Social Information and Gender Differences in Competitive Preferences

Christopher Roby*

The University of Texas at Dallas

September 26, 2016

Abstract

I examine the effect of selected social information on gender differences in selection into a competitive environment using a simple addition task. Participants perform this task in multiple rounds and are then asked to select into a competitive or non-competitive pay scheme. Prior to choosing their pay schemes, participants are shown selected results about average performance and choices in a similar experiment. I find that the inclusion of selected social information eliminates the extant gender gap in selection into a competitive environment in every treatment. The reduction in gender gap is not due to greater efficiency of choices by men or women, even though inefficient under entry is mostly eliminated. Rather, the inclusion of feedback causes men and women to select into a competitive pay scheme in a similar manner, which thusly removes the gender gap. I find that the existence, or lack, of a tournament entry gap is consistent with participants' competitive preferences in all but one treatment.

*I would like to thank Sherry Li for all of her invaluable help, guidance, and encouragment on this project. Vineetha Sadavsan, Mrunal Hadke, and Miryam Ahmadi for their help running experimental sessions and assistance in the LBOE lab. The Labratory for Behavioral Operations and Economics for the use of their laboratory. Shuo Yang for all of his assistance throughout this project. All participants at the 2015 Economic Science Association North American Meeting in Dallas, Texas for their valuable input. The Negotiation Center for financial support. Participants at the Economics Department Brown Bag Seminar at UT Dallas. Participants at the LBOE Seminar at UT Dallas. Wendy Lee for her help with zTree Programming. Holger Rau and Anya Samek for sharing versions of their zTree program. Contact: The University of Texas at Dallas, 800 West Cambell Road, GR 31 (room GR 2.822), Richardson, TX 75080. email: cxr113330@utdallas.edu

1 Introduction

Despite a wealth of advancements for women in the labor market during the last half-century, women are still vastly underrepresented in high earning executive and managerial positions. Using a data set from US firms, Bertrand and Hallock (2001) find that women only account for about 2.5% of high level executives and earn 45% less than men as of 1997. Even when accounting for differences in ability, this gap remains; Bertrand and Hallock conclude that women are ”virtually absent” from the corporate world. When looking at specifically CEO’s, Wolfers (2006) finds that women only account for 1.3% of CEO’s from 1992-2004. Even with these bleak figures, things seem to be improving in the United States. As of 2010, women comprise 25.5% of CEO’s, yet only earn about 80% of what their male counterparts in the same positions do (Solis and Hall, 2011). If differences in ability are not an issue, then we are left with myriad explanations as to why these differences both exist, and persist. A more recent explanation for these differences in outcomes is the existence of differences in preferences for competition and competitive environments (which many higher paying executive jobs inhabit). It is natural to imagine that if men and women have dissimilar preferences for competitive work environments, then this could help explain why this gender gap exists.

Using an experimental structure similar to Niederle and Vesterlund (2007), I investigate the role of selected social information in reducing or eliminating any gender gap in selection into a competitive environment. This feedback is selected average performance information from a related experiment. In three treatments, I assess the effect of gender neutral and gendered feedback on selection into a winner take all tournament. While the inclusion of social information does not lead to perfect sorting based on ability, it does lead to an elimination of the gender gap in each of the treatments, due to the increased tournament entry of women, even when it is not optimal to do so. While some trepidation exists about the ability of experiments like this to actually explain differences in competitive preferences, I find a relationship between the existence of a statistically significant gender gap in tournament entry and significant differences in competitive preferences between men and women in all but one of the treatments.

Much of the extant research on gender differences is limited to the exploration of their existence in different environments (Niederle and Vesterlund, 2011). This line of research is important in acknowledging the differential challenges women face relative to men, which are largely due to societal structures and socialization. However, few studies talk about ways to close the gender gap, and those that do rely on mechanisms which are often unrealistic. This paper provides a simple way to eliminate the gender gap in competitive selection using social information. This type of feedback encourages high performing women to enter competitive environments at the rate that they should based on their performance. The remainder of the paper proceeds as follows: section 2 reviews relevant literature about gender differences and social information, section 3 explains the design of the experiment and the three treatments, section 4 presents the results of the experiment, and section 5 concludes.

2 Literature Review

The study of gender differences in outcomes and behavioral preferences is not new. Evolutionary psychologists have long studied how the socialization of children can lead to differences in attitudes by adulthood. An illustrative example of this is "the play styles that both sexes adopt. Boys often play games which are considered to be competitive interactions governed by rules aimed at specific goals. While girls are more involved in 'play' which is considered a cooperative activity in which there is no winner and no clear endpoint" (Cambell, 2013)[1]. Economists focus heavily on gender differences in labor market outcomes and wage differentials. Recently the focus has broadened to include a rigorous study of gender differences in the preferences which underlie economic decision making. Research has assessed gender differences in altruism (Andreoni and Vesterlund, 2001), risk aversion (Eckel and Grossman, 2002; Eckel and Grossman, 2008), cooperation in negotiation (Eckel, de Olivera, and Grossman, 2008), selfishness in dictator games (Eckel and Grossman, 1998), and more recently, competitive preferences have come into focus.

[1]Cambell, (2013) provides a thorough review of the evolutionary psychology literature on gender.

Psychology looks at the degree to which nature or nurture forces can account for gender differences. Some economic research adds to this discussion by showing that cross cultural differences in behavior of men and women exists. Gneezy, Leonard, and List (2009) perform a simple competitive task in both a patriarchal and a small Matrilineal society. They find that in the patriarchal society, men have a much greater preference for competition than women. In the Matrilineal society, women have a much greater preference for competition than men. The matrilineal women are actually as competitive as the patriarchal men, which suggests that preference for competition has strong ties to socialization and gender norms in a society.

Niederle and Vesterlund (2007) assess gender differences in competitive preferences using a simple experiment meant to mimic selection into competitive jobs. Participants perform a simple addition task and select into either a competitive or non-competitive payment scheme. Niederle and Vesterlund find that while there are no gender differences in performance in the task, there are significant gender differences in selection of competitive payment schemes for the task. This gender difference is driven by both women under selecting the competitive pay scheme and men over selecting it, based on their performance. Many successive works sought to replicate this result and modify the experiment to assess the robustness of the findings Dargnies, (2009) Sutter and Rutzler, (2010), Mayr et al. (2012), and Samek, (2013) show that these differences both exist and persist across the life cycle.[2] More recent research seeks to modify the informational environment of the experiment to try to reduce or eliminate the gender gap in tournament entry. Brandts, Groenert, and Rott (2012) find that intergenerational advice is able to eliminate the gender gap in choices, mostly through the increased entry of high performing women into the competition. Niederle, Segal, and Vesterlund (2012) investigate the role of including ex-ante affirmative action quotas on competitive selection. They find that the inclusion of these quotas increases selection efficiency of high performing women, thus eliminating any gender differences. Wozniak, Harbaugh, and Mayr (2014) find that including information about relative performance induces more efficient competitive pay scheme choices and eliminates the gender gap in competitive entry.

[2] Both Croson and Gneezy (2009) and Niederle and Vesterlund (2011) provide excellent reviews of many of these works.

Why should we expect that social information can be an effective way to reduce or eliminate gender differences in competitive preferences? The role of social information to act as an influential nudging factor in experiments with low stakes is well known. This type of information encourages donations to charity when it is known that other contribute (Frey and Meier, 2004; Martin and Randal, 2008; Croson and Shang, 2008; Shang and Croson, 2009), increases contribution likelihood in public goods games (Fischbacher, Gachter, and Fehr, 2001; Potters, Sefton, and Vesterlund, 2005), increases contribution to a movie rating website (Chen et al. 2010), and can increase payoffs for both parties when social history is included in the trust game (Berg, Dickhaut, and McCabe 1995). Coffman, Featherstone, and Kessler (2014) indicate that social information is an influential factor in altering behavior in high stakes decisions as well. They find that including social information about a previous year's Teach for America cohort makes an individual more likely to accept a teaching job after training and return for a second year.

In a similar vein, this experiment seeks to modify the informational environment using a new mechanism, social information. This is selected social information from a previous experiment, specifically, Wozniak, Harbaugh, and Mayr (2014). This type of intervention is important since it is relatively simple and cheap to provide. This is in contrast to some of the interventions above. Getting a department, company, industry, or government to agree to affirmative action policy is difficult, and not without dissent and time costs. An example of the legal difficulty affirmative action policies are met with is the 2003 case Gratz v. Bollinger. In this case the supreme court struck down the explicit use of affirmative action policies in college admissions (McBride, 2003). Other interventions are unrealistic. If these types of experiments are meant to mimic job entry, then it is highly unrealistic, even illegal in many cases, for relative ranking information among applicants to be disseminated. These issues lead to a desire for an easier and more cost effective type of information than can eliminate the gender gap in competitive preferences which leads to differential selection.

The aim of this paper is to investigate how the inclusion of social information affects optimal sorting of participants into competitive and non-competitive pay schemes. A benefit of the social information used here is that it is relatively cheap and realistic, in that there are real world examples of the use of this type of information. Average information about a subgroup is not always fully informative (as relative performance information may be), it can be cheaply provided and helpful when making decisions. Imagine a high school senior who is in the process of applying for college. While they may know their relative ranking within your high school, they will certainly not know their relative ranking among other college applicants. The most commonly available information (aside from college and major rankings) about their entry cohort is left to average performance statistics provided by the colleges themselves. These are statistics like average GPA, SAT/ ACT scores, demographic characteristics, the percentage of enrolled students in each major, etc. While this information is not gendered as in the social information I provide, it is a type of average performance information. There are a variety of factors that will ultimately go into a decision to apply to and enter college, but this type of information can help nudge a student into selecting a set of schools which best suits them.

I provide participants with average performance of both genders, their respective tournament entry decisions, and their results with respect to winning the tournament prior to making their competitive choice. Participants are informed that this data is selected to avoid any deception. To test the effect of including social information, I use three treatments. One treatment shows that women and men performed equally on average (and reveals their differential competitive entry choices), one which shows that men outperform women on average, and the third treatment shows women outperform men on average. I find that the inclusion of social information eliminates the gender gap in choices in all of the treatments. This result is driven mostly by women increasing their rate of tournament entry. Not all of this increase in entry is payoff maximizing, as the resulting choices by women are just as inefficient as their male counterparts. There has been discussion recently about whether these types of experiments are actually measuring competitive preferences, or just the impact of social norms. I find that there are significant gender differences in competitive preferences whenever there are significant differences in tournament selection between genders in all but one treatment. This suggests that differences in competitive preferences are weakly consistent with gender differences in tournament entry decisions.

3 Experimental Design

The structure of this experiment is based closely on the design in Niederle and Vesterlund (2007). Here, the baseline condition will be identical to the aforementioned paper, and the treatment conditions will be variations of this design. I use a between subjects design for this experiment. The experiment was conducted at the University of Texas at Dallas using the subject pool of the Laboratory for Behavioral Operations and Economics (LBOE). Participants were recruited using ORSEE (Grenier, 2003), an online recruitment system which allowed pre-screening to make sure the sample was gender balanced. The task is to add up five sets of two digit numbers for a period of five minutes without using a calculator. These numbers are randomly generated using zTree software (Fischbacher, 2007). A participant submits an answer and another problem appears on the screen, along with an indication of whether or not they solved the previous problem correctly. Participants earn a $5.00 dollar show-up fee, and an additional $5.00 for completing the experiment. One of the four rounds was randomly selected for payment, in addition to the flat rate payments. The payment round was selected prior to the sessions using a fair four-sided die. The average payout was $16.74 dollars. Experimental sessions took, on average, 30 minutes to run. The majority of participants were majoring in either engineering, computer science, or a business/ related field (like accounting, management, etc.). The average GPA for participants is 3.57. See Table 9 in Appendix A for a complete breakdown of summary statistics for the participants. The baseline condition proceeds as follows:

In Round 1 participants perform the addition task under a piece rate scheme where they earn $.50 cents per problem they solve correctly (a 'Round' in this experiment is equivalent to the 'Tasks' in Niederle and Vesterlund 2007). In Round 2 participants perform under a four person winner-take-all tournament. Individuals are randomly assigned into four person groups with two men and two women each. Participants are told explicitly that they are sitting in the same row as their group members. The individual who performs the best in the four person group earns $2.00 per problem solved correctly, while the other members of the group receive nothing. In the event of a tie, the winner of the tournament was chosen randomly through the zTree (Fishbacher, 2007) program. In Round 3, participants are able to choose which of the piece rate or tournament pay schemes they would like to perform the addition task under. If they select the piece rate, they receive $.50 cents per problem solved correctly. If participants choose the tournament, their performance is evaluated relative to the performance of the members of their respective group from the Round 2 tournament. As long as the participants solve more problems correctly than the Round 2 winner and select the tournament, then they will receive $2.00 dollars per correct problem [3]. Because of this, the Round 3 choice is actually an individual decision task. For Round 4, participants are asked to choose which payment scheme they would like applied to their Round 1 performance. The addition task is not performed in this round. If they select the piece rate, they receive $.50 cents per problem solved correctly, which results in the same profit as if Round 1 were selected for payment. If tournament is chosen, then participants will receive $2.00 dollars for each correct problem if they performed the best in their group in Round 1, otherwise they will receive no payment. A Round 4 decision will not affect the payment of any other group member and does not depend on the entry decisions of others. Round 4 is used to determine if general factors such as risk or feedback aversion by themselves cause a gap in tournament entry (Niederle and Vesterlund, 2007). After Round 4 is completed, participants are asked to guess how they ranked relative to their four person group, the average number of problems solved correctly in each of the first two rounds by the session, and to provide their preference for competition on a scale of 1-7 (7 is the highest and 1 is the lowest).

Three treatments were conducted to assess the effect of social information on choices in Round 3. Feedback is given prior to making the decision in Round 3 (after the instructions of how the Round will proceed are read). The intention of these treatments is to see if social information showing that men and women perform similarly (or differently) can alter decisions and remove

[3] This design leaves open the possibility that all four members in a group can win the tournament in Round 3. It is possible that none of the members win the tournament. Neither of these situations occurred in this experiment.

Table 1: Experimental Design

Treatments	Information Intervention	Sessions	Participants
Baseline (Niederle and Vesterlund, 2007)	None	10	40
Gender Neutral (GN)	Men and Women on average perform the same. Men choose tournament more, but their win percentage is lower,	10	40
Men as High Performers (MHP)	Men on average perform better than women. Men select and win the tournament more.	10	40
Women as High Performers (WHP)	Women on average perform better than men. Women select and win the tournament more	10	40

Participants are in groups of 4, with two men and two women in each group. The social information is read aloud to all participants before they choose their remuneration scheme in Round 3.

any gender gap in tournament entry. It is important to note that the inclusion of this feedback does not change the nature of the choice with respect to how a participant's payoff may affect the payoffs of other participants. It is still an individual decision task, as before. However, the choice will be slightly modified because participants are now presented with different information. This can change their perspective on the choice, from both the numerical and gendered information presented in each of the treatments. The feedback information itself is exogenous to this experiment since it comes from a different experiment and is constructed to test my hypotheses. This type of feedback gives no indication of how an individual performed relative to other participants in their group, but it may give them a benchmark of performance by which to estimate which payment scheme to choose in this round. Participants are informed that the feedback is selected so as to avoid any deception.

3.1 Treatments

Treatment 1 is the Gender Neutral Treatment (GN). This treatment is constructed to show that men and women performed similarly on average. Participants are informed that men and women both solve 11 problems. They are then provided with information as to how many men and women selected the tournament in Round 3, and the number who won[4]. The average subject in Round 3 of this experiment solved 12.26 problems correctly. So by showing this lower average performance information, individuals may be enticed to enter the tournament at a high rate, regardless of gender.

Hypothesis 1:(a)In the presence of social information, relative to the baseline condition, more women should enter the tournament because of the feedback, thus eliminating the gender gap in tournament entry. Men will enter the tournament at a similar rate to the baseline condition. The hypotheses for these treatments need to address how over and under entry may contribute to these changes in decision making for men and women.[5] (b)Inefficient under entry by women will decrease or even be eliminated. High performing women (who do not enter the tournament enough in the baseline condition) should be enticed into entering the tournament when they should in the presence of the relatively low performance feedback in this treatment. (c)Under entry for men will be low, similar to the baseline. Since the average performance provided in this treatment is on the low side, we should see an increase in over entry across genders, because those who should not enter the tournament do in the presence of social information.

Treatment 2 is the Men as High Performers Treatment (MHP). The intent of this treatment is to exploit any notion or stereotype that men may be better than women at math, even though I find no evidence of this. Participants are shown that men solved 14.5 problems on average, and women solved 10.63 problems on average. Then information about tournament choices in Round 3 is provided in the feedback.

Hypothesis 2: (a)This should lead women away from the tournament, regardless of perfor-

[4]See Appendix B for verbiage and layout of the feedback

[5]Under entry is selecting the piece rate when it would have been optimal to choose the tournament. Over entry is choosing the tournament when it would have been optimal to select the piece rate. Here, optimal means the payoff maximizing choice based on a participant's performance.

Table 2: Mean Performance by Round and Treatment

	Treatment	Round 1	Round 2	Round 3
(Women)	Base	9.55	11.30	11.95
(Men)	Base	9.35	11.45	11.75
	GN	10.25	11.20	12.20
	GN	10.95	12.45	13.45
	MHP	9.30	10.95	11.85
	MHP	9.50	11.00	11.35
	WHP	9.75	11.65	12.15
	WHP	11.70	13.25	13.4
	Overall	9.71	11.28	12.04
	Overall	10.38	12.04	12.49

Table reports mean performance by gender and treatment in each Round where the addition task is performed. For the cumulative distributions see Figures 1A-1D in Appendix A.

mance. (b)Men will be undeterred in their decisions in the presence of the feedback, and enter the tournament at at least the same rate as in the baseline condition. The general expectation is that the results will be similar to the baseline condition in terms of tournament entry decisions.

Treatment 3 is the Women as High Performers Treatment (WHP). This treatment looks at how participants respond to receiving information that women are the high performers in the math task. Subjects are informed that women solve 14.33 problems on average, and men solve 10.64. Again, tournament selections and the number of individuals that won is provided.

Hypothesis 3: (a)Seeing that women are high performers will induce a larger number of women to enter the tournament. (b)This will cause men to shirk tournament entry because they are shown that men perform worse on average.

4 Results

One possibility for why a gender difference exists in these choices is that there are differences in performance between men and women. If there are gender differences in performance of the addition task, then we are unable to say anything about the effect of the included feedback. First, I will compare performance across all treatments to see if there are any differences in performance between men and women. Table 2 shows that in Round 1 the average number of problems solved is 9.71 for women and 10.38 for men. This difference is not significant using a Mann-Whitney test (z=0.277). The gender difference in Round 2 is insignificant as well(z=0.293), where women solve 11.28 problems correctly and men solve 12.04 on average. Even in the third round, where women solve 12.04 problems and men solve 12.49 problems on average, there remains no statistical difference in performance (z=0.691). The conclusions are the same using a two sided t-test.

While the mean values are informative about performance, the values from the WHP treatment are further apart than the other treatments. In this instance, the medians are informative. In Round 1 men have a median of 11.5 and women have a median of 10.5. This difference is insignificant using a Pearson Chi Squared test (p=0.191). In Round 2 men have a median of 13 and women have a median of 12; this difference is insignificant (p=0.514). Round 3 remains insignificant (p=0.191) even with men having a median of 13.5 and women of 11.5. Figures 1A-1D in Appendix A show the CDF's of performance for men and women in Round 2 of each treatment. There are no significant differences in performance across rounds or treatments, whether at the mean, median, or distributional level, so I can rule out differences in performance as an explanation for any gender gap that may be found.

With regards to performance, it is worthwhile to assess how the variance in performance changes in the face of social information. The social information could be surprising or unexpected, and lead to a higher variance in the treatments. It is worth noting that the information provided to participants is typically incorrect. Men and women perform similarly on simple math tasks like these (Hyde, Fennema, and Lamon, 1990), contrary to the performance feedback provided in the WHP and MHP treatments. In the absence of feedback a gender gap in tournament selection emerges, unlike the information provided in the GN treatment. Table 10 in Appendix A shows performance variances by gender and round for each treatment. Men's variance is about 1.5 to 2 times higher than women's, but variance is relatively stable across all treatments. Usually variance is higher in Rounds 1 and 2 (consistent with there being a significant improvement in performance from Round 1 to Round 2), but by Round 3 the variance is near or at its lowest point in any of the rounds. The only time the Round 3 variance is considerably higher than in

Table 3: Tournament Entry Across Rounds and Treatments

	Treatment	Round 3 PR	Round 3 T	Round 4 PR	Round 4 T
(Women)	Base	80%	20%**	85%	15%**
(Men)	Base	30%	70%**	55%	45%**
	GN	40%	60%	55%	45%
	GN	25%	75%	45%	55%
	MHP	40%	60%	60%	40%
	MHP	25%	75%	70%	30%
	WHP	45%	55%	75%	25%
	WHP	25%	75%	60%	40%

***P* < 0.05, ****P* < 0.01. *PR indicates a participant chose the piece rate, T is choosing the tournament*

Round 2 is in the MHP treatment. Women increase their variance from 8.89 in Round 2 to 12.97 in Round 3, and men increase from 13.36 in Round 2 to 16.13 in Round 3. This constitutes a 31% increase in variance for women and a 17% increase in variance for men. Outside of noting that men have larger performance variances than women, there does not seem to be a considerable effect on performance in the presence of social information.

Next, we need to look at choices in Rounds 3 and 4 to see if there is a gender gap, and if it remains in the treatments. Table 3 shows the frequency of tournament entry by gender in both Rounds 3 and 4. In the baseline condition for Round 3 men choose the tournament 70% of the time and women only choose the tournament 20% of the time. This is a statistically significant difference using a Mann-Whitney test (z=0.001). In the GN treatment the gender gap is eliminated as women enter the tournament 60% of the time and men enter it 75% of the time (z=0.317). I somewhat surprisingly find an identical, and not significant (z=0.317), result in the MHP treatment, where men enter the tournament 75% of the time and women enter the tournament 60% of the time. Results are similarly insignificant in the WHP treatment. Men enter the tournament 75% of the time and women enter 55% of the time (z=.190). It is important to note that men are hardly changing their behavior with respect to tournament entry. The reduction in the gender gap is due to women entering the tournament at a higher rate in the face of feedback. This is consistent with previous research that finds that women's behavior in the lab is more sensitive to interventions which may influence a participant's belief about relative performance (Niederle and Vesterlund 2011). There are significant differences in tournament selection in Round 4 for the baseline treatment only. Men select the tournament 45% of the time, while women only select the tournament 45% of the time (z=0.04).

Result: This replicates the result from Niederle and Vesterlund (2007) that there are gender differences in both the Round 3 and Round 4 decisions. However, these gender differences are eliminated in the presence of social information.

Relative to the baseline treatment, is the increase in entry for women in the treatments statistically significant? Comparing the baseline condition to the both the GN and MHP treatment (since women enter the tournament at the same rate in both treatments), the increase of women's tournament entry from 20% to 60% is significant at the 95% level using a Mann-Whitney test (z=0.011). Unsurprisingly, the tournament increase between the baseline and GN treatment of 70% to 75% is statistically insignificant (z=0.727). Men increase entry in this manner for each treatment, all of which are insignificant. Comparing the baseline to the WHP treatment, the increase from 20% tournament entry to 55% entry is statistically significant at the 95% level as well (z=0.024). With this in mind, and the information in Table 3, it is evident that the treatments were successful in eliminating the gender gap in tournament entry due to significant increases in the tournament selection of women in the treatments.

Assessing the choices in Round 3 paints a clear picture that the treatments succeeded in eliminating the gender gap. Yet, there are a variety of covariates we need to account for in a regression analysis first. Since the outcomes of interest (either the Round 3 or Round 4 choice) are just binary variables, I will use a probit model with clustering at the individual level[6]. The dependent variable in the table below is the participant's choice in Round 3. It is equal to 1 if tournament was chosen and zero if piece rate was chosen. Table 5 reports marginal effects from two probit regressions. All treatments are pooled, and treatment effects are reported by interaction terms that separate the effect of each treatment. This is essentially Differences-in-Differences analysis, where the first difference is gender and the second difference is the treatment.

[6]Clustering at the group level produces similar results.

Table 4: Round 3 Decision by Treatment (Probit, ME)

Columns:	(1)	(2)
Female	-0.469***	-0.427***
	(0.136)	(0.130)
GN	0.052	-0.002
	(0.146)	(0.134)
MHP	0.052	0.145
	(0.146)	(0.145)
WHP	0.052	0.019
	(0.146)	(0.139)
GN x fem	0.323	0.368**
	(0.203)	(0.171)
MHP x fem	0.323	0.196
	(0.203)	(0.197)
WHP x fem	0.281	0.328*
	(0.204)	(0.189)
Improve Round 2		0.024*
		(0.014)
Rank Round 1		-0.068**
		(0.033)
GPA		0.037
		(0.048)
Age		-0.013
		(0.014)
Rankguess Round 2		-0.148***
		(0.047)
Competitive Preference		-0.011
		(0.023)
N	160	160
Ps. R^2	0.095	0.221
ClustVar	Individual	Individual

*Dependent variable is choice in Round 3. It is equal to 1 if tournament is chosen and zero if piece rate is chosen. Marginal Effects and Standard Errors from a probit regression are reported here. ***P*< 0.01, ***P*< 0.05, **P*< 0.10. *Clustering is at the individual level.*

Female is an indicator variable for gender and is equal to 1 if the participant is a female and zero if they are a male. This variable will show the gender effect in the baseline condition. GN is an indicator variable equal to 1 if the participant was in the GN treatment, and zero otherwise. MHP is an indicator variable equal to 1 if the participant was in the MHP treatment, and zero otherwise. WHP is an indicator equal to 1 if a participant was in the WHP treatment, and zero otherwise. GN x fem is a composite indicator variable which is equal to 1 if the participant was a woman and in the GN treatment, and zero otherwise. MHP x fem is a composite indicator variable that is equal to 1 if the participant was a woman and in the MHP treatment, and zero otherwise. WHP x fem is a composite indicator, equal to 1 if a participant is a woman in the WHP treatment, and zero otherwise. Improve Round 2 is a control variable to account for the fact that participants perform almost unilaterally better in the Round 2 tournament than the Round 1 piece rate. It is the simple difference between a participant's Round 2 and Round 1 performance. Rank Round 1 is the participant's rank in Round 1 and is meant to account for any effect that rank may have on tournament entry. Age and GPA are the self-reported age and GPA of the participants. Rankguess Round 2 is the participant's guessed rank in Round 2. This guess is elicited after the experiment is completed, and is intended to control for confidence when making the Round 3 decision. Competitive Preference is the individual's preference for competition on a scale of 1-7, where 7 is the highest level of preference for competition. Table 4 reports marginal effects for the probit regressions.

What we see in Table 4 is that after conditioning on performance and other covariates, the gender difference in tournament entry remains significant in the baseline condition.

Result: In the full regression in column 2 being a woman decreases the probability of tournament entry by 42.7% in the baseline condition. This constitutes a large gender gap, about twice as large as the any previous result (other works find this difference to be on the order of 12% to 20% Niederle and Vesterlund, 2011).

If including social information can remove a gender gap this large, then it shows both the strength of this intervention and the potential usefulness for a mechanism such as this to alleviate these gender differences. A few of the remaining covariates are significant in the regression. While the focus of this paper is not explicitly on the role of confidence or rank effects in these experiments, these results are in line with expectation as to their roles. Both Rank Round 1 and Rankguess Round 2 are negative and statistically significant, meaning that a person who has a higher rank in Round 1 (associated with a smaller number on a scale of 1-4) or guesses a higher rank for themselves in Round 2 is going to have a higher probability of entering the tournament. It is not surprising that the Rankguess Round 2 coefficient is highly significant because confidence about an individual's relative ranking matters when deciding to enter the tournament.

To assess the effect of the treatments on the gender difference we need to look at the combination of the female variable in column 2 with each of the MHP x fem, GN x fem, and WHP x fem variables. This analysis is necessary to see if the combination of the female variable and each of the three treatment interaction variables are equal to zero. If I can reject that the combination of these variables are equal to zero, then there is a gender effect which remains in the treatments, suggesting that they may not have eliminated the gender gap.

Result:I am able to marginally reject that there is any remaining gender effect in both all 3 treatments: GN treatment (p=0.641), the MHP treatment (p=0.175), and the WHP treatment (z=0.457). This confirms that even after all covariates are accounted for, the treatments were successful in eliminating the gender gap in tournament entry. This is in line with hypothesis 1a, which says that the GN treatment will cause women to increase their rate of tournament entry relative to the baseline.

Why are the results for the MHP treatment contrary to hypotheses 2a and 2b? And why are the WHP results contrary to hypothesis 3b? One possibility is the gendered information content of these two treatments may have some effect on confidence through reactions that participants may have. The MHP treatment is set to show that men are better at solving math problems (and WHP to show women perform better), even though there is no evidence that suggests that men and women perform differently on simple math tasks (Hyde, Fennema, and Lamon, 1990). One reason that women may enter the tournament more in this treatment is because they feel that the gendered information is a challenge for them to either outperform the feedback or try to prove it wrong. Another possibility is that it is simply the lack of differences in competitive preferences in this treatment that drives this result. A third possibility is that participants were keying on the numerical aspects of the feedback, not the gendered aspects. A participant receives their performance feedback in the form of the number of problems they solved correctly in a certain Round. Since this is the only feedback about their own performance they receive prior to the social information intervention, it is reasonable to consider that the numerical average performance information would have been the most salient to assess the prospect of tournament entry. The results show that social information drives more women to enter the tournament, regardless of the gendered nature of the feedback information.

4.1 What is Driving Changes in Tournament Selection in Treatments?

Just seeing that the treatment conditions resulted in an elimination of gender differences in choices is not enough to explain the hypotheses completely. To do this, I need to determine the role of over and under entry into the tournament in Round 3 in relation to the reduction of the gender gap. Over entry is when a participant should not enter the tournament based on their performance in Round 2, but does so. This is calculated using a participant's relative ranking in Round 2. If a participant was ranked first in their group, then their optimal decision (in terms of maximizing payoffs) is to enter the tournament. Participants ranked second through fourth should select the piece rate. If a participant was tied for first, their rank is still considered first, and it is optimal for them to enter the tournament, even though they may not have actually won the Round 2 tournament. Under entry is a similar concept, except that under entry occurs when a participant is ranked first and does not choose the tournament. Both over and under entry are inefficient behaviors with respect to participant sorting. If sorting were perfectly efficient based on payoff maximization, a participant would know their ability type, relative to others, and sort accordingly. Inefficiency is created partly by withholding relative performance information when making the tournament entry decision in Rounds 3 and 4. However, as is discussed in the introduction, if the desire is to replicate entry into competitive jobs, it is highly unlikely that an individual will know their relative performance on a job interview, rather just whether they were awarded the job or not.

In previous experiments women tend to under enter the tournament and men tend to over enter the tournament in Round 3. This is what creates the gender difference in tournament selection. In the baseline, this pattern continues as 4 out of 5 women under enter and 9 out of 14 men over enter in the baseline condition. Table 5 shows the results for over and under entry.

Result: In the GN treatment, inefficient under entry is almost completely eliminated for men, and is non-existent for women. This follows with hypotheses 1b and 1c. With respect to over entry, men behave similarly to the baseline, but women severely increase over entry such that 8 out of 16 women are over entering.

In the MHP treatment, we see similar results to the Gender Neutral treatment. Both men and women have low levels of under entry and high levels of over entry. This is contrary to hypothesis 2b because we still see women entering the tournament at a high rate in this treatment. For the WHP treatment, women enter the tournament at a higher rate than in the baseline, resulting in 4 out of 9 women over entering the tournament. Inefficient under entry is severely decreased, with 4 of 11 women under entering. Men still enter the tournament at a high rate though, and their behavior is similar to the baseline with respect to over and under entry, contrary to hypothesis 3b.

Result: The result of the inclusion of social information causes men and women to behave similarly with regard to tournament entry, instead of presenting a gendered effect in the MHP or WHP treatment in line with the feedback information.

Overconfidence is usually a variable of interest in these types of studies. One explanation for why gender differences may exist is that men are more overconfident about activities that typically fall under a stereotype of the ”male domain”, like math tasks. It is important to assess whether there is a gender difference in overconfidence, and if there are differences between the treatment and control conditions. Overconfidence is defined by a participant guessing a higher rank than their actual rank in Round 2. I find no gender differences in confidence in any of the treatments. This is surprising, given that every previous work on this matter finds that significant differences in over-

Table 5: Over and Under Entry into the Tournament

Under Entry	Baseline	GN	MHP	WHP
(Women)	4/5 (80%)	0/4 (0%)	1/6 (16.7%)	4/11 (33%)
(Men)	1/6 (16.7%)	1/7 (14.3%)	1/6 (16.7%)	1/12 (8.3%)
Over Entry	**Baseline**	**GN**	**MHP**	**WHP**
(Women)	3/15 (20%)	8/16 (50%)	7/14 (50%)	4/9 (44%)
(Men)	9/14 (64%)	9/13 (69%)	10/14 (71%)	4/8 (50%)

Over and under entry are calculated using a participant's group rank from Round 2. Under entry is selecting the piece rate when it would have been optimal to choose the tournament. Over entry is choosing the tournament when it would have been optimal to select the piece rate. What drives the tournament entry gap in the baseline condition (and in previous works) is that men over enter into the tournament at a high rate and women under enter at too high of a rate.

confidence exist. I include controls for confidence levels in the regression since confidence may play an idiosyncratic role in a participant's decision making, but there is no systematic difference in overconfidence between genders.[7]

Recently, it has come into question as to whether these types of experiments are actually measuring differences in preferences for competition, or something like the role of socialization and societal gender norms. To assess this, participants were asked to provide their preference for competition on a scale of 1-7, where 7 is the highest preference for competition. Table 6 shows the results of these elicitations. In the baseline condition, there is a gender difference at the 95% level (z=0.033). This gender difference does not exist in the GN and MHP treatments. There is a gender difference in competitive preferences at the 99% level (z=0.0102)in the WHP treatment, despite the lack of a gender gap in the WHP treatment.

Result: This suggests there is a weak relationship between competitive preferences and gender differences in tournament selection in Round 3.

It is interesting to note that the vast majority of participants, the lowest being 75 percent, report having a preference for competition of at least 4. This may help explain why we see such high tournament entry by both genders in the treatment conditions.

[7]For more detailed results concerning overconfidence, see Appendix A

	Treatment	Average	Mean if Enter R3 T	Percent Confidence >= 4
(Women)	Base	5**	5.5	75%
(Men)	Base	6**	6.2	95%
	GN	6	5.8	95%
	GN	6.45	6.4	95%
	MHP	5.6	5.6	90%
	MHP	5.5	5.3	85%
	WHP	5.45**	5.82	85%
	WHP	6.6**	6.73	100%

Table 6: Competitive Preferences

**$P < 0.05$, ***$P < 0.01$. *Competitive preferences are on a scale of 1 to 7, where 7 is the highest preference for competition. Participants are asked to provide their preference for competition following Round 4.*

4.2 Round 4 Results

Table 8 in Appendix A reports marginal effects. The dependent variable is the participant's choice in Round 4. It is equal to 1 if the tournament is chosen, and zero if the piece rate is chosen. All control variables are identical to the above section. I do find gender differences in tournament selection in Round 4 in the baseline condition, even after accounting for control variables. Looking at column 2, we can see that the gender difference in the baseline is significant and says that a woman has a 29% lower probability of selecting the tournament in Round 4. As in Table 5, a participant's guessed rank in Round 2 is significant in the expected direction. A participant who guessed that they had a higher rank (associated with a lower number) is more likely to choose the tournament.

Again, it important to assess if there are gender effects that spill over into the treatments in Round 4, even though the treatments are not intended to impact the Round 4 decision. In both treatments I am able to marginally reject that there is a gender effect in the GN treatment (p=0.619), the MHP treatment (p=0.237), and the WHP treatment (p=0.443). There is no significant gender difference in Round 4 decisions in any of the treatments. 90 percent of participants who selected the tournament in Round 4 selected the tournament in Round 3, whereas 54 percent of participants who selected the Round 3 tournament selected the Round 4 tournament.

Result: In both cases, selecting the tournament in one Round increases your likelihood of selecting it in the next round, suggesting there are spillover effects when tournament entry is high.

5 Conclusion

Information about your performance relative to others in a variety of scenarios is very difficult, even impossible, to get. A more realistic type of information to provide is social information that can be helpful in deciding whether to enter into a more or less competitive scenario. In this paper I find that including social information eliminates the gender gap in tournament entry, using a simple real effort task. The reduction of the gender gap is not due to greater efficiency of tournament selection, but that women significantly increase tournament entry, even when it may not be a payoff maximizing choice. While I account for overconfidence, I find no systematic gender differences in overconfidence any of the treatments. I find that there is a weak relationship between tournament entry decisions and competitive preferences. As an extension of this work, it would be interesting to see how social information that was more rudimentary would fare in reducing the competitive gender gap, and how it could be extended into different environments. Buser, Niederle, and Oosterbeek (2014) find that gender differences in competitive preferences can account for up to 20% of the decision to enter a more prestigious and competitive career path. This suggests that there is an enormous potential benefit for a simple mechanism, like the social information used here, which can alleviate the effects of gender differences in competitive preferences.

6 Bibliography

[1]Andreoni, James and Lise Vesterlund. Which is the Fair Sex? Gender Differences in Altruism. Quarterly Journal of Economics. Vol. 116, No. 1 (2001), 293-312.

[2]Ball, Sheryl, Catherine Eckel, Phillip J. Grossman, and William Zane Status in Markets. The Quarterly Journal of Economics. Vol. 116, No. 1 (2001), 161-188.

[3] Berg, Joyce, John Dickhaut, and Kevin McCabe. "Trust, Reciprocity, and Social History." Games and Economic Behavior. Vol 10 (1995), 122-142.

[4]Bertrand, Marianne, and Kevin F. Hallock. The Gender Gap in Top Corporate Jobs. Industrial and Labor Relations Review. Vol. 55 (2001), 3-21.

[5]Brandts, Jordi, Valeska Groenert, and Christina Rott. The Impact of Advice on Womens and Mens Selection into Competition. Unpublished Manuscript. 2012

[6]Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. Gender, Competitiveness, and Career Choices. Quarterly Journal of Economics. Vol. 129, No. 3 (2014), 1409-1447.

[7]Campbell, Anne. A Mind of Her Own: The Evolutionary Psychology of Women. 2nd ed. Oxford: Oxford University Press, 2013.

[8]Chen, Yan, F. Maxwell Harper, Joseph Konstan, and Sherry X. Li. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens." American Economic Review. Vol. 100, No. 4 (2010), 1358-1398.

[9]Coffman, Lucas C., Clayton R. Featherstone, and Judd B. Kessler. "Can Social Information Affect What Job You Choose and Keep?" Working Paper, 2014.

[10]Croson, Rachel, and Jen (Yue) Shang. "The Impact of Downward Social Information on Contribution Decisions." Experimental Economics Vol. 11 (2008), 221-233.

[11]Croson, Rachel, and Uri Gneezy. Gender Differences in Preferences. Journal of Economic Literature. Vol. 47, No. 2 (2009), 448-474.

[12]Dreber, Anna, Emma von Essen, and Eva Ranehill. Gender and Competition in Adolescence: Task Matters. Experimental Economics. Vol. 17, No. 1 (2014), 154-172.

[13]Eckel, Catherine C., and Philip J. Grossman, Are Women Less Selfish Than Men?: Evidence from Dictator Games. The Economic Journal 108(448): 726- 735. May, 1998.

[14]Eckel, Catherine C. and Phillip J. Grossman. Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecasting Gambles. Journal of Economic Behavior and Organization. Vol. 68 (2008), 1-17.

[15]Eckel, Catherine C., Angela C. M. de Olivera and Philip J. Grossman. Gender and Negotiation in the Small: Are Women (Perceived To Be) More Cooperative than Men? Negotiation Journal 24(4): 429-445. October 2008.

[16]Eckel, Catherine and Phillip J. Grossman. Sex Differenced and Statistical Stereotyping in Attitudes toward Financial Risk. Journal of Evolution and Human Behavior. Vol. 23 (2002),

281-295.

[17]Eckel, Catherine C. and Sascha Fullbrunn. Thar SHE Blows? Gender, Competition, and Bubbles in Experimental Asset Markets. American Economic Review. Vol. 102, No. 2 (2015), 906-920.

[18]Eckel, Catherine C., and Philip J. Grossman, The Relative Price of Fairness: Gender Differences in a Punishment Game, Journal of Economic Behavior and Organization. Vol. 30, No. 2 (1996), 143-158.

[19]Fischbacher, Urs. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." Experimental Economics. Vol. 10, (2007) 171-178.

[20]Fischbacher, Urs, Simon Gachter, and Ernst Fehr. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." Economic Letters. Vol. 71, No. 3 (2001), 397-404.

[21]Frey, Bruno S., and Stephen Meier. "Social Comparisons and Pro-Social Behavior: Testing 'Conditional Cooperation' in a Field Experiment." American Economic Review. Vol 94, No. 5 (2004), 1717-1722.

[22]Gneezy, Uri, Kenneth L. Leonard, and John A. List. Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. Econometrica. Vol. 77, No. 5 (2009), 1637-1664.

[23]Gobillon, Laurent, Dominique Meurs and Sbastien Roux. Estimating Gender Differences in Access to Jobs. Journal of Labor Economics. Vol. 33, No. 2 (2015), 317-363.

[24]Goldin, Claudia. A Grand Gender Convergence: Its Last Chapter. American Economic Review. Vol. 104, No. 4 (2014), 1091-1119.

[25] Grenier, Ben. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." Journal of the Economic Science Association. Vol 1, No. 1 (2015), 114-125.

[26]Heinz, Mattias, Hans-Theo Normann, and Holger A. Rau. How Gender Differences in Competitiveness may cause a Gender-Wage gap: Experimental Evidence. Working Paper, 2013.

[27]Hyde, Janet S., Elizabeth Fennema, and Susan J. Lamon. Gender Differences in Mathematics Performance: A Meta-Analysis. Psychological Bulletin. Vol. 107, No. 2 (1990), 139-155.

[28]Manning, Alan, and Farzad Saidi. Understanding the Gender Pay Gap: Whats Competition Got To Do With It? Industrial and Labor Relations Review. Vol. 63, No. 4 (2010), 681-698.

[29]Martin, Richard, and John Randal. "How is Donation Behavior Affected by the Donations of Others?" Journal of Economic Behavior and Organization. Vol. 67, No. 1 (2008), 228-238.

[30]Masclet, David, Emmanuel Peterle, and Sophie Larribeau. Gender Differences in Tournament and Flat-Wage Schemes: An Experimental Study. Journal of Economic Psychology. Vol. 47 (2015), 103-115.

[31]Mayr, Ulrich, David Wozniak, Casey Davidson, David Kuhns, and William T. Harbaugh. "Competitiveness Across the Life Span: The Feisty Fifties." Psychology and Aging. Vol. 27, No. 2 (2011), 278-285.

[32] McBride, Alex. "Grutter v. Bollinger and Gratz v. Bollinger (2003)." http://www.pbs.org/wnet /supremecourt/future/landmark_grutter.html. [accessed April 23, 2016].

[33]Niederle, Muriel, and Lise Vesterlund. Do Women Shy Away From Competition? Do Men Compete Too Much? Quarterly Journal of Economics. Vol. 122, No. 3 (2007), 1067-1101.

[34]Niederle, Muriel, and Lise Vesterlund. Explaining the Gender Gap in Math Test Scores: The Role of Competition. Journal of Economic Perspectives. Vol. 24, No. 2 (2010), 129-144.

[35]Niederle, Muriel and Lise Vesterlund. Gender and Competition. Annual Review of Economics. Vol. 3 (2011), 601-630.

[36]Niederle, Muriel, Carmit Segal, and Lise Vesterlund."How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness." Management Science. Vol. 59, No. 1 (2012), 1-16.

[37]Potters, Jan, Martin Sefton, and Lise Vesterlund. "After You - Endogenous Sequencing in Voluntary Contribution Games." Journal of Public Economics. Vol. 89 (2005), 1399-1419.

[38]Samek, Anya C. "Is there a Gender Gap in Preschoolers' Competitiveness? An Experiment in the US." Journal of Economic Behavior and Organization. Vol. 92 (2013), 22-31.

[39]Shang, Jen, and Rachel Croson. "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods." The Economic Journal.

Vol. 119 (2009), 1422-1439.

[40]Shurchkov, Olga. Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints. Journal of the European Economic Association. Vol. 10, No. 5 (2012), 1189-1213.

[41]Solis, Hilda, and Keith Hall. Women in the Labor Force: A Databook. Technical Report no. 1034. US Department of Labor and US Bureau of Labor Statistics, Washington, DC. 2012.

[42]Wolfers, Justin. "Diagnosing Discrimination: Stock Returns and CEO Gender." Journal of the European Economic Association. Vol 4, No. 2/3, Papers and Proceedings of the Twentieth Annual Congress of the European Economic Association (2006), 531-541.

[43]Wozniak, David, William T. Harbaugh, and Ulrich Mayr. The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Preferences. Journal of Labor Economics. Vol. 32, No. 1 (2014), 161-198

7 Appendix A: Additional Results and Figures

Figure 1A: Round 2 Performance CDF in Baseline Condition

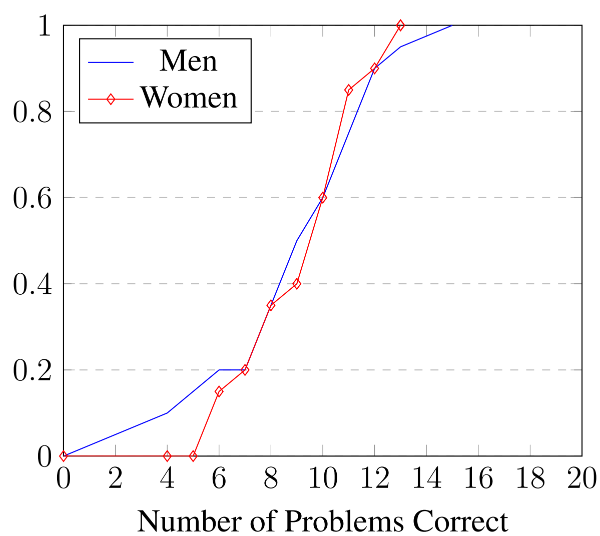


Figure 1B: Round 2 Performance CDF in GN Treatment

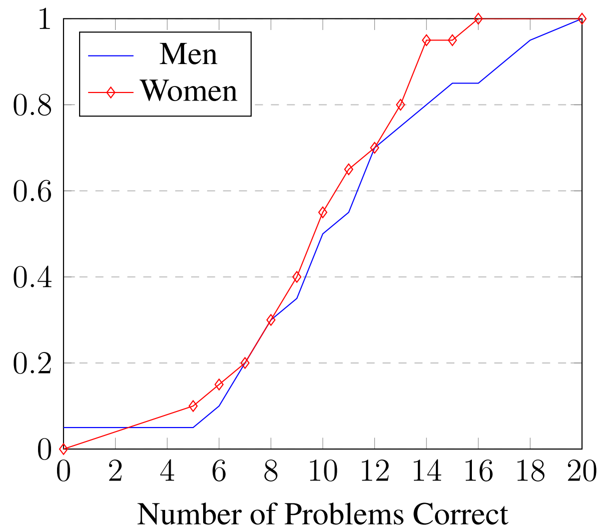


Figure 1C: Round 2 Performance CDF in MHP Treatment

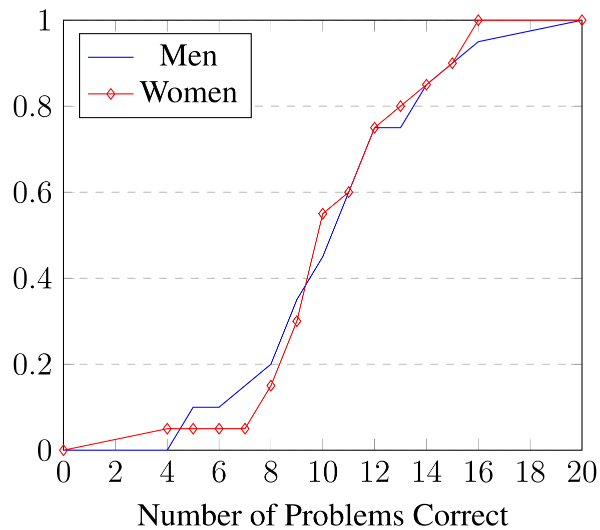
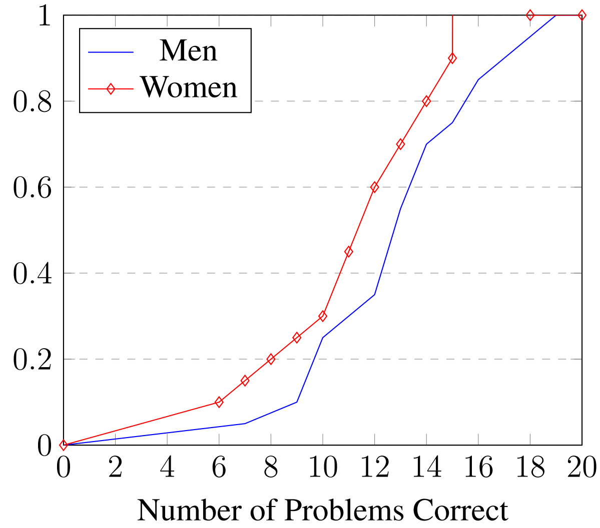


Figure 1D: Round 2 Performance CDF in WHP Treatment



Overconfidence Results: Both men and women in all treatments of this experiment are actually ”underconfident”, which means they guess a rank lower than their actual rank in Round 2. In the baseline condition only 17.5% of participants are overconfident about their relative rank, while 50% are underconfident. In the GN treatment only 7.5% of participants are overconfident, while 55% are underconfident. In the MHP treatment 15% of participants are overconfident and 40% are underconfident. For the WHP treatment 60% are underconfident and 7.5% are overconfident. Overall, there are no significant gender differences in overconfidence and underconfidence in any of the treatments. So while I will include controls for overconfidence in the regression analysis, there is little difference that is necessary to control for here between genders or treatments. Participants are quite inaccurate when it comes to guessing their rank. For the entire sample, 63 percent of participants submit incorrect guesses of their ranks.

Table 7: Round 4 Decision by Treatment (Probit, ME)

Columns:	(1)	(2)
Female	-0.325**	-0.290**
	(0.153)	(0.153)
GN	0.09	0.109
	(0.142)	(0.139)
MHP	-0.142	-0.137
	(0.145)	(0.138)
WHP	-0.045	-0.077
	(0.143)	(0.136)
GN x fem	0.236	0.223
	(0.211)	(0.205)
MHP x fem	0.422**	0.460**
	(0.209)	(0.205)
WHP x fem	0.175	0.186
	(0.217)	(0.203)
Improve Round 2		-0.024
		(0.017)
Rank Round 1		-0.055
		(0.034)
GPA		-0.059
		(0.063)
Age		0.028**
		(0.014)
Rankguess Round 2		-0.122**
		(0.054)
Competitive Preference		-0.017
		(0.025)
N	160	160
Ps. R^2	0.05	0.13
ClustVar	Individual	Individual

*Dependent variable is choice in Round 4. It is equal to 1 if tournament is chosen and zero if piece rate is chosen. Marginal Effects and Standard Errors from a probit regression are reported here. ***P<0.01, **P<0.05, *P<0.10. Clustering is at the individual level.*

Table 8: Summary Statistics of Participants

Major	Men	Women
Business and related fields	18	24
Computer Science	52	32
Engineering	7	16
Other	3	8
Race/ Ethnicity	Men	Women
African American	1	1
East Asian	33	46
South Asian	43	29
Hispanic	1	2
Middle Eastern	1	1
Caucasian	1	1
Mean GPA	Men	Women
	3.53	3.61
Mean Age	Men	Women
	24.4	23.75
N	80	80

Business and related fields includes majors like accounting, finanace, supply chain management, and management. Engineering includes all subsets of engineering (except civil engineering). Other majors is composed of biology, biotechnology, cognitive science, applied cognition and neuroscience, healthcare studies, and mathematics majors.

Table 9: Round 3 Optimal Tournament Entry

	Baseline	GN	MHP	WHP
Women	5	4	6	11
Men	6	7	6	12

Optimal tournament entry is calculated using a participant's rank in Round 2. If a participant was ranked first, then it is optimal for them to select the tournament. If they were ranked 2-4, they should choose the piece rate. If a participant was tied for first, their rank is still considered first, and it is optimal for them to enter the tournament, even though they may not have actually won the Round 2 tournament.

Table 10: Performance Variance by Round and Treatment

	Treatment	Round 1	Round 2	Round 3
(Women)	Base	4.89	3.17	5.94
(Men)	Base	8.87	21.84	14.93
	GN	10.09	11.50	10.80
	GN	21.10	16.16	14.57
	MHP	8.53	8.89	12.97
	MHP	11.42	13.36	16.13
	WHP	8.61	10.87	7.39
	WHP	9.80	10.20	8.04
	Overall	7.85	8.35	8.95
	Overall	13.30	15.58	13.82
	Treat Only	8.93	10.16	10.06
	Treat Only	14.48	13.67	13.45

Table reports performance variance by gender and treatment in each Round where the addition task is performed. Treat Only is the average variance of the treatment conditions only.

8 Appendix B: Experimental Instructions

The instructions for the baseline treatment are as follows. The baseline treatment is identical to the one in Niederle and Vesterlund (2007). However, the instructions are slightly different. After the baseline instructions are introduced, the instructions for different treatments are shown in the way that they modify the baseline instructions. The instructions are as they appeared on the zTree screen for participants. All of the social information is from the math task data in Wozniak, Harbaugh, and Mayr (2014).

8.1 Baseline Treatment Instructions

Welcome

Today, you are asked to complete four different rounds of this experiment. None of these rounds will take more than 5 minutes.

At the end of the experiment, you will receive a $5 show up fee and $5 for completing the experiment; in addition, we will randomly select one of the rounds and pay you based on your performance in that round.

The method we use to determine your earnings varies among rounds. Before each round, we will describe in detail how your payment is determined. At the end of the experiment, you will be asked to come to the back of the room individually, where you will receive your payment.

For each round you will be asked to calculate the sum of five randomly chosen two-digit numbers. On your computer screen you will see five sets of two-digit numbers, which you are tasked to add together.

You will be given 5 minutes to calculate the correct sum of as many of these problems as you are able to.You cannot use a calculator to determine this sum; however, you are welcome to write the numbers down using the provided scratch paper. You submit an answer, after entering it into the white text box, by clicking the submit button with your mouse.

When you enter your answer, the computer will inform you whether your answer is correct or not, and then another problem will appear on your screen. Your answers to the problems are anonymous. At the end of the 5 minute round, you will be informed of the total number of correct problems you solved.

Remember, the method we use to determine earnings varies across rounds, but the task will not.

In this experiment, there may be periods of waiting time. To minimize this, please do not forget to push the 'OK' or 'Submit' button when you are finished with a screen; otherwise everyone may be waiting for you.

If there are any questions throughout the experiment, please raise your hand quietly and an experimenter will assist you.

Round 1

If Round 1 is randomly selected for payment, then you get 50 cents per problem you solve correctly in the 5 minutes. Your payment does not decrease if you provide an incorrect answer to a problem. We henceforth refer to this payment as the piece rate payment.

Round 2

For this round, your payment depends on your performance relative to that of a group of other participants. Each group consists of four people; your group will be those who are seated in the same area as you.

If Round 2 is randomly selected for payment, then your earnings depend on the number of problems you solve compared to the three other people in your group. The individual who correctly solves the largest number of problems will receive $2 per correct problem, while the other participants will receive no payment. We refer to this as the tournament payment.

At the end of this 5 minute round you will still see the number of correct problems you will solve, but you will not be informed of how you did in the tournament until all four rounds have been completed. If there are ties, the winner will be randomly determined.

Round 3

Now, you will get to choose which of the two previous payment schemes you prefer to apply to your performance on the third round, the piece rate payment or the tournament payment.

If Round 3 is randomly selected for payment, then your earnings for this round are determined as follows:

If you choose the piece rate, you receive 50 cents per problem you solve correctly.

If you choose the tournament, your performance will be evaluated relative to the performance of the members of your group from the Round 2 tournament. The Round 2 tournament is the one you just completed. If you correctly solve more problems than the individual from your group who won the tournament in Round 2, then you will receive $2 per correct problem.

You will receive no earnings for this round if you choose the tournament and do not solve more problems correctly now, than the others in your group did in the Round 2 tournament. You will still be notified of the number of problems you completed correctly at the end of the 5 minute round.

Example: Say the members of Group ZY solved 8, 10, 12, and 9 problems correctly in Round 2. Say one of the members of Group ZY chooses the tournament to be applied to their performance in Round 3. For the tournament (and to earn $2 per correct problem) this member needs to solve at least 13 problems to win outright. If this member solves 12 problems and ties with the previous winner, the result will be decided randomly.

You will not be informed of how you did in the tournament until all four tasks have been completed. If there are ties, the winner will be randomly determined

Round 4

You do not have perform any addition task for the fourth and final round of the experiment. Instead, you may be paid for the number of problems you solved in Round 1.

However, you now have to choose which payment scheme you want applied to the number of problems you solved. You can either choose to be paid according to the piece rate, or according to the tournament.

If the fourth round is the one selected for payment, then your earnings for this round are determined as follows:

If you choose the piece rate, you receive 50 cents per problem you solved in Round 1. This will result in the exact same profit as Round 1.

If you choose the tournament, your performance will be evaluated relative to the performance of the other three participants of your group in the Round 1 piece rate. If you correctly solved more problems in Round 1 than they did, you receive $2 per correct problem.

You will receive no earnings for this task if you choose the tournament and did not solve more problems correctly in Round 1 than the other members of your group.

The next screen will remind you how many problems you solved correctly in Round 1, and then will ask you to choose whether or not you want the piece rate or tournament applied to your performance.

8.2 Gender Neutral Treatment

The modification for this treatment, in terms of the instructions, comes during the Round 3 instructions section. An extra paragraph to preface the feedback was added, in addition to a separate screen for to show the participants the social information prior to their remuneration scheme decision in Round 3. The italicized paragraph is identical in each of the treatment conditions, so I will only present it once.

Round 3

Now, you will get to choose which of the two previous payment schemes you prefer to apply to your performance on the third round, the piece rate payment or the tournament payment.

If Round 3 is randomly selected for payment, then your earnings for this round are determined as follows:

If you choose the piece rate, you receive 50 cents per problem you solve correctly.

If you choose the tournament, your performance will be evaluated relative to the performance of the members of your group from the Round 2 tournament. The Round 2 tournament is the one you just completed. If you correctly solve more problems than the individual from your group who won the tournament in Round 2, then you will receive $2 per correct problem.

You will receive no earnings for this round if you choose the tournament and do not solve more problems correctly now, than the others in your group did in the Round 2 tournament. You will still be notified of the number of problems you completed correctly at the end of the 5 minute round.

Example: Say the members of Group ZY solved 8, 10, 12, and 9 problems correctly in Round 2. Say one of the members of Group ZY chooses the tournament to be applied to their performance in Round 3. For the tournament (and to earn $2 per correct problem) this member needs to solve at least 13 problems to win outright. If this member solves 12 problems and ties with the previous winner, the result will be decided randomly.

You will not be informed of how you did in the tournament until all four tasks have been completed. If there are ties, the winner will be randomly determined

Before making your decision, you will receive some additional information about today's experiment which you may find helpful. Everyone in the room has the same information on their screen. After you read this information, you will be asked to choose whether you want the piece rate or the tournament applied to your performance.

(After seeing this, participants then were shown the social information for the specific treatment.)

An experiment very similar to this one has already been run at another university, in a laboratory like this one. What follows on this screen is selected performance information from those sessions.

Average Number of Problems Correct for Men in Round 3: 11

Of these 12 men, 10 chose the tournament in Round 3. Four of these men won the tournament.

Average Number of Problems Correct for Women in Round 3: 11

Of these 12 women, 3 chose the tournament in Round 3. All three of these women won the tournament.

8.3 Men as High Performers Treatment

The social information screen is as follows:

An experiment very similar to this one has already been run at another university, in a laboratory like this one. What follows on this screen is selected performance information from those sessions.

Average Number of Problems Correct for Men in Round 3: 14.5

Of these 12 Men, 8 chose the tournament in Round 3. Four of these men won the tournament.

Average Number of Problems Correct for Women in Round 3: 10.63

Of these 12 Women, 3 chose the tournament in Round 3. One of these women won the tournament.

8.4 Women as High Performers Treatment

The social information screen is as follows:

An experiment very similar to this one has already been run at another university, in a laboratory like this one. What follows on this screen is selected performance information from those sessions.

Average Number of Problems Correct for Men in Round 3: 10.64

Of these 12 Men, 5 chose the tournament in Round 3. One of these men won the tournament.

Average Number of Problems Correct for Women in Round 3: 14.33

Of these 12 Women, 7 chose the tournament in Round 3. Five of these women won the tournament.