

# **LUCK OR SKILL: HOW WOMEN AND MEN RESPOND TO NOISY FEEDBACK**

## **ONLINE APPENDIX (NOT FOR PUBLICATION)**

Gauri Kartini Shastry

Olga Shurchkov

Lingjun “Lotus” Xia

Wellesley College

Wellesley College

Massachusetts Institute of  
Technology

### [Contents](#)

- A. SCREENSHOTS OF INSTRUCTIONS FOR THE MAIN EXPERIMENT
- B. PROTOCOL FOR THE OPTIONAL ADDITIONAL FEEDBACK TREATMENT
- C. TASK SELECTION CONSIDERATIONS AND PILOT SURVEYS
- D. ANALYSIS WITH SAMPLE RESTRICTED TO WAVES 1, 2, AND 3 ONLY
- E. ANALYSIS WITH SAMPLE RESTRICTED TO WAVES 1 AND 4 ONLY
- F. ADDITIONAL ANALYSIS
- G. ROBUSTNESS CHECKS

## A. Screenshots of Instructions for the Main Experiment



### **Participation is voluntary**

It is your choice whether or not to participate in this research. If you choose to participate, you may change your mind and quit the study at any time. Refusal to participate or stopping your participation will involve no penalty or loss of benefits to which you are otherwise entitled.

### **What is the purpose of this research?**

The purpose of this research is to understand how individuals perform in certain types of analytical tasks.

### **How long will I take part in this research?**

Your participation will take approximately 10 minutes to complete.

### **What can I expect if I take part in this research?**

As a participant, you will answer a series of questions.

### **What are the risks and possible discomforts?**

If you choose to participate, the effects should be comparable to those you would experience from viewing a computer monitor for 10 minutes and using a mouse or keyboard.

### **Are there any benefits from being in this research study?**

In addition to payment, the study includes analytical questions that may be interesting to consider.

### **Will I be compensated for participating in this research?**

You will receive a base payment of \$0.5 if you finish the survey and pass the attention checking questions. You will also receive bonus based on your performance in this survey. On average, you will get \$1.6 of bonus, in addition to the base payment. Note that if you fail to answer the attention checking questions correctly, you will not receive any compensation.

### **If I take part in this research, how will my privacy be protected? What happens to the information you collect?**

Your data will be kept completely anonymous by the assignment of random participant numbers. Your data will be linked only to your participant number; there is no way of matching any data collected to your name. When the research is completed, the raw data will be stored on a password-protected computer, accessible only to the researchers.

### **If I have any questions, concerns or complaints about this research study, who can I talk to?**

The researcher for this study is Olga Shurchkov who can be reached at 781-283-2984, 106 Central St., Wellesley, MA 02481, [olga.shurchkov@wellesley.edu](mailto:olga.shurchkov@wellesley.edu). This research has been reviewed by the Institutional Review Board of Wellesley College. If you have any questions about your rights as a participant in this study or any concerns or complaints, please contact the chair of the Wellesley College IRB, Nancy Marshall, [nmarshall@wellesley.edu](mailto:nmarshall@wellesley.edu).

### **Consent**

You may wish to print this page for your records.

I would like to participate in this study

I would like to leave this study



**Here is your 5-digit ID number for this HIT: 14816**

Please write the ID number down. It is absolutely essential that you keep this ID with you throughout this HIT. This is the identifier we use to pay you for your participation.

You need to enter the 5-digit ID number above in **TWO** different places:

- (1) On the last page of this survey.
- (2) On the MTurk page for this HIT.



---

This survey consists of two rounds. Each round contains a section of problem solving and a few open-ended questions.

You will be paid a base payment of **\$0.5** for your participation in this study, which will take about 10 minutes. You will also be paid a **bonus** based on your performance in the problem solving sections. Detailed payment scheme will be revealed after you complete each section of problem solving.

Now, round 1 begins.

In the beginning of the section, you will carefully read the instructions and an example of the kind of questions you will see in this section.

Then you will have **2.5 minutes** to complete as many problems as you can. Time will run out automatically. You will not be able to advance to the next section before the end of the 2.5 minutes.

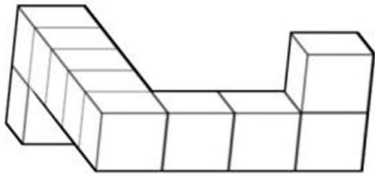
After the 2.5-minute problem solving period, we will ask you to answer a few follow-up questions.

There is an **attention checking question** in this section. If you fail to answer it correctly, you will not be able to continue with the rest of the experiment, and you will not receive any compensation for participating.

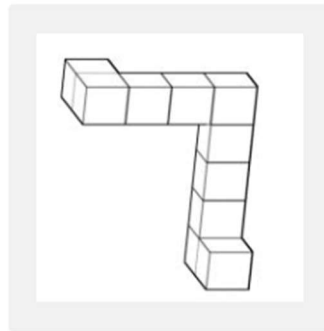
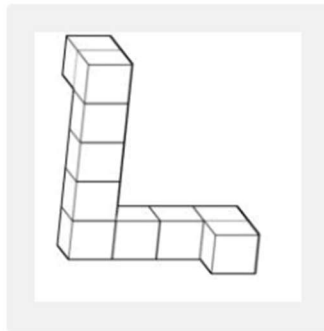
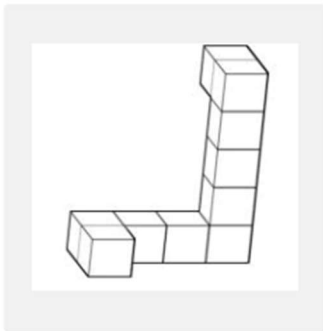


Below is an example of the type of problems you will see in this problem solving section.

For each question, you will see a target shape at the top, and three choices. Select the choice that is a **rotated version** of the target one at the top.



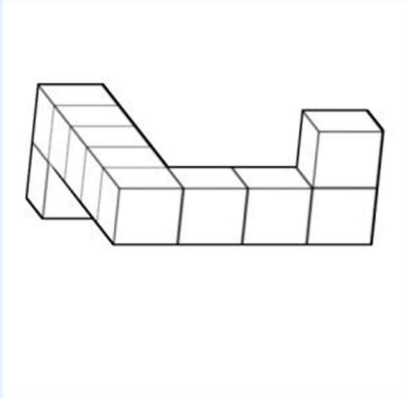
PRACTICE: Select the shape below that is a rotated version of the one at the top (these shapes, made up of ten blocks, are similar to the ones that will follow).



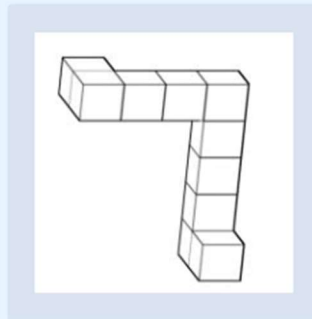
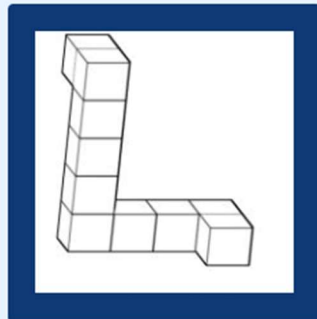
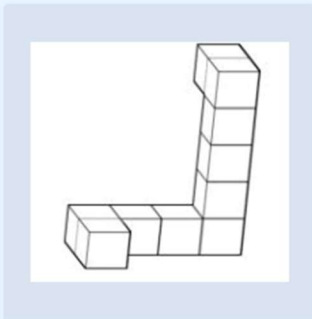
Below is an example of the type of problems you will see in this problem solving section.

For each question, you will see a target shape at the top, and three choices. Select the choice that is a **rotated version** of the target one at the top.

Please answer this question.



PRACTICE: Select the shape below that is a rotated version of the one at the top (these shapes, made up of ten blocks, are similar to the ones that will follow).



---

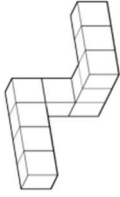
Now, you will have **2.5 minutes** to solve as many problems as possible in this section. Time will run out automatically. You will not be able to advance to the next section before the end of the 2.5 minutes.



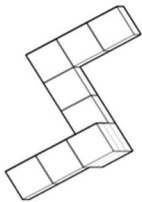
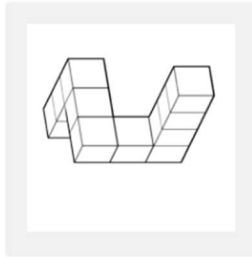
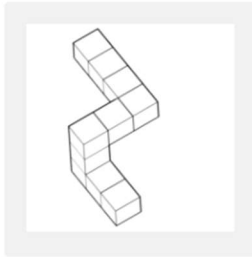
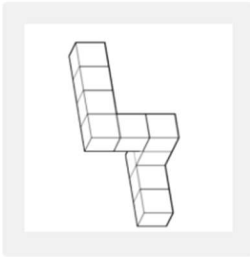
---

0130

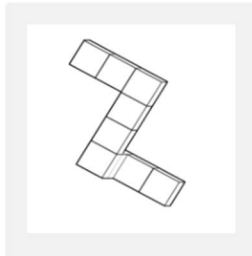
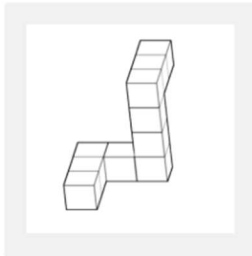
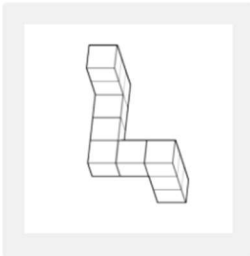
---



Select the shape that is a rotated version of the one above.



Select the shape that is a rotated version of the one above.



---

followed by more MRT questions below...



This is a question to ensure that you are paying attention. If you do not answer it correctly, you will not receive compensation for taking the survey. Please choose the most accurate description of your current activity.

I am jogging outside.

I am sleeping.

I am swimming.

I am playing the piano.

I am taking a survey.

>>

---

How many questions do you think you solved correctly in this section?



As alluded to in the beginning of this section, your bonus will be calculated based on your performance in the problem solving section. Here is how we will calculate your bonus payment:

We previously ran this test on a similar group and got the performance. We will randomly pair you with a participant from that group, and your payment will be as follows:

- If your total score is **higher** than your match's, you get **20 cents** for each correct answer.
- If your total score is **lower** than your match's, you get **15 cents** for each correct answer.

For example, if your score is 3 in the problem solving section, while your randomly chosen match's score is 2, then you will get a bonus of 60 cents (20 cents \* 3 correct answers). On the other hand, if your score is 1 and your match's score is 2, then you will get a bonus of 15 cents (15 cents \* 1 correct answer).



Our matching process has randomly matched you with a participant from the other group. Your score has been compared with his/hers, and your payment is shown below.

Please click on the number to indicate that you have seen this payment.

0

Unknown Gender Condition

OR

Our matching process has randomly matched you with a male participant from the other group. Your score has been compared with his, and your payment is shown below.

Please click on the number to indicate that you have seen this payment.

0

Male Match Condition

OR

Our matching process has randomly matched you with a female participant from the other group. Your score has been compared with hers, and your payment is shown below.

Please click on the number to indicate that you have seen this payment.

0.15

Female Match Condition

>>

Do you think your payment is an **above average** payment or a **below average** payment ?

above average

below average



We have calculated the average payment of the group of participants who have previously completed the same test. Compared with the average payment, your payment is **below average**.

Forced feedback condition  
(if payment below average)

>>

We have calculated the average payment of the group of participants who have previously completed the same test. Compared with the average payment, your payment is **above average**.

Forced feedback condition  
(if payment above average)

>>

Now you have an opportunity to find out if your payment is above the average payment of the other group. The knowledge of your **relative payment** will inform your decision in the next round and influence your payment.

Please select the **maximum amount** you are willing to pay to receive this information. Please note that, if you select a certain amount, we automatically assume you are willing to pay any price that is lower than your selection. For example, if you select 10 cents, we automatically assume that you are also willing to pay 5 cents.

Then, we will randomly draw a price from these five options. If this price is less than or equal to your maximum, then the price will be subtracted from your payment, and you will receive information about your relative payment. On the other hand, if this price is greater than your maximum, then you will not get any information, and your payment will be unaffected.

Please note that, in a rare instance, your price of information may be greater than your final bonus payment, in which case, the difference will be deducted from your base payment.

5 cents

10 cents

15 cents

20 cents

25 cents

Optional feedback condition

>>

---

The randomly selected amount is: 15

You have indicated that you are NOT willing to pay this much.

Optional feedback condition  
(if bid is below feedback price)

>>

---

The randomly selected amount is: 10

You have indicated that you are willing to pay this much.

Optional feedback condition  
(if bid is above feedback price)

>>

---

We have calculated the average payment of the group of participants who have previously completed the same test. Compared with the average payment, your payment is above average.

Optional feedback condition  
(if bid is above feedback price  
and payment above average)

>>

---

We have calculated the average payment of the group of participants who have previously completed the same test. Compared with the average payment, your payment is below average.

Optional feedback condition  
(if bid is above feedback price  
and payment below average)

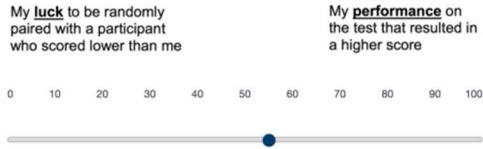
>>



On a scale from 0 to 100, please rate the relative importance of your own test score and your random match's score in contributing to your overall payment.

Example:

For example, if you choose 55 as shown in the example, you attribute 55% of your payment to your **performance** on the test that resulted in a higher score, and 45% of your payment to your **luck** of being randomly paired with a participant who scored lower than you.



Now, move the slider to indicate the relative importance of your own test score and your random match's score in contributing to your overall payment.



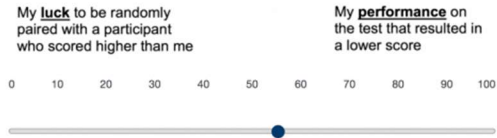
Attribution question  
if success



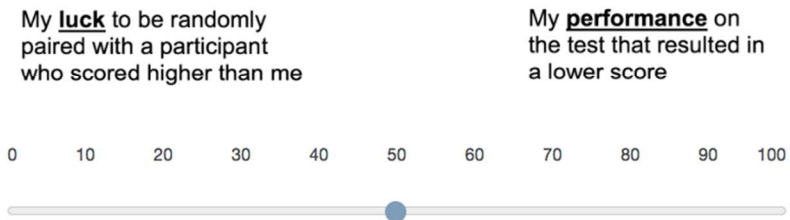
On a scale from 0 to 100, please rate the relative importance of your own test score and your random match's score in contributing to your overall payment.

Example:

For example, if you choose 55 as shown in the example, you attribute 55% of your payment to your **performance** on the test that resulted in a lower score, and 45% of your payment to your **luck** of being randomly paired with a participant who scored higher than you.



Now, move the slider to indicate the relative importance of your own test score and your random match's score in contributing to your overall payment.



Attribution question if failure



Now, Round 2 begins.

You will have 2.5 minutes to complete another section of the same kind of problems as in the first round.

Before you work on the problems, you will have an opportunity to choose what kind of payment scheme you want to use in this round.

**Option 1:**

We randomly pair you with a participant from the other group.

- If your total score is higher than your match's, you get **25 cents** for each correct answer.
- If your total score is lower than your match's, you get **10 cents** for each correct answer.

**Option 2:**

You get **17.5 cents** for each correct answer, regardless of anyone else's score.

Your bonus payment for this section will be determined depending on what option you choose.

Which payment scheme do you want to use?

Option 1

Option 2

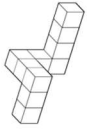


---

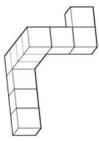
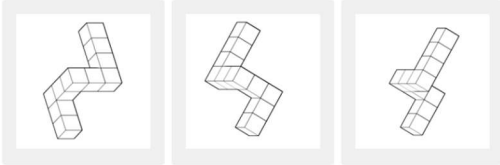
You have chosen your payment scheme for this round. The second section of problem solving starts on next page.



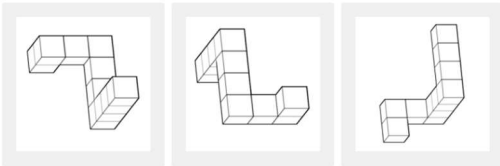
0219



Select the shape that is a rotated version of the one above.



Select the shape that is a rotated version of the one above.



...and more MRT questions below

You have completed all of the questions. Since you selected option 1 as your payment scheme, our matching process has randomly paired you with a participant from the other group. For your information, the score of your random match is shown below.

Please click on the number to indicate that you have seen the number.

1

You have completed all of the questions. We will calculate your bonus payment in this round and add it to your bonus payment in round 1. The final payment will be transferred to you within seven days.

Please continue to next page to complete a short questionnaire.

>>

If Option 1 is chosen

You have completed all of the questions. We will calculate your bonus payment in this round and add it to your bonus payment in round 1. The final payment will be transferred to you within seven days.

Please continue to next page to complete a short questionnaire.

>>

If Option 2 is chosen

---

Please rank from 1-10 how you see yourself:

Are you generally a person who avoids taking risks or are you fully prepared to take risks?

(1 is unwilling to take risks and 10 is fully prepared to take risks)

1      2      3      4      5      6      7      8      9      10



Do you think men or women are more likely to get a high score in this task?

men

women

What is your gender?

Male

Female

Other/ Do not wish to disclose

What is your age in years?

What is the highest level of education you have completed?

Less than high school

High school or GED

Some College

2-year college degree (Associate)

4-year college degree (BA, BS)

Master's degree (MA, MS)

Doctoral degree (PhD)

Professional degree (MD, JD, DDS, etc.)

---



Are you of Hispanic origin or descent, such as Mexican, Puerto Rican, Cuban, or other Spanish background?

Yes

No

---

Which of the following best describes your race?

White

Black or African American

Native American

Asian

Asian-Pacific Islander

Other/ Do not wish to disclose

Which of the following best describes your annual household income before taxes?

Less than \$10,000

\$10,000 - \$19,999

\$20,000 - \$29,999

\$30,000 - \$39,999

\$40,000 - \$49,999

\$50,000 - \$74,999

\$75,000 - \$99,999

\$100,000 - \$149,999

\$150,000 - \$249,999

\$250,000 - \$499,999

\$500,000 and over

>>

## B. Protocol for the Optional Additional Feedback Treatment

In this condition, participants had an opportunity to give up some of their payment in exchange for feedback on relative payment. Specifically, participants indicated their willingness to pay to receive feedback from a list of options ranging from 5 cents to 25 cents. The computer then randomly drew a price from the list. If the participant indicated that she was willing to pay the randomly drawn price, then the price was subtracted from her payment, and the rest of the design was as in the Additional Feedback condition. If the participant indicated that she was not willing to pay, then she did not receive any additional feedback, and the rest of the design was as in the Limited Feedback condition.

### C. Task Selection Considerations and Pilot Surveys

Pilot surveys were designed to help us identify a task that would satisfy the following three criteria. The choice of the MRT task was based on three criteria. First, we looked for a task that cannot be easily “cracked” by the AMT subjects who, unlike lab subjects, have access to the internet and calculators. Second, we preferred a skill-based task to a menial task, such as the slider task (Gill and Prowse 2012), because of the more natural applications to real-world contexts where gender gaps are the greatest. Finally, our ideal task was either gender-neutral or stereotyped to suit men.<sup>1</sup>

In order to find a suitable task, we conducted two waves of pilot surveys that tested for gender differences in performance, confidence, and gender perceptions of three potential candidate tasks: Mental Rotation Task (MRT); “find the median” task; and a “pattern” task (also known as MPT or the matrix test).

In particular, in both waves, participants completed three sets of questions: Mental Rotation Task (MRT) questions; “find the median” questions; and “missing shape” questions (see examples below). Within each section, the order of questions was randomly assigned. Each question section lasted 5 minutes. Participants were not able to advance to the next page before the end of the 5 minutes.

After each section, participants were asked to state how many questions they thought they had solved correctly and which gender (female or male) is more likely to perform well on this task. Specifically, we asked:

*“How many questions do you think you solved correctly in this section?”*

*“Do you think men or women are more likely to get a high score in this section? Please select one:           Men                   Women”*

The only difference between the two waves was the payment scheme. In the first wave, we paid participants for their participation: each participant received \$5 regardless of how many questions they solved correctly. In the second wave, we paid participants based on their relative performance: after each section, we randomly paired a participant from the first wave (P1) with a participant in the second wave (P2). If P2 correctly solved more questions than P1, then we paid P2 \$0.2 per correct answer in the MRT and \$8 per correct answer in the pattern task. If P2 solved fewer questions than P1, then we paid P2 \$0.15 per correct answer in the MRT and \$6 per correct answer in the pattern task. We ensured that the average payoffs across the two waves are comparable.

---

<sup>1</sup> Shurchkov (2012) points out that gender gaps in competitiveness are particularly pronounced in tasks which are perceived to favor men, such as the task we use in this study. Because the purpose of this experiment is to capture the effects of noisy feedback on women who find themselves in male-stereotyped environments that produce the largest gender gaps, we chose a task that would most likely model this type of domain. In a separate pilot study, we experimented with a stereotypically female-favoring task, namely, the anagram task. However, we found that in many cases participants were able to use online search engines to achieve perfect scores.

Throughout, we included questions that checked respondents' attention. At the end of the first wave and the second wave, we also asked a few demographic questions (see below).

Based on the first pilot wave, only the MRT produced significant gender differences in gender perception and confidence (Table C1). None of the three tasks yielded a statistically significant difference in actual performance in wave 1. However, in a second pilot survey which used a tournament payment scheme (pilot wave 2) and used a different set of AMT subjects, the pattern test resulted in a significant gender gap in performance, while the MRT did not (Table C2).

Table C1: Summary Statistics for Wave 1

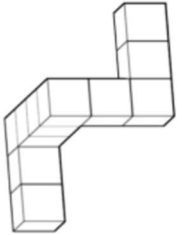
	Average Score	Average Confidence	Share Reporting Gender X is Better
<u>MRT Test</u>			
Male	8.59	9.83	0.65
Female	7.15	7.80	0.35
t-test p-value	(0.1276)	(0.0463)	(0.0306)
<u>Pattern Test</u>			
Male	19.66	13.72	0.41
Female	18.40	13.20	0.59
t-test p-value	(0.3357)	(0.3948)	(0.2017)
<u>Find the Median Test</u>			
Male	17.24	13.52	0.51
Female	19.40	11.45	0.49
t-test p-value	(0.2580)	(0.3398)	(0.8881)
Number of observations	50	50	50

Table C2: Summary Statistics for Wave 2

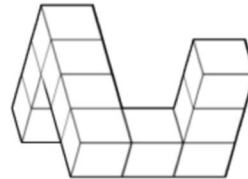
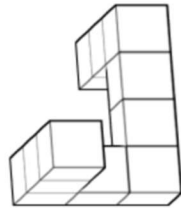
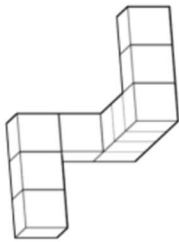
	Average Score	Average Confidence	Share Reporting Gender X is Better
<u>MRT Test</u>			
Male	7.24	7.52	0.57
Female	6.60	5.08	0.43
t-test p-value	(0.3983)	(0.0034)	(0.3319)
Number of observations	51	51	51
<u>Pattern Test</u>			
Male	20.57	16.10	0.44
Female	17.11	11.50	0.56
t-test p-value	(0.0657)	(0.0231)	(0.4016)
Number of observations	50	50	50

## MRT

In each question, you will see a target shape at the top, and three choices. Select the choice that is a **rotated version** of the target one at the top.<sup>2</sup>



Select the shape that is a rotated version of the one above.



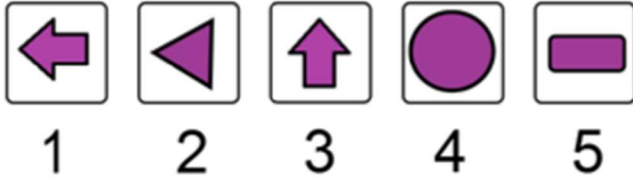
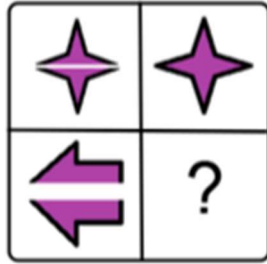
In this example, the answer is the third choice.

## Missing shape puzzle

In each problem, you will see a sequence of shapes. Your job is to fill in the question mark with a shape from the choices.

---

<sup>2</sup> The MRT questions we used in our experiment were slightly different from the original MRT (Vandenberg and Kuse 1978). In the original MRT, there are four choices for each target shape. Exactly two of the choices are correct. Participants get 1 point for each correct choice and lose 1 point for each wrong choice. In order to reduce the difficulty level, we took out one of the correct choices for each target shape, and removed the penalty for incorrect choices.



In this example, the answer is the first choice

**Find the median task [Note that this task was only piloted in the first wave]**

In each problem, you will see a grid that contains 9 numbers. Your job is to identify the median of the 9 numbers.

What is the median? Imagine you lined up the 9 numbers in order, from smallest to largest. The median would be the number in the middle -- the number that is greater than exactly 4 of the numbers and less than exactly 4 of the numbers.

2.34	0.16	3.52
4.58	7.14	2.86
1.64	8.27	6.82

In this example, 3.52 is the median.

**Demographic questions:**

1. **What is your gender?**

- a. Male

- b. Female
- c. Other/Do not wish to disclose

**2. What is your age in years?**

**3. What is the highest level of education you have completed?**

- a. Less than high school
- b. High school or GED
- c. Some college
- d. 2-year college degree (Associates)
- e. 4-year college degree (BA, BS)
- f. Master's degree (MA, MS)
- g. Doctoral degree (PhD)
- h. Professional degree (MD, JD, DDS, etc.)

**4. Are you of Hispanic origin or descent, such as Mexican, Puerto Rican, Cuban, or other Spanish background?**

- a. Yes
- b. No

**5. Which of the following best describes your race?**

- a. White
- b. African-American or Black
- c. Asian
- d. Native Hawaiian or Other Pacific Islander
- e. Native American
- f. Other/Do not wish to disclose

**6. Which of the following best describes your annual household income before taxes?**

- a. Less than \$10,000
- b. \$10,000 - \$19,999
- c. \$20,000 - \$29,999
- d. \$30,000 - \$39,999
- e. \$40,000 - \$49,999
- f. \$50,000 - \$74,999
- g. \$75,000 - \$99,999
- h. \$100,000 - \$149,999
- i. \$150,000 - \$249,999
- j. \$250,000-\$499,999
- k. \$500,000 and over



D. Analysis with Sample Restricted to Waves 1, 2, and 3 Only

Figure D1: Gender Differences in Tournament Entry in the Limited Feedback and Additional Feedback Treatments, on Average and by Information Condition about the Gender of Random Match (Waves 1-3 Equivalent of Figure 3)

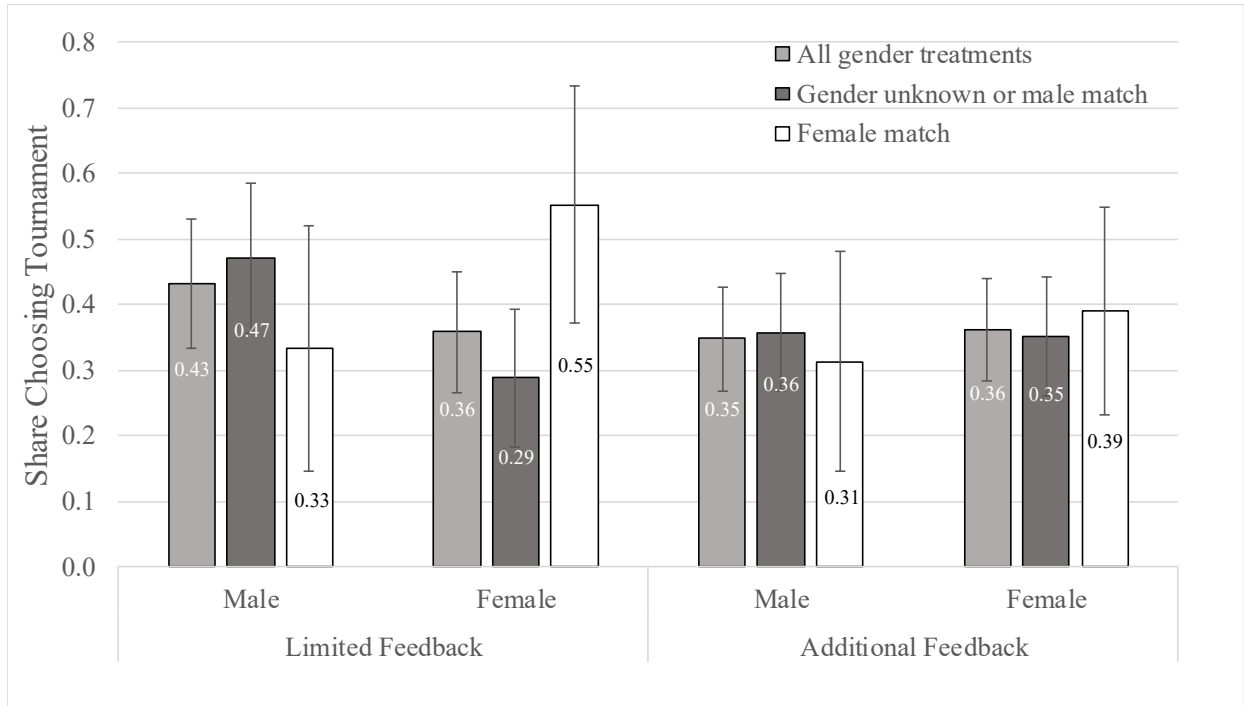


Figure D2: Local Linear Regression of Tournament Entry on Score in Round 1 by Gender and Feedback Treatment Condition (Waves 1-3 Equivalent of Figure 4)

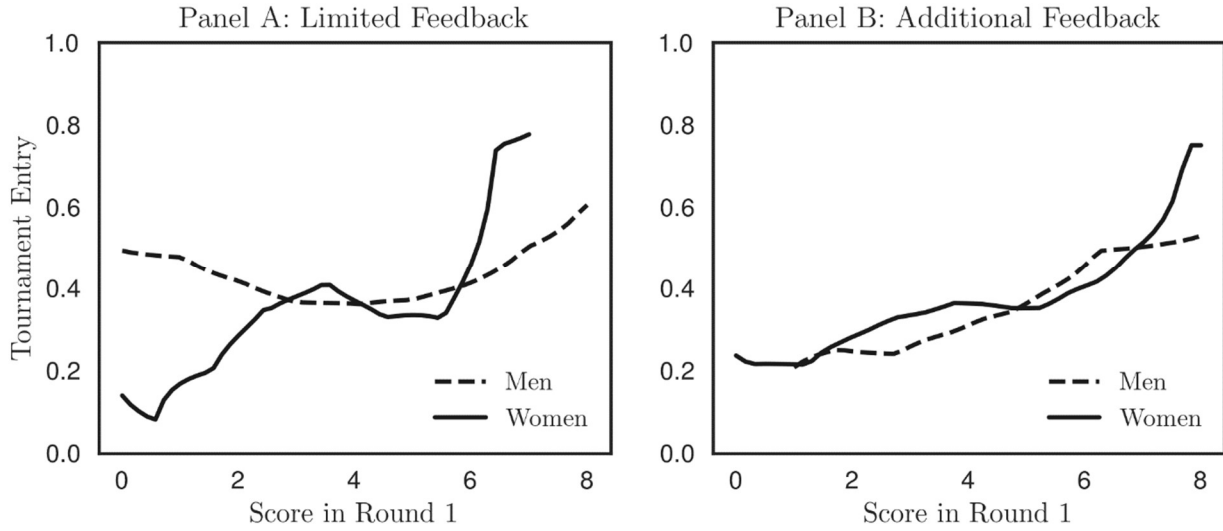


Figure D3: Local Linear Regression of Tournament Entry on Score in Round 1 by Gender and Type of Additional Feedback Received (Waves 1-3 Equivalent of Figure 5)

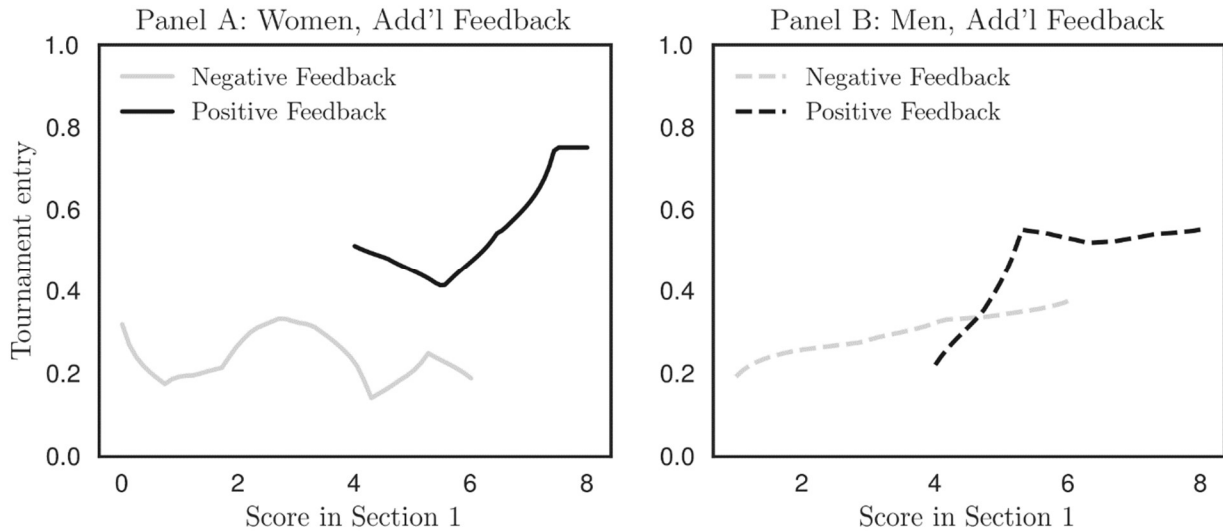


Figure D4: Gender Differences in Performance in Tournament (Round 2) in the Limited Feedback and Additional Feedback Treatments, on Average and by Round 1 Performance (Waves 1-3 Equivalent of Figure 6)

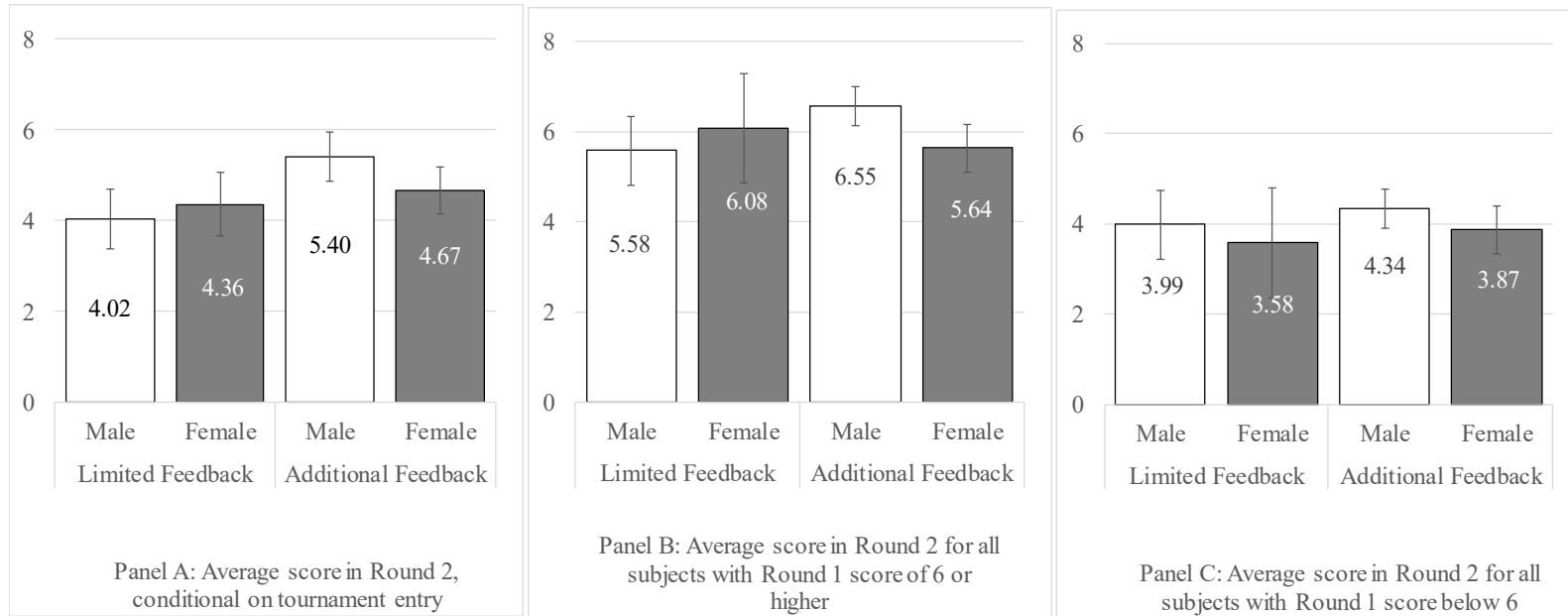


Table D1: Mean Comparisons of Gender Differences in Behavioral Traits  
(Waves 1-3 Equivalent of Table 2)

Variable	Male	Female	Diff
Average score in Round 1	4.30	3.79	0.511***
Average bonus in Round 1	0.83	0.71	0.113***
Score Confidence	4.29	3.53	0.758***
Proportion self-evaluating below average	0.41	0.64	-0.226***
Self-Reported Risk Preference	5.85	4.56	1.287***
Number of obs.	238	253	

Notes: Significance levels \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Summary statistics are based on data from no feedback and forced feedback conditions. Risk elicitation occurred as part of the post-experiment questionnaire and therefore took place post-treatment.

Table D2: Average Attribution to Own Ability by Gender and Type of Feedback Outcome  
(Waves 1-3 Equivalent of Table 3)

Type of Feedback	Male	Female	Difference
Negative surprise (Self-Evaluation of Payment Above Average, but Payment Actually Below)	55.4	78.1	-22.76***
Positive reinforcement (Self-Evaluation of Payment Above Average and Payment Actually Above)	77.7	64.1	13.65**
Positive surprise (Self-Evaluation of Payment Below Average but Payment Actually Above)	45.3	58.1	-12.83
Negative reinforcement (Self-Evaluation of Payment Below Average and Payment Actually Below)	74.1	70.8	3.31

Notes: Summary statistics are based on data from waves 1 and 2 of the experiment, additional feedback condition only (wave 3 did not contain the attribution question). Significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table D3: Determinants of Tournament Entry Decision by Treatment (Waves 1-3 Equivalent of Table 4)

Samples	Limited Feedback			Additional Feedback		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.0646 (0.0700)	-0.254 (0.163)	-0.231 (0.160)	0.0196 (0.0570)	0.118 (0.135)	0.195 (0.132)
Male x Score in Round 1		0.0118 (0.0278)	-0.00500 (0.0274)		0.0700*** (0.0218)	0.0789*** (0.0217)
Female x Score in Round 1		0.0669** (0.0263)	0.0558** (0.0271)		0.0527** (0.0221)	0.0568** (0.0230)
Dependent variable mean	0.392	0.392	0.392	0.354	0.354	0.355
Observations	204	204	204	288	288	287
R-squared	0.0160	0.0445	0.0685	0.0126	0.0655	0.134

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of the match. Columns 3 and 6 control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table D4: Gender Differences in the Effect of Receiving Negative Feedback on Attribution and Tournament Entry (Waves 1-3 Equivalent of Table 5)

Sample	All		Negative Self-Evaluation		Positive Self-Evaluation	
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Dep Var: Tournament Entry</b>						
Female	0.106 (0.0924)	0.127 (0.0879)	0.244 (0.153)	0.227 (0.156)	0.0616 (0.120)	0.0944 (0.111)
Male x Negative Feedback	-0.148* (0.0844)	-0.0550 (0.104)	-0.0680 (0.127)	-0.0597 (0.152)	-0.107 (0.125)	-0.00868 (0.151)
Female x Negative Feedback	-0.283*** (0.0884)	-0.106 (0.115)	-0.236* (0.128)	-0.158 (0.164)	-0.332** (0.137)	-0.147 (0.167)
Dependent variable mean	0.354	0.355	0.295	0.297	0.417	0.417
F-test of equality of interactions (p-value)	0.243	0.656	0.340	0.590	0.188	0.403
F-test of female = female X neg fbk (p)	0.0155	0.180	0.0566	0.165	0.0797	0.309
Observations	288	287	149	148	139	139
R-squared	0.0545	0.135	0.0747	0.132	0.0476	0.155
<b>Panel B: Dep Var: Attribution</b>						
Female	-10.62** (5.381)	-10.53* (5.555)	11.91 (10.10)	9.800 (10.39)	-14.67** (6.006)	-14.39** (6.375)
Male x Negative Feedback	-7.656 (5.117)	-6.606 (7.639)	27.54*** (9.864)	21.83* (11.81)	-24.41*** (5.758)	-22.91** (9.078)
Female x Negative Feedback	11.08** (5.440)	12.30* (6.917)	11.32 (8.368)	5.514 (11.22)	12.99* (6.984)	14.39 (8.987)
Dependent variable mean	67.64	67.56	67.60	67.43	67.69	67.69
F-test of equality of interactions (p-value)	p < 0.01	p < 0.01	0.170	0.177	p < 0.01	p < 0.01
F-test of female = female X neg fbk (p)	0.0241	0.0216	0.970	0.801	0.0209	0.0315
Observations	190	189	97	96	93	93
R-squared	0.0571	0.0653	0.131	0.143	0.221	0.222

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants in the additional feedback treatment. Panel B omits wave 3 since it did not contain the attribution question. Even numbered columns control for score, risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table D5: Gender Differences in Sorting into Tournament Entry in Response to Negative Feedback (Waves 1-3 Equivalent of Table 6)

Dep Var: Samples	Tournament Entry				Attribution	
	Limited Feedback		Additional Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)	(3)	(4)
Female	-0.397*	-0.392*	0.210	0.270	34.42**	33.45*
	(0.231)	(0.224)	(0.140)	(0.166)	(15.47)	(17.30)
Male x Score Bin 2	-0.0186	-0.00703	0.339***	0.373***	35.70***	34.46***
	(0.239)	(0.234)	(0.0940)	(0.131)	(11.77)	(12.22)
Male x Score Bin 3	-0.161	-0.163	0.369**	0.405**	35.03***	34.77***
	(0.248)	(0.242)	(0.142)	(0.166)	(12.00)	(12.46)
Male x Score Bin 4	0.0576	0.0551	0.358	0.455*	0	0
	(0.295)	(0.288)	(0.238)	(0.261)	(.)	(.)
Female x Score Bin 2	0.265**	0.300***	0.160	0.198	7.875	7.587
	(0.112)	(0.110)	(0.151)	(0.141)	(11.88)	(12.23)
Female x Score Bin 3	0.229*	0.278**	-0.00940	0.0229	2.957	2.979
	(0.123)	(0.125)	(0.181)	(0.169)	(12.64)	(13.05)
Female x Score Bin 4	0.964***	0.977***	-0.0332	0.00305	5.251	5.352
	(0.0986)	(0.0990)	(0.206)	(0.187)	(14.21)	(14.45)
Dependent variable mean	0.355	0.355	0.272	0.274	68.56	68.43
F-test of male interactions (p-value)	0.696	0.660	p < 0.01	0.0355	0.0100	0.0170
F-test of female interactions (p-value)	p < 0.01	p < 0.01	0.395	0.293	0.874	0.895
F-test of equality of interactions (p-value)	0.0136	p < 0.01	0.244	0.237	0.0580	0.143
Observations	141	141	169	168	121	120
R-squared	0.109	0.131	0.0463	0.108	0.158	0.161

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants who received negative feedback in the additional feedback treatment and participants who would have received negative feedback in the limited feedback treatment. Columns 5 and 6 omit wave 3 since it did not contain the attribution question. Even numbered columns control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



E. Analysis with Sample Restricted to Waves 1 and 4 Only

Figure E1: Gender Differences in Tournament Entry in the Limited Feedback and Additional Feedback Treatments, on Average and by Information Condition about the Gender of Random Match (Waves 1 and 4 Equivalent of Figure 3)

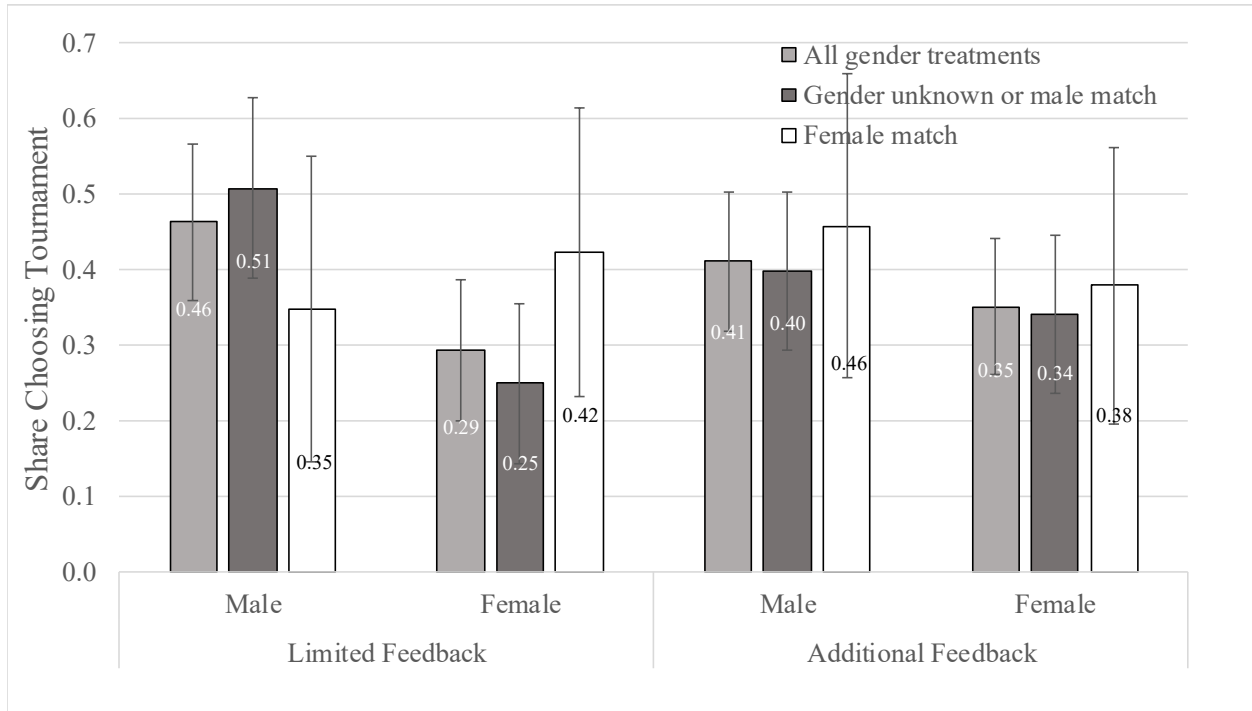


Figure E2: Local Linear Regression of Tournament Entry on Score in Round 1 by Gender and Feedback Treatment Condition (Waves 1 and 4 Equivalent of Figure 4)

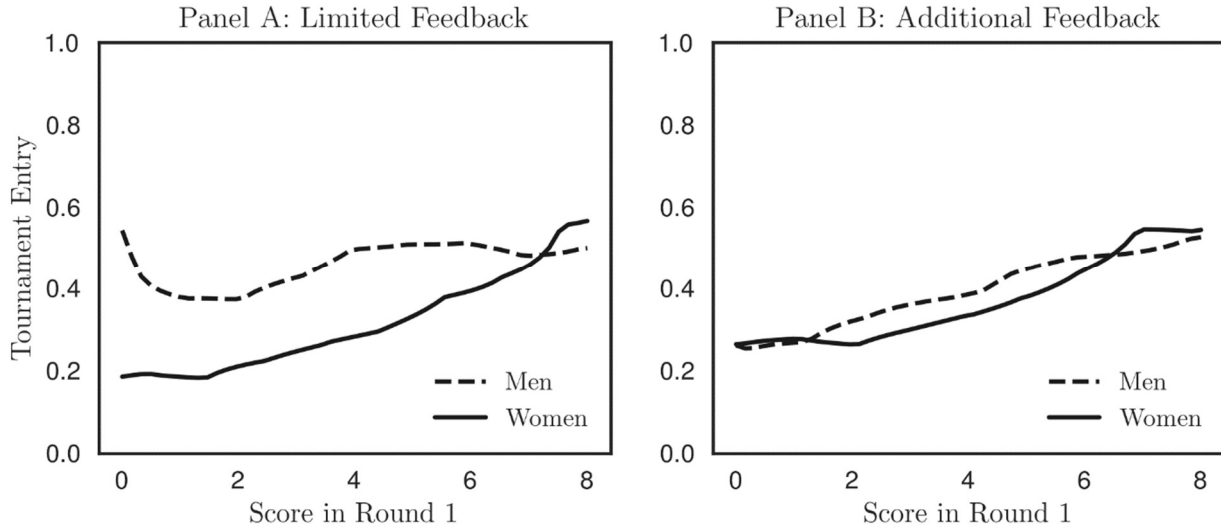


Figure E3: Local Linear Regression of Tournament Entry on Score in Round 1 by Gender and Type of Additional Feedback Received (Waves 1 and 4 Equivalent of Figure 5)

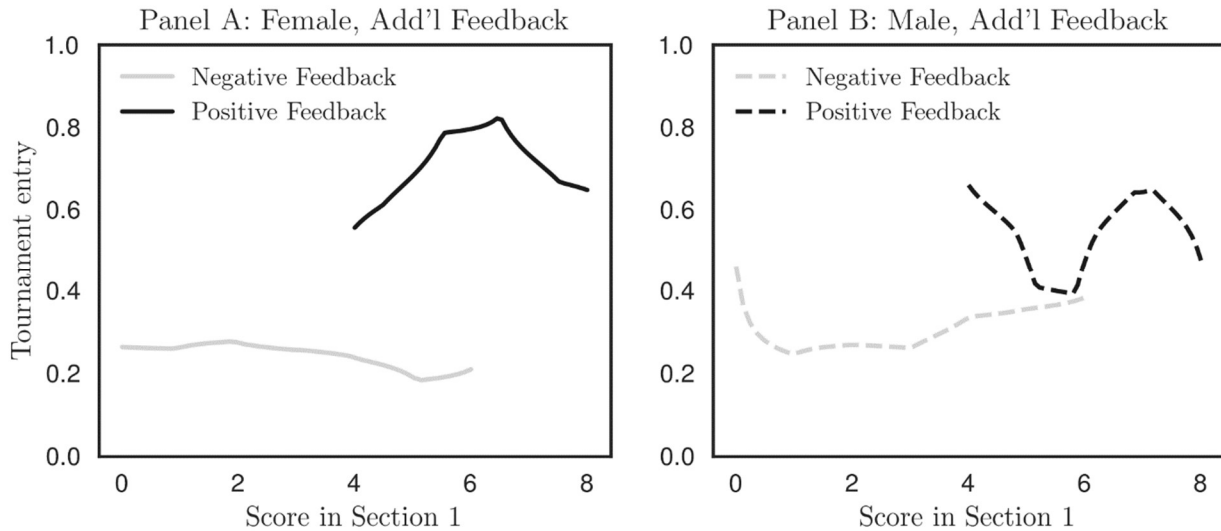


Figure E4: Gender Differences in Performance in Tournament (Round 2) in the Limited Feedback and Additional Feedback Treatments, on Average and by Round 1 Performance (Waves 1 and 4 Equivalent of Figure 6)

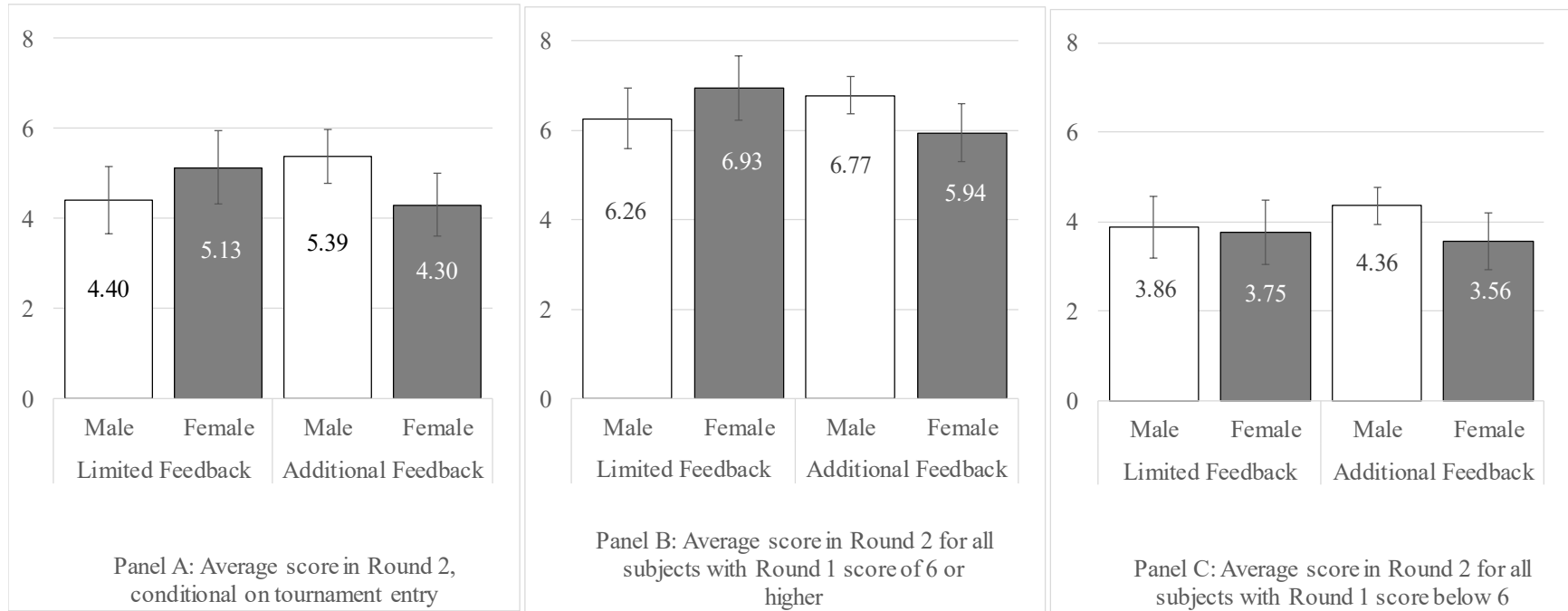


Table E1: Mean Comparisons of Gender Differences in Behavioral Traits  
(Waves 1 & 4 Equivalent of Table 2)

Variable	Male	Female	Diff
Average score in Round 1	4.36	3.73	0.628***
Average bonus in Round 1	0.84	0.70	0.142***
Score Confidence	4.39	3.61	0.785***
Proportion self-evaluating below average	0.51	0.66	-0.149***
Self-Reported Risk Preference	5.45	4.43	1.019***
Number of obs.	187	216	

Notes: Significance levels \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Summary statistics are based on data from no feedback and forced feedback conditions. Risk elicitation occurred as part of the post-experiment questionnaire and therefore took place post-treatment.

Table E2: Average Attribution to Own Ability by Gender and Type of Feedback Outcome (Waves 1 & 4 Equivalent of Table 3)

Type of Feedback	Male	Female	Difference
Negative surprise (Self-Evaluation of Payment Above Average, but Payment Actually Below)	54.3	72.0	-17.73***
Positive reinforcement (Self-Evaluation of Payment Above Average and Payment Actually Above)	73.0	66.5	6.44
Positive surprise (Self-Evaluation of Payment Below Average but Payment Actually Above)	66.2	56.8	9.43
Negative reinforcement (Self-Evaluation of Payment Below Average and Payment Actually Below)	71.3	72.3	-1.01

Notes: Summary statistics are based on data from waves 1 and 4 of the experiment, additional feedback condition only. Significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table E3: Determinants of Tournament Entry Decision by Treatment (Waves 1 & 4 Equivalent of Table 4)

Samples	Limited Feedback			Additional Feedback		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.167** (0.0725)	-0.323* (0.175)	-0.281 (0.171)	-0.0525 (0.0657)	-0.0676 (0.155)	-0.0739 (0.154)
Male x Score in Round 1		0.0184 (0.0294)	0.0107 (0.0302)		0.0456* (0.0245)	0.0247 (0.0247)
Female x Score in Round 1		0.0607** (0.0286)	0.0570** (0.0272)		0.0602** (0.0253)	0.0568** (0.0265)
Dependent variable mean	0.370	0.370	0.370	0.380	0.380	0.384
Observations	184	184	184	221	221	219
R-squared	0.0312	0.0576	0.128	0.0249	0.0671	0.162

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of the match. Columns 3 and 6 control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table E4: Gender Differences in the Effect of Receiving Negative Feedback on Attribution and Tournament Entry (Waves 1 & 4 Equivalent of Table 5)

Sample	All		Negative Self-Evaluation		Positive Self-Evaluation	
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Dep Var: Tournament Entry</b>						
Female	0.0978 (0.117)	0.169 (0.111)	0.0474 (0.194)	0.0609 (0.170)	0.136 (0.147)	0.242 (0.154)
Male x Negative Feedback	-0.241** (0.102)	-0.173 (0.124)	-0.267* (0.150)	-0.215 (0.181)	-0.161 (0.158)	-0.126 (0.191)
Female x Negative Feedback	-0.399*** (0.105)	-0.347*** (0.131)	-0.284* (0.170)	-0.229 (0.186)	-0.477*** (0.147)	-0.440** (0.193)
Dependent variable mean	0.380	0.384	0.325	0.331	0.449	0.449
F-test of equality of interactions (p-value)	0.261	0.191	0.938	0.942	0.118	0.106
F-test of female = female X neg fbk (p)	0.0145	0.0133	0.322	0.346	0.0200	0.0220
Observations	221	219	123	121	98	98
R-squared	0.105	0.188	0.0770	0.152	0.135	0.230
<b>Panel B: Dep Var: Attribution</b>						
Female	-8.592 (5.792)	-7.608 (5.903)	-9.971 (10.61)	-9.358 (10.22)	-8.837 (6.571)	-7.419 (7.124)
Male x Negative Feedback	-7.948 (4.860)	-6.881 (7.285)	5.745 (7.516)	10.68 (10.32)	-25.77*** (6.394)	-25.70*** (9.705)
Female x Negative Feedback	9.758* (5.698)	11.12 (6.976)	17.05* (10.04)	22.86** (10.45)	3.563 (7.424)	2.960 (9.485)
Dependent variable mean	68.08	68.00	69.65	69.55	66.10	66.10
F-test of equality of interactions (p-value)	0.0158	0.0153	0.347	0.293	p < 0.01	p < 0.01
F-test of female = female X neg fbk (p)	0.0868	0.0897	0.169	0.0771	0.326	0.460
Observations	221	219	123	121	98	98
R-squared	0.0381	0.0549	0.0371	0.0656	0.172	0.185

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants in the additional feedback treatment. Even numbered columns control for score, risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table E5: Gender Differences in Sorting into Tournament Entry in Response to Negative Feedback (Waves 1 & 4 Equivalent of Table 6)

Dep Var: Samples	Tournament Entry				Attribution	
	Limited Feedback		Additional Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)	(3)	(4)
Female	-0.379 (0.280)	-0.395 (0.283)	-0.313 (0.286)	-0.298 (0.292)	-14.13 (12.83)	-13.06 (10.99)
Male x Score Bin 2	-0.111 (0.276)	-0.0976 (0.283)	-0.286 (0.264)	-0.284 (0.269)	-11.14 (9.385)	-9.442 (8.959)
Male x Score Bin 3	-0.189 (0.285)	-0.194 (0.293)	-0.196 (0.285)	-0.258 (0.287)	-6.649 (9.922)	-3.460 (9.639)
Male x Score Bin 4	-0.0155 (0.323)	-0.0170 (0.326)	-0.231 (0.344)	-0.302 (0.358)	-45.50*** (13.68)	-38.42*** (13.59)
Female x Score Bin 2	0.0756 (0.133)	0.123 (0.123)	0.0876 (0.158)	0.0759 (0.167)	14.15 (10.35)	15.02 (9.353)
Female x Score Bin 3	0.155 (0.139)	0.215 (0.132)	-0.0269 (0.179)	-0.0307 (0.184)	11.50 (11.23)	12.56 (10.36)
Female x Score Bin 4	0.890*** (0.117)	0.924*** (0.115)	-0.0831 (0.206)	-0.0840 (0.200)	11.17 (14.15)	12.83 (13.32)
Dependent variable mean	0.302	0.302	0.272	0.276	68.47	68.37
F-test of male interactions (p-value)	0.835	0.820	0.701	0.771	p < 0.01	0.0154
F-test of female interactions (p-value)	p < 0.01	p < 0.01	0.694	0.720	0.596	0.463
F-test of equality of interactions (p-value)	p < 0.01	p < 0.01	0.484	0.660	0.0287	0.0293
Observations	129	129	147	145	147	145
R-squared	0.0883	0.106	0.0362	0.0733	0.132	0.153

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants who received negative feedback in the additional feedback treatment and participants who would have received negative feedback in the limited feedback treatment. Even numbered columns control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



## F. Additional Analysis

In this appendix, we perform additional analysis to support the claims made in the main paper. The first section presents additional summary statistics. The second section demonstrates that the tournament entry patterns in wave 3 are similar to those in the other waves, suggesting that posing the attribution question did not affect tournament entry. The third section addresses the concern that our design allows savvy subjects to back out their score from payment information. The remaining sections present estimates from linear regressions that include wave fixed effects and treatment indicators for gender of opponent (male/female in KG or unknown in UG), and in some cases include non-linear specifications.

F1: Additional Summary Statistics

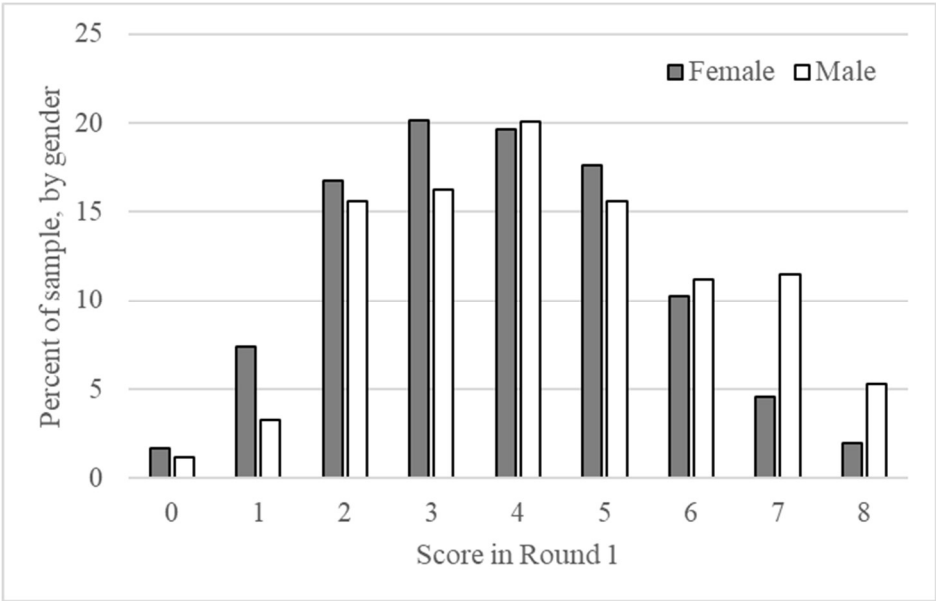
Table F1.1: Summary of Demographic Characteristics and Balance Tests

Variables	Wave 1			Wave 2	Wave 3			Wave 4		
	Limited Feedback	Additional Feedback	LF vs AF p-value	Additional Feedback	Limited Feedback	Additional Feedback	LF vs. AF p-value	Limited Feedback	Additional Feedback	LF vs. AF p-value
age	42.23	40.13	0.15	38.2	36.85	36.06	0.64	38.84	39.40	0.69
female	0.59	0.55	0.59	0.5	0.48	0.46	0.77	0.51	0.49	0.73
education										
Less than high school	0.00%	0.00%		0.00%	1.00%	0.00%	0.32	0.00%	0.00%	
High school or GED	13.46%	7.92%	0.21	21.59%	8.00%	8.16%	0.97	10.00%	13.56%	0.45
Some College	25.96%	15.84%	0.08	26.14%	24.00%	22.45%	0.80	23.75%	29.66%	0.36
2-year college degree	9.62%	14.85%	0.25	12.50%	7.00%	18.37%	0.02	10.00%	11.86%	0.68
4-year college degree	35.58%	46.53%	0.11	29.55%	47.00%	40.82%	0.38	45.00%	34.75%	0.15
Master's degree	10.58%	11.88%	0.74	7.95%	10.00%	7.14%	0.48	10.00%	7.63%	0.56
Professional degree	3.85%	2.97%	0.68	1.14%	3.00%	1.02%	0.32	1.25%	1.69%	0.80
Doctoral degree	0.96%	-1.2E-17	0.40	1.14%	0.00%	2.04%	0.15	0.00%	0.85%	0.41
income										
Less than \$10,000	3.85%	4.95%	0.74	4.55%	6.00%	9.18%	0.40	2.50%	5.08%	0.37
\$10,000 - \$19,999	11.54%	7.92%	0.36	6.82%	8.00%	10.20%	0.59	8.75%	5.08%	0.31
\$20,000 - \$29,999	15.38%	7.92%	0.11	18.18%	10.00%	12.24%	0.62	6.25%	13.56%	0.10
\$30,000 - \$39,999	6.73%	13.86%	0.11	15.91%	13.00%	16.33%	0.51	18.75%	11.86%	0.18
\$40,000 - \$49,999	16.35%	7.92%	0.07	11.36%	10.00%	8.16%	0.66	5.00%	11.86%	0.10
\$50,000 - \$74,999	22.12%	18.81%	0.56	21.59%	27.00%	23.47%	0.57	35.00%	31.36%	0.59
\$75,000 - \$99,999	13.46%	10.89%	0.57	9.09%	17.00%	10.20%	0.17	13.75%	12.71%	0.83
\$100,000 - \$149,999	6.73%	21.78%	0.00	7.95%	7.00%	8.16%	0.76	8.75%	5.08%	0.31
\$150,000 - \$249,999	3.85%	5.94%	0.40	4.55%	2.00%	2.04%	0.98	1.25%	3.39%	0.35
\$250,000 - \$499,999	0.00%	0.00%	1.00	0.00%	0.00%	0.00%		0.00%	0.00%	
race										
Asian	6.73%	11.88%	0.14	6.82%	7.00%	10.20%	0.42	7.50%	8.47%	0.81
Black or African American	4.81%	7.92%	0.35	3.41%	8.00%	7.14%	0.82	5.00%	8.47%	0.35
Native American	0.96%	0.00%	0.23	0.00%	2.00%	2.04%	0.98	1.25%	0.85%	0.78
White	84.62%	78.22%	0.20	89.77%	83.00%	77.55%	0.34	83.75%	81.36%	0.67
Other/ Do not wish to disclose	2.88%	1.98%	0.67	0.00%	0.00%	3.06%	0.08	2.50%	0.85%	0.35

Table F1.2: Demographic Characteristics Balance by Gender

Variables	Male	Female	p-value
age	37.30	40.75	0.00
education			
Less than high school	0.00%	0.25%	0.33
High school or GED	11.69%	11.91%	0.92
Some College	20.52%	27.05%	0.03
2-year college degree	12.47%	11.66%	0.73
4-year college degree	43.12%	38.71%	0.21
Master's degree	9.09%	8.19%	0.65
Professional degree	2.08%	1.74%	0.73
Doctoral degree	1.04%	0.50%	0.38
income			
Less than \$10,000	5.71%	5.96%	0.89
\$10,000 - \$19,999	9.61%	6.45%	0.10
\$20,000 - \$29,999	13.51%	11.66%	0.44
\$30,000 - \$39,999	14.55%	12.90%	0.50
\$40,000 - \$49,999	11.69%	9.68%	0.36
\$50,000 - \$74,999	21.82%	28.04%	0.04
\$75,000 - \$99,999	10.65%	13.65%	0.20
\$100,000 - \$149,999	8.83%	9.18%	0.86
\$150,000 - \$249,999	3.64%	2.23%	0.24
\$250,000 - \$499,999	0.00%	0.25%	0.33
race			
Asian	9.35%	5.96%	0.07
Black or African American	5.71%	6.95%	0.48
Native American	1.04%	0.74%	0.66
White	82.86%	84.12%	0.63
Other/ Do not wish to disclose	1.04%	2.23%	0.19

Figure F1.1: Distribution of Score in Round 1 by Gender



## F2: Tournament Entry Patterns from Wave 3 (No Attribution Question)

We excluded the attribution question from wave 3 in light of the concern that posing the question might prime tournament entry and bias our results. We find no evidence of differences in tournament entry patterns in wave 3 as compared to the other waves, but the small sample size in wave 3 and the fact that most of the negative surprises in our study come from the first wave preclude us from saying anything more conclusive about whether posing the question primes tournament entry. Note however that all our main regressions include wave fixed effects.

Figure F2.1 replicates Figure 3 using only data from wave 3. The gender gap is not statistically significant in either condition, but facing a female opponent increases selection into tournament for women ( $p = 0.096$ ) and decreases selection into tournament for men (but not significantly), as before. Also as before, the largest gender gap exists when the gender of the opponent is known to be male. This gender gap (when facing a male opponent) is eliminated in the additional feedback condition, but interestingly, a reverse (but not statistically significant) gender gap emerges when participants face an opponent known to be female or of unknown gender.

Table F2.1 replicates Table 4 using only data from wave 3. The sign of the gender gap in the limited and additional feedback conditions are the same as in Table 4 but none of the coefficients are statistically significant. We find similarly signed, but not statistically significant, point estimates suggesting that women sort correctly into tournament in both conditions but men sort correctly only in the additional feedback condition.

Figure F2.2 demonstrates that the patterns we saw in Figure 5 are replicated in the wave 3 data, although the small sample size adds noise. The small sample size in wave 3 precludes us from replicating the remaining results since they require splitting the sample further (for example, by positive or negative feedback).

Table F2.1: Determinants of Tournament Entry Decision by Treatment (Wave 3 Equivalent of Table 4)

Samples	Limited Feedback			Additional Feedback		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.0206 (0.101)	-0.118 (0.235)	-0.0851 (0.236)	0.0777 (0.100)	-0.0500 (0.217)	0.0399 (0.218)
Male x Score in Round 1		0.00222 (0.0376)	-0.0191 (0.0381)		0.0271 (0.0353)	0.0538 (0.0354)
Female x Score in Round 1		0.0302 (0.0431)	0.0135 (0.0456)		0.0646* (0.0380)	0.0764 (0.0470)
Dependent variable mean	0.430	0.430	0.430	0.357	0.357	0.357
Observations	100	100	100	98	98	98
R-squared	0.0134	0.0186	0.0414	0.00744	0.0376	0.102

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of the match. Columns 3 and 6 control for risk and confidence. Significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Figure F2.1: Gender Differences in Tournament Entry in the Limited Feedback and Additional Feedback Treatments, on Average and by Information Condition about the Gender of Random Match (Wave 3 Equivalent of Figure 3)

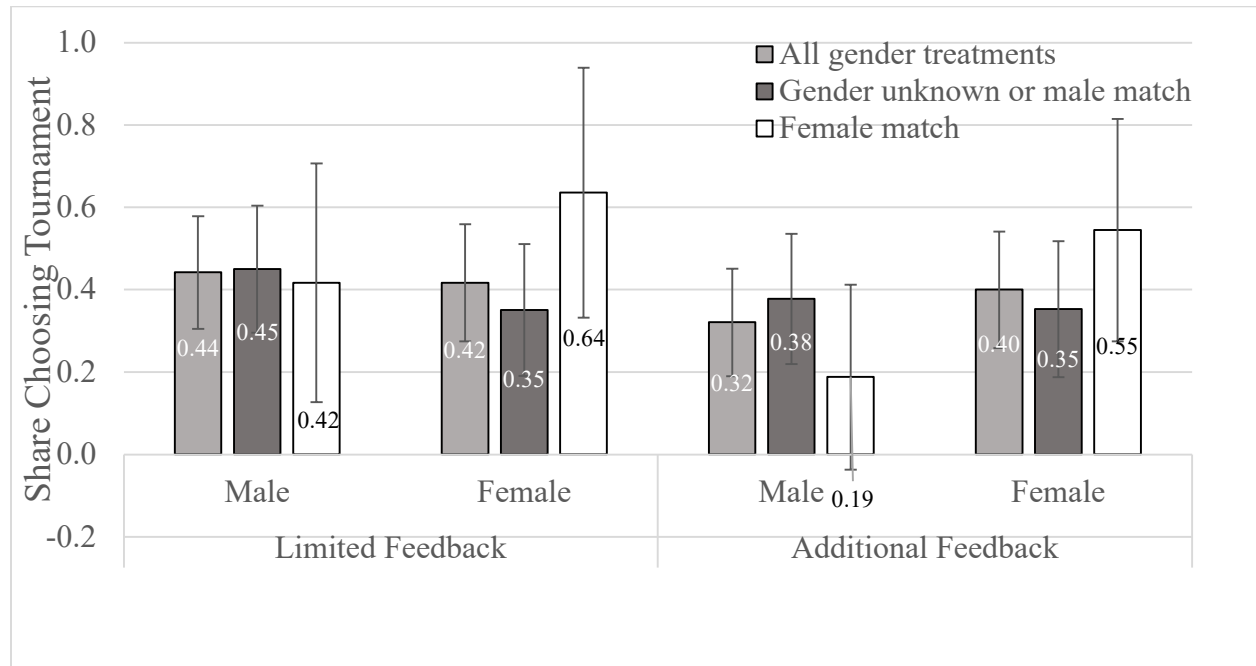
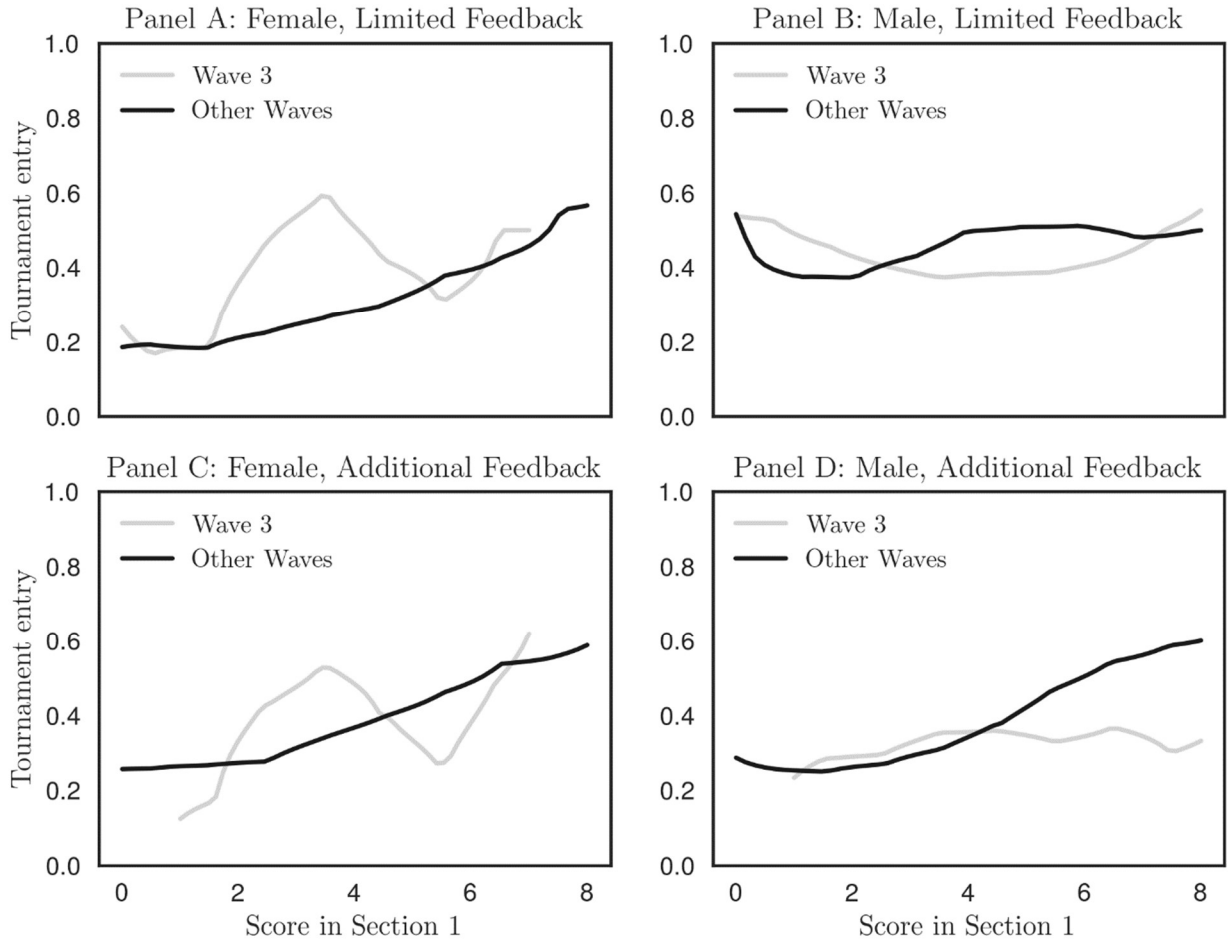


Figure F2.2: Comparison of Local Linear Regression of Tournament Entry on Score in Round 1 in Wave 3 and All Other Waves, by Gender and Feedback Treatment



### F3: Discussion of the Information Embedded in the Piece-rates Used to Compute Payment in Round 1

If the majority of subjects were using the piece-rates to deduce their score from whether their payment was divisible by 0.20 or by 0.15, we would expect to observe strategic behaviors depending on their bonus payment. This could explain some of our results if, for example, men are better at interpreting the signal from the payment – the additional feedback treatment would then be more informative for women than for men. Table F3.1 shows that neither men nor women respond to whether or not their payment was divisible by 0.15. Using only data from the limited feedback treatment, when the only feedback is the payment, we regress tournament entry (Columns 1 and 2) and whether participants perceived their payments to be below average (Columns 3 and 4) on their bonus payment or score, a female indicator variable, and whether or not their payment was divisible by 0.15 for both men and women separately. None of the coefficients are statistically significant. If anything, women are less likely to enter tournament when their payment is divisible by 0.15 and men are more likely.

In addition, Figures F3.1 and F3.2 present the fraction of subjects who entered the tournament and the fraction with a negative self-evaluation for each possible payment from the first round for those in the limited feedback treatment. Consider those who earned \$0.30 (got 2 questions right and lost the match-up) and those who earned \$0.40 (got 2 questions right, but won the match-up). Those who lost the tournament are slightly less likely to consider their payment below average and while they are also less likely to enter the tournament, the difference is very small. Similarly, consider those who earned \$0.75 and \$1, all of whom got 5 questions right, but some won the match-up and others lost it. Tournament entry and self-evaluation actually go in the opposite direction – those who lost the match-up are more likely to have a positive self-evaluation and enter the tournament. In results not shown, we break these responses by gender. The samples get very small, but the results suggest that, if anything, men are less likely to respond as if they lost the tournament when their payment is divisible by 0.15.

Attribution tells a similar story, but the sample size gets very small since we have to separate the sample by their self-evaluation (the attribution question wording depends on whether they perceived their payment to be above or below average). Furthermore, looking at the participants in the additional feedback group, participants respond more strongly to the type of feedback than



to whether or not their payment was divisible by \$0.15, but the numbers in each cell are too small to infer much.

Table F3.1: Participant Responses to Divisibility of Payment by 0.15

Dep Var:	Tournament Entry		Negative Self-Evaluation	
	(1)	(2)	(3)	(4)
Bonus from Round 1	0.148 (0.0924)		-0.323*** (0.0846)	
Score in Round 1		0.0321* (0.0189)		-0.0726*** (0.0169)
Female	-0.0351 (0.0895)	-0.0340 (0.0895)	0.158* (0.0887)	0.155* (0.0891)
Male X Divisible by 15	0.105 (0.0963)	0.0977 (0.0937)	0.0237 (0.0933)	0.0339 (0.0900)
Female X Divisible by 15	-0.0179 (0.0902)	-0.0266 (0.0871)	-0.0840 (0.0898)	-0.0693 (0.0867)
Observations	284	284	284	284
R-squared	0.0350	0.0363	0.128	0.138

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to the limited feedback group. Significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Figure F3.1: Fraction of Participants Entering the Round 2 Tournament for Each Possible Bonus Payment in Round 1

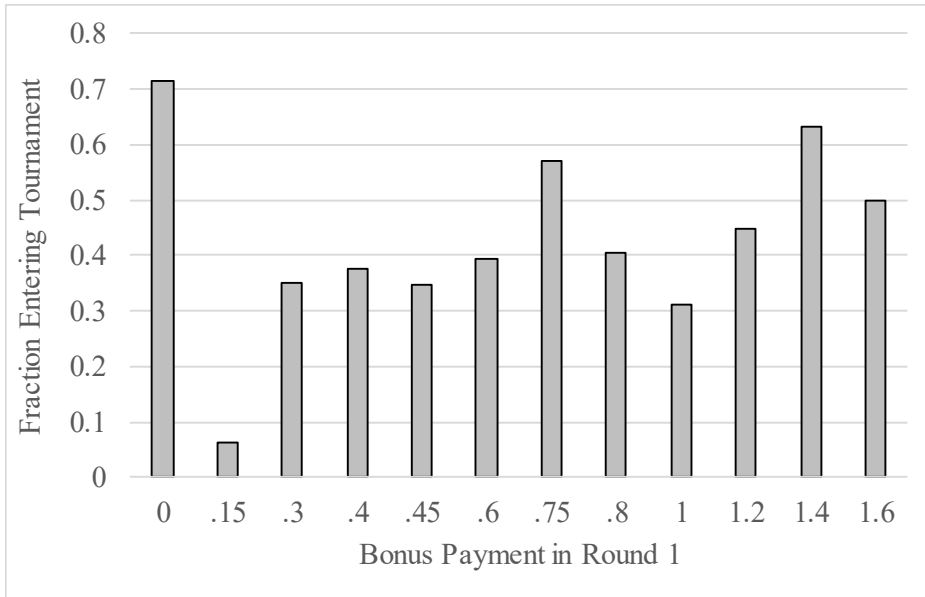
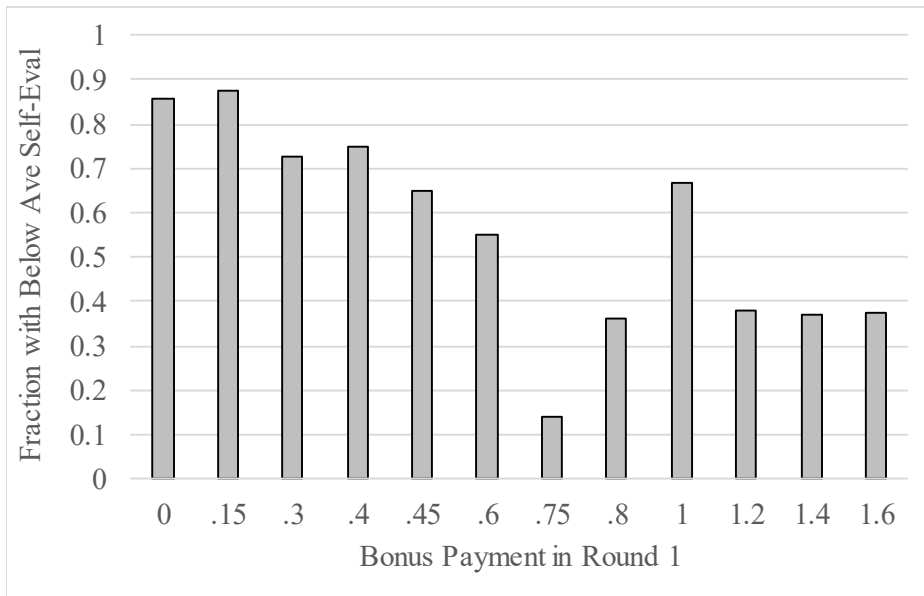


Figure F3.2: Fraction of Participants with a Below-Average Self-Evaluation for Each Possible Bonus Payment in Round 1



#### F4: Confirming Gender Differences in Behavioral Traits

Table F4.1 provides OLS estimates with wave fixed effects and controls for gender of the match treatments to confirm that women are, on average, significantly less confident than men both in terms of score (Column 1) and in terms of self-evaluation of payment relative to the average (Column 3). Columns 2 and 4 show that the gender gap in both confidence measures decreases somewhat but remains significant if we control for performance in round 1. The finding is robust to including demographic controls (see Appendix Table G1).

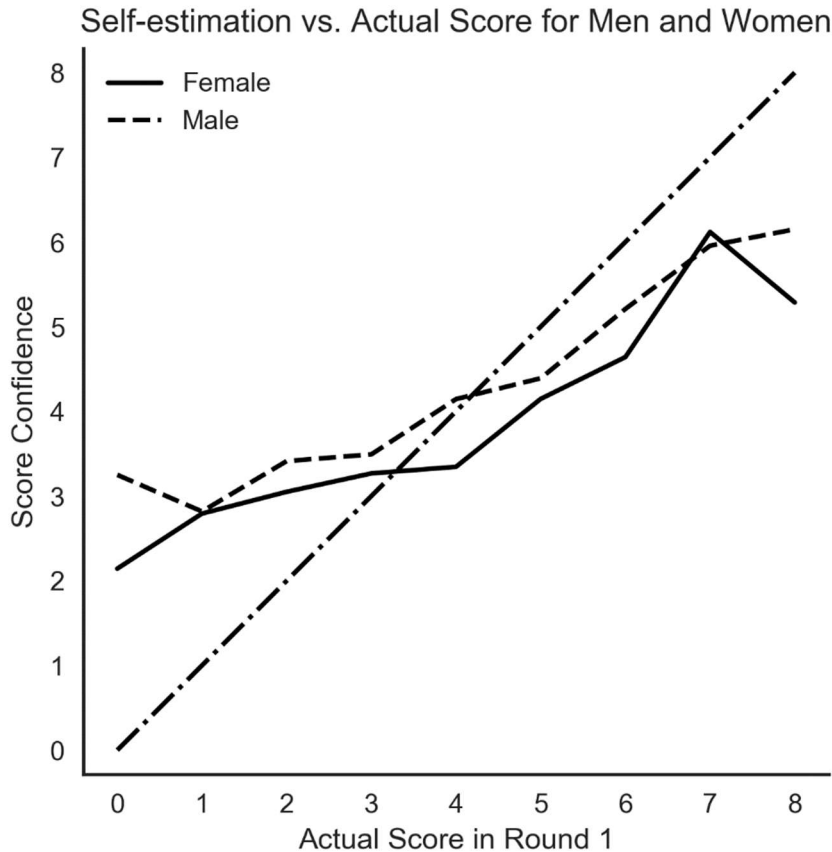
A simple comparison of actual and predicted score for men and women reveals that women are directionally under-confident in our experiment (mean confidence of 3.61 relative to actual average score of 3.77 but the t-test p-value is 0.130). Men on the other hand correctly predict their scores, on average. Figure F4.1 plots the relationship between participants' expected and actual scores for the entire distribution of ability. We make three observations: 1) Men are systematically more confident about their score ( $p < 0.01$ ) than women, conditional on getting the same score. 2) Men with median performance (solved 4 out of 8 questions) on average correctly estimate their scores. Women at the median, on the other hand, underestimate their score (p-value of 0.0004). 3) Participants of both genders with higher-than-median performance tend to underestimate their score. Participants of both genders with lower-than-median performance tend to overestimate their score.

Table F4.1: OLS Estimates of Gender Gaps in Confidence and Self-Evaluation

Dependent variable	Score Confidence (Self-Reported Score in Round 1)		Self-evaluation of Payment to be Below Average	
	(1)	(2)	(3)	(4)
Female	-0.714*** (0.149)	-0.447*** (0.133)	0.168*** (0.0375)	0.116*** (0.0360)
Score in Round 1		0.484*** (0.0368)		-0.0946*** (0.00844)
Dependent variable mean	3.960	3.960	0.548	0.548
Observations	691	691	691	691
R-squared	0.0367	0.238	0.0408	0.161

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. In Columns 1 and 2, score-confidence is measured on a scale of 0 to 10 (the highest possible score in Round 1 is 8). Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Figure F4.1: Comparison of Actual and Predicted Score in Round 1 by Gender



## F5: Effect of Feedback on Tournament Entry Along the Entire Ability Distribution

Table F5.1 breaks up the analysis in Table 4 by ability bins (with the lowest bin, those scoring 0 or 1 out of 8 as the omitted category, the second bin including those scoring below average, either 2 or 3 out of 8, the third bin including those who scored approximately average, either 4 or 5 out of 8, and the highest bin including those score 6 or above out of 8).

Table F5.1: Determinants of Tournament Entry Decision by Treatment, Nonlinear Specifications

Samples	Limited Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)
Female	-0.360*	-0.340*	-0.0953	-0.0383
	(0.202)	(0.196)	(0.196)	(0.210)
Male x Score Bin 2	-0.0630	-0.0599	0.00740	0.000290
	(0.198)	(0.195)	(0.178)	(0.188)
Male x Score Bin 3	-0.0860	-0.138	0.123	0.120
	(0.194)	(0.190)	(0.178)	(0.188)
Male x Score Bin 4	0.0123	-0.0523	0.248	0.214
	(0.201)	(0.198)	(0.181)	(0.196)
Female x Score Bin 2	0.151	0.173	0.154	0.159
	(0.109)	(0.106)	(0.115)	(0.123)
Female x Score Bin 3	0.218**	0.217**	0.256**	0.227*
	(0.109)	(0.106)	(0.117)	(0.126)
Female x Score Bin 4	0.386***	0.372***	0.377***	0.349**
	(0.145)	(0.137)	(0.130)	(0.142)
Dependent variable mean	0.391	0.391	0.381	0.383
F-test of male interactions (p-value)	0.835	0.787	0.0436	0.103
F-test of female interactions (p-value)	0.0520	0.0503	0.0177	0.0769
Observations	284	284	407	405
R-squared	0.0458	0.0903	0.0574	0.122

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. Even numbered columns control for risk and confidence. Significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Once again, we observe that, in the limited feedback condition, women with the lowest scores are marginally less likely to enter the tournament than similar men (Column 1, row 1). The interactions reveal that selection into tournament increases with ability for women. Relative to the

bottom group, women in the second bin are 15 percentage points more likely to choose tournament, while women in the two highest groups are 22 and 39 percentage points more likely to compete, respectively. These interactions are jointly significant, as seen in the p-values at the bottom of the table (p-values around 0.05). As we saw in Figure 4, there appears to be no relationship between score and tournament entry for men. In fact, men in the higher groups are directionally, but not significantly, less likely to compete as compared to those at the bottom and the coefficients on all male interaction terms are not jointly significant. As in Table 4, the difference in sorting between men and women is not statistically significant – the three interactions for men are not jointly different from the three interactions for women – but the interaction terms for men and women do differ marginally for the top two bins (p-values 0.102 and 0.072, respectively, in Column 2).

Columns 3 and 4 confirm that additional feedback changes the sorting pattern for men from the slight, not significant, U-shape to an increasing trend. The test of joint significance of interactions relative to the omitted lowest scorers for men now produces p-values of 0.04 and 0.10 (with controls). Columns 3 and 4 also confirm the results reported in Table 4 that feedback does not affect tournament entry behavior for women since they continue to sort correctly. As discussed further in the paper, these results mask heterogeneity in women’s response to the type of feedback they receive. Specifically, equally capable women sort differently after receiving negative feedback than after receiving positive feedback.

These results using these score bins are robust to using quartiles instead, with individuals scoring below average (4 out of 8) in the bottom quartile, those scoring exactly average in the second, those scoring 5 out of 8 in the third quartile and those scoring 6 or higher in the top quartile (see Appendix Table G4).

F6: Effect of Negative Feedback on Tournament Entry and Attribution in the AF Condition Relative to the LF Condition

Table F6.1 provides support for the pattern in Figure 5 by estimating linear regressions with wave and match gender condition fixed effects that split up the sample by gender and compare subjects in the additional feedback condition to those in the limited feedback condition. All columns control for risk and confidence, although the results are robust to omitting these controls.

Table F6.1: Determinants of Tournament Entry Decision by Type of Feedback Received

Samples Type of Potential Feedback	Women			Men		
	All	Negative		All	Negative	
	(1)	(2)	(3)	(4)	(5)	(6)
Additional Feedback	-0.0335 (0.115)	0.138 (0.144)	0.0480 (0.137)	-0.281** (0.130)	-0.320* (0.179)	-0.226 (0.264)
Limited Feedback x Score in Round 1	0.0470** (0.0228)	0.0618* (0.0359)		-0.00132 (0.0221)	-0.0189 (0.0403)	
Add'l Feedback x Score in Round 1	0.0603*** (0.0199)	0.0103 (0.0327)		0.0430** (0.0185)	0.0405 (0.0399)	
Limited Feedback x Score Bin 2			0.183* (0.103)			-0.0399 (0.201)
Limited Feedback x Score Bin 3			0.209* (0.121)			-0.146 (0.219)
Limited Feedback x Score Bin 4			0.959*** (0.100)			0.00845 (0.270)
Add'l Feedback x Score Bin 2			0.181 (0.124)			0.0293 (0.202)
Add'l Feedback x Score Bin 3			0.0634 (0.157)			0.0498 (0.226)
Add'l Feedback x Score Bin 4			0.0190 (0.183)			0.0764 (0.304)
Dependent variable mean	0.361	0.269	0.269	0.412	0.350	0.350
F-test of equality of interactions (p-value)	0.646	0.263	p < 0.01	0.106	0.257	0.879
Observations	352	223	223	337	177	177
R-squared	0.106	0.0526	0.0903	0.138	0.0719	0.0709

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for risk, confidence and gender of match treatment. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Columns 1 and 4 of Table F6.1 include all female and male participants, respectively, and replicate our results from Table 4. Women in our sample sort correctly into tournament with or without additional feedback on relative payment (Column 1). On the other hand, men who do not receive feedback on relative standing do not sort positively based on score (row 2 of Column 4), while men who do receive this additional feedback exhibit a strong positive sorting relationship (row 3 of Column 4).

Restricting the sample to the women who receive or would have received negative feedback on relative standing reveals an interesting result (Columns 2 and 3). Women who would have received negative feedback, yet did not because they were in the limited feedback group (row 2 of Column 2), still sort based on score. As compared to Column 1, the coefficient is larger in magnitude, although statistically significant only at 10% as the sample size shrinks. More importantly, comparable women in the additional feedback group no longer sort correctly (row 3 of Column 2). Controlling for score in bins, instead of linearly, Column 3 demonstrates that this is driven by women at the top of the score distribution, although we see marginally significant differences in tournament entry along the entire distribution: higher-performing women in the limited feedback group are significantly more likely to enter, while higher-performing women who receive negative feedback are not. The p-value at the bottom of the table, testing equality of the three interactions for the two treatment conditions (that is, testing three hypotheses where the first is Limited feedback X Score Bin 2 = Additional feedback X Score Bin 2 and the second and third are the corresponding hypotheses for bins 3 and 4), confirms that women sort into tournament entry differently when they receive negative feedback than when they receive no additional feedback.

We replicate this analysis for men (Columns 5-6). As with all feedback in Column 4, men who receive negative feedback seem to respond by correcting their tournament sorting. Even though we lose power, row 3 of Column 5 shows that even negative feedback produces a directionally positive sorting based on score. When we break up the effect into ability bins for men, we observe once again that entry increases with score for men who actually receive the negative feedback, although the F-test of equality of the three interaction terms does not reject equality. Tables G5, G6 and G7 confirm that the results in Table F6.1 are qualitatively the same if we include demographic controls, use a logit specification, or use score quartiles instead of bins, respectively.

Next, we revisit the analysis in Table 4 in the paper, focusing on the participants who would have received negative feedback. That is, we compare the subjects who receive negative feedback in the additional feedback treatment group and those who would have received negative feedback in the limited feedback group. Table F6.2 presents the linear specifications of the same kind of regressions we show in Table 6 in the paper.



Table F6.2: Gender Differences in Sorting into Tournament Entry in Response to Negative Feedback

Dep Var: Samples	Tournament Entry				Attribution	
	Limited Feedback		Additional Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)	(3)	(4)
Female	-0.335** (0.168)	-0.335** (0.166)	0.0991 (0.149)	0.126 (0.151)	-2.243 (9.555)	0.101 (9.072)
Male x Score in Round 1	-0.0157 (0.0400)	-0.0169 (0.0394)	0.0423 (0.0389)	0.0478 (0.0391)	-2.928 (2.348)	-1.493 (2.281)
Female x Score in Round 1	0.0491 (0.0365)	0.0573 (0.0363)	0.00419 (0.0330)	0.0104 (0.0326)	0.519 (1.991)	0.940 (1.948)
Dependent variable mean	0.341	0.341	0.274	0.277	68.89	68.81
F-test of equality of interactions (p-value)	0.211	0.143	0.416	0.427	0.246	0.378
R-squared	0.0649	0.0834	0.0213	0.0475	0.0455	0.0747

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants who received negative feedback in the additional feedback treatment and participants who would have received negative feedback in the limited feedback treatment. Columns 5 and 6 omit wave 3 since it did not contain the attribution question. Even numbered columns control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

The first four columns use tournament entry as the dependent variable, as in Tables 4 and 6, while Columns 5 and 6 use attribution. The linear specifications in the table below directionally replicate the results in Table 4, but note that we lose power due to a smaller number of observations (only those receiving negative feedback are included). In particular, women sort correctly into the tournament without receiving additional feedback (Columns 1 and 2), but not when they receive negative feedback on relative standing (Columns 3 and 4). On the other hand, men do not sort correctly into tournament with limited feedback (Columns 1 and 2) but do when they receive additional feedback (Columns 3 and 4). Columns 5 and 6 show that men with higher scores are more likely to attribute the negative feedback to luck, while women with higher scores are more likely to attribute the negative feedback to lack of ability, although neither coefficient is statistically significant and they are not significantly different from each other (p-value at the bottom of the table).

## G. Robustness Checks

Figure G1: Local Linear Regression of Tournament Entry on Score in Round 1 by Gender and Type of Additional Feedback Participants Would Have Received (Sample Restricted to Limited Feedback Condition)

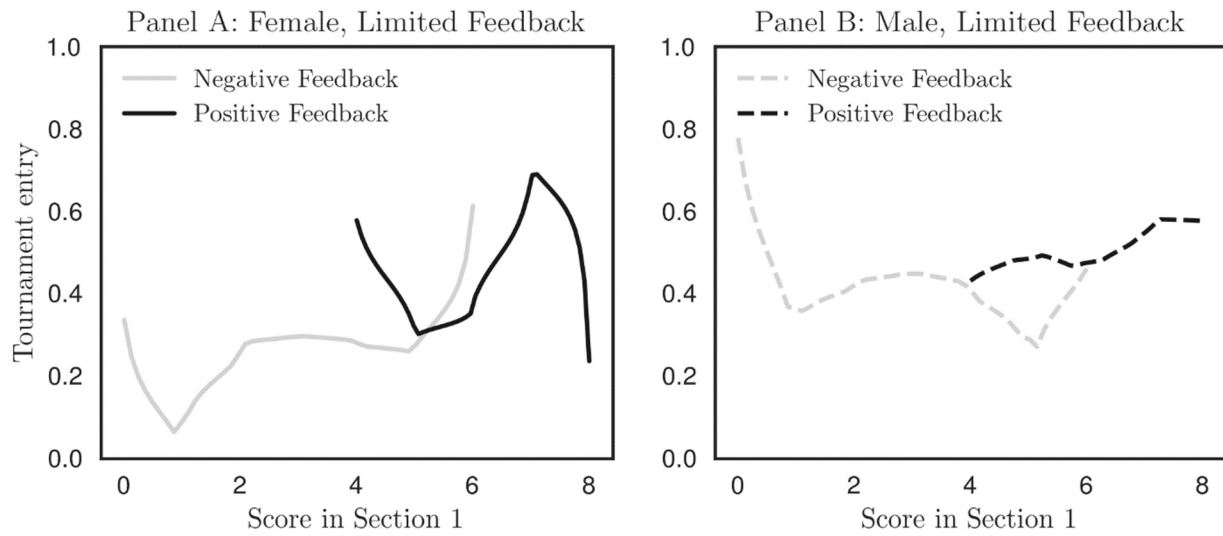


Figure G2: Gender Differences in Bonus Payment in Tournament (Round 2) in the Limited Feedback and Additional Feedback Treatments, on Average and by Round 1 Performance

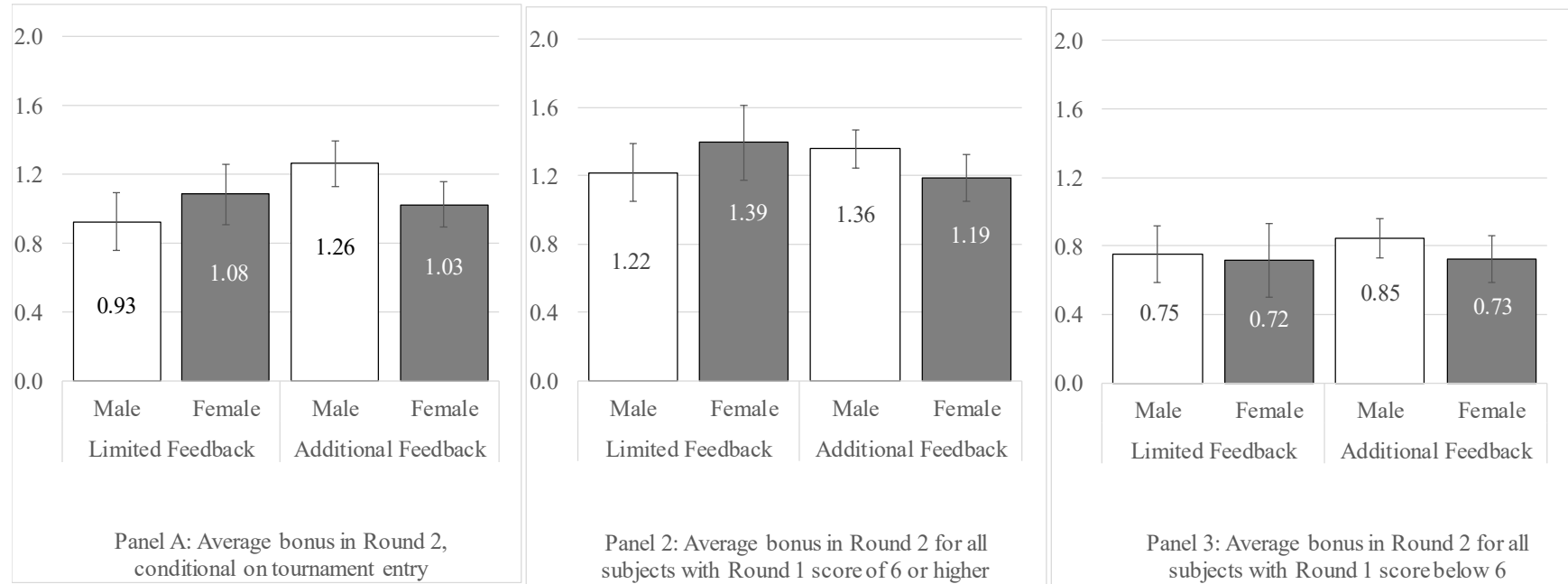


Table G1: OLS Estimates of Gender Gaps in Behavioral Traits With Demographic Controls

Dependent variable	Score Confidence (Self-Reported Score in Round 1)		Self-evaluation of Payment to be Below Average	
	(1)	(2)	(3)	(4)
Female	-0.724*** (0.154)	-0.453*** (0.137)	0.163*** (0.0390)	0.112*** (0.0373)
Score in Round 1		0.490*** (0.0378)		-0.0924*** (0.00867)
Dependent variable mean	3.963	3.963	0.547	0.547
Observations	687	687	687	687
R-squared	0.0628	0.262	0.0721	0.183

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of the match. Demographic controls include age and fixed effects for education category, income bracket and race. In Columns 1 and 2, score-confidence is measured on a scale of 0 to 10 (the highest possible score in Round 1 is 8). Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G2: Determinants of Tournament Entry Decision by Treatment with Demographic Controls

Samples	Limited Feedback			Additional Feedback		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A</b>						
Female	-0.109*	-0.257*	-0.217	0.0166	0.0613	0.122
	(0.0611)	(0.141)	(0.137)	(0.0505)	(0.122)	(0.120)
Male x Score in Round 1		0.0106	-0.00134		0.0634***	0.0623***
		(0.0233)	(0.0229)		(0.0192)	(0.0200)
Female x Score in Round 1		0.0505**	0.0426*		0.0627***	0.0602***
		(0.0216)	(0.0221)		(0.0197)	(0.0211)
Dependent variable mean	0.390	0.390	0.390	0.383	0.383	0.383
Observations	282	282	282	405	405	405
R-squared	0.110	0.125	0.156	0.0524	0.106	0.175

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of the match. Columns 3 and 6 control for risk and confidence. Demographic controls include age and fixed effects for education category, income bracket and race. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G3: Logit Model of Determinants of Tournament Entry Decision by Treatment

Samples	Limited Feedback			Additional Feedback		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A</b>						
Female	-0.488** (0.248)	-1.134* (0.627)	-1.037* (0.626)	0.0188 (0.207)	0.0943 (0.562)	0.311 (0.596)
Male x Score in Round 1		0.0454 (0.0915)	-0.0210 (0.0953)		0.275*** (0.0836)	0.283*** (0.0930)
Female x Score in Round 1		0.219** (0.109)	0.186* (0.111)		0.299*** (0.0878)	0.315*** (0.101)
Dependent variable mean	0.391	0.391	0.391	0.381	0.381	0.383
Observations	284	284	284	407	407	405

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects. Columns 3 and 6 control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G4: Determinants of Tournament Entry Decision by Treatment and Quartile

Samples	Limited Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)
Female	-0.190** (0.0907)	-0.154* (0.0923)	0.0257 (0.0713)	0.0922 (0.0730)
Male x Score Quartile 2	-0.0446 (0.117)	-0.0917 (0.116)	0.0146 (0.0900)	0.0267 (0.0868)
Male x Score Quartile 3	-0.0157 (0.137)	-0.0887 (0.139)	0.235** (0.103)	0.231** (0.0950)
Male x Score Quartile 4	0.0654 (0.114)	-0.00991 (0.115)	0.243*** (0.0847)	0.219** (0.0876)
Female x Score Quartile 2	0.170* (0.101)	0.171* (0.0998)	0.0518 (0.0915)	0.0185 (0.0880)
Female x Score Quartile 3	0.0220 (0.101)	-0.0240 (0.101)	0.212** (0.0985)	0.182* (0.101)
Female x Score Quartile 4	0.268** (0.125)	0.229* (0.120)	0.250*** (0.0939)	0.222** (0.0965)
Dependent variable mean	0.391	0.391	0.381	0.383
F-test of male interactions (p-value)	0.848	0.816	0.00831	0.0185
F-test of female interactions (p-value)	0.0945	0.0932	0.0220	0.0668
Observations	284	284	407	405
R-squared	0.0466	0.0934	0.0675	0.131

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. Even numbered columns control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G5: Determinants of Tournament Entry Decision by Type of Feedback Received With Demographic Controls

Samples Type of Potential Feedback	Women			Men		
	All	Negative		All	Negative	
	(1)	(2)	(3)	(4)	(5)	(6)
Additional Feedback	-0.0203 (0.116)	0.103 (0.143)	0.0908 (0.137)	-0.253* (0.141)	-0.197 (0.217)	-0.111 (0.309)
Limited Feedback x Score in Round 1	0.0427* (0.0228)	0.0427 (0.0351)		-0.00218 (0.0225)	0.00139 (0.0412)	
Add'l Feedback x Score in Round 1	0.0536** (0.0214)	0.00636 (0.0343)		0.0374* (0.0202)	0.0281 (0.0491)	
Limited Feedback x Score Bin 2			0.237** (0.103)			0.0354 (0.211)
Limited Feedback x Score Bin 3			0.165 (0.116)			-0.0582 (0.236)
Limited Feedback x Score Bin 4			1.052*** (0.164)			0.104 (0.266)
Add'l Feedback x Score Bin 2			0.178 (0.131)			0.0119 (0.239)
Add'l Feedback x Score Bin 3			0.0249 (0.168)			0.0369 (0.276)
Add'l Feedback x Score Bin 4			0.0101 (0.187)			0.0271 (0.366)
Dependent variable mean	0.361	0.269	0.269	0.412	0.349	0.349
F-test of equality of interactions (p-value)	0.708	0.423	0.00000385	0.170	0.647	0.932
Observations	352	223	223	335	175	175
R-squared	0.186	0.175	0.222	0.187	0.140	0.142

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for risk, confidence and gender of match treatment. Demographic controls include age and fixed effects for education category, income bracket and race. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



Table G6: Logit Model of the Determinants of Tournament Entry Decision by Type of Feedback Received

Samples Type of Potential Feedback	Women			Men		
	All	Negative		All	Negative	
	(1)	(2)	(3)	(4)	(5)	(6)
Additional Feedback	-0.138 (0.620)	0.727 (0.799)	0.318 (1.035)	-1.366** (0.615)	-1.487* (0.835)	-1.023 (1.177)
Limited Feedback x Score in Round 1	0.232** (0.114)	0.328* (0.195)		-0.00927 (0.0962)	-0.0798 (0.170)	
Add'l Feedback x Score in Round 1	0.284*** (0.0983)	0.0595 (0.178)		0.205** (0.0901)	0.197 (0.196)	
Limited Feedback x Score Bin 2			1.147 (0.772)			-0.151 (0.800)
Limited Feedback x Score Bin 3			1.290 (0.841)			-0.624 (0.895)
Limited Feedback x Score Bin 4			0 (.)			0.0401 (1.086)
Add'l Feedback x Score Bin 2			1.045 (0.807)			0.152 (0.972)
Add'l Feedback x Score Bin 3			0.388 (1.029)			0.253 (1.076)
Add'l Feedback x Score Bin 4			0.0900 (1.356)			0.372 (1.403)
Dependent variable mean	0.361	0.269	0.262	0.412	0.350	0.350
F-test of equality of interactions (p-value)	0.716	0.270	3.32e-29	0.0887	0.246	0.879
Observations	352	223	221	337	177	177
R-squared						

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for risk, confidence and gender of match treatment. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G7: Determinants of Tournament Entry Decision by Type of Feedback Received and Quartile

Samples Type of Potential Feedback	Women	Men
	Negative	
	(1)	(2)
Additional Feedback	0.0502 (0.0736)	-0.171* (0.0957)
Limited Feedback x Score Quartile 2	0.0808 (0.120)	-0.0652 (0.149)
Limited Feedback x Score Quartile 3	0.0441 (0.160)	-0.386*** (0.108)
Limited Feedback x Score Quartile 4	0.810*** (0.0836)	0.0364 (0.205)
Add'l Feedback x Score Quartile 2	-0.170 (0.118)	-0.120 (0.145)
Add'l Feedback x Score Quartile 3	0.147 (0.232)	0.146 (0.160)
Add'l Feedback x Score Quartile 4	-0.127 (0.148)	0.0494 (0.228)
Dependent variable mean	0.269	0.350
F-test of equality of interactions (p-value)	0.000000218	0.0137
Observations	223	177
R-squared	0.0817	0.0864

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for risk, confidence and gender of match treatment. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G8: Gender Differences in the Effect of Receiving Negative Feedback on Attribution and Tournament Entry, With Demographic Controls

Sample	All		Negative Self-Evaluation		Positive Self-Evaluation	
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Dep Var: Tournament Entry</b>						
Female	0.0972 (0.0764)	0.134* (0.0713)	0.136 (0.128)	0.136 (0.121)	0.0825 (0.107)	0.147 (0.101)
Male x Negative Feedback	-0.198*** (0.0728)	-0.1000 (0.0888)	-0.198* (0.110)	-0.158 (0.132)	-0.113 (0.123)	-0.0647 (0.142)
Female x Negative Feedback	-0.300*** (0.0756)	-0.145 (0.0975)	-0.245** (0.109)	-0.146 (0.146)	-0.411*** (0.127)	-0.275* (0.151)
Dependent variable mean	0.383	0.383	0.329	0.329	0.446	0.446
F-test of equality of interactions (p-value)	0.314	0.645	0.756	0.936	0.0807	0.201
F-test of female = female X neg fbk (p)	0.00343	0.0515	0.0714	0.217	0.0166	0.0486
Observations	405	405	219	219	186	186
R-squared	0.105	0.180	0.156	0.202	0.130	0.238
<b>Panel B: Dep Var: Attribution</b>						
Female	-9.239** (4.235)	-7.689* (4.425)	-1.952 (8.159)	-1.601 (8.562)	-12.27** (5.107)	-10.77** (5.415)
Male x Negative Feedback	-5.618 (4.191)	-1.837 (6.389)	11.94 (7.377)	15.92 (9.819)	-24.19*** (5.863)	-23.18*** (8.601)
Female x Negative Feedback	11.90*** (4.346)	15.69*** (5.413)	17.84*** (6.584)	23.05*** (8.396)	6.250 (6.149)	7.798 (7.519)
Dependent variable mean	67.62	67.62	67.95	67.95	67.24	67.24
F-test of equality of interactions (p-value)	0.00375	0.00383	0.536	0.455	0.000362	0.000396
F-test of female = female X neg fbk (p)	0.00632	0.00306	0.132	0.0707	0.0664	0.0892
Observations	307	307	167	167	140	140
R-squared	0.101	0.111	0.192	0.210	0.280	0.292

Notes: Robust standard errors in parentheses. All specifications are based on data from waves 1, 2 and 4 of the experiment, additional feedback condition only (wave 3 did not contain the attribution question), and include wave fixed effects and controls for gender of match treatment. Even numbered columns control for score, risk, and confidence. Demographic controls include age and fixed effects for education category, income bracket and race. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G9: Logit Model of the Effect of Receiving Negative Feedback on Tournament Entry

Sample	All		Negative Self-Evaluation		Positive Self-Evaluation	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.377 (0.304)	0.594* (0.315)	0.515 (0.496)	0.557 (0.507)	0.364 (0.398)	0.645 (0.425)
Male x Negative Feedback	-0.851*** (0.310)	-0.394 (0.430)	-0.981** (0.482)	-0.664 (0.638)	-0.423 (0.457)	-0.156 (0.612)
Female x Negative Feedback	-1.342*** (0.326)	-0.716 (0.450)	-1.047** (0.442)	-0.561 (0.633)	-1.707*** (0.554)	-1.203* (0.714)
Score in Round 1		0.191* (0.103)		0.158 (0.162)		0.162 (0.141)
Dependent variable mean	0.381	0.383	0.326	0.329	0.446	0.446
F-test of equality of interactions (p-value)	0.254	0.474	0.918	0.878	0.0591	0.148
F-test of female = female X neg fbk (p)	0.00184	0.0387	0.0586	0.238	0.0107	0.0530
Observations	407	405	221	219	186	186
<b>R-squared</b>						

Notes: Robust standard errors in parentheses. All specifications are based on data from waves 1, 2 and 4 of the experiment, additional feedback condition only (wave 3 did not contain the attribution question), and include wave fixed effects and controls for gender of match treatment. Even numbered columns control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G10: Gender Differences in Sorting into Tournament Entry in Response to Negative Feedback, With Demographic Controls

Dep Var: Samples	Tournament Entry				Attribution	
	Limited Feedback		Additional Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)	(3)	(4)
<b>Panel A</b>						
Female	-0.335*	-0.325*	0.0852	0.139	-0.383	2.264
	(0.180)	(0.179)	(0.172)	(0.172)	(10.31)	(9.607)
Male x Score in Round 1	-0.00691	-0.00918	0.0245	0.0351	-2.752	-1.119
	(0.0425)	(0.0425)	(0.0427)	(0.0435)	(2.525)	(2.483)
Female x Score in Round 1	0.0422	0.0484	-0.0110	-0.00703	-0.363	0.160
	(0.0357)	(0.0371)	(0.0352)	(0.0348)	(2.281)	(2.233)
Dependent variable mean	0.339	0.339	0.277	0.277	68.81	68.81
F-test of equality of interactions (p-value)	0.354	0.276	0.489	0.416	0.467	0.677
Observations	174	174	224	224	176	176
R-squared	0.188	0.200	0.0994	0.123	0.143	0.172
<b>Panel B</b>						
Female	-0.414**	-0.408**	-0.143	-0.102	-13.35	-12.29
	(0.199)	(0.202)	(0.229)	(0.234)	(12.41)	(10.76)
Male x Score Bin 2	-0.0234	-0.0114	-0.00659	0.00184	-11.25	-9.164
	(0.207)	(0.211)	(0.208)	(0.213)	(8.590)	(8.099)
Male x Score Bin 3	-0.0416	-0.0389	-0.00840	-0.00186	-9.239	-5.466
	(0.231)	(0.234)	(0.237)	(0.242)	(10.04)	(9.926)
Male x Score Bin 4	0.0408	0.0379	-0.0607	-0.00962	-41.19***	-32.43**
	(0.261)	(0.259)	(0.302)	(0.319)	(14.56)	(14.01)
Female x Score Bin 2	0.226**	0.260**	0.173	0.176	11.15	12.13
	(0.113)	(0.107)	(0.121)	(0.131)	(10.33)	(9.458)
Female x Score Bin 3	0.203	0.245*	0.0323	0.0256	10.01	10.71
	(0.128)	(0.131)	(0.162)	(0.168)	(11.74)	(11.03)
Female x Score Bin 4	0.969***	0.981***	-0.101	-0.0975	4.458	5.984
	(0.202)	(0.209)	(0.218)	(0.213)	(15.04)	(14.08)
Dependent variable mean	0.339	0.339	0.277	0.277	68.81	68.81
F-test of male interactions (p-value)	0.985	0.988	0.996	1.000	0.0460	0.116
F-test of female interactions (p-value)	p < 0.01	p < 0.01	0.228	0.245	0.712	0.616
F-test of equality of interactions (p-value)	0.0177	0.0151	0.686	0.649	0.187	0.186
Observations	174	174	224	224	176	176
R-squared	0.217	0.229	0.114	0.138	0.189	0.209

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants who received negative feedback in the additional feedback treatment and participants who would have received negative feedback in the limited feedback treatment. Columns 5 and 6 omit wave 3 since it did not contain the attribution question. Even numbered columns control for risk and confidence. Demographic controls include age and fixed effects for education category, income bracket and race. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G11: Logit Model of Gender Differences in Sorting into Tournament Entry in Response to Negative Feedback

Dep Var: Samples	Tournament Entry			
	Limited Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)
<b>Panel A</b>				
Female	-1.622** (0.815)	-1.638** (0.807)	0.486 (0.768)	0.613 (0.794)
Male x Score in Round 1	-0.0615 (0.168)	-0.0686 (0.169)	0.208 (0.192)	0.240 (0.198)
Female x Score in Round 1	0.259 (0.196)	0.298 (0.192)	0.0218 (0.171)	0.0606 (0.174)
Dependent variable mean	0.341	0.341	0.274	0.277
F-test of equality of interactions (p-value)	0.186	0.125	0.421	0.451
Observations	176	176	226	224
<b>Panel B</b>				
Female	-1.853* (1.061)	-1.879* (1.029)	-0.795 (1.096)	-0.672 (1.184)
Male x Score Bin 2	-0.183 (0.801)	-0.164 (0.809)	-0.0459 (0.920)	0.0915 (0.986)
Male x Score Bin 3	-0.449 (0.883)	-0.490 (0.889)	0.0699 (1.032)	0.111 (1.086)
Male x Score Bin 4	0.175 (1.066)	0.152 (1.060)	0.0672 (1.349)	0.287 (1.434)
Female x Score Bin 2	1.003 (0.834)	1.121 (0.785)	0.979 (0.711)	1.128 (0.808)
Female x Score Bin 3	1.139 (0.899)	1.338 (0.849)	0.189 (0.962)	0.276 (1.016)
Female x Score Bin 4	0 (.)	0 (.)	-0.169 (1.319)	-0.0283 (1.328)
Dependent variable mean	0.333	0.333	0.274	0.277
F-test of male interactions (p-value)	0.889	0.870	0.997	0.998
F-test of female interactions (p-value)	0.427	0.272	0.311	0.298
F-test of equality of interactions (p-value)	0.436	0.316	0.531	0.555
Observations	174	174	226	224

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants who received negative feedback in the additional feedback treatment and participants who would have received negative feedback in the limited feedback treatment. Even numbered columns control for risk and confidence. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table G12: Gender Differences in Sorting into Tournament Entry in Response to Negative Feedback by Quartile

Dep Var: Samples	Tournament Entry				Attribution	
	Limited Feedback		Additional Feedback		Additional Feedback	
	(1)	(2)	(3)	(4)	(3)	(4)
<b>Panel B</b>						
Female	-0.190** (0.0917)	-0.177* (0.0936)	0.0167 (0.0713)	0.0488 (0.0728)	7.033 (4.583)	7.281 (4.901)
Male x Score Quartile 2	-0.0199 (0.145)	-0.0377 (0.147)	-0.0799 (0.143)	-0.117 (0.146)	-3.497 (8.563)	-1.669 (9.235)
Male x Score Quartile 3	-0.396*** (0.100)	-0.368*** (0.106)	0.135 (0.165)	0.134 (0.156)	8.279 (6.779)	10.01 (6.812)
Male x Score Quartile 4	0.0635 (0.206)	0.0506 (0.202)	0.0433 (0.226)	0.0598 (0.226)	-34.18*** (11.41)	-28.33** (11.31)
Female x Score Quartile 2	0.0735 (0.124)	0.0885 (0.123)	-0.172 (0.117)	-0.199* (0.117)	-3.644 (6.658)	-3.977 (6.649)
Female x Score Quartile 3	0.0233 (0.155)	0.0424 (0.159)	0.0857 (0.221)	0.137 (0.216)	11.53 (10.41)	13.64 (11.31)
Female x Score Quartile 4	0.801*** (0.0930)	0.789*** (0.0885)	-0.151 (0.154)	-0.150 (0.144)	0.857 (10.36)	1.779 (10.16)
Dependent variable mean	0.341	0.341	0.274	0.277	68.89	68.81
F-test of male interactions (p-value)	p < 0.01	p < 0.01	0.750	0.627	p < 0.01	0.0131
F-test of female interactions (p-value)	p < 0.01	p < 0.01	0.384	0.231	0.582	0.515
F-test of equality of interactions (p-value)	p < 0.01	p < 0.01	0.858	0.845	0.138	0.215
Observations	176	176	226	224	178	176
R-squared	0.0974	0.110	0.0349	0.0662	0.112	0.134

Notes: Robust standard errors in parentheses. All specifications include wave fixed effects and controls for gender of match treatment. The sample is restricted to participants who received negative feedback in the additional feedback treatment and participants who would have received negative feedback in the limited feedback treatment. Columns 5 and 6 omit wave 3 since it did not contain the attribution question. Even numbered columns control for risk and confidence.

Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.