00 (0.09) |
| | Task 2 Score | 3.36 (0.13) | 3.09 (0.13) | 3.23 (0.09) |
| | Task 3 Score | 3.57 (0.14) | 3.55 (0.14) | 3.56 (0.10) |
| | Confidence 1 | 0.56 (0.04) | 0.60 (0.05) | 0.58 (0.03) |
| | Confidence 2 | 0.55 (0.04) | 0.55 (0.05) | 0.55 (0.03) |
| | Confidence 3 | 0.29 (0.04) | 0.48 (0.05) | 0.38 (0.03) |
| | Risk Attitudes | 5.05 (0.18) | 6.00 (0.19) | 5.50 (0.14) |
| | Income ($) | 39,031 (2,761) | 39,370 (2,726) | 39,194 (1,939) |
| **Total** | N | N=496 | N=498 | N=994 |
| | Age | 36.58 (0.52) | 33.70 (0.42) | 35.14 (0.35) |
| | Task 1 Score | 3.06 (0.07) | 2.89 (0.06) | 2.98 (0.05) |
| | Task 2 Score | 3.34 (0.07) | 3.14 (0.07) | 3.24 (0.05) |
| | Task 3 Score | 3.61 (0.07) | 3.45 (0.07) | 3.53 (0.05) |
| | Confidence 1 | 0.50 (0.02) | 0.56 (0.02) | 0.53 (0.02) |
| | Confidence 2 | 0.61 (0.02) | 0.57 (0.02) | 0.59 (0.02) |

| | | | |
|---|---|---|---|
| Confidence 3 | 0.46 (0.02) | 0.55 (0.02) | 0.51 (0.02) |
| Risk Attitudes | 5.11 (0.10) | 5.84 (0.10) | 5.48 (0.07) |
| Income ($) | 37,137 (1,353) | 39,337 (1,329) | 38,239 (948) |

*Notes: Standard errors in parentheses. Confidence 1 is the incentivized belief question on scoring better in Task 2 than in Task 1 (1 if Yes, 0 if No). Confidence 2 is the incentivized belief question on scoring better in Task 3 than in Task 2 (1 if Yes, 0 if No). Confidence 3 is the incentivized belief question on scoring better than the other person in Task 2 (1 if Yes, 0 if No). Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Income is the self-reported personal earnings in the last year. It takes on the value 1 if the income is less than $10,000; 2 if income is within the interval from $10,000 to $20,000; 3 if $20,000 to $30,000; 4 if $30,000 to $40,000; 5 if $40,000 to $50,000; 6 if $50,000 to $60,000; 7 if $60,000 to $70,000; 8 if $70,000 to $80,000; 9 if $80,000 to $90,000; 10 if $90,000 to $100,000; 11 if over $100,000. While calculating the average, we take the mid-point of each interval. Mid-point of income over $100,000 is estimated as $125,000.*

**Table D2: Proportion (in percentages) Who Chooses Tournament Pay in Task 3**

| | **Women** | **Men** | **Total** | ***p-values for gender difference*** | |
|---|---|---|---|---|---|
| | | | | t-test for Proportions | Fisher's Exact test |
| **Other** | 27.8 (4.2) | 40.0 (4.3) | 34.3 (3.0) | p=0.045 | p=0.059 |
| **Other, Same Gender** | 21.9 (3.7) | 34.1 (4.2) | 28.0 (2.8) | p=0.029 | p=0.037 |
| **Other, Same Ability** | 30.6 (4.2) | 33.3 (4.3) | 32.0 (3.0) | p=0.653 | p=0.682 |
| **Self** | 35.7 (4.2) | 31.1 (4.3) | 33.5 (3.0) | p=0.446 | p=0.501 |
| **Total** | 29.0 (2.0) | 34.7 (2.1) | 31.9 (1.5) | p=0.054 | p=0.057 |

*Notes: Standard errors in parentheses.*

**Table D3: Risk Preferences**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test | Wilcoxon test |
| **Other** | 5.41 (0.22) | 5.78 (0.20) | 5.60 (0.15) | p=0.222 | p=0.243 |
| **Other, Same Gender** | 5.00 (0.21) | 5.84 (0.20) | 5.42 (0.15) | p=0.004 | p=0.005 |
| **Other, Same Ability** | 5.02 (0.22) | 5.76 (0.21) | 5.39 (0.15) | p=0.016 | p=0.025 |
| **Self** | 5.05 (0.18) | 6.00 (0.19) | 5.50 (0.14) | p=0.000 | p=0.000 |
| **Total** | 5.11 (0.10) | 5.84 (0.10) | 5.48 (0.07) | p=0.000 | p=0.000 |

*Notes: Standard errors in parentheses. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk".*

**Table D4: Confidence**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test for Proportions | Fisher's Exact test |
| **Other** | 0.53 (0.05) | 0.61 (0.04) | 0.57 (0.03) | p=0.223 | p=0.246 |
| **Other, Same Gender** | 0.48 (0.04) | 0.50 (0.04) | 0.49 (0.03) | p=0.755 | p=0.803 |
| **Other, Same Ability** | 0.53 (0.04) | 0.63 (0.04) | 0.58 (0.03) | p=0.143 | p=0.155 |
| **Other-treatments combined** | 0.51 (0.03) | 0.58 (0.03) | 0.55 (0.02) | p=0.085 | p=0.091 |
| **Self** | 0.55 (0.04) | 0.55 (0.05) | 0.55 (0.03) | p=0.947 | p=1.000 |
| **Total** | 0.52 (0.02) | 0.57 (0.02) | 0.55 (0.02) | p=0.128 | p=0.143 |

*Notes: Standard errors in parentheses. Confidence is a binary variable that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments).*

**Table D5: Improvement in Task Scores (in percentages) Between Task 1 and Task 2, by Treatment and Gender**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test | Wilcoxon test |
| **Other** | 15.40 (6.34) | 33.80 (8.32) | 25.23 (5.36) | p=0.087 | p=0.409 |
| **Other, Same Gender** | 13.63 (5.65) | 17.89 (7.00) | 15.75 (4.49) | p=0.636 | p=0.789 |
| **Other, Same Ability** | 30.35 (7.12) | 21.88 (7.20) | 26.16 (5.06) | p=0.404 | p=0.437 |
| **Self** | 11.30 (4.78) | 25.25 (8.30) | 18.01 (4.71) | p=0.139 | p=0.759 |
| **Total** | 17.58 (3.01) | 24.78 (3.86) | 21.20 (2.45) | p=0.142 | p=0.899 |

*Notes: Improvement in score between Task 1 and Task 2. Standard errors in parentheses. Score improvement is calculated by ([Task2Score/Task1Score] – 1)*100.*

**Table D6: Probit Regressions for the Tournament Entry Decision (marginal effects from probit)**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | (Other) | (Other) | (Self) | (Self) |
| Female | -0.125** | -0.090* | 0.051 | 0.079 |
| | (0.06) | (0.05) | (0.06) | (0.06) |
| Confidence | | 0.239*** | | 0.126** |
| | | (0.05) | | (0.06) |
| Risk | | 0.045*** | | 0.030** |
| | | (0.02) | | (0.01) |
| N | 245 | 245 | 248 | 248 |
| Pseudo R^2 | 0.022 | 0.146 | 0.005 | 0.033 |

| | (5) | (6) | (7) | (8) |
|---|---|---|---|---|
| | (Other, Same Gender) | (Other, Same Gender) | (Other, Same Ability) | (Other, Same Ability) |
| Female | -0.121** | -0.092* | -0.028 | 0.024 |
| | (0.05) | (0.05) | (0.06) | (0.06) |
| Confidence | | 0.255*** | | 0.281*** |
| | | (0.05) | | (0.05) |
| Risk | | 0.026** | | 0.039*** |
| | | (0.01) | | (0.01) |
| N | 257 | 257 | 244 | 244 |
| Pseudo R^2 | 0.016 | 0.142 | 0.001 | 0.134 |

*Notes: This table is the probit analysis version of Table 2 – Panels B and C in the main paper. Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments). Significance: ***p<0.01 **p<0.05 *p<0.1*

**Table D7: OLS Regressions for the Difference in Difference Estimates in Tournament Entry Decision**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Female | -0.039 | 0.043 | | |
| | (0.04) | (0.06) | | |
| Treatment (Other) | | 0.089 | -0.171 | -0.108* |
| | | (0.06) | (0.11) | (0.06) |
| Female*Treatment | | -0.166* | | |
| | | (0.09) | | |
| Risk | | | 0.027* | |
| | | | (0.01) | |
| Risk*Treatment | | | 0.032* | |
| | | | (0.02) | |
| Confidence | | | | 0.116* |
| | | | | (0.06) |
| Confidence*Treatment | | | | 0.200** |
| | | | | (0.08) |
| Constant | 0.333*** | 0.289*** | 0.144 | 0.268*** |
| | (0.05) | (0.06) | (0.10) | (0.06) |
| N | 493 | 493 | 493 | 493 |
| R-square | 0.002 | 0.010 | 0.050 | 0.062 |

*Notes: Regressions are for the pooled "Other" and "Self" treatments from the online experiment. Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments). Significance: ***p<0.01 **p<0.05 *p<0.1*

## Appendix E: Analysis for the Pooled Data

**Table E1: Confidence for combined three "Other" treatments, Pooled Data**

| | Women | Men | Total | *p-values for gender difference* | |
|---|---|---|---|---|---|
| | | | | t-test for Proportions | Fisher's Exact test |
| **Other treatments** | 0.54 (0.02) | 0.61 (0.02) | 0.57 (0.02) | p=0.032 | p=0.037 |
| **Self** | 0.55 (0.04) | 0.57 (0.04) | 0.56 (0.03) | p=0.747 | p=0.830 |
| **Total** | 0.54 (0.02) | 0.60 (0.02) | 0.57 (0.01) | p=0.047 | p=0.054 |

*Notes: Data is pooled from laboratory and online experiment. Standard errors in parentheses. Comparison of Confidence in the Self-treatment to the combined three Other-treatments. Confidence is a binary variable that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments).*

**Table E2: OLS Regressions for the Tournament Entry Decision, Pooled Data**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | (Other) | (Other) | (Self) | (Self) |
| Female | -0.146*** | -0.097** | -0.009 | 0.044 |
| | (0.05) | (0.05) | (0.05) | (0.05) |
| Confidence | | 0.251*** | | 0.097* |
| | | (0.05) | | (0.05) |
| Risk | | 0.041*** | | 0.048*** |
| | | (0.01) | | (0.01) |
| Constant | 0.432*** | 0.017 | 0.372*** | 0.012 |
| | (0.04) | (0.07) | (0.04) | (0.09) |
| N | 345 | 345 | 352 | 352 |
| R-square | 0.066 | 0.180 | 0.010 | 0.063 |

*Notes: Regressions are for the pooled "Other" and "Self" treatments from the online and laboratory experiments. Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for the normalized task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments). Significance: ***p<0.01 **p<0.05 *p<0.1*

**Table E3: OLS Regressions for the Difference in Difference Estimates in Tournament Entry Decision, Pooled Data**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Female | -0.077** | -0.010 | | |
| | (0.04) | (0.05) | | |
| Treatment (Other) | | 0.071 | -0.085 | -0.135*** |
| | | (0.05) | (0.10) | (0.05) |
| Female*Treatment | | -0.135* | | |
| | | (0.07) | | |
| Risk | | | 0.043*** | |
| | | | (0.01) | |
| Risk*Treatment | | | 0.015 | |
| | | | (0.02) | |
| Confidence | | | | 0.101** |
| | | | | (0.05) |
| Confidence*Treatment | | | | 0.214*** |
| | | | | (0.07) |
| Constant | 0.403*** | 0.366*** | 0.114 | 0.309*** |
| | (0.03) | (0.04) | (0.07) | (0.04) |
| N | 697 | 697 | 697 | 697 |
| R-square | 0.030 | 0.035 | 0.078 | 0.076 |

*Notes: Regressions are for the pooled "Other" and "Self" treatments from the online and laboratory experiments. Dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for the normalized task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments). Significance: ***p<0.01 **p<0.05 *p<0.1*