ERASMUS UNIVERSITY ROTTERDAM

HOW DO MEN AND WOMEN RESPOND TO GOOD AND BAD COMPETITION?

AN EMPIRICAL STUDY ON HOW COMPETITIVE BEHAVIOR DEPENDS ON THE PERCEIVED SKILL LEVEL OF YOUR OPPONENTS

Abstract

We examine whether men and women differ in their motivation to compete when varying the skill level of the competition. We conducted an online experiment, where respondents had to solve three tasks, in a non-competitive environment and in a competitive environment. For the third task, respondents could choose which one to repeat. Respondents were randomly put in either the control group, where there was no information about the skill level of the competition, or in one of the treatment groups where respondents were either encouraged by competing against bad opponents or discouraged with good opponents. The results showed that men and women were equally competitive, controlling for risk aversion and overconfidence. Encouragement increased the likeliness to compete by 14% relative to the control group, where this effect was 6% stronger for women. Discouragement decreased the likeliness to compete by almost 13%, and this effect was 5.5% stronger for men. In the discussion, suggestions are made on how to apply these insights in a work environment, in order to decrease the gender gap in high level jobs.

HANNAH BARTELS

23-04-2022

Supervisor: Jan Stoop

Table of Contents

1. Introduction

Competition. Something that is present in all parts of society, and something that people are often confronted with for the first time at a young age. Whether it is a game at home, or a running competition at school, research has shown that children start to experience competitive feelings from the age of (on average) 5 (Klein, 2014). The reason is that children learn to categorize, and therefore they can determine for themselves in what category they fall; the better or the worser group.

When time passes, and people get older, competition still plays a maybe even more significant role in our society and is present in all kinds of environments; social, economic, cultural. Who has the biggest car in the street, which friend earns the most money, or even who is the happiest in their love life. Often, competition is unconscious, and the people involved are not even aware of the feeling. However, competition can have a huge impact on important decisions that you make in life, both positive and negative (DiMenichi & Tricomi, 2015).

1.1 Competition at work

Another environment where competition plays an important role is the work environment, and here competition is probably even more visible compared to for example the social environment. People strive to do better, mostly because doing better means earning more money and reward is one of top factors that affect employee motivation (Chadwick, 2019).

The extent of competitiveness can differ per industry and will probably be more extreme in some industries and almost non-existent in others. For example, technology, and specifically software development, the competition is high because of the number of applications and employees are expected to work extremely long hours to advance further in the company (Palmer, 2021).

Consultancy is another industry where the competition is high, also because this industry is very individually oriented. For example, HR Consultants often work on their own for clients, and compete against other consultants in the firm on who earns the most commission. In these types of companies, competition can take various shapes or forms. There can be certain quotas, where you will receive bonusses when you pass a certain number. Often, there will be company trips or holidays where only the top employees can attend, based on for example their sales. The most common example of direct competition in the workplace is making promotion, as employees often compete directly with colleagues to earn this particular place.

1.2 Gender gap in high level jobs

Competition in business and the ambition to make a promotion, is something that is discussed often in literature as well, and more particular, in the discussion about why men in general promote more compared to women. This phenomenon, often referred to as the gender gap in high level jobs, can be explained by multiple factors and competition is one of these factors (Niederle & Vesterlund, 2007).

Fact is that men on average act in higher level jobs compared to women and are more often promoted than women. A study conducted by McKinsey found that for every 100 men that are promoted to manager, only 85 women had the same privilege in corporate America. As a result, women represented only 38% of the manager-level positions, whereas men represented 62% in 2020 (McKinsey, 2021).

In January 2022, WOMEN Inc launched a campaign called “Mijn naam is Peter” (WOMEN Inc., 2022). This campaign, with the goal to decrease gender inequality in the workplace, brought to the attention that there are more CEOs at listed companies in the Netherlands that are called Peter than there are female CEOs in the Netherlands. Only 12% of the people in the Board of Directors for listed companies are women.

1.3 Possible explanations

Why does this gender gap occur repeatedly? Literature and online articles come up with various reasons, ranging from the potential relative lack of long-term career commitment for women (Bertrand & Hallock, 2001) to the argument that men offer characteristics that would be more suitable for leading positions than women (Lown, n.d.). Niederle & Vesterlund (2007) offer a different explanation: men are more competitive than women. Since promotion, and the ambition to get higher up in a company comes from competitive feelings, women might not be as motivated as men to promote. As a result, men are more likely to achieve these higher-level jobs than women.

If this is the case and competitiveness, or the lack of competitiveness for women, is part of the reason that men fulfill higher roles, there is a great opportunity here to try and reduce the gender gap by using this insight.

Everybody responds differently to competition and the way you frame the competitive aspect might have a different effect on men and women. If you can find a way to influence women in a proper way by using competition, that will motivate them instead of discourage them, it might lead to more women wanting high level jobs, and eventually it might decrease the gender gap in high level jobs.

That is what we will try to achieve in this research paper, and we will especially focus on the perceived skill level of the opponents, and how this influences the intrinsic motivation of men and women. If we can find differences, it could be used to make suggestions on how to use competition in a positive way to motivate women in the workplace, that will lead to a smaller gender gap.

2. Theoretical Framework

2.1 Men more competitive than women?

Not every individual reacts the same to competition, and personal characteristics such as gender and age can influence how someone experiences this feeling of competition. A lot of research has been done about the difference between men and women and their competitiveness.

When it comes to the decision to compete or not, researchers found that men are more likely to enter a tournament than women. Niederle and Vesterlund (2007) conducted an experiment where subjects could choose if they wanted to perform a task under a piece rate (non-competitive) or in a tournament (competitive). After performing both tasks, subject could choose which of the compensation schemes they wanted to apply on their next task. 73% of the men chose the tournament, whereas only 35% of the women chose the tournament, even when there was gender difference in their performance. This shows that men are more like to compete and might be evidence that men are more competitive.

The same results were found by Gupta, Pulsen and Villeval (2005) and by Vandegrift and Yavas (2009), who conducted a similar experiment. This experiment also found that even when controlling for skill level of the subjects, men were more likely to participate in the competition than women.

Secondly, research has shown that men´s performance improves when they are competing, compared to when they are performing in a non-competitive environment. This effect does not necessarily happen for women. This gender difference already starts at a young age, as found by Gneezy and Rustichini (2004). In this experiment, elementary school students were asked to run a 40 meters track, alone in the first round and the second time against another student that performed similarly in the first round. The results showed that boys performed significantly better in the second, competitive round compared to the first, non-competitive round, whereas girls ran slower in the second round.

These findings play on to similar research conducted a year earlier (Gneezy, Niederle & Rustichini, 2003). In this experiment, subjects were asked to solve mazes, both in a non-competitive and a competitive environment. Again, the researchers found that while men and women perform equally in the non-competitive situation, men outperform women significantly in the competitive environment.

Finally, there has also been done research on the extent of competitiveness depending on the sex of the opponent. Women tend to feel more intimidated when their opponent is a man, compared to a same gender game (Vermeulen, Castellar & Van Looy, 2014). In their study, women experienced more stress when competing against men than against women. However, this study did not report significant differences in competitiveness between same gender or opposite gender games.

This difference is found in a study conducted by Sutter et al. in 2009. It appears that gender pairing has a significant effect on behaviour. In a same gender game, subjects showed a lot more competition and retaliation compared to opposite gender games, for both men and

women. Supporting evidence was found by Gupta et al. (2005). In their experiment, men competed more against other men than they did against women. However, this evidence was not found for women.

Another interesting study was conducted in The Netherlands, which focused on competition incentives in sales (Delfgaauw et al., 2009). A short-term sales competition was introduced in 128 stores, where a team consisted of a team manager and several employees. All employees of the same store, including the store manager, received a bonus when their store sales increased the most. Even though the researchers did not find significant differences between men and women and their competitiveness, they did find an interesting interaction: in stores with a male manager, the sales grew with the share of male employees, and the same holds for female managers. In other words, the more employees of the same sex as the manager in a store, the more the competition incentive worked. This is an interesting finding, as previous studies already showed that competing against the same gender increases competitiveness, but this study showed that competing with the same gender also has a positive effect on competitiveness.

2.2 Overconfidence and risk aversion

There is clear evidence that men are more competitive than women; men are more likely to enter a tournament and perform better in a tournament setting. But there can be multiple reasons why men would be more likely to enter a tournament, besides the fact that they are just more competitive. Two of these reasons could be overconfidence and risk aversion (Van Veldhuizen, 2017), and again literature has found significant differences between men and women.

For one, men tend to be more overconfident than women. Psychology has already established that men are often more overconfident than women and multiple researchers have applied these findings on economic situations. In trading, men hurt themselves by trading too much because they are overconfident (Barber & Odean, 2001). Trading reduces the net turnover for men on average by 0.94 percentage point more than women, and although how much men and women trade can partly be explained by risk aversion, this drop in net turnover is an effect of overconfidence.

Bengtsonn et al. (2005) conducted a study with students and their exams. They found that even though men and women scored similarly in their exam, men were more confident in their answers than women. This does not however prove that men are more overconfident, it only shows that men are more confident.

Correll investigated gender differences in biased self-assessments (overconfidence), and the influence on career choices (2001). It appeared that for a mathematical task, men perceived their own skill level to be higher compared to women with the same math test scores and grades. Cho (2017) replicated these findings and found that women are not less confident in math, but that men are more overconfident. In conclusion, literature provides extensive evidence that men are more overconfident than women.

Risk aversion, or risk seeking behaviour, can also influence whether people enter a tournament or not. According to the existing literature, women tend to be more risk averse

than men. There has been a lot of research on gender differences in risk aversion and even though not all experiments are consistent, the majority of the studies show that women are more risk averse than men, in particular when it comes to low stakes and in abstract gamble experiments (Eckel & Grossman, 2002, 2008; Holt & Laury, 2002). Eckel and Grossman conducted an experiment in 2002, where subjects had to select one of five gambles, where the first gamble was a sure bet, and as the bet number increased, so did the risk level. Women were four times as likely to choose the sure bet, and only one-third as likely to choose the riskiest bet. The average chosen bet was significantly less risky for women than for men, proving that women are more risk averse than men. In 2009, Borghans et al. conducted a study on risk aversion and ambiguity aversion and confirmed previous findings that women are more risk averse.

2.3 Skill level of opponent

We have seen that in general, men are more likely to enter a tournament than women. This is influenced by factors such as overconfidence and risk aversion, but these are all internal factors. However, there will be other external factors that also influence the likeliness of competing, such as the gender or other characteristics of the competition. The effect of the gender of the competition is already discussed, but how about the perceived skill level of your competition?

It appears that people adjust their behaviour based on competitors experience level (Slonim, 2005). A high perceived skill level of the competition can undermine the intrinsic motivation of an individual. Rogers and Feller (2016) refer to this phenomenon as 'discouragement by peer excellence.' In this study, students were asked to review the work of fellow students, and exposure to exemplary peer performance caused a lot of students to quit the course. The performance of others seems unreachable and resulted in a drop of motivation. These findings were later replicated by Cho (2019), but only found the discouraging effect of peer excellence for traditional students[1]. This could imply that the effect of, in this case, discouragement, can differ per individual, and maybe also per gender.

In contrary to discouragement because of high perceived skill level of the competition, one can also be encouraged when they get the feeling they are doing a good job. Recognition plays a huge role in this, as 70% of employees say that motivation would improve when recognition from their managers increased as well (Madison, 2017). Doing good is almost always relative to others, and encouragement can therefore not only come from managers, but also from the fact that you are doing a better job than other people.

2.4 Research question and hypotheses

Someone's competition strategy can depend on the perceived skill level of the competition. Playing against a "good or experienced" competitor will discourage a player and decrease the intrinsic motivation to compete. On the other hand, competing against "bad or inexperienced" competitor will encourage a player and increase the intrinsic motivation to compete. These conclusions are supported in existing literature, as well as the difference

[1] The main factor that determines whether students are traditional or non-traditional is their age. Traditional students often are under 25, whereas non-traditional students are on average older, or work (parttime), are financially independent or have children.

between men and women in standard competitive situations. But until now, there has not been done research on the effect of discouragement and encouragement in competition for men and women separately. That is why this article focuses on the following research question:

"How does the skill level of the opponent affect competitiveness for men and women?"

To answer this research question, a difference between competitiveness between men and women must be found. Literature shows that men are more competitive than women (Niederle & Vesterlund, 2007; Gupta, Pulsen & Villeval, 2005; Vandegrift & Yavas, 2009), and therefore men are often more likely to enter a tournament. Overconfidence and risk aversion also play a role (Van Veldhuizen, 2017), and therefore we must control for these effects. Taking this into account, the first hypothesis that needs be tested is:

H1: Men are more competitive compared to women, taking overconfidence and risk aversion into account.

To investigate the effect of the skill level of the opponent on competitiveness, we have formulated two hypotheses, one considering encouragement and one considering discouragement. Since literature on these effects specified on gender is limited, it is difficult to form expectations. Literature does however show that men are more overconfident than women (Barber & Odean, 2001; Bengtsonn et al, 2005; Correll, 2001; Cho, 2017), and because of this we would expect them to react stronger to encouragement, i.e. competing against "bad" people. Therefore, the following hypotheses formed:

H2: Encouragement (i.e. competing against 'bad' opponents) increases the likeliness of competing, and this effect is stronger for men than women.

On the other hand, we expect that discouragement will decrease the likeliness of competing. Literature showed that women experience on average more stress from competing (Vermeulen, Castellar & Van Looy, 2014), and therefore we expect that discouragement has a stronger effect on women.

H3: Discouragement (i.e. competing against 'good' opponents) decreases the likeliness of competing, and this effect is stronger for women than men.

3. Methodology

3.1 Resemblances and differences between original and this study

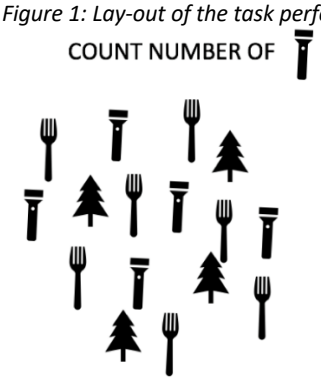
To gain insight on the hypotheses above, and to answer our research question, we performed a quantitative research based on previous studies by Niederle and Vesterlund (2007). The goal of their study, as well as in this study, is to test the difference in competition between men and women. The experiment consists of 3 tasks, where the first task is noncompetitive, and the second task is competitive. In task 3, respondents must decide for themselves what task, competitive or non-competitive, they wish to repeat. The decision men and women make in this third round is used as an indication of their competitiveness.

Where this study differentiates itself from the original, is the fact that we also test the effect of encouragement and discouragement on competition and how these effects differ between men and women. Respondents will be discouraged by informing them that they will play against the three best players from a previous conducted pilot in case they decide to play the competitive option. The encouraged respondents will play against the three worst players from the pilot. The purpose of this treatment is to test whether men or women respond differently to this kind of (loss of) motivation.

This experiment also uses a different kind of task; in the original experiment the researchers asked their respondents to add up five two-digit numbers. The experiment was conducted at the University of Pittsburgh Experimental Economics Laboratory, and the respondents were supervised by the researchers during the tasks. This experiment, however, is conducted online, and therefore without any supervision. To limit the possibility of cheating by using tools such as a calculator, we needed to select a task that could not increase performance by using such tools. Secondly, the needed to be ¨gender neutral¨; there shouldn't be any significant differences between men and women in pure performance, regardless of any competition, risk aversion and overconfidence.

Taking these requirements into account, we chose to let respondents count specific figures within a picture where multiple figures are displayed (Figure 1). This task meets the requirements, as there is no possibility to cheat, and there is no significant gender difference in performance, as we have performed a t-test with our pilot data. As can be seen in the output in Table 13 in Appendix A.2, there was no significant performance difference found between genders.

Figure 1: Lay-out of the task performed



3.2 Survey structure

The survey was created with the program Qualtrics. It consists of multiple parts; an introduction, the actual experiment (three tasks) and finally some questions that will be used to test for overconfidence and risk aversion (task 4 and 5). The survey flow and instructions can be found in Appendices A.3 and A.4.

The introduction explains the survey in the following way:

"The survey will consist of 5 TASKS, and with each task, money can be earned. The money that can be earned will differ among the tasks and will be explained at the beginning of each task.
At the end of this survey, 10 subjects will be randomly selected who will receive the actual money that they have earned with 1 OF THE 5 TASKS. What task will be paid out, will also be selected randomly. So, make sure you try your best, as you can win actual money!"

Part of the respondents will receive real money after finishing the survey. Research has shown that monetary incentives help to motivate respondents to perform well (Bonner & Sprinkle, 2002), and for the validity of the experiment, it is desirable if the participants perform up to their abilities. Therefore, we pay a portion of the respondents money by bank transfer. Following up on this, it is important that the participants answer all questions truthfully. We need the participants to be honest about their perceived performances instead of giving the desired or expected answers, and a monetary incentive will help to achieve this. That is the reason that in the introduction, the experiment is presented as five equal tasks, instead of three performance tasks and two separate questions, and one out of five tasks is paid out instead of one out of three. Finally, only one task will be paid out per respondent, and not all. The main reason is that it will limit the costs of the experiment, and secondly, we diminish the chance that performance and answers from certain tasks will be used to hedge against decisions in other tasks (Niederle & Vesterlund, 2007).

This introduction is followed by some general questions about gender, age, and highest level of completed education. The questions about age and gender were forced since this information is important for the research. After these general questions, task 1 began. To emphasize ones more, this first task is noncompetitive, there are no other respondents involved. The respondents were told that they would earn 1 euro per correct answer, in case this task would be selected for actual payout. This payment scheme is referred to as the **piece rate payment scheme** (Niederle & Vesterlund, 2007), and from now on we will use this term.

Before the actual first task started, the respondents were given a practice question to make sure that they understood their task. Then, a timer started, and the respondents had 45 seconds to answer as many questions as possible. Participants could see the timer at the top of the page, but if they scrolled down to the next questions, it was not in their sights the entire time. After 45 seconds, the page was automatically closed, and respondents were redirected to the second round. Note that there was no feedback provided about the respondents' absolute performances, respondents only have information based on their own perception/experience because feedback might influence the remainder of the experiment.

In the second, competitive round, respondents played against subjects from a previously conducted pilot (Appendix A.1). Before conducting this survey, a shorter survey was distributed and filled in by 10 respondents. The pilot consisted of one 45-seconds task, the same as we have seen in round 1 of the actual experiment. To have an incentive to perform, 10% of the respondents (randomly selected) would receive 1 euro per correct answer. More elaborate instructions and the data from this pilot can be found in Appendix A.1. The respondents from this pilot new serve as the opponents in the actual experiment, and in round 2, respondents play against 3 randomly selected subjects from the pilot. This randomization was done with a Wheel of Names[2], where 3 respondents were selected to serve as the random pilot group.

In case the respondent outperformed the subjects from the pilot, he or she would receive four euros per correct answer (a 25% winning chance with 4 euros per answer offers the same expected payoff as one answer under the piece-rate payment scheme). But if someone from the pilot outperformed them, the respondents would not receive anything. This payment scheme is referred to as **a competitive tournament scheme** (Niederle & Vesterlund, 2007). After 45 seconds, the page was automatically closed again. Participants have no information about their relative performance compared to the competition, since this might bias the effort and performance in task 3.

In round 3, the actual treatment started. Without the respondents knowing, all have been assigned to one of following groups: the control group, the encouragement group and the discouragement group. The control group would again compete against three randomly selected subjects from the pilot, whereas the discouragement and encouragement group competed against the three best and worst subjects from the pilot, respectively. The participants were unaware of the treatment they received. After receiving information about their competition, they had to decide between repeating round 1 under the piece rate payment scheme, or round 2 under the competitive tournament scheme. Task 3 then continued the same as the previous 2 rounds.

Task 4 serves as a control for overconfidence. As mentioned previously, entering a tournament can also be influenced by overconfidence and risk aversion, and therefore we need to control for this. To do this, the respondents were asked to guess their performance in task 2 by rating themselves compared to their competition (Appendix A.4 provides a more elaborate explanation). This question was asked after task 3, because otherwise the question and moment of self-reflection might influence the decision made in task 3. In case the respondent would correctly guess their performance, and this task is selected for actual payout, the respondent will receive five euros, and nothing otherwise.

To control for risk aversion, we incorporated task 5 that was based on a study by Eckel and Grossman (2008) (Figure 5, Appendix A.4). Respondents were asked to choose one of five bets, where the first bet was a certain 8 euros and the following bets increased in risk. Bet number 5, the riskiest one, had a 50% chance of winning 24 euros, and nothing otherwise. The lower the chosen bet, the more risk averse the respondent is considered to be.

[2] Randomly chose three names using the following link: https://wheelofnames.com

At the end of the survey, respondent could leave their email address if they wished to have a chance at the actual payout, that will be done by bank transfer.

3.3 Descriptive statistics

This online survey was distributed via an anonymous link, that was send to several groups of people through Whatsapp, Facebook or email. Additionally, the survey was posted on SurveySwap[3]. As a result, the respondents varied in age and gender, but females in the age group 20 – 25 are overrepresented and ultimately, the sample is not perfect representation of the Dutch population.

In total, 161 respondents filled in the survey. Of those, 8 were unfinished and 4 responses were invalid for different reasons (Appendix A.5). In the end, 149 responses could be used for the analysis.

Table 1 shows the distribution of all respondents over the different groups. There were 45 respondents in the control group, and 54 in both treatment groups. All groups consisted of more men than women, due to the fact that in total, 85 women took place in this experiment and 64 men did.

Table 1: Descriptive statistics per treatment per gender

	Female	Male	Total
Control	26	19	45
Discouragement	31	21	52
Encouragement	28	24	52
Total	85	64	149

The average age of the respondents is 28 years. Almost half of the respondents completed their Bachelors, and 38 of the respondents have already finished their Masters as well (Table 14, Appendix A.6).

3.4 Description of the variables

3.4.1 Overconfidence

We looked at the number of correct responses in task 2, and the self-perceived ranking of the respondents. In case the respondent was correct, the respondent was labeled "Correct" in *overconfident,* i.e. overconfidence as a categorical variable. There were also respondents with the same number of correct answers as someone in the pilot. For example, the respondent answered 8 questions correctly, which is a shared second place with the people from the pilot. In this case, both ranking 2 and 3 were considered correct. When the respondents' self-perceived ranking was too low compared to their actual position, they were labeled "Underconfident". Logically, when the perceived ranking was too high, respondents were labeled as "Overconfident". Table 5 in section 4.2 provides more details specifically on gender differences in overconfidence.

[3] A website where students share their surveys with each other, and by earning credits by filling in the survey of others, people will fill in your survey as well (https://surveyswap.io/sign-up)

overdum is the binary variable for overconfidence. All respondents that were either correct or underconfident are assigned a value of 0, and only the overconfident respondents are assigned a value of 1. *overconfident* is the categorical version, where there is still a distinction between the three levels.

3.4.2 Risk aversion

The variable *riskaversion* is based on the decision respondents have made in the final task, where they had to select one out of 5 bets that differed in risk level. The categories range from 1 (bet number 1) to 5 (bet number 5). This variable would classify as an ordinal variable, as the categories do have a clear order, but the distance between all categories is not completely measurable.

3.4.3 Education

The variable *educcat* is a categorical variable, that has five categories (1 = Elementary school, 2 = High school, 3 = MBO (practical education), 4 = Bachelor's degree and 5 = Master's degree). We have also created a dummy variable for education, to make a distinction between respondents that have received higher education and the ones who have not. This is also easier to interpret in our analysis. The variable *educdum* takes a value of 1 if the respondent finished either a Bachelor's or Master's degree, and 0 otherwise.

Table 2: Description of variables

sex	Binary variable, that either takes the value 1 for female or 2 for male
competitive	Binary variable, that either takes the value 0 if the respondent chose the piece rate scheme, and 1 if the respondent chose the tournament scheme
overdum	Binary variable, that either takes the value of 0 if the respondent is not overconfident, and 1 if the respondent is overconfident.
overconfident	Categorical variable, from 0 (=underconfident), 1 (= correct) to 2 (=overconfident).
educcat	Categorical variable, ranging from category 1 (elementary school) till 5 (Master's degree).
educdum	Binary variable, that either takes the value of 0 if the respondent has not finished higher education, and 1 otherwise.
age	Continuous variable, indicating the age of every respondent
riskaversion	Categorical variable, that can take the value of 1 for bet number 1, and rise up to 5 for bet number 5.
treatment	Categorical variable, where category 1 is the control group, category 2 the discouragement group and 3 the encouragement group
task1correct	Number of correct answers in task 1

task2correct	Number of correct answers in task 2
task3correct	Number of correct answers in task 3
discouragement	Binary variable, that either takes the value of 0 is the respondent is in the control group, and 1 in the discouraged treatment group. All respondents that are in the encouragement group are omitted.
encouragement	Binary variable, that either takes the value of 0 is the respondent is in the control group, and 1 in the encouraged treatment group. All respondents that are in the discouragement group are omitted.

4. Results

4.1 Performance in the piece rate and tournament scheme – Task 1 and 2

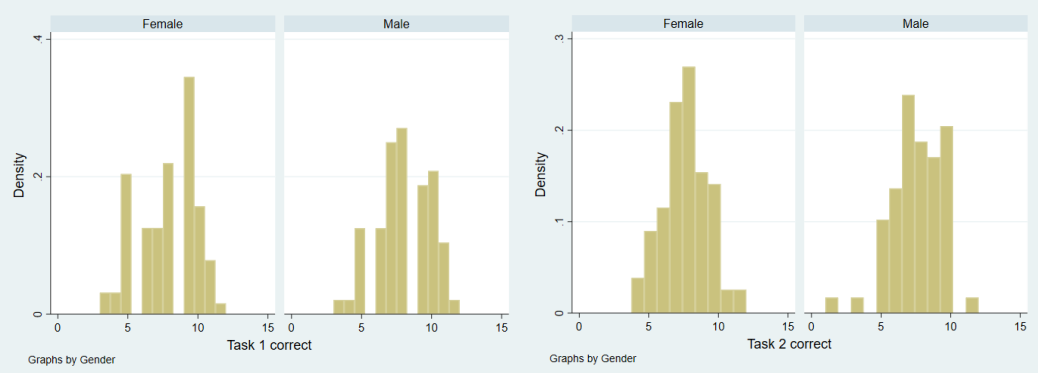
4.1.1 Gender differences for each task

First, we want to test if there is no difference in performance between men and women, as that might influence their decision in round 3 if this would be the case. The pilot data did not show any differences between men and women's performance, however we still want to test if this also holds in the actual experiment, with a bigger sample.

Table 3: Mean number of correct answers per gender per payment scheme/task

	Female	Male	Total
Piece rate scheme (Task 1)	7.76	7.95	7.85
Tournament scheme (Task 2)	7.75	7.67	7.71

Figure 2: Histogram number of correct answers per gender per task



Under the piece rate scheme, the average number of correct answers is 7.76 for women and 7.95 for men. Using a two-sided t-test with equal variances, we can confirm that there is no significant performance difference between men and women in task 1, as the p-value of our alternative hypothesis that the difference is not equal to zero is 0.5733 (Table 16, Appendix B.1).

Under the tournament scheme, the average number of correct answers is 7.75 for women and 7.67 for men. We will perform a two-sided t-test with unequal variances, as we have found that the variances significantly differ at a 10% significance level between the two groups (p-value of 0.1050) (Table 17, Appendix B.1). With a p-value of 0.7565, we can say that there is also no performance difference in task 2 between genders (Table 18, Appendix B.1).

4.1.2 Performance differences over tasks

Conclusively, there is no significant difference in pure task performance between men and women, both for the piece rate as in a competitive setting. Secondly, we could look at the difference in performance between tasks, i.e., are respondents performing better in for example a competitive setting? Based on literature, we would expect performance to

increase in the tournament scheme compared to the piece rate, as this is the case in the original experiment, as well as what can be seen in other literature. This could be explained by learning, or because of the competitive character of the task.

In other words, our alternative hypothesis is that the number of correct answers is larger in task 2 than in task 1 *(diff = mean Task1correct - mean Task2correct < 0*). The variances differ significantly as $p < 0.1$ (Table 19, Appendix B.1), and therefore we will use a two-sample t-test with unequal variances.

If we look at the statistical output in Table 20 in Appendix B.1, we can see that there is no statistical evidence to support this alternative hypothesis. To the contrary, the mean correct answers is higher for the piece rate scheme than for the tournament scheme (not significant though, p-value is 0.2747). This is still an interesting finding, as you would expect our respondent to learn from the first task, to be more familiar with the task. A potential explanation could be that respondents felt pressured in task 2, and that this had a negative impact on their performance.

This difference in performance between tasks is almost completely caused by men, which is also quite interesting. The mean correct answers for men is almost significantly higher in task 1 than in task 2 (p-value = 0.1959) (Table 21, Appendix B.1). Based on the literature, you would expect that especially men's performance would increase in a competitive setting, but these results prove otherwise.

In conclusion, this experiment has found no gender difference in pure performance for each individual task. We did however find that competing had a negative impact on the men's performance, whereas the women did not show this behavior. This contradicts what literature has found so far, as most papers show that men's performance increases more in a competitive setting than the performance of women.

4.2 Hypothesis 1 - Gender differences in tournament entry

Men are more competitive compared to women, taking risk aversion and overconfidence into account

4.2.1 Statistical analysis

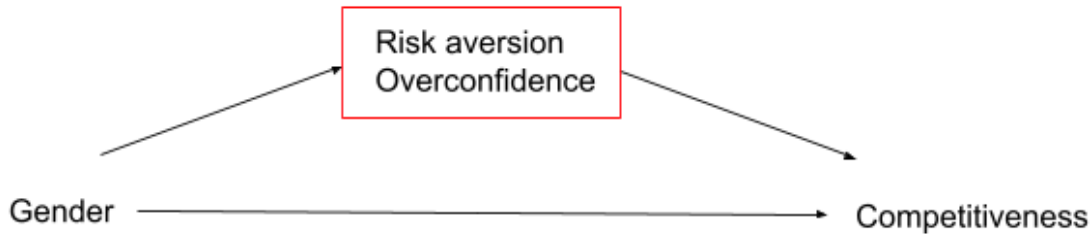
In order to test this, we first look at the different level of competitiveness between men and women in Table 4. Of the 85 women that took part in this experiment, 57 chose the competitive option. This is 67%. 64 men took part, and of them, 43 chose the competitive option, which is also 67%. This indicates that in this sample, men and women are relatively equal competitive in the absence of other variables, such as risk aversion and overconfidence. This is confirmed when we test if there is a correlation between the choice made in task 3 and gender, which is only 0.0014 (Table 22, Appendix B.2).

Table 4: Numbers and percentages per gender on choice in task 3

	Female	Male	Total
Piece rate	28 (32.94%)	21 (32.81%)	49
Tournament	57 (67.06%)	43 (67.19%)	100
Total	85 (100%)	64 (100%)	149

Before we perform a regression to test the relationship between gender and competitiveness, it is important to find out if there are other variables correlated with both our independent and dependent variables (Figure 3). When these variables are not included in the model, omitted variable bias could appear. Note that if we would find a significant relationship between gender and competitiveness, this would not be a causal relationship, as gender is not randomized. Therefore, there would not be a causal relationship, however we could still get an idea of the effect of gender on competitiveness.

Figure 3: Graphical explanation of potential mediators



Firstly, existing literature has shown that men are more overconfident compared to women. More than half of the men in this sample are overconfident about their own performance, 53.1%, whereas for women this percentage is lower, 38.8% (Table 5). This relationship is confirmed when we test the correlation between sex and overconfidence as a dummy variable, where we see a correlation of 0.1423 (Table 23, Appendix B.2). This means being male and being overconfident are positively correlated. Therefore, we could say that that overconfidence probably has a mediating effect on the relationship between gender and competitiveness and should be included in the regression.

Table 5: Overconfidence per gender (per category, % of gender that were in certain category)

	Female	Male	Total
Underconfident	11 (13%)	9 (14.1%)	20
Correct	41 (48.2%)	21 (32.8%)	62
Overconfident	33 (38.8%)	34 (53.1%)	67
Total	85 (100%)	64 (100%)	149

Secondly, we expect that there is a relationship between risk aversion and gender, and also between risk aversion and competitiveness. Existing literature shows that women tend to be more risk averse than men, and we can see in Table 24 in Appendix B.2 that there is minor

correlation between the two. If we look at Table 6, we can clearly see that the percentage of men that chose bet number 5, the riskiest bet, is significantly higher than the percentage of women that chose bet number 5 (39.1 and 22.6 percent, respectively). Therefore, we include risk aversion in the regression as well, as we expect risk aversion has a mediating effect.

Table 6: Descriptive statistics risk aversion (per category, % of gender that chose this option)

	Female	Male	Total
Bet number 1	20 (23.8%)	15 (23.4%)	35
Bet number 2	19 (22.6%)	13 (20.3%)	32
Bet number 3	15 (17.9%)	6 (9.4%)	21
Bet number 4	11 (13.1%)	5 (7,8%)	16
Bet number 5	19 (22.6%)	25 (39.1%)	44
Total	84 (100%)	64 (100%)	148

Our dataset also contains information about the age and education level of our respondents. Existing literature suggests that age can have a significant effect on competitiveness, and that competitiveness increases until your fifties, and decreases thereafter (Ultich et al., 2012). Education could also have an impact (Baumann & Winzar, 2014), but on both topics the literature is not abundant. However, it could be interesting to incorporate these variables in our regression.

However, we expect that education and age are strongly correlated, as younger respondents have not finished their education yet and will automatically have a lower education level compared to the older respondents. We tested this and see an increase in age has a positive influence on education (Table 26, Appendix B.2). To avoid multicollinearity, one of the two variables must be excluded from the regression. Because there is a higher correlation between competitiveness and age than competitiveness and education (Tables 27 and 28, Appendix B.2), we decided to include this variable in the regression and leave education out.

Taking this all into account, we will use the following regression to find support for or reject hypothesis 1:

$$competitive = \beta_0 + \beta_1 sex + \beta_2 riskaversion + \beta_3 overconfident + \beta_4 Age + \varepsilon$$

The output of the regression can be found in Table 7, and more detailed in Table 28 in Appendix B.2.

Table 7: Regression output hypothesis 1

competitive	Coefficient	t	P > | t |
sex	- 0.0246538	-0.32	0.748
riskaversion	0.0906113	3.84	0.000***
overconfident	0.0167919	0.29	0.771
age	- 0.0006304	-0.20	0.840
constant	0.4269164	2.42	0.017**

* = 10% significance level
** = 5% significance level
*** = 1% significance level

4.2.2 Gender differences in competitiveness

This regression confirms what we already expected; there are no significant differences between men and women and their competitiveness, even if we take risk aversion and overconfidence into account. If we purely look at the coefficient of gender, *sex*, we could say that the probability of competing versus not-competing decreases by 2.47% for a man compared to women, ceteris paribus, what would imply that women are slightly more competitive. However, this coefficient is not significant (p-value > 0.1) (Table 7) and 2.5% is an almost neglectable effect.

4.2.3 Competitiveness and risk aversion

We do however find a significant relationship between risk aversion and competitiveness. The coefficient of 0.0906 indicates that with one higher chosen bet, the probability that the respondent selects the tournament scheme increases on average by 9.06%, ceteris paribus. This effect is significant as our p-value is 0.000. The less risk averse or the more risk seeking a respondent is, the bigger the chance that the respondent will choose to compete.

If we take a closer look at risk aversion, it is clear that women are more risk averse than men. This is already discussed before in this paper, and also in the existing literature. To formally test this, we use a two-sided t-test with unequal variances, as the variances differ significantly between men and women (p = 0.0019) (Table 30, Appendix B.2). As the output in Table 31 in Appendix B.2 shows, there is a (almost) significant difference between men and women, where men often choose the higher, riskier bet (p = 0.1245).

4.2.4 Competitiveness and overconfidence

The regression also looks at the relationship between overconfidence and competitiveness. Even though the coefficient of overconfidence is not significant (p = 0.771), it is still interesting to interpret the coefficient. 0.0167919 indicates that when a respondent increases one unit in the level of overconfidence, compared to the baseline underconfident, the probability that the respondent chooses the tournament scheme increases on average by 1.68%, ceteris paribus. In other words, when the respondent for example has a correct self-perceived performance, compared to a respondent that thinks lower of him- or herself,

he or she is more likely to compete. However, this effect is quite small and insignificant, so these findings are not that relevant.

There is however an almost significant difference between genders when it comes to overconfidence when we formally test this. With a p-value of 0.1277, we could say that men are more overconfident than women (Table 33, Appendix B.2).

4.2.5 Competitiveness and age

Interestingly, there is no real age effect. The (non-significant) coefficient of -0.0006304 (Table 7) indicates that age does not have an impact on competitiveness, since the probability of competing decreases 0.063% per year, ceteris paribus, which is neglectable.

In conclusion, there is no significant relationship between gender and competitiveness, and therefore we reject hypothesis 1. We do however find that men tend to be more overconfident and risk seeking than women.

4.2.6 Probit model

In the original paper by Niederle and Vesterlund, a probit model was used to test the effect of gender on the choice in task 3. In Appendix B.2, Table 34 the results of this test can be found. The test confirms what we already know; there is no significant gender difference in competition.

4.3 Hypothesis 2 – Gender differences in encouragement

Encouragement (i.e. competing against 'bad' opponents) increases the likeliness of competing, and this effect is stronger for men than women.

As a reminder, encouragement/competing against bad opponents means that the subject is competing against the 3 respondents who performed the worst in the previously conducted pilot. Every respondent that was assigned to the encouragement group, was informed about the fact that they were competing against these 3 people and thereafter had to decide to compete under the tournament scheme, or to play under the piece rate scheme.

Firstly, we want to test whether encouragement has a treatment effect in general, regardless of gender. We will compare the encouragement group with the control group, as they did not receive any form of treatment.

Table 8: Mean competitive per treatment and gender

	Treatment			
Sex	Control	Discouragement	Encouragement	Total
Female	0.65384615	0.5483871	0.82142857	0.67058824
Male	0.68421053	0.52380952	0.79166667	0.671875
Total	0.66666667	0.53846154	0.80769231	0.67114094

Table 9: Competition difference relative to control per treatment and gender

	Difference for women relative to control	Difference for men relative to control	Difference between men and women
Encouragement	0.16758242	0.10745614	0.0601263
Discouragement	-0.10545915	-0.16040101	0.0549419

Looking at Table 8, we can see that the mean of the variable competition is 0.667 for the control group, i.e. 66.7% of the respondents in the control group have chosen the competitive option (tournament scheme) and 33.3% the piece rate scheme. In the encouragement group, the mean of *competitive* is 0,807, meaning that 80.7% of the respondents chose the competitive option. This indicates that encouragement increases the likeliness of competing (by 14%), as the first part of hypothesis 2 states as well.

We performed a t-test to formally test if the difference between competition in the control group and encouraged group is significantly different. In the results in Table 36 in Appendix B.3, we can see that the mean competition is significantly less in the control group than the encouragement group, at a 10% (and almost 5%) significance level. This proves the first part of hypothesis 2, showing that encouragement increases the competition/likeliness to compete.

As for the second part of our hypothesis, we are looking for a gender difference in this encouragement effect. The hypothesis states that the effect of encouragement is stronger for men than for women, i.e. the difference between the control and encouragement group in competitiveness should be larger for men.

We can see that there is a clear difference between men and women (Table 9), however the effect is the other way around; women respond more to the treatment than men. Women are 16.76% more likely to compete when encouraged than in the control group (0.8214 - 0.6538), whereas for men this increase is only 10.75% (0.7917 - 0.6842).

This 6% difference is an indication of a treatment effect, but to test if there is a significant difference, we performed a difference in difference test. The output can be found in Table 37 in Appendix B.3. As the p-value of the sex#encouragement variable is 0.740, the treatment effect is not significant. One potential reason for this underpowered effect is the lack of observations, as there are only 97.

Even though there is not a significant effect, we can still see that women will probably respond stronger to the encouragement compared to men. Encouragement significantly increases the likeliness to compete for all respondents, but this effect is probably stronger for women. This means we did not find support for hypothesis 2, as we have found the effect is the other way around as we have stated before.

4.4 Hypothesis 3 – Gender differences in discouragement

Discouragement (i.e. competing against 'good' opponents) decreases the likeliness of competing, and this effect is stronger for women than men.

Hypothesis 3 is very similar to hypothesis 2, only we are now researching the effect of discouragement. As a reminder, discouragement means competing against the 3 best performing subjects from the pilot. Again, the respondents who were randomly assigned to the discouragement group were informed about their competition and had to decide to compete against these people or to play under a piece rate scheme.

We again first want to see if there is a general effect of discouragement, and we will compare all the discouraged respondents with all the respondents in the control group.

As we can see in Table 8, the mean of variable competition is 0.5385 in the discouragement group, meaning that 53.85% of the respondents who were discouraged chose the competitive option, whereas 66.7% of the respondents in the control group selected the tournament scheme. This is a decrease of almost 13%, indicating that the discouragement has a negative impact on the likeliness of competing.

Again, we performed a t-test to see if the difference in competition between the control and discouraged group is significant. As can be seen Table 38 in Appendix B.4, the mean competition is significantly lower in the treatment group than in the control group at a 10% significance level (p-value is 0.1015). This proves the first part of hypothesis 3, that discouragement has a negative effect on the likeliness to compete.

The second part of hypothesis 3 states that the discouragement effect will be stronger for women than for men, i.e. the difference in competition between the control and discouragement group will be larger for women. Our results for hypothesis 2 already showed that women show a larger encouragement effect, but this does not say anything yet about the effect of discouragement.

As can be seen Table 9, competitiveness for women decreases by approximately 10.5% when they are discouraged compared to the women in the control group (0.5484 - 0.6538) . For men, their likeliness to compete decreases by 16% compared to the control group (0.5238 - 0.6842).

This 5.5% difference indicates a treatment effect that is stronger for men than for women, but we again tested this with a difference in difference test (Table 39, Appendix B.4). The difference is not significant, as the coefficient of sex#discouragement has a p-value of 0.789. However, we can still see that there is an indication that the discouragement effect is larger for men. In conclusion, we found that discouragement has a negative significant effect on the likeliness to compete for all respondents, and this effect is probably stronger for men than for women. We did not find support for hypothesis 3, as we have found that again, the effect is the contrary.

5. Conclusion and Discussion

5.1 Conclusion

The goal of this research paper was to find an answer to the following question: “How does the skill level of the opponent affect competitiveness for men and women?”. Three hypotheses, focusing on competitive differences between men and women in general, and a focus on encouragement and discouragement tried to help answer this question.

As for the first hypothesis, we surprisingly did not find support that men are more competitive than women. Based on the elaborate existing literature, we expected that men would be more likely to enter a tournament, taking overconfidence and risk aversion into account. We have however seen that the chances of men and women entering the tournament setting are equal. We did however find evidence that men are more overconfident than women, and the experiment also shows that women tend to more risk averse than men.

As for the encouragement and discouragement effect, we have found interesting results, even though they are the opposite as what we would have expected. But since the literature on the effect of skill level of the opponent on competition was not extensive, it was difficult to form well considered expectations.

As expected, we have found significant evidence that encouraging people, i.e. informing them that their competition is “bad”, motivates people to compete, and discouraging people, i.e. informing them that their competition is “good”, will decrease the urge to compete.

Interestingly, we have found that the encouragement effect is larger for women, and that they are 6% more likely than men to enter a competition when knowing they have “bad” competition compared to having no knowledge about their competition (control).

For the discouragement effect, we have found that men respond heavier, and will be 5.5% more likely than women to step out of the competition when they have strong competition compared to when they have zero knowledge about their competition.

Unfortunately, neither of these findings are significant. Still, these findings offer great opportunity to make suggestions about how competition and the skill level of opponents can be optimally used in business to motivate women (and potentially men as well) to compete for higher jobs.

5.2 Implications

We can make suggestions on how to apply these findings in practice. Firstly, we have seen that encouragement has a positive effect on competitiveness in general. So, if you want to motivate all employees, you could emphasize the fact that they have a good chance at winning. When there are actual numbers available, it will help to personalize the encouragement by informing an employee that he or she is performing at the top and has a realistic chance at winning.

The opposite might, although less likely, occur as well. If you want to discourage employees for some reason, you should inform the employee about the extensive and heavy competition, and that way they will probably quit sooner.

These suggestions can also be applied to situations where you want to motivate or discourage a certain gender. In some industries there are now mandatory women's quotas, and a company could for example need a certain percent of the board to be female. That might be a real challenge, certainly in industries where women are underrepresented, or just not as ambitious as men. In that situation, personalized competition based on gender might work to motivate women, for example for a promotion.

In this case, you can encourage women by informing them about the relative position, of course only when it is good. This can be done with a ranking, saying that for example they are in the top 3 of the team and if they keep going like this, they will win. In some situations, you could maybe even say that they have a higher chance of succeeding because they are female, but this is a very sensitive subject and would first need further research before one can say whether this positively influences women.

As a company, you could also unnotably want to discourage men, when you inform them that there are others that are equal or better than them. Men might give up sooner than women, potentially scared to lose. However, it might feel a bit unethical to discourage men like this based on their gender, and it would be wise to be cautious when you do this as a company.

5.3 Limitations

We have made some interesting suggestions, but as for every research, this research paper comes with some limitations, that will have an effect on the validity, both internal and external.

First of all, due to the relatively low number of observations, the findings are not significant. In total, the experiment consisted of 149 respondents, but these respondents were divided into a lot of groups. There was a control, encouragement and discouragement group, and we looked at men and women separately. This is a total of 3 x 2 = 6 groups, and therefore some groups had limited respondents (the control group only consisted of 19 men). As a result, it was almost impossible to find significant results.

Secondly, the question can be raised on how trustworthy and representable this data set is. As can be read in the literature section, there has been done extensive research on the difference between men and women and their competitiveness, and almost all research shows that men are significantly more competitive than women. However, we tested this for hypothesis 1, but have not found any support for this. A possible explanation would be that only 10% of the respondents were actually paid out, and that might influence the motivation of the respondents.

It can also be that our sample consists of relatively competitive women, which also has an effect on the other conclusions in the paper. It raises the question whether the findings in the research paper are applicable or representable for the real world, as the basic fact of more competitive men was not found in this sample.

Finally, the suggestions made can only be used in limited situations. For example, you need a clear ranking of the opponents to be able to say how well or bad you are doing compared to the rest. However, the findings in this research paper about discouragement and encouragement could maybe be suitable for other areas as well, but more research had to be done on this.

5.4 Future research

There are endless possibilities for future research, as the gender gap in high level jobs, but also in a lot of other areas, is a hot topic and very important to research further. It is first important to find out if the findings in this research paper actually hold in a larger experiment.

The first suggestion would be to repeat the experiment with a more representable and bigger sample. Not only were men and women evenly competitive and is this probably not true in the real world, the average age of our respondents was 28. In the future, you could do something similar with a broader range of respondents, and in a field experiment where every respondent receives their payoff in real money. In this experiment, you might find support that men are more competitive indeed. It would then be interesting to see if the encouragement effect is still then still stronger for women and the discouragement effect is still stronger for men.

You could research to what extent the way you frame the skill level of the competition matters for the competition. For example, you could test if a ranked list with all employees has a different effect than just saying one is in the top 50%. Also, these findings are only applicable to situations where such a ranking exists, but competition can be used in other ways in the workplace as well in order to motivate women to get higher up.

Finally, competitive preferences are only one factor in the equation why men serve higher level jobs than women. There has been a lot of research done on the gender gap in high level jobs, as mentioned previously, and more recently, Grossman together with some other professors found something else that influences the gender gap in leading roles (Grossman et al., 2019).

Grossman found that even if women want to promote or compete for higher level jobs, society might not agree with this. According to the research, there is a general gender bias against female leaders. Factors that influence leader success, such as competitiveness, were taken out, but results still showed that people perceive women to be a less successful leader.

The task to decrease the gender gap in high level jobs is twofold; on the one hand, women need to have the same motivation as men, and competition and the implications done in this research paper might help to facilitate this. But on the other hand, there needs to be a culture in which women are accepted as leaders just as much as men are, and that might be just as big of a task.

All in all, we have come a long way in decreasing gender differences already, both in the workplace and in other parts of society, but without a doubt there is still a long road ahead to achieve complete gender equality.

6. Reference list

Barber, B. M., & Odean, T. (1998). Boys will be Boys: Gender, Overconfidence, and Common Stock Investment. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.139415

Baumann, C., & Winzar, H. (2014). The role of secondary education in explaining competitiveness. *Asia Pacific Journal of Education*, *36*(1), 13–30. https://doi.org/10.1080/02188791.2014.924387

Bengtsson, C., Persson, M., & Willenhag, P. (2005). Gender and overconfidence. *Economics Letters*, *86*(2), 199–203. https://doi.org/10.1016/j.econlet.2004.07.012

Bertrand, M., & Hallock, K. F. (2001). The Gender Gap in Top Corporate Jobs. *ILR Review*, *55*(1), 3–21. https://doi.org/10.1177/001979390105500101

Bonner, S. E., & Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research. *Accounting, Organizations and Society*, *27*(4–5), 303–345. https://doi.org/10.1016/s0361-3682(01)00052-6

Borghans, L., Golsteyn, B., Heckman, J. J., & Meijers, H. (2009). Gender Differences in Risk Aversion and Ambiguity Aversion. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.1351149

Chadwick, G. (2019, July 10). *5 Factors that affect employee motivation*. Penguins. Retrieved April 13, 2022, from https://www.penguins.co.uk/incentive-travel-blog/5-factors-that-affect-employee-motivation

Cho, S. Y. (2017). Explaining Gender Differences in Confidence and Overconfidence in Math. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.2902717

Cho, K. W. (2019). Exploring the dark side of exposure to peer excellence among traditional and nontraditional college students. *Learning and Individual Differences*, *73*, 52–58. https://doi.org/10.1016/j.lindif.2019.05.001

Correll, S. (2001). Gender and the Career Choice Process: The Role of Biased Self-Assessments. *American Journal of Sociology*, *106*(6), 1691–1730. https://doi.org/10.1086/321299

Delfgaauw, J., Dur, R., Sol, J., & Verbeke, W. (2009). Tournament Incentives in the Field: Gender Differences in the Workplace. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1489218

DiMenichi, B. C., & Tricomi, E. (2015). The power of competition: Effects of social motivation on attention, sustained physical effort, and learning. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01282

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, *23*(4), 281–295. https://doi.org/10.1016/s1090-5138(02)00097-1

Eckel, C. C., & Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, *68*(1), 1–17. https://doi.org/10.1016/j.jebo.2008.04.006

Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results*, *1*, 1060–1073.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074. https://doi.org/10.1162/00335530360698496

Gneezy, U., & Rustichini, A. (2004). Gender and Competition at a Young Age. *American Economic Review*, *94*(2), 377–381. https://doi.org/10.1257/0002828041301821

Grossman, P. J., Eckel, C., Komai, M., & Zhan, W. (2019). It pays to be a man: Rewards for leaders in a coordination game. *Journal of Economic Behavior & Organization*, *161*, 197–215. https://doi.org/10.1016/j.jebo.2019.04.002

Gupta, N. D., Poulsen, A., & Villeval, M. (2005). Do (wo)men prefer (non-) competitive jobs? *Economic Science Association, World Meeting.*

Holt, C. A., Laury, S. K. (2002). Risk Aversion and incentive effects. *American Economic Review, 92* (5), 1644-1655

Klein, T. P. (2014). *How Toddlers Thrive*. Touchstone.

LeanIn.Org & McKinsey & Company. (2021, September). *Women in the Workplace 2021*.

Lown, G. (n.d.). *Society says men are more suitable in leadership roles*. Kantar. Retrieved April 13, 2022, from https://www.kantar.com/inspiration/equality/society-says-men-are-more-suitable-in-leadership-roles

Mayr, U., Wozniak, D., Davidson, C., Kuhns, D., & Harbaugh, W. T. (2012). Competitiveness across the life span: The feisty fifties. *Psychology and Aging*, *27*(2), 278–285. https://doi.org/10.1037/a0025655

Madison, B. (2017, June 12). *Workers willing to leave a job if not praised enough*. Employee Benefit News. Retrieved April 23, 2022, from

https://www.benefitnews.com/news/workers-willing-to-leave-a-job-if-not-praised-enough

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. https://doi.org/10.1162/qjec.122.3.1067

Palmer, J. (2021, May 19). *3 Industries Where Talent Competition is Fierce Right Now*. Skywalk Group. Retrieved April 13, 2022, from https://skywalkgroup.com/blog/post/3-industries-where-talent-competition-is-fierce-right-now

Rogers, T., & Feller, A. (2016). Discouraged by Peer Excellence. *Psychological Science*, *27*(3), 365–374. https://doi.org/10.1177/0956797615623770

Slonim, R. L. (2005). Competing Against Experienced and Inexperienced Players. *Experimental Economics*, *8*(1), 55–75. https://doi.org/10.1007/s10683-005-0437-3

Sutter, M., Bosman, R., Kocher, M. G., & Van Winden, F. (2009). Gender pairing and bargaining—Beware the same sex! *Experimental Economics*, *12*(3), 318–331. https://doi.org/10.1007/s10683-009-9217-9

Turner, C. (2014, August 5). *When Kids Start Playing To Win*. NPR. Retrieved April 13, 2022, from https://www.npr.org/sections/ed/2014/08/05/331412567/when-kids-start-playing-to-win?t=1649844239343

Vandegrift, D., & Yavas, A. (2009). Men, women, and competition: An experimental test of behavior. *Journal of Economic Behavior & Organization*, *72*(1), 554–570. https://doi.org/10.1016/j.jebo.2009.06.003

Veldhuizen, V. R. (2017). *EconStor: Gender differences in tournament choices: Risk preferences, overconfidence or competitiveness?* ECONSTOR. Assessed at 20th of December 2021, van https://www.econstor.eu/handle/10419/162513

Vermeulen, L., Núñez Castellar, E., & Van Looy, J. (2014). Challenging the Other: Exploring the Role of Opponent Gender in Digital Game Competition for Female Players. *Cyberpsychology, Behavior, and Social Networking*, *17*(5), 303–309. https://doi.org/10.1089/cyber.2013.0331

WOMEN Inc. (2022, March 7). *Mijn naam is Peter*. Retrieved April 13, 2022, from https://www.womeninc.nl/actueel/mijn-naam-is-peter

7. Appendices

7.1 Appendix A – Methodology section

A.1 Pilot instructions

Link to the pilot: https://erasmusuniversity.eu.qualtrics.com/jfe/form/SV_5urAHjZaG1k0S2y

Introduction:

¨This survey is part of my master thesis for Behavioral Economics and will act as a pilot for the actual research. You will be asked to perform a short task. At the end of the pilot, 10% of the subjects will be randomly selected who will receive the money they have earned with the task. So make sure that you try your best, as you can win real money!
Thank you in advance for your corporation. It will take approximately 2 minutes to complete the survey.¨

Followed by questions about age (fill in any number), gender (multiple choice question with options male, female or non-binary) and the highest level of education you completed (multiple choice question with the options elementary school, high school, MBO (practical education, Bachelor´s degree or Master´s degree).
The task is explained in the following way:

¨In this task, you will have 45 seconds to correctly answer as many questions as possible. Every question will show a picture with multiple figures, and you are asked to count how often a specific figure is displayed in the picture. To test if you understand the question, the following question is a practice question. The figure you have to count can be found at the top of the picture every time.¨

The participants first had to answer a practice question, to get an impression of what the questions will look like.

Participants received feedback for this practice question, by telling them the correct answer is 6. Then, they got an explanation about the actual task:

¨The task will now begin. You will have 45 seconds to answer as many questions as possible. When you think you answered correctly, you can scroll down for the next question. After 45 seconds, the page will automatically be closed. In case you are selected for the actual pay out, you will receive 1 EURO PER CORRECT ANSWER. ¨

Participants could see the timer at the start, but if they scrolled down to the next question, the timer disappeared from their screen. In other words, participants were not fully aware of the time left while performing their task. After 45 seconds, the page was automatically closed, and the participants were thanked for their participation. No feedback was provided about their performance.

A.2 Pilot Data

Table 10: Descriptive statistics and data pilot

Respondent number	Age	Gender	Education	Number correct
Respondent 1	22	Female	Bachelor	11
Respondent 2	21	Female	Bachelor	8
Respondent 3	23	Male	Bachelor	11
Respondent 4	54	Female	Master	6
Respondent 5	55	Male	Master	7
Respondent 6	54	Female	Master	8
Respondent 7	18	Female	High School	8
Respondent 8	20	Female	High School	10
Respondent 9	22	Female	Bachelor	10
Respondent 10	25	Male	Master	6

Table 11: Pilot Ranking Best to Worst

Ranking:	Respondent number	Number correct
1	Respondent 1	11
	Respondent 3	11
3	Respondent 8	10
	Respondent 9	10
5	Respondent 2	8
	Respondent 6	8
	Respondent 7	8
8	Respondent 5	7
9	Respondent 4	6
10	Respondent 10	6

Table 12: Composition of Pilot Groups (random, bad and good group)

	Random Group	Bad Group	Good Group
1	Respondent 3 (11 answers)	Respondent 10 (6 answers)	Respondent 1 (11 answers)
2	Respondent 7 (8 answers)	Respondent 4 (6 answers)	Respondent 3 (11 answers)
3	Respondent 10 (6 answers)	Respondent 5 (7 answers)	Respondent 8 (10 answers)

Table 13: T-test to test the performance difference between genders in the pilot

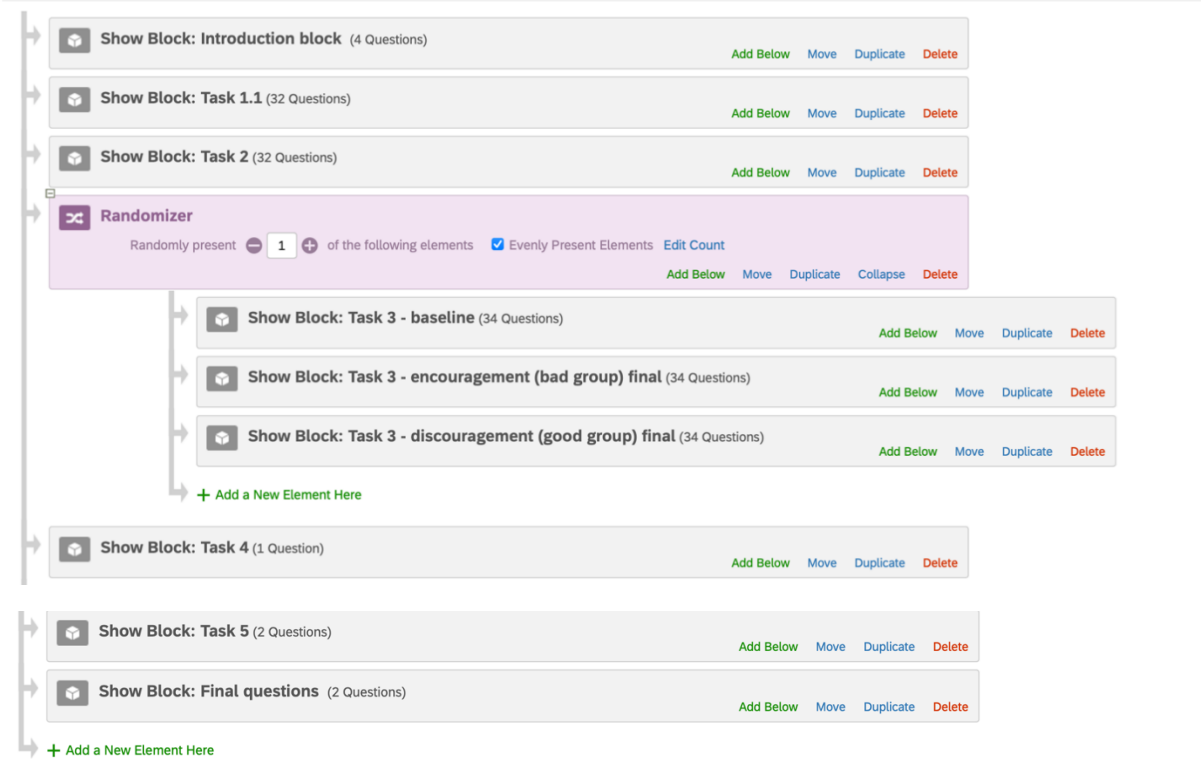
Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Female	7	8.714286	0.6441785	1.704336	7.138038	10.29053
Male	3	8	1.527525	2.645751	1.427589	14.57241
Combined	10	8.5	0.6009252	1.900292	7.140613	9.859387
Diff		0.7142857	1.367753		-2.439758	3.86833

diff = mean (Female) – mean (Male)		t = 0.5222
H0: diff = 0		Degrees of freedom = 8
Ha: diff < 0	Ha: diff ! = 0	Ha: diff > 0
P = 0.6922	P = 0.6157	P = 0.3078

A.3 Survey flow

Link to the survey:
https://erasmusuniversity.eu.qualtrics.com/jfe/form/SV_6AsdcPtdQcxoqP4

Figure 4: Survey flow



A.4 Survey instructions

The survey started with an introduction:

¨This survey is part of my master thesis for Behavioral Economics. Thank you in advance for your corporation. It will take approximately 5 minutes to complete this survey. The survey will consist of 5 TASKS, and with each task, money can be earned. The money that can be earned will differ among the tasks, and will be explained at the beginning of each task. At the end of this study, 10 subjects will be randomly selected who will receive the actual money they have earned with 1 OF THE 5 TASKS. What task will be paid out, will also be selected randomly. So make sure that you try your best, as you can win real money!¨

Followed by questions about age (fill in any number), gender (multiple choice question with options male, female or non-binary) and the highest level of education you completed

(multiple choice question with the options elementary school, high school, MBO (practical education, Bachelor´s degree or Master´s degree).

Task 1 began with an explanation:

¨In this first task, you will have 45 seconds to correctly answer as many questions as possible. Every question will show a picture with multiple figures, and you are asked to count how often a specific figure is displayed in the picture.¨

Task 1 contained the same practice question as the pilot, and all questions had the same layout as the example in Appendix A.1.

Task 2 was explained in the following way:

¨You will again have 45 seconds to identify the correct number of icons for every picture. But in this round, you will compete against 3 other people. We conducted a pilot in which 10 subjects performed the same task, and you are competing against 3 subjects that are randomly chosen from the pilot group. In case you answer the most questions correctly compared to the other 3, you will receive 4 EUROS PER CORRECT ANSWER, but in case someone else will do better, you will receive NOTHING.¨

After the page was automatically closed again, task 3 was explained:

¨In this task, you can choose to repeat either task 1 or 2. As a reminder:

- *In task 1, there was no competition and you received 1 euro per correct answer*
- *In task 2, you played against 3 other people from the pilot, and in case you had the most correct answers, you received 4 euro per answer and 0 otherwise.¨*

Part of the explanation was different for every treatment group:

- Control group: In case you want to play the competitive round, you will again play against 3 randomly selected subjects from the pilot.
- Discouragement group: In case you want to play the competitive round, you will now play against THE 3 BEST SUBJECTS FROM THE PILOT. So, from the 10 subjects who played the game already, you will compete against the 3 subjects who answered the most amount of questions correct.
- Encouragement group: In case you want to play the competitive round, you will now play against THE 3 WORST SUBJECTS FROM THE PILOT. So, from the 10 subjects who played the game already, you will compete against the 3 subjects who answered the least amount of questions correct.

This explanation was followed by a multiple-choice question, where option A was a non-competitive round (as in round 1) or a competitive round (as in round 2). Round 3 again took 45 seconds, whereafter the page was closed.

Task 4 was explained:

¨How do you think you performed compared to the competition in task 2, so the first time you played against other people? If you think you answered most questions correct, rank yourself first and so on. In case you ranking is correct, and this task is selected for actual money, you will receive 5 EUROS. But if you do not rank yourself correctly, you receive NOTHING. ¨

The participants could drag ¨Yourself¨ to the desired position.

The final task, task 5, was introduced with following explanation:

¨You are asked to choose one of the 5 bets below in the table. If this task is selected for actual money, you will receive the bet you chose. So, when you have chosen bet number 1, you will definitely receive 8 euros. But if you chose bet number 3, you have a 50% chance of receiving 16 euros, and a 50% chance of receiving 4 euros. ¨

Figure 5: Overview of the bets in task 5

Bet number	A		B	
1	50%	€8	50%	€8
2	50%	€12	50%	€6
3	50%	€16	50%	€4
4	50%	€20	50%	€2
5	50%	€24	50%	€0

Participants could choose one of the five bets. Finally, the participants were thanked for their participation and could leave their email address.

¨Thank you so much for your participation in this survey, I highly appreciate it. In case you would like to get a shot at getting your answers paid out for real, please leave your email address below. I will send you an email if you have been selected, so that I can arrange to get the money to you. ¨

A.5 Survey data & invalid responses

The data sets are separate Excel files and can be downloaded here.

Four responses were considered to be invalid and were therefore deleted from the sample. Three of those because the respondents answered (almost) all questions in each task, in a timespan ranging from 15 till 30 seconds. This is impossible to do as a normal person, and because we suspect the respondents cheated in some way, we deleted these responses entirely from the sample.

One respondent answered for almost every question in every task 1 number above the correct answer. We suspect that he or she included the example after the ¨Count number of¨ sentence in their counting and misunderstood the task. For this reason, the subject is deleted from the sample.

A.6 Descriptive statistics

Table 14: Statistics on completed education sample

	Female	Male	Total
Elementary school	3	0	3
High School	22	10	32
MBO (practical education)	2	4	6
Bachelor's degree	39	31	70
Master's degree	19	19	38
Total	85	64	149

7.2 Appendix B – Results section

B.1 Performance under piece rate and tournament scheme

Table 15: F-test to test for equal variances in task 1 between genders

F (2, 141)	0.14
Prob > F	0.8710

Table 16: T-test to test the performance difference between genders in task 1

Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Female	85	7.764706	0.2236989	2.062402	7.319856	8.209556
Male	64	7.953125	0.2444114	1.955291	7.464708	8.441542
Combined	149	7.845638	0.1648741	2.012545	7.519826	8.171449
Diff		-0.1884191	0.3338423		-0.8481694	0.4713311

diff = mean (Female) – mean (Male)	t = -0.5644	
H0: diff = 0	Degrees of freedom = 147	
Ha: diff < 0	**Ha: diff ! = 0**	Ha: diff > 0
P = 0.2867	**P = 0.5733**	P = 0.7133

Table 17: F-test to test for equal variances in task 2 between genders

F (2, 141)	2.29
Prob > F	0.1050

Table 18: T-test to test the performance difference between genders in task 2

Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Female	85	7.752941	0.192947	1.778883	7.369245	8.136637
Male	64	7.76525	0.2441973	1.953578	7.168261	8.144239
Combined	149	7.711409	0.151567	1.850111	7.411894	8.010924
Diff		0.0966912	0.3112248		-0.5190919	0.7124743

diff = mean (Female) – mean (Male)	t = 0.3107	
H0: diff = 0	Satterthwaite's degrees of freedom = 128.619	
Ha: diff < 0	**Ha: diff ! = 0**	Ha: diff > 0
P = 0.6217	**P = 0.7565**	P = 0.3783

Table 19: F-test to test for equal variances between tasks

F (2, 141)	2.83
Prob > F	0.0627*

* = 10% significance level
** = 5% significance level
*** = 1% significance level

Table 20: T-test to test the performance difference between tasks

Variable	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Task1correct	149	7.845638	0.1648741	2.012545	7.519826	8.171449
Task2correct	149	7.711409	0.151567	1.850111	7.411894	8.010924
Combined	298	7.778523	0.1118568	1.930948	7.558391	7.998656
Diff		**0.1342282**	0.2239554		-0.3065311	0.5749875

diff = mean (Task1correct) – mean (Task2correct)	t = 0.5994	
H0: diff = 0	Satterthwaite's degrees of freedom = 293.928	
Ha: diff < 0	Ha: diff ! = 0	**Ha: diff > 0**
P = 0.7253	P = 0.5494	**P = 0.2747**

Table 21: T-test to test the performance difference between tasks, sorted by gender
Sex = Female

Variable	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Task1correct	85	7.764706	0.2236989	2.062402	7.319856	8.209556
Task2correct	85	7.752941	0.192947	1.778883	7.369245	8.136637
Combined	170	7.758824	0.1472703	1.92017	7.468097	8.04955
Diff		**0.0117647**	0.2954145		-0.5715295	0.5950589

diff = mean (Task1correct) – mean (Task2correct)	t = 0.0398	
H0: diff = 0	Satterthwaite's degrees of freedom = 164.455	
Ha: diff < 0	Ha: diff ! = 0	Ha: diff > 0
P = 0.5159	P = 0.9683	P = 0.4841

Sex = Male

Variable	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Task1correct	64	7.953125	0.2444114	1.955291	7.464708	8.441542
Task2correct	64	7.65625	0.2441973	1.953578	7.168261	8.144239
Combined	128	7.804688	0.1725712	1.952421	7.4632	8.146175
Diff		**0.296875**	0.3454986		-0.3868565	0.9806065

diff = mean (Task1correct) – mean (Task2correct)		t = 0.8593
H0: diff = 0		Satterthwaite's degrees of freedom = 126
Ha: diff < 0	Ha: diff ! = 0	**Ha: diff > 0**
P = 0.8041	P = 0.3918	**P = 0.1959**

B.2 Hypothesis 1

Table 22: Correlation between sex and competition

Observations = 149	sex	competitive
sex	1	
competitive	0.0014	1

Table 23: Correlation between sex and overconfidence

Observations = 149	sex	overdum
sex	1	
overdum	0.1423	1

Table 24: Correlation between sex and risk aversion

Observations = 149	sex	Riskaversion
sex	1	
riskaversion	0.0968	1

Table 25: Correlation between age and education

Observations = 149	Age	educcat
age	1	
educcat	0.3264	1

Table 26: Regression age on education

educcat	Coefficient	Standard Error	t	P-value	95% confidence interval	
age	0.029523	0.0070507	4.19	0.000***	0.0155891	0.0434569
Constant	2.889867	0.2177628	13.27	0.000***	2.459517	3.320217

* = 10% significance level
** = 5% significance level
*** = 1% significance level

Table 27: Correlation between competitiveness and age

Observations = 149	competitive	Age
competitive	1	
age	-0.0221	1

Table 28: Correlation between competitiveness and education

Observations = 149	competitive	educcat
competitive	1	
educcat	-0.0062	1

Table 29: Regression output hypothesis 1

competitive	Coefficient	Standard Error	t	P-value	95% confidence interval	
sex	-0.0246538	0.0765753	-0.32	0.748	-0.1760196	0.126712
riskaversion	0.0906113	0.0236205	3.84	0.000***	0.0439209	0.1373016
overconfident	0.0167919	0.057693	0.29	0.771	-0.0972493	0.1308332
age	-0.0006304	0.0031209	-0.20	0.840	-0.0067995	0.0055386
constant	0.4269164	0.1766519	2.42	0.017**	0.0773	0.7761028

* = 10% significance level
** = 5% significance level
*** = 1% significance level

Table 30: F-test to test for equal variances between gender in risk aversion

F (2, 143)	6.54
Prob > F	0.0019***

* = 10% significance level
** = 5% significance level
*** = 1% significance level

Table 31: T-test to test the risk aversion difference between genders

Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Female	84	2.880952	0.1628109	1.492187	2.557128	3.204777
Male	64	3.1875	0.2087793	1.670234	2.770288	3.604712
Combined	148	3.013514	0.1293436	1.573533	2.7579	3.269127
Diff		-0.3065476	0.2647568		-0.8304452	0.2173499

diff = mean (Female) – mean (Male)		t = -1.1578
H0: diff = 0		Sattertwaite's degrees of freedom = 127.213
Ha: diff < 0	Ha: diff ! = 0	Ha: diff > 0
P = 0.1245	P = 0.2491	P = 0.8755

Table 32: F-test to test for equal variances between gender in overconfidence

F (2, 143)	0.08
Prob > F	0.9194

Table 33: T-test to test the overconfidence difference between genders

Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Female	85	1.258824	0.073246	0.6752943	1.113166	1.404481
Male	64	1.390625	0.0907891	0.7263124	1.209198	1.572052
Combined	149	1.315436	0.0572095	0.6983305	1.202383	1.428489
Diff		-0.1318015	0.1154543		-0.3599662	0.0963632

diff = mean (Female) – mean (Male)		t = -1.1416
H0: diff = 0		Degrees of freedom = 147
Ha: diff < 0	Ha: diff ! = 0	Ha: diff > 0
P = 0.1277	P = 0.2555	P = 0.8723

Table 34: Output probit model

competitive	Coefficient	Standard Error	z	P-value	95% confidence interval	
sex	-0.0654387	0.2247607	-0.29	0.771	-0.5059615	0.3750841
riskaversion	0.2644064	0.0726111	3.64	0.000***	0.1220913	0.4067216
overconfident	0.0498763	0.1576735	0.32	0.752	-0.2591581	0.3589107
age	-0.0021747	0.0088847	-0.24	0.807	-0.0195883	0.0152389
constant	-0.2306936	0.4734232	-0.49	0.626	-1.158586	0.6971988

* = 10% significance level

** = 5% significance level

*** = 1% significance level

B.3 Hypothesis 2

Table 35: Number of respondents per treatment per gender

	Female	Male	Total
Control	26	19	45
Discouragement	31	21	52
Encouragement	28	24	52
Total	85	64	149

Table 36: T-test to test competitive difference between control and treatment group (encouragement)

Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Control group	45	0.6666667	0.0710669	0.4767313	0.5234407	0.8098926
Encouragement group	52	0.8076923	0.0551869	0.3979586	0.6968999	0.9184847
Combined	97	0.742268	0.0446405	0.4396578	0.6536574	0.8308787
Diff		-0.1410256	0.0888134		-0.3173425	0.0352912

diff = mean (Control) – mean (Encouragement)		t = -1.5879
H0: diff = 0		Degrees of freedom = 95
Ha: diff < 0	Ha: diff ! = 0	Ha: diff > 0
P = 0.0578**	P = 0.1156	P = 0.9422

* = 10% significance level
** = 5% significance level
*** = 1% significance level

Table 37: Output difference in difference test for gender differences in encouragement

competitive	Coefficient	Standard Error	P – value	95% confidence interval	
sex	0.0303644	0.1329855	0.820	-0.2337185	0.2944472
encouragement	0.1074561	0.1353048	0.429	-0.1612325	0.3761447
sex#encouragement	0.0601263	0.180854	0.740	-0.2990139	0.4192664
constant	0.6234818	0.2002157	0.002***	0.2258931	1.02107

* = 10% significance level
** = 5% significance level
*** = 1% significance level

B.4 Hypothesis 3

Table 38: T-test to test competitive difference between control and treatment group (discouragement)

Group	Observations	Mean	Standard Error	Standard Deviation	95% confidence interval	
Control group	45	0.6666667	0.0710669	0.4767313	0.5234407	0.8098926
Discouragement group	52	0.5384615	0.0698066	0.5033822	0.398319	0.678604
Combined	97	0.5979381	0.0500425	0.4928614	0.4986046	0.6972717
Diff		0.1282051	0.1000121		-0.070344	0.3267543

diff = mean (Control) – mean (Discouragement)		t = 1.2819
H0: diff = 0		Degrees of freedom = 95
Ha: diff < 0	Ha: diff ! = 0	Ha: diff > 0
P = 0.8985	P = 0.2030	**P = 0.1015***

* = 10% significance level
** = 5% significance level
*** = 1% significance level

Table 39: Output difference in difference test for gender differences in discouragement

competitive	Coefficient	Standard Error	P – value	95% confidence interval	
sex	0.0303644	0.1497856	0.840	-0.2670802	0.327809
discouragement	-0.160401	0.1571343	0.310	-0.4724385	0.1516365
sex#discouragement	0.0549419	0.2052049	0.789	-0.3525542	0.4624381
constant	0.6234818	0.2255091	0.007***	0.1756655	1.071298

* = 10% significance level
** = 5% significance level
*** = 1% significance level