Perception Matters: The Role of Task Gender Stereotype on Confidence and Tournament Selection

Brianna Halladay
University of California, Santa Barbara

November 22, 2016

Click here for the most current version

Abstract

It has been widely documented that women avoid competition even when they can benefit from potential rewards, suggesting there exists a gender difference in preference for competition. While a difference in preference for competition across genders is one explanation for the observed gender gap, it may be the case that a gender difference in beliefs about future performance are driving this observed gender gap. That is, it is possible women's forecast of their future performance in a tournament is lower than men's forecast of their future performance in a tournament. Using a laboratory experiment, I analyze differences in tournament entry using both a male stereotype task and a female stereotype task where performance is similar across genders for both tasks. I find that while women enter the tournament at significantly lower rates than men under the male stereotype task, more women than men enter the tournament under the female stereotype task. Additionally, greater female confidence and lower male confidence in the female stereotype task drives this observed difference in entry rates across tasks. These results suggest the effect of competitiveness on gender is not exclusively about a difference in preference for competition, but rather more consistent with a difference in beliefs about future performance in a tournament.

JEL Classifications: C91, J16, J22

1 Introduction

A gender gap in competitiveness is well documented in the literature. There are two possible channels that may be responsible for the observed gap. It is possible that this gender gap is driven mainly by women's distaste for competition. However, it is also possible that this gender gap is driven mainly by a difference in beliefs about future performance across genders. In their seminal paper, Niederle and Vesterlund (2007) note, "However to the extent that there are gender differences in the participants' beliefs about their future performance and that these influence tournament entry, our study incorrectly attributes such an effect to men and women having different preferences for performing in a competition." Currently, there is little research analyzing which channel is responsible for the gender gap in competitiveness. This study will serve as a first step in distinguishing between these two possible channels.

I use a controlled laboratory setting to analyze tournament selection using a task that carries a female gender stereotype. While much of the previous research has employed tasks carrying either a male gender stereotype or a "neutral" gender stereotype, my study adds to the body of research by using a task that previous research suggests carries a female stereotype.

This experiment exploits the gender stereotype of two tasks: an arithmetic task gender stereotyped in favor of men and a facial emotion recognition task gender stereotyped in favor of women. While both men and women have the capacity to excel at math, society traditionally views math as a male domain. Lundeberg et al. (1994) find evidence that in certain settings, specifically math, men were more confident than women. Nosek et al. (2002) echo this finding in their series of studies. In particular, the authors consistently find that women have more negative views of math than men both implicitly and explicitly.

Neurological research substantiates the belief that women are more intuitive and emotional beings than men. In tasks where men and women assess the approachability of an individual's face, interestingly, men take significantly longer to make judgments despite no significant difference in accuracy found (Hall et al., 2012). Also, using fMRI data, Hall et al. (2012) find that men show significantly more brain activity than women when making the facial judgments. As a control, brain activity was consistent across both genders when making determinations of sex based on a photograph, suggesting that social decisions are more taxing and complex for men. In an online study for the Edinburgh International Science Festival involving more than 15,000 people, 77% of women judged themselves are highly intuitive while only 58% of men judged themselves are highly intuitive. Additionally, women are better at expressing their emotions and are therefore viewed as more empathic than men. A researcher from the Greater Good Science Center, Emiliana Simon-Thomas (2007), notes, "We all know the stereotype: Women are better than men at taking other people's perspectives, feeling their pain, and experiencing compassion for them." In a study conducted in Germany, 63% of Germans believed women had better intuition about selecting a romantic partner. A majority of subjects also believed that women's intuition was overall better in regards to decisions about one's personal life. Gigerenzer et al. (2013) conclude that this confirms the validity of the common gender stereotype that women are between at inferring people's intentions.

While there are many factors that may influence an individual's decision to enter a competition in the field, the control of the laboratory allows me to isolate differences in behavior across task gender stereotype, absent concerns about discrimination for example. I test whether men enter a tournament more often than women when the task carries a female stereotype as opposed to a male stereotype.

For the purpose of this study, I replicated the study by Niederle and Vesterlund (2007) and added an additional treatment. Niederle and Vesterlund (2007) use a math task in their study, which can be thought of as carrying a male stereotype. For the additional treatment, I follow the experimental design of Niederle and Vesterlund (2007), but use a task that carries a female stereotype. My experiment uses a between-subjects design such that no subject participates in more than one treatment. Subjects participate in groups of four, two men and two women. In both treatments, subjects complete the task under a piece-rate compensation scheme and then under a tournament compensation scheme. Next, subjects can select which compensation scheme they prefer to determine compensation in the third round. In the final round, subjects choose whether they want their piece-rate performance paid according to a piece-rate compensation scheme or a tournament compensation scheme. While subjects are aware of their own absolute performance, at no point until the end of the experiment do subjects learn about their relative performance. Subjects are randomly paid for one of the four rounds. Additionally, subjects report beliefs about relative performance, participate in a Gneezy and Potters (1997) risk assessment, and complete the Cognitive Reflection Test (Frederick, 2005).

By comparing tournament entry rates across gender and across tasks, I determine if the task affects the tournament entry decision. Specifically, I analyze how female tournament selection responds to a task carrying a female stereotype compared to a task carrying a male stereotype. Higher tournament entry rates for men than women in both tasks is evidence for the distaste for competition channel, while observing more tournament entry for women in the female stereotype task than men provides evidence for the gender difference in beliefs about future performance channel.

The main finding is, controlling for performance, men enter the tournament significantly more than women in the male stereotype task, yet more women than men enter the tournament when the task carries a female stereotype, though this difference is not statistically significant. Additionally, female (male) confidence in relative performance appears higher when the task carries a female (male) stereotype and male (female) confidence in relative performance appears lower when the task carries a female (male) stereotype. Further, my findings in the math task replicate the findings of Niederle and Vesterlund (2007).

The paper proceeds as follows. Section 2 provides an overview of previous research related to the my research question. Section 3 identifies my experimental design while section 4 explores my hypotheses and predictions. I present my results in section 5 and discuss the implications of these results in section 6. Section 7 concludes.

2 Literature Review

In studies where a competitive environment is introduced and subjects do not have the ability to opt out of the competition, differences in changes in performance across genders depend on task and opponent gender. Gneezy et al. (2003) use a maze task to assess responses to competition. They find that while male performance in the task rises significantly in the presence of competition, female performance remains static when competing against men. However, in single sex tournaments, women did respond to competition. Gneezy and Rustichini (2004b) assess responses to competition among children. The authors observe no difference in running speed across boys and girls in the noncompetitive environment, but do observe a significant increase in the running speed of boys after introducing a competition. The running speed of girls was unaffected by the introduction of the competition. Replicating and extending the work by Gneezy et al. (2003), Günther et al. (2010) find that men respond to competition while women do not in the maze task (which they denote the male stereotype task), men and women both respond to competition in the word task (which they denote the neutral stereotype task), and women respond to competition while men do not in the memory/distraction task (which they denote the female stereotype task). It is not clear why this is considered a female stereotype task, nor do the authors have any evidence about gender differences in response to competition when subjects are given the opportunity to opt out of the tournament. Replicating and extending Gneezy and Rustichini (2004b), Dreber et al. (2011) find that there is no difference in responses to competition across genders for any of the tasks selected (running, jumping rope, and dancing). I find that in both tasks, men and women increase performance when the competitive environment is introduced. Additionally, there is no gender difference in this increase in performance in either task. Another important question to ask is how subjects respond when given the opportunity to opt out of a competition.

Gneezy and Rustichini (2004a) use two tasks varying in task gender stereotype. The authors selected a sports task as the male task and a verbal task as the female task. Men select into the tournament more often than women in both tasks but the gender gap is smaller with the female task. Given that subjects observe the gender of their selected opponent prior to making their compensation scheme decision, Gupta et al. (2005) find that men choose the competition more often than women. They suggest that risk aversion primarily drives the choice for women. Despite similar performances in the task for men and women, Niederle and Vesterlund (2007) find that men enter the tournament twice as often as women. The authors documented a preference for competition, but did not distinguish whether it is driven by distaste for competition or a belief about future performance in the competition. My study will help to distinguish whether the observed gender gap is arising out of a gender difference in preference for competition or a gender difference in beliefs about future performance. Again, despite similar performance, in the arithmetic task, I find that more men enter the tournament than women while in the facial emotion task, more women than men enter the tournament.

Recent research has provided evidence about environments that aid in closing the gender gap in response to competition. Using a maze task, Niederle and Yestrumskas (2008) assess gender differences in preferences for performing harder tasks in order to potentially increase one's payoff. Men select the harder task significantly more often than women despite no differences in performance and beliefs about relative performance. Given a more flexible choice such that subjects can adjust their choice between rounds, there is no longer a gender gap in selection of the harder task. In a study assessing differences in responses to competition across a matrilineal society and a patrilineal society, Gneezy et al. (2009) find that women choose the tournament more often than men in the matrilineal society while men choose the tournament more often than women in the patrilineal society. Team competition is another environment in which the gender gap in response to competition appears to eliminate the gender gap in response to competition (Dargnies, 2010; Kuhn and Villeval, 2014). Evidence provided by Booth and Nolen (2012) suggests that "nurture" may play a significant role in the commonly observed gender gap in tournament entry decisions. In a study assessing the tournament entry decisions of girls from single-sex schools and girls from coed schools, the authors find that girls from single-sex schools choose the competition at the same rate as boys. Using the math task from Niederle and Vesterlund (2007), Niederle et al. (2013) implement an affirmative action quota in the laboratory and find that women significantly increase their tournament entry while men significantly decrease their tournament entry. Cassar et al. (2016) find that creating tournament prizes that benefit one's child such as workplace daycare as opposed to monetary prizes erased the gender gap in tournament selection. Halladay (2016) finds that women will perform in a competition just as well as men in the presence of negative emotions towards the opponent.

A few studies have attempted to analyze the effect of a non-male stereotype task on tournament selection. Grosse et al. (2014) use a neutral stereotype task and two male stereotype tasks. The male stereotype tasks include a sports task where subjects throw tennis balls in a bucket and the math task as in Niederle and Vesterlund (2007). The neutral stereotype task involved a list of five words that a subject needed to place in order to form a grammatically correct sentence. While the authors observe difference in tournament selection rates across genders in the male stereotype tasks, there is no observed difference in the neutral task. However, these results are a bit concerning given the complex nature of the experimental design and the fact that performance was not similar across genders for the selected tasks. To avoid this concern, I have chosen two tasks that ex-ante have similar performance across genders.

In a study perhaps most closely related to my research, Dreber et al. (2014) utilize both a male stereotype task and a female stereotype task in order to assess gender differences in self-selection into tournaments. The authors choose a math task similar to the one used in Neiderle and Vesterlund (2007) for the male stereotype task and a crossword-like puzzle for the female stereotype task. While math tasks undoubtedly carry a male dominant perception, gender perception is less clear for verbal tasks. IT is not clear that verbal tasks carry a female stereotype. A meta-analysis by Hyde and Linn (1988) of gender differences in verbal tasks finds an insignificant effect size of +0.11 in favor of females while other studies mentioned above have used verbal tasks as "neutral". Additionally, previous studies assessing gender differences in self-concept in verbal tasks find no gender difference (Hyde et al., 1990; Meece et al., 1982; Skaalvik and Skaalvik, 2004). Also, by selecting a math and a verbal task in conjunction with a within-subjects design, there are certain problematic interactions regarding how individuals perceive their math versus verbal ability. Because subjects are likely to view math and verbal abilities as dichotomous and are completing both tasks, the selection of these two specific task categories can influence subjects' choices and behavior. The authors find that men select into the tournament more than women in both the math and verbal tasks but the difference in tournament selection is only significant in the math task. While this result is interesting, given the concerns above, more research is warranted. My study will help to clarify the validity of these results absent the concerns mentioned above.

3 Experimental Design

The experimental design follows almost exactly that of Niederle and Vesterlund (2007). Instructions were adapted from the instructions available on the Lise Vesterlund's website and available in the appendix.

The experiment takes place in four main stages, lasting five minutes each followed by a few simple tasks. The stages differ by the compensation scheme while the task remains the same across all stages. Participants only learn of the compensation scheme differences immediately before each stage.

There are two treatments differing by the task. One task is identical to the math task of Niederle and Vesterlund (2007) while the other task requires subjects to identify facial emotions in a series of images.

In the math task subjects must add up a series of five randomly selected two-digit numbers. Subjects have five minutes to correctly answer as many problems as they can. Subjects cannot use a calculator, but can use the provided scratch paper.

The facial emotion task utilizes professionally classified images obtained from The Great Good Science Center at the University of California, Berkeley. In each stage subjects view a series of 15 images, each appearing on the screen for two seconds, and subjects attempt to correctly select the depicted emotion out of four options. Subjects have 20 seconds to submit their answer after each image is displayed. Subjects know that the same emotions can repeat but the same image will never appear more than once.

In both tasks, subjects immediately observe whether the answer submitted was correct or not while also generating a new series of numbers or image depending on the treatment. Consistent with Niederle and Vesterlund (2007), in each stage, subjects continually observe their absolute performance, but subjects never observe their relative performance until the end of the experiment.

In the first stage, payment follows a piece-rate compensation scheme. For each correct answer, subjects earn 50 cents.

In the second stage, payment follows a tournament compensation scheme. Payment depends on the performance of the other three subjects in a participant's group. If a participant submits more correct answers than her other three group members, she receives $2 for each correct answer, and nothing otherwise.

The third stage gives subjects the opportunity to select which compensation scheme they want applied to their performance in the third stage. If a subject selects the piece-rate compensation scheme, she receives 50 cents per correct answer in the third stage. If a subject selects the tournament compensation scheme, her earning depend on her performance in the third stage relative to the performance of her other three group members in the second stage tournament. If the subject submits more correct answers in the third stage than her other three group members did in the second stage tournament, she receives $2 per correct answer, and nothing otherwise. Niederle and Vesterlund (2007) note that they designed this stage specifically as an individual choice, independent of the choices of the other three group members. They also note that by comparing a subjects Task 3 performance to others' Task 2 performance, subjects need only base their choice on their beliefs about relative performance in a tournament and not beliefs about who will or will not enter the tournament. Lastly, Niederle and Vesterlund (2007) explain that an observed gender gap in the choices of subjects in Task 3 indicated a gender difference in preference for performing in competitive environments.

The fourth, and final main stage gives subjects another opportunity to select a compensation scheme. However, in the fourth stage, subjects do not need to complete any additional tasks. Task 4 was designed to distinguish between a difference in preference for competition across genders or perhaps differences in overconfidence, risk, or feedback aversion across genders. Subjects decide whether they want their performance from the first stage piece-rate round to be submitted to a tournament or to remain piece-rate. If a subject selects the piece-rate compensation scheme, she receives 50 cents per correct answer. If a subject selects the tournament compensation scheme, her earnings depend on her performance in the first stage relative to the performance of her other three group members in the first stage. If the subject submits more correct answers in the first stage than her other three group members did, she receives $2 per correct answer, and nothing otherwise. Before submitting their choices, subjects are reminded of their first stage performance.

At the conclusion of the fourth stage, subjects report their beliefs about their relative performance. First subjects guess their rank in the first stage piece-rate task and then subjects guess their rank in the second stage tournament. Guesses are incentivized. Subjects earn $1 for each correct guess.

Additionally, subjects complete a risk elicitation. The risk elicitation follows the adaptation of the Gneezy and Potters (1997) method utilized by Charness and Gneezy (2010). Subjects have 200 tokens ($2) to keep or invest in a risky project. Subjects can choose to invest any integer value of their tokens between zero and 200, inclusive. The risky project has a 50% chance of success. If the project is successful, a subject receives 2.5 to 1 on the invested amount. If the project fails, whatever is not invested is kept. We randomly draw a number between one and four. If we draw a one or a two, the project is successful. If we draw a three or a four, the project is a failure.

After observing the risk elicitation results, subjects answer three logical questions based on the Cognitive Reflection Test (Frederick, 2005). Subjects have five minutes to answer all three questions. This experiment uses three questions in the same spirit of the Cognitive Reflection Test, but not the exact three proposed by Frederick (2005) because of a concern that a significant portion of the subject pool had been previously exposed to the Cognitive Reflection Test. The three questions used follow those of Gillen, Snowberg, and Yariv (2015).Subjects can earn 50 cents per correct answer. After the five minutes elapse, subjects are asked how many of the three questions they believe they answered correctly and also what percent of subjects they believe answer more questions correctly than they did.

One of the four main stages is randomly selected for payment. In addition to payment from the randomly selected stage, subjects earn a $5 show-up fee, payment for correct guesses about relative performance, payment from the results of the risk elicitation, and payment from the Cognitive Reflection Test.

The experiment was programmed in zTree (Fischbacher, 2007). All sessions took place in the Experimental and Behavioral Economics Laboratory at the University of California, Santa Barbara. The University's ORSEE system was used to recruit subjects. Three or four groups of four participants participated in each session. Subjects were seated four per row, each row consisting of two males and two females although gender was never explicitly discussed. Groups consisted of the other participants sitting in the same row as a subject. A total of 154 students participated, 76 subjects (38 males and 38 females) in the math task, and 72 subjects (36 males and 36 females) in the facial emotion task. No subject participated in more than one session or more than one treatment. Average earnings were $14.48 and each session lasted approximately an hour.

4 Hypotheses

This experiment addresses three main hypotheses about how women and men respond to competitive environments.

Both experimental tasks, the arithmetic task and the facial emotion task, involve very little prior knowledge. There is no general consensus that one gender is better than another at either task but each task carries a gender stereotype. In their selection of this specific arithmetic task, Niederle and Vesterlund (2007) note that in a meta-analysis of over 100 gender studies, there was no observed difference in male and female performance on computational math tasks. In a study assessing gender differences in facial emotion recognition, Hoffmann et al. (2010) find that when the emotional stimuli is highly expressive, men and women perform equally well on the task. Additionally, Hoffmann et al. (2010) point out evidence from previous research is mixed, some studies report no gender difference while others suggest a female advantage. The studies that report a female advantage provide very small mean effect sizes. Therefore, it seems very reasonable to hypothesize that there will be no gender difference in performance for either task. Equal performance across genders is a very important task characteristic for the purpose of this study, allowing me to rule out gender differences in tournament entry decisions arising from differing performance levels. To summarize:

Hypothesis 1: There is no gender difference in performance for either task.

There appears to be a significant confidence gap between men and women. The Harvard Business Review mentions a Hewlett-Packard study that found women would only apply for a promotion when they felt they met 100% of the qualifications while men would apply when they met at least 60% of the qualifications. However, because evidence discussed above alludes to the fact that women are more confident in their ability to harness their "female intuition" as opposed to their math ability, I hypothesize that women will be more confident in the facial emotion task. Succinctly,

Hypothesis 2: Women are more confident than men in the facial emotion task, while men are more confident than women in the arithmetic task.

If women are more confident in their relative ability in the facial emotion task, it is logical to hypothesize that women will therefore be more willing to enter a tournament in the facial emotion task as opposed to the math task. Additionally, in their experiment, Charness et al. (2016) find that differences in confidence across genders can explain the gender difference in tournament entry rates. Claude Steele (1997) details how an individual who identifies with a specific domain, in particular a domain that carries a relevant stereotype, places themselves in a potentially self-threatening position. As a result, in many instances, individuals will either avoid the identification all together or disidentify with the relevant domain. In the arithmetic task, a female who selects the tournament (therefore identifying with math) is thus in a position to possibly validate the existing stereotype that math is for men if she fails to win the tournament. Consequently, females avoid the tournament in the arithmetic task. Nosek et al. (2002) point out that because math is so heavily aligned with men, the stronger the association, the more positively men view math, hence one can expect to view higher tournament entry rates for men in the arithmetic task. With the facial emotion task, the stereotype threat is absent for women and therefore I expect to see increased tournament selection. Explicitly:

Hypothesis 3: Women (men) will select into the tournament more often than men (women) in the facial emotion task (math task), not controlling for confidence.

5 Results

All statistical tests performed in the following analysis are two-tailed unless otherwise specified.

Finding 1: Regardless of task, there is no difference in male and female performance under either the piece-rate and tournament compensation schemes.

Consistent with Niederle and Vesterlund (2007), men and women perform equally well in the arithmetic task under both the piece-rate and tournament compensation schemes. Men solve an average of 9.55 tasks under the piece-rate compensation scheme while women solve 8.58 tasks under the piece-rate compensation scheme. The null hypothesis that these averages are equal cannot be rejected by a two-sided test (p=0.233). The results are similar under the tournament compensation scheme. Men solve an average of 11.42 tasks while women solve 10.13 tasks. Again, the null hypothesis that these averages are equal cannot be rejected by two-sided test (p=0.183). Additionally, using a Kolmogorov–Smirnov test there is no difference in the distribution of male and female performance on the arithmetic task under either the piece-rate and tournament compensation schemes (p=0.449 and p=0.287).

Furthermore, there is no difference in male and female performance in the facial emotion task in either the piece-rate or tournament compensation schemes. Men correctly identify an average of 8.19 and 9.72 emotions under the piece-rate and tournament compensation schemes, respectively. Similarly, women correctly identify 8.08 and 9.77 emotions under the piece-rate and tournament compensation schemes, respectively. The differences between these figures are not significant using a two-sided test of averages (p=0.806 and p=0.894) and there is also no statistically significant difference in the distribution of male and female performance under either compensation scheme using a Kolmogorov-Smirnov test (p=0.413 and p=0.965). Figure 1 and Figure 2 depict the cumulative distribution functions for men and women in both tasks under both the piece-rate and tournament compensation schemes.

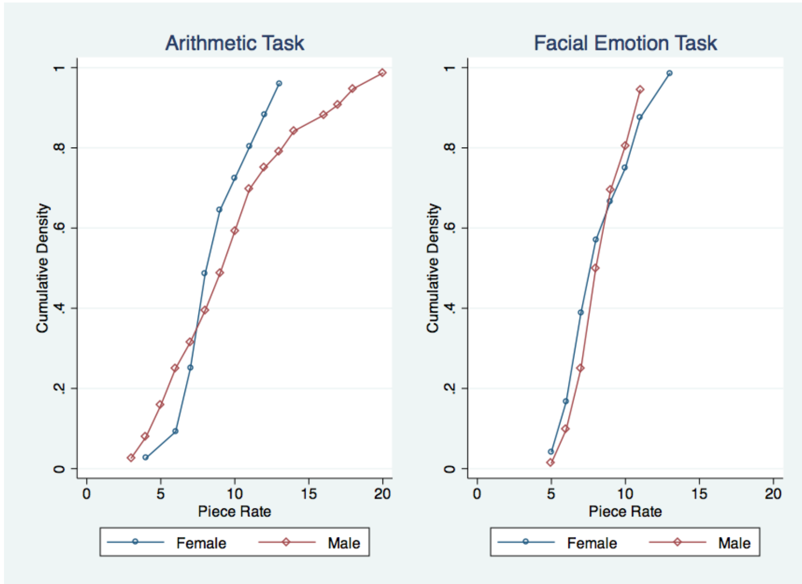

Figure 1: There is no difference in male and female performance in either treatment under the piece-rate compensation scheme.

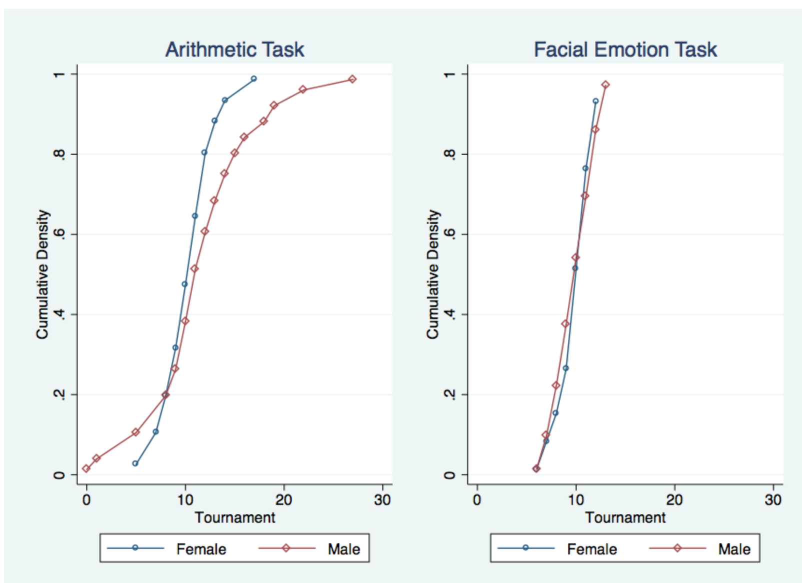

Figure 2: There is no difference in male and female performance in either treatment under the tournament compensation scheme.

Finding 2: In both tasks, both men and women perform significantly better in the tournament than in the piece-rate task. Additionally, for both tasks, there is no gender difference in the increase in performance between the piece-rate and tournament compensation schemes.

In both tasks, male and female performances across the piece-rate and tournament compensation schemes are highly correlated. Additionally, in both tasks, both genders perform significantly better in the Task 2 tournament (p<0.000 for all four one-sided hypothesis tests). Furthermore, the improvement in performance from the piece-rate to the tournament does not differ across genders for either task (p=0.546 for the arithmetic task and p=0.727 for the facial emotion task). The distributions of performance improvements across gender are also not statistically different for either task (p=0.651 and p=0.615).

Because we are unable to reject the null hypothesis that performance is equal across genders for both tasks, we can investigate how tournament selection differs across genders and compare these results to those of Niederle and Vesterlund (2007) without too much concern that any observed differences are due to gender differences in performance.

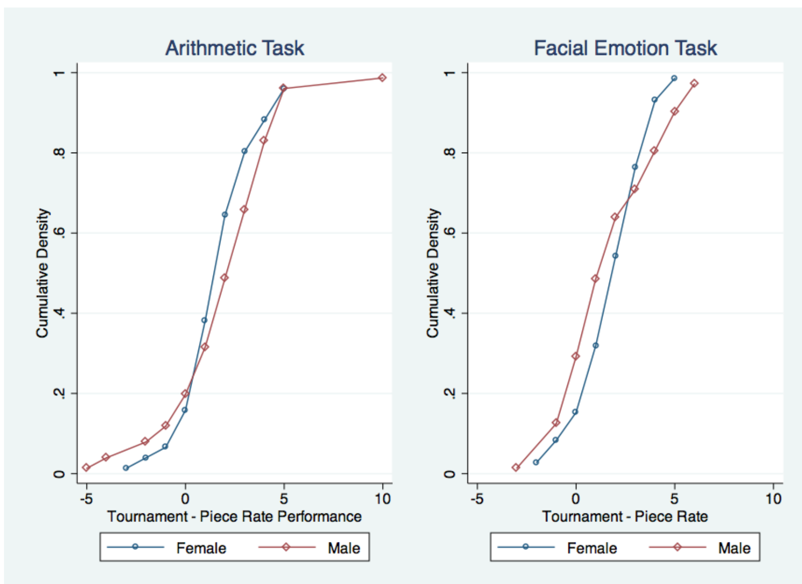

Figure 3: There is no difference in the increase in male and female performance between the piece-rate and tournament compensation schemes in either task.

Finding 3: In the arithmetic task, males enter the Task 3 tournament significantly more than females despite the similar Task 1 and Task 2 performances.

Given the similar performances by men and women under both the piece-rate and tournament compensation schemes in either task, one might expect, absent a concern about confidence, tournament entry decisions should also not vary by gender. In contrast to this logical expectation, Niederle and Vesterlund (2007) observe males choosing to enter the tournament significantly more than females. In the arithmetic task, I find a similar result. While 66% (25/38) of males choose to enter the tournament in Task 3, only 37% (14/38) of females choose to enter the tournament in Task 3 despite the similar Task 1 and Task 2 performances. This difference in tournament entry rates for the arithmetic task is highly significant using a one-tailed test of proportions (p=0.006).

Finding 4: In the facial emotion task, 45% more women than men choose to enter the Task 3 tournament.

In the facial emotion recognition task, the tournament selection results contrast with those of the arithmetic task. More women than men choose to enter the tournament in the facial emotion recognition task. While 31% (11/36) of men choose to enter the Task 3 tournament, 45% (16/36) of women choose to enter the Task 3 tournament, a nearly 50% increase. Strikingly, 45% more women than men choose to enter the tournament in the facial emotion task. A one-tailed test of proportions (consistent with the directional hypothesis) gives p = 0.112, which is not significant but is certainly in the right direction.

Figure 3 depicts tournament entry proportions by gender and task. Running a probit regression of Task 3 compensation choice as a function of a treatment dummy variable, a gender dummy variable, and an interaction of these two variables allows a test of a test of differences-in-differences between male and female tournament selection in Task 3. Each of the reported marginal effects is highly significant suggesting there is something systematically different about tournament selection decisions across tasks (p=0.003, p=0.012, and p=0.009, respectively). A joint hypothesis test about the effect of gender is significant as well as a joint hypothesis test about the treatment effect (p=0.020 and p=0.008, respectively).

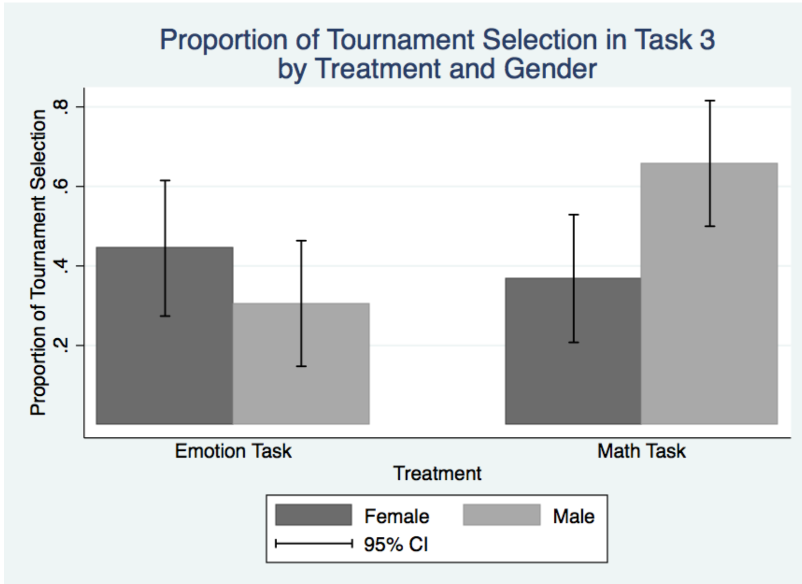

Figure 4: While significantly more males than females enter the tournament in the arithmetic task, there is no statistically-significant difference in tournament entry rates across genders in the facial emotion task.

Finding 5: Tournament entry decisions are not driven by performance differences across gender for either the arithmetic task or the facial emotion task.

Niederle and Vesterlund (2007) do not find significant evidence that tournament entry decisions are the result of a difference in performance for those subjects opting into the tournament and those selecting piece-rate across genders. They do report a marginal result for a difference in tournament performance between men selecting into the tournament and men selecting piece-rate. Similarly, for the arithmetic task, I do not find significant evidence that a difference in performance is responsible for the tournament selection decision. I find two marginally-significant results. One is the difference in tournament performance for women who select into the tournament and women who select the piece-rate (p=0.080). A second is the difference in the increase in performance from Task 1 to Task 2 for men who enter the tournament versus men who select piece-rate (p=0.066). In the facial emotion task, there is no evidence that performance differences drive tournament entry differences for either gender.

Finding 6: While gender is a significant predictor of the tournament entry decision in the arithmetic task, gender does not significantly effect the tournament entry decision in the facial emotion task.

In a probit regression, Niederle and Vesterlund (2007) find that gender was the only significant predictor of the tournament entry decision, controlling for tournament performance and the increase in performance from piece-rate to tournament. Table 1 presents my probit regression results. In the arithmetic task, I find very similar results, with again, only the gender dummy variable yielding a significant coefficient. However, in the facial emotion task, none of the three variables are significant, including gender.

Table 1: Probit of Task 3 Tournament Choice

	Arithmetic Task	Facial Emotions Task
	Coefficient (p-value)	Coefficient (p-value)
Female	-0.691**	0.356
	(0.023)	(0.245)
Tournament Performance	0.051	-0.027
	(0.261)	(0.788)
Tournament - Piece Rate Performance	0.101	0.127
	(0.140)	(0.156)

P-values in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Additionally, while past performance does not appear to be driving the difference in tournament entry, future performance also does not correlate with the tournament entry decision. There is no evidence of a difference in Task 3 performance for men and women in either task regardless of the tournament entry decision (p=0.165, p=0.196, p=0.633, p=0.620).

Finding 7: Gender differences in confidence across tasks appear to help explain the difference in tournament entry rates across genders.

In regards to beliefs about relative performance, while Niederle and Vesterlund (2007) find that both men are women are overconfident, specifically that men are more overconfident than women, I find that only men are overconfident in both tasks using a Fisher's exact test of independence between the distribution of guess rank and actual rank (p=0.000 and p=0.030). A Fisher's exact test of the independence between the distribution of guessed rank and actual rank for women in the arithmetic task and the facial emotion task yield p=0.226 and p=0.847, respectively. It does not appear that women are overconfident in either task. Women appear to be calibrating their beliefs about their performance correctly, while men fail to calibrate their beliefs correctly in either tasks.

Niederle and Vesterlund (2007) observe that the guesses of men and women differ significantly. Using a Fisher's exact test of independence of the distributions for men and women, I also observe that the guesses of own rank by men and women differ significantly at the 10% level (p=0.070) in the arithmetic task, but the guesses of men and women do not differ significantly in the facial emotion task (p=0.376). Additionally, while there is little evidence that the distribution of guesses of women differs across tasks (p=0.131), there is mild evidence that the distribution of guesses for men differs across tasks (p=0.093). Specifically, men have lower confidence in the facial emotion task. While 50% of men guess the top rank in the arithmetic task, only 31% of men selected the top rank in the facial emotion task. A one-sided test of a difference in these proportions yields a significant result (p=0.044). In the arithmetic task, 24% of women guess the top rank while 42% of women guess the top rank in the facial emotion task. A one-sided test of a difference in proportions is significant (p=0.049). There does appear to be some evidence that men are losing confidence while women are gaining confidence in the facial emotion task.

Table 2 presents the results from an ordered probit regression of guessed tournament rank (guessed ranks of 1, 2, or 3, excluding guesses of 4) as a function of gender, tournament performance, and the increase in performance from piece-rate to tournament. Both both tasks, I find an insignificant coefficient on the gender dummy variable, suggesting that once controlling for performance, confidence is similar across genders for either task. Furthermore, when running an ordered probit of guessed tournament rank conditioning on gender and controlling for both tournament performance and the increase in performance from Task 1 to Task 2, as well as including a task dummy variable and an interaction between female and the task dummy variable, I find evidence of a significant difference in confidence across tasks. There does appear to be a significant treatment effect for women. In particular, the coefficient on the interaction term is significant at the 10% level (p=0.075). Figure 4 and Figure 5 provide additional qualitative evidence that task affects both male and female confidence. Specifically, male confidence falls in the facial emotion task while female confidence increases.

Table 3 provides separate probit regressions predicting tournament selection as a function of gender, tournament performance, the increase in performance from the piece-rate to the tournament, and guessed tournament rank, a proxy for confidence. While it appears that the gender gap in tournament selection cannot be explain by beliefs about relative performance in the arithmetic task, once controlling for confidence, there exists no gender gap in tournament entry in the facial emotion task. Due to the finding that there is a treatment effect on confidence across tasks, I tested a pooled probit regression allowing for the treatment effect as well as an interaction between the

Table 2: Ordered Probit of Task 2 Tournament Guessed Ranks

	Arithmetic Task	Facial Emotions Task	Pooled
	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)
Female	0.337	-0.417	-0.440
	(0.300)	(0.153)	(0.122)
Tournament Performance	-0.118**	-0.287**	-0.150***
	(0.031)	(0.21)	(0.001)
Tournament - Piece Rate Performance	-0.190**	-0.212***	-0.224***
	(0.043)	(0.008)	(0.000)
Treatment			-0.491
			(0.104)
Treatment*Female			0.725*
			(0.075)

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01

treatment and guessed tournament rank. With this pooled regression, the gender coefficient is no longer significant (p=0.184), therefore the results suggest that the difference in tournament entry rates across tasks are driven by a difference in confidence across tasks.

Table 3: Probit of Task 3 Tournament Choice

	Arithmetic Task		Facial Emotions Task		Pooled
	Coefficient (p-value)		Coefficient (p-value)		Coefficient (p-value)
	(1)	(2)	(1)	(2)	(1)
Female	-0.692**	-0.674**	0.253	0.057	-0.308
	(0.033)	(0.040)	(0.417)	(0.864)	(0.184)
Tournament Performance	0.027	0.021	-0.007	-0.15	0.008
	(0.589)	(0.685)	(0.948)	(0.238)	(0.864)
Tournament - Piece Rate Performance	0.059*	0.045	0.105	0.031	0.017
	(0.051)	(0.619)	(0.246)	(0.754)	(0.789)
Guessed Tournament Rank		-0.142		-0.774***	-0.665***
		(0.592)		(0.003)	(0.003)
Treatment					-0.364
					(0.554)
Treatment*Guessed Tournament Rank					0.419
					(0.193)

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01

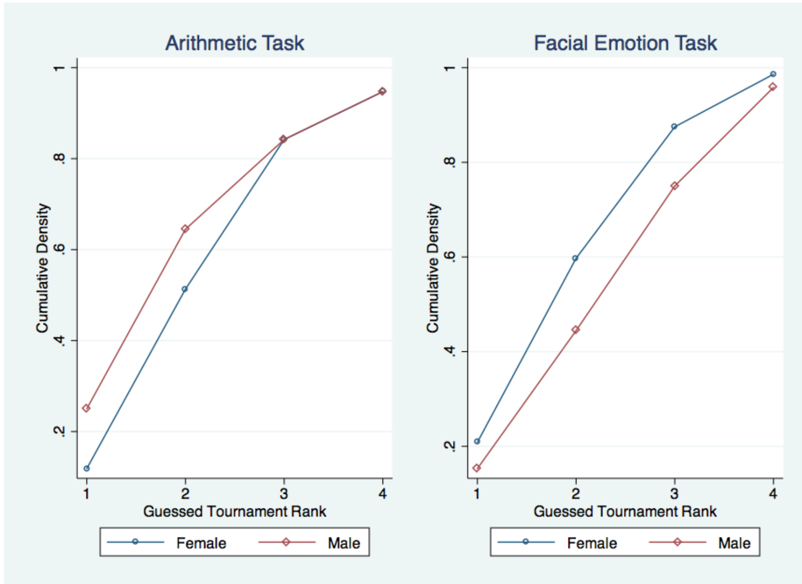

Figure 5: While the guess distribution for males lies above that of females in the arithmetic task, it appears that the guess distribution for females always lies above the distribution for males in the facial emotion task.

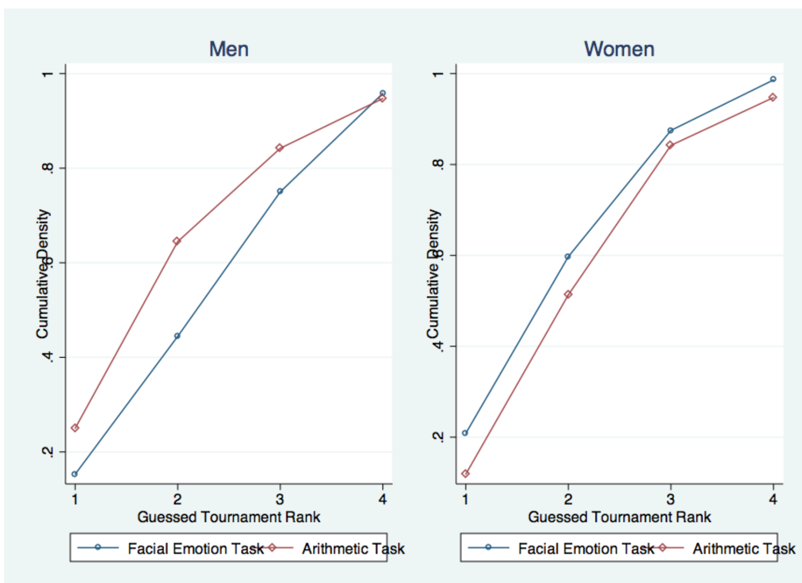

Figure 6: The guess distribution for males in the arithmetic task lies above that of males in the facial emotion task, while it appears that the guess distribution for females in the facial emotion task lies above that of females in the arithmetic task.

Finding 8: Controlling for risk preference eliminates the gender difference in tournament entry rates in the arithmetic task.

Entering a tournament requires a subject to undertake some level of risk. Research suggests that levels of risk aversion differ for men and women. Subsequently, the differing tournament entry rates may be due to a difference in risk preferences. Women invest significantly less than men. While men invest 69% of their endowment, women invest roughly 51% of their endowment (p=0.000). This difference is also significant using a Wilcoxon Rank-Sum test (p=0.000).

The probit regressions of Table 4 suggest the observed difference in tournament entry rates in the arithmetic task mostly vanish when controlling for risk preferences. A one-sided hypothesis test for the coefficient on female suggests the gender gap still partially remains (p=0.057). Additionally, I cannot reject the null hypothesis that there the investment percentage among men who choose the tournament and men who select piece-rate for either task is equal (p=0.473 and p=0.153). Similarly, I cannot reject the null hypothesis that the investment percentage among women who choose the tournament and women who select piece-rate for either task is equal (p=0.672 and p=0.104). Because of these results, risk preference does not seem to be the main driving force behind the difference in tournament entry rates across tasks.

Table 4: Probit of Task 3 Tournament Choice Controlling for Risk Preference

	Arithmetic Task			Facial Emotions Task		
	Coefficient (p-value)			Coefficient (p-value)		
	(1)	(2)	(3)	(1)	(2)	(3)
Female	-0.694**	-0.581	-0.581	0.254	0.389	0.196
	(0.033)	(0.113)	(0.113)	(0.417)	(0.239)	(0.578)
Tournament Performance	0.027	0.022	0.018	-0.007	0.029	-0.097
	(0.589)	(0.674)	(0.734)	(0.948)	(0.801)	(0.459)
Tournament - Piece Rate Performance	0.059	0.065	0.055	0.105	0.084	0.012
	(0.506)	(0.464)	(0.559)	(0.246)	(0.375)	(0.905)
Percent Invested		0.471	0.406		1.525**	1.349**
		(0.499)	(0.573)		(0.014)	(0.040)
Guessed Tournament Rank			-0.105			-0.721***
			(0.700)			(0.007)

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Finding 9: While there is a significant differences in tournament selection in Task 4 in the arithmetic task, there is difference in tournament selection in Task 4 in the facial emotion task.

In Task 4, subjects can select whether to submit their Task 1 piece-rate performance to a tournament or to a piece-rate compensation scheme in the event that Task 4 is the task randomly selected for payment. Subjects do not need to complete any additional tasks in Task 4 so the choice does not involve the prospect of additional effort. In the arithmetic task, 45% of men select the tournament while only 21% of women select the tournament. In the facial emotion task, 25% of both men and women select the tournament. Using a Fisher's exact test, men and women differ in their Task 4 choice in the arithmetic task (p=0.050) but not in the facial emotion task (p=1.000).

Consistent with Niederle and Vesterlund (2007), there is no difference in piece-rate performance for women who select the tournament versus women who select piece-rate in the arithmetic task (9.125 and 8.43, p=0.468). However, there is a difference in piece-rate performance for men who select the tournament versus men who select the piece-rate in the arithmetic task (11.76 and 7.76, p=0.004). In the facial emotion task, there is no difference in performance for individuals who select the tournament versus the piece-rate compensation scheme for either gender (p=0.692 and p=0.864). These results provide evidence that performance is not driving the tournament selection decision.

Finding 10: Differences in confidence across genders can explain the difference in tournament entry rates in Task 4.

Beliefs about relative performance in Task 1 are consistent with beliefs about relative performance in Task 2. This suggests that the competitive environment is not affecting beliefs about relative performance. A Fisher's exact test again shows that while men are overconfident in both tasks (p=0.003 and p=0.020), women are overconfident in neither task (p=0.225 and p=0.193). Again, women are calibrating correctly while men fail to accurately calibrate beliefs about relative performance. I observe that the distribution of guesses of men and women differ significantly in the arithmetic task (p=0.000), but the guesses of men and women do not differ significantly in the facial emotion task (p=0.326). However, there is evidence that the distribution of guesses for men differs across tasks (p=0.000) while the distributions of guess for women does not differ across tasks (p=0.880). While it appears that men are overconfident in both tasks, I do not observe the difference in tournament entry rates for men and women across both tasks.

Table 5 presents results from an ordered probit regression of guessed piece-rate rank (guessed ranks of 1, 2, or 3, excluding guesses of 4) on gender and piece-rate performance. I find a significant coefficient on the gender dummy variable, suggesting that despite controlling for performance, there remains a gender difference in confidence for both tasks. In particular, women guess significantly lower ranks than men in the arithmetic task, while women guess significant higher ranks than men in the facial emotion task. Furthermore, running an ordered probit regression of guessed piece-rate rank, conditioning on gender and controlling for piece-rate performance, I find evidence of a difference in male confidence across tasks but not a difference in female confidence across tasks. Specifically, the highly significant coefficient on the treatment variable in column 3 of Table 5 illustrates that men are significantly less confident in the facial emotion task than the arithmetic task.

Table 5: Ordered Probit of Task 1 Piece Rate Guessed Ranks

	Arithmetic Task	Facial Emotions Task	Males	Females
	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)
Female	1.050***	-0.545*		
	(0.000)	(0.075)		
Piece Rate Performance	-0.201***	-0.278***	-0.274***	-0.156**
	(0.000)	(0.000)	(0.000)	(0.019)
Treatment			-1.299***	0.204
			(0.000)	(0.469)

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 6 presents probit regression results that suggest controlling for confidence eliminates the observed gender gap in tournament selection.

Table 6: Probit of Task 4 Tournament Choice

	Arithmetic Task		Facial Emotions Task	
	Coefficient (p-value)		Coefficient (p-value)	
	(1)	(2)	(1)	(2)
Female	-0.567*	-0.323	0.001	-0.231
	(0.076)	(0.354)	(0.997)	(0.503)
Piece-Rate Performance	0.138***	0.077	0.018	-0.086
	(0.005)	(0.178)	(0.831)	(0.380)
Guessed Piece-Rate Rank		-0.567**		-0.574**
		(0.033)		(0.029)

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Despite controlling for confidence, risk, and feedback aversion through the decision to submit one's piece-rate score to a tournament and beliefs about relative performance in the tournament, Table 7 illustrates that there is still a gender gap in tournament performance in the arithmetic task, significant at the 10% level. While Niederle and Vesterlund (2007) take this as evidence that women and men differ in their preference for competition, this statement needs clarification. This gender gap is not observed in the facial emotion task and is eliminated when controlling for the treatment effect on confidence. A restatement of Niederle and Vesterlund (2007) is that a difference in preference for competition across genders is only part of the story. A difference in beliefs about future performance appears to significantly affect tournament selection decisions across genders.

Table 7: Probit of Task 3 Tournament Choice

	Arithmetic Task			Facial Emotions Task		
	Coefficient (p-value)			Coefficient (p-value)		
	(1)	(2)	(3)	(1)	(2)	(3)
Female	-0.694**	-0.674**	-0.622*	0.254	0.057	0.083
	(0.033)	(0.040)	(0.062)	(0.417)	(0.864)	(0.806)
Tournament Performance	0.027	0.021	0.013	-0.007	-0.150	-0.128
	(0.589)	(0.685)	(0.814)	(0.948)	(0.238)	(0.319)
Tournament - Piece Rate Performance	0.059	0.045	0.079	0.105	0.031	0.049
	(0.506)	(0.619)	(0.429)	(0.246)	(0.754)	(0.635)
Guessed Tournament Rank		-0.142	-0.068		-0.774***	-0.750***
		(0.592)	(0.341)		(0.003)	(0.004)
Submitting the Piece Rate			0.359			0.443
			(0.341)			(0.242)

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01

6 Discussion

My data confirm that altering subject beliefs through the perception of the task significantly affects tournament entry rates. While more men than women enter the tournament with the male stereotype task, more women than men enter the tournament with the female stereotype task, though this difference is not significant. This observed difference seems to be working through the confidence channel. My results provide evidence that male confidence decreases while female confidence increases in the female stereotype task leading to similar tournament entry rates across genders. I find evidence that a gender difference in beliefs about future performance may be driving the observed gender gap in tournament selection as opposed to a gender difference in pure preference for competition.

Performance in both the arithmetic task and the facial emotion task are consistent across genders regardless of compensation scheme. Because we do not observe a significant difference in performance across genders for either task, this observed difference in tournament entry rates does not appear to be driven by a performance difference, further supporting my results.

One might be concerned that while men were more confident than women in the arithmetic task as expected, there was no difference in male and female confidence in the facial emotion task. Why do we not observed overconfidence for women in the facial emotion task? As Lundeberg et al. (1994) succinctly put it, "...the problem may not be that women necessarily lack confidence, but that in some cases men have too much confidence especially when they are wrong!" Additionally, Lundeberg et al. (1994) note that women are much more able to calibrate their confidence than men. Because men are unable to fully calibrate their confidence, it may look like they are just as confident as women, but the real result lies in the increase in the confidence of women through correct calibration.

Ideally, one would hope women would enter a tournament more than men when the task is in their favor, but as Lundeberg et al. (1994) note, men are consistently overconfident when they have no reason to be. The facial emotion task carries a very strong female gender stereotype, but it may take a task with an even stronger female stereotype for female tournament entry rates to exceed that of males. While it might be possible to find a more female-stereotyped task, any such task would be most useful if there is equal performance across genders.

7 Conclusion

Much research suggests that women avoid competition even when it can be beneficial and many researchers conclude that women differ in their preference for competition compared to men. My study explores the potential that another channel may be yielding the observed gender gap in tournament selection: a gender difference in beliefs about future performance.

My findings suggest that the observed difference in behavioral responses to competition among men and women is not due to a difference in preference for competition, but rather a difference in beliefs about future performance, evidence through the response to changing the task gender stereotype. Women are opting out of competitions when the selected task carries a connotation of male dominance, an environment where women would carry lower beliefs about future performance. I find that when the task carries a female stereotype, more women than men choose to enter the tournament. It appears an increase in female confidence and decrease in male confidence is driving this result.

Despite the fact that women earn roughly 57% of awarded bachelor's degrees, women occupy only 21 of the CEO positions for S&P 500 companies. While women are clearly capable of top positions, we are still not observing women seeking out these positions at rate similar to men. This study suggests that this observed difference may result from the gender stereotype of CEO positions. A Google image search of the word "CEO" produces only three images of females in the first 100 results.

The same thinking follows for the lack of women in STEM fields. Research suggests men can outnumber women as much as four to one in computer science classes while in non-STEM courses, women generally outnumber men. As discussed above, math and science are predominately viewed as male fields and given my results, it is again not surprising to observe this disparity. This study suggest s that policy interventions regarding gender-stereotyped tasks might indeed be useful and that more research is needed.

References

Booth, A. & Nolen, P. (2012). Choosing to Compete: How Different are Boys and Girls? Journal of Economic Behavior and Organization, 81, 542-555.

Cassar, A., Wordofa, F., & Zhang, Y. J. (April 25, 2016). Competing for the benefit of offspring eliminated the gender gap in competitiveness. Proceedings of the National Academy of Sciences, doi: 10.1073/pnas.1520235113.

Charness, G. & Gneezy, U. (2010). Portfolio Choice and Risk Attitudes: An Experiment. Economic Inquiry, 48 (1) 133-146.

Connor, S. (2005, April 12). The myth of female intuition exploded by fake smile test. Retrieved July 7, 2015.

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. Journal of Economic Literature , 47 (2), 448-474.

Dargnies, MP. (2012). Men Too Sometimes Shy Away from Competition: The Case of Team Competition. Management Science, 58 (11), 1982-2000.

Datta Gupta, N., Poulsen, A., Villeval, M.C. (2013). Gender Matching and Competitiveness: Experimental Evidence. Economic Inquiry, 51 (1), 816-835.

Dreber, A., von Essen, E., & Ranehill, E. (2014). Gender and Competition in Adolescence: Task Matters. Experimental Economics, 17, 154-172.

Dreber, A., von Essen, E., & Ranehill, E. (2011). Outrunning the Gender Gap–Boys and Girls Compete Equally. Experimental Economics, 14, 567-582.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics (10), 171-178.

Frederick, S. (2005). Cognitive Reflection and Decision Making. The Journal of Economic Perspectives, 19 (4), 25-42.

Gigerenzer, G., Galesic, M., & Garcia-Retamero, R. (2013). Stereotypes About Men's and Women's Intuitions: A Study of Two Nations. Journal of Cross-Cultural Psychology, 45 (1), 62-81.

Gillen, B., Snowberg, E., & Yariv, L. (2015). Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study. Unpublished manuscript, National Bureau of Economic Research Working Paper #21517.

Gneezy, U., Leonard, KL., & List, J. (2009) Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. Econometrica, 77, 1637-1664.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. The Quarterly Journal of Economics , 118 (3), 1049-1074.

Gneezy U., & Potters J. (1997) An Experiment on Risk Taking and Evaluation Periods. The Quarterly Journal of Economics, 112 (2), 631-645.

Gneezy, U., & Rustichini, A. (2004). Executives versus Teachers. Unpublished manuscript, Rady School of Management, San Diego, California.

Gneezy, U., & Rustichini, A. (2004). Gender and Competition at a Young Age. American Economic Review, 94 (2), 377-381.

Grosse, N., Riener, G., & Dertwinkel-Kalt, M. (2014). Explaining Gender Difference in Competitiveness: Testing a Theory on Gender-Task Stereotypes. Unpublished manuscript, Department of Economics, University of Mannheim, Mannheim, Germany.

Günether, C., Ekinci, NA., Schwieren, C., & Strobel, M. (2010). Women Can't Jump? An Experiment on Coompetitive Attitudes and Stereotype Threat. Journal of Economic Behavior and Organization, 75, 395-401.

Hall J, Philip RCM, Marwick, K Whalley HC, Romaniuk L, et. al. (2012) Social Cognition, the Male Brain and the Autism Spectrum. PLoS ONE 7(12): e49033.

Halladay, B. (2016). Gender, Emotions, and Tournament Performance in the Laboratory. Unpublished manuscript, Department of Economics, University of California, Santa Barbara, Santa Barbara, California.

Hoffman, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, HC. (2010). Women Recognize only Subtle Facial Emotions Better than Men. Acta Psychologica, 135, 278-283.

Hyde, J., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta analysis. Psychology of Women Quarterly, 14, 299-324.

Hyde, J. & Linn, M. (1988). Gender Differences in Verbal Ability: A Meta-Analysis. Psychological Bulletin, 104 (1), 53-69.

Kuhn, P. & Villeval, MC. (2015). Are Women More Attracted to Co-Operation Than Men? The Economic Journal, 125 (582), 115-140.

Lundeberg, MA., Fox, PW., & Punccohar, J. (1994). Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judegements. Journal of Educational Psychology, 86 (1), 114-121.

Meece, J., Parsons, J., Kaczala, C., Goff, S., & Futterman, R. (1982). Sex differences in math achievement: Toward a model of academic choice. Psychological Bulletin, 91, 324-348.

Mohr, T. (2014, August 25). Why Women Don't Apply for Jobs Unless They're 100% Qualified. Retrieved July 7, 2015.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. Management Science, 59 (1), 1-16.

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? The Quarterly Journal of Economics , 122 (3), 1067-1101.

Niederle, M., & Vesterlund, L. (2011). Gender and Competition. Annual Review of Economics, 3, 601-630.

Niederle, M., & Yestrumskas, A. (2008). Gender Differences in Seeking Challenges: The Role of Institutions. Unpublished manuscript, National Bureau of Economics Research Working Paper #13922.

Nosek, B., Banaji, M., & Greenwald, A. (2002). Math = Male, Me = Female, Therefore Math ≠ Me. Journal of Personality and Social Psychology, 83 (1), 44-59.

Simon-Thomas, E. (2007). Are Women More Empathic Than Men?. Retrieved November 18, 2016.

Skaalvik, S. & Skaalvik, E. (2004). Gender Differences in Math and Verbal Self-Concept, Performance Expectations, and Motivation. Sex Roles, 50 (3), 241-252.

Steele, CM. (1997). A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. American Psychologist, 52 (6), 613-629.

Appendix

Math Task Treatment Instructions

WELCOME

In the experiment today you will be asked to complete four different tasks. None of these will take more than 5 minutes. At the end of the experiment we will randomly select one of the tasks and pay you based on your performance in that task. Once you have completed the four tasks we determine which task counts for payment by drawing a number between 1 and 4. The method we use to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined.

Your total earnings from the experiment are the sum of your payment for the randomly selected task and a $5 show up fee. At the end of the experiment you will be paid in private.

TASK 1 – Piece Rate

For Task 1 you will be asked to calculate the sum of five randomly chosen two-digit numbers. You will be given 5 minutes to calculate the correct sum of a series of these problems. You cannot use a calculator to determine this sum, however you are welcome to write the numbers down and make use of the provided scratch paper. You submit an answer by clicking the submit button with your mouse. When you enter an answer the computer will immediately tell you whether your answer is correct or not. Your answers to the problems are anonymous.

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

Please do not talk with one another for the duration of the experiment. If you have any questions, please raise your hand.

TASK 2 - Tournament

As in Task 1 you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However for this task your payment depends on your performance relative to that of a group of other participants. Each group consists of four people, the three other members of your group are located in the same row as you.

If Task 2 is the one 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 receive no payment. We refer to this as the tournament payment. You will not be informed of how you did in the tournament until the end of the experiment. If there are ties the winner will be randomly determined.

Please do not talk with one another. If you have any questions, please raise your hand.

TASK 3 - Choice

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

If Task 3 is the one randomly selected for payment, then your earnings for this task 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 other three participants of your group in the previous task (Task 2-tournament). The Task 2-tournament is the one you just completed. If you correctly solve more problems than they did in Task 2, then you receive four times the payment from the piece rate, which is $2 per correct problem. You will receive no earnings for this task if you choose the tournament and do not solve more problems correctly now, than the others in your group did in the Task-2 tournament. You will not be informed of how you did in the tournament until the end of the experiment. If there are ties the winner will be randomly determined.

Please choose whether you want the piece rate or the tournament applied to your performance. You will then be given 5 minutes to calculate the correct sum of a series of five randomly chosen two-digit numbers.

Please do not talk with one another. If you have any questions, please raise your hand.

TASK 4 – Submit Piece Rate

You do not have to add any numbers for the fourth task of the experiment. Instead, your score from Task 1 - Piece Rate will be used according to your choice for payment. You can either choose to be paid according to the piece rate, or according to the tournament.

If the fourth task is the one selected for payment, then your earnings for this task are determined as follows. If you choose the piece rate you receive 50 cents per problem you solved in Task 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 Task 1-piece rate. If you correctly solved more problems in Task 1 than they did then you receive four times the earnings of the piece rate, which is equivalent to $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 Task 1 than the other members of your group.

Above, we remind you how many problems you correctly solved in Task 1. Please choose whether you want the piece rate or the tournament applied to your performance.

Please do not talk with one another. If you have any questions, please raise your hand.

Facial Emotion Task Treatment Instructions

WELCOME

In the experiment today you will be asked to complete four different tasks. None of these will take more than 5 minutes. At the end of the experiment we will randomly select one of the tasks and pay you based on your performance in that task. Once you have completed the four tasks we determine which task counts for payment by drawing a number between 1 and 4. The method we use to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined.

Your total earnings from the experiment are the sum of your payment for the randomly selected task and a $5 show up fee. At the end of the experiment you will be paid in private.

TASK 1 – Piece Rate

For Task 1 you will be shown 15 photographs depicting individual's faces. For each image, you will be asked to identify the emotion depicted on the individual's face. The emotions in the images have been professionally classified by psychologists doing research in this field. The images will be projected on your computer screen for a very short period of time (2 seconds). After the image is shown, you will be given four options from which to select the correctly displayed emotion. You will have 20 seconds to submit your answer. You submit an answer by clicking the submit button with your mouse. When you enter an answer the computer will immediately tell you whether your answer is correct or not. Your answers to the problems are anonymous. While you will never see the same image twice, the same emotion may be repeated.

If Task 1 is the one randomly selected for payment, then you get 50 cents per emotion you correctly identify in Task 1. Your payment does not decrease if you provide an incorrect answer to an image. We refer to this payment as the piece rate payment.

Please do not talk with one another for the duration of the experiment. If you have any questions, please raise your hand.

TASK 2 – Tournament

As in Task 1 you will be shown 15 photographs of individual's faces. You will again be trying to correctly identify the emotions depicted on the faces. For this task your payment depends on your performance relative to that of a group of other participants. Each group consists of four people. The three other members of your group are located in the same row as you.

If Task 2 is the one randomly selected for payment, then your earnings depend on the number of emotions you correctly identify compared to the three other people in your group. The individual who correctly identifies the largest number of emotions will receive $2 per correct emotion, while the other participants receive no payment. We refer to this as the tournament payment. You will not be informed of how you did in the tournament until the end of the experiment. If there are ties the winner will be randomly determined.

Please do not talk with one another. If you have any questions, please raise your hand.

TASK 3 – Choice

As in the previous two tasks, you will again be shown 15 photographs of individual's faces and asked to identify the depicted emotions. However you will now get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task.

If Task 3 is the one randomly selected for payment, then your earnings for this task are determined as follows.

If you choose the piece rate you receive 50 cents per correctly identified emotion. If you choose the tournament your performance will be evaluated relative to the performance of the other three participants of your group in the previous task (Task 2-tournament). The Task 2-tournament is the one you just completed. If you correctly identify more emotions than they did in Task 2, then you receive four times the payment from the piece rate, which is $2 per correct answer. You will receive no earnings for this task if you choose the tournament and do not correctly identify more emotions now, than the others in your group did in the Task-2 tournament. You will not be informed of how you did in the tournament until the end of the experiment. If there are ties the winner will be randomly determined.

Please choose whether you want the piece rate or the tournament applied to your performance. The images will begin on the screen once everyone makes their selection.

Please do not talk with one another. If you have any questions, please raise your hand.

TASK 4 – Submit Piece Rate

You do not have to identify any emotions for the fourth task of the experiment. Instead, your score from Task 1 - Piece Rate will be used according to your choice for payment. You can either choose to be paid according to the piece rate, or according to the tournament.

If the fourth task is the one selected for payment, then your earnings for this task are determined as follows. If you choose the piece rate you receive 50 cents per correct answer in Task 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 Task 1-piece rate. If you correctly identified more emotions in Task 1 than they did then you receive four times the earnings of the piece rate, which is equivalent to $2 per correct answer. You will receive no earnings for this task if you choose the tournament and did not correctly identify more emotions in Task 1 than the other members of your group.

Above, we remind you how many problems you correctly solved in Task 1. Please choose whether you want the piece rate or the tournament applied to your performance.

Please do not talk with one another. If you have any questions, please raise your hand.