Erasmus University Rotterdam

Erasmus School of Economics

MSc thesis program Behavioural Economics

Gender differences in the effect of time pressure and task difficulty on performance and competitive behaviour

Name student: Kayleigh Kraijo

Student ID number: 507499

Supervisor: Jan Stoop

Second assessor: Chen Li

Date final version: 19-07-2022

1. ABSTRACT

This research investigates possible gender differences in performance and competitiveness, considering time pressure and different task difficulties. Previous literature has shown that there are no conclusive gender differences in performance under time pressure. Yet, it is proven that there are gender differences in competitive behaviour. Men seem to be more competitive under time pressure. Also, when considering task difficulty, men are still more willing to compete, however, the gender gap in competition is bigger for easy tasks than for difficult tasks.

In this paper, gender differences are investigated by using an online survey experiment in which participants have to answer math questions in multiple rounds. They have to perform a piece-rate task, a competition task and for the final task they are given a choice between the two. Results show that there are no significant gender differences in performance, however, there is a difference found when looking at competitiveness. Men appear to be more competitive when it concerns low time pressure. Also, no matter if questions are easy or difficult, men are more competitive than women. These results show that there are indeed gender differences and a gender gap in competitive behaviour. To close this gender gap, it should be taken into account that men and women might act differently in a comparable situation, to be able to reach the fairest and most equal environment for every individual.

Keywords: gender, competition, performance, time pressure, task difficulty

2. TABLE OF CONTENTS

3. INTRODUCTION

"*Pay gap between men and women fails to improve*", "*Men still dominate managerial positions*" and "*Why women are more burned out than men*" are several headlines from the last five years (BBC, 2021a; The Guardian, 2017; BBC, 2021b). There are still enormous gender differences in multiple aspects on the work floor. These are just some examples, the gender pay gap, the domination of male managers and higher stress levels for female employees.

Focusing mostly on this last gender difference, this can be the result of multiple factors. Women are known to divide their time between family and work, more than men do. Their work restrictions depend on their children and husband, making them have to adapt their work time to the needs of their family (Maume, 2006). At the workplace, the circumstances also influence stress levels. Professional women, defined as working women, turn out to experience more daily stress than housewives and men in general (Haynes & Feinleib, 1980). A male-dominated environment also makes female workers more subject to chronic stress (Nelson & Quick, 1985).

3.1 Relevance and research question

Even though it is known that female workers experience higher stress levels than their male colleagues, it is unknown how both genders perform under stress and how their behaviour changes in such circumstances. Also, having to perform tasks with high or low difficulty might influence behaviour for men and women in a dissimilar way. This research can help fill this research gap, as it investigates if men and women act differently under stress and with different task difficulties. Alongside this addition to the literature of gender differences, this research can be societally relevant. Results of this research can help inform companies under which circumstances female and male workers can perform best. By making this division, companies can help both their workers with having an environment that helps them be most achieving, but it is also helpful for the company itself, to reach the best outcome possible.

The research question that follows from this, is:

> *"What are the gender differences in performance and competitive behaviour when performing under time pressure and with multiple task difficulties?"*

As the research question suggests, this paper does not primarily focus on the gender differences in performance but will also investigate gender differences in competitive behaviour under time pressure and with multiple task difficulties.

3.2 Studies on time pressure and task difficulty

First, literature regarding similar experimental research will be investigated to gain a greater understanding of gender differences within multiple aspects. For example, women generally are to suffer more stress than men, and they cope with this using their emotions more than men do, according to Matud (2004). Other research also shows that women know higher stress levels than men, but when they have to perform a certain task, stress levels for both men and women rise (Herrero, Saldaña, Rodriguez, & Ritzel, 2012). Relating competition back to the work floor, it is found that male-led teams are more responsive to competition when a big part of the team is male, whereas female-led teams are more responsive to competition when a bigger part of the team is female (Delfgaauw, Dur, Sol, & Verbeke, 2013). These papers show that gender differences can be found anywhere.

As for gender differences in overall performance, previous studies are inconclusive. Starting with verbal tasks, the overall gender difference appears to be small (Petersen, 2018). However, female students are relatively better in writing than reading, even though they score better than male students in both tasks. Next to verbal tasks, it has been found that in the last few years, the mathematic scores of both men and women have increased (Ding, Song, & Richardson, 2006). However, the average score of women is higher than that of men. Finally, research from Voyer (2011) shows that gender differences in mental rotation – rotating mental pictures of 2D or 3D objects in your mind (Wikipedia, 2022) – are bigger when there are time constraints. Also, the gender difference linearly increases with the time constraint.

The main aspect that will be investigated in this research, is the difference in competitive behaviour between men and women. Niederle and Vesterlund (2007) are at the beginning of researching this gender difference. They find that when considering only performances, there are no significant gender differences. Yet, when participants can choose between a non-competing task and a competing task, twice as many men as women choose to compete. In their paper from 2011, the authors further explain this gender difference. Niederle and Vesterlund (2011) find that the difference depends mostly on differences in how men and women see competition, and how self-confident they are in competition.

Datta Gupta, Poulsen, and Villeval (2005) find the same results regarding competitiveness. They also conclude than men indeed choose more competition than women. The choice of competition for both men and women does not rely on beliefs of relative ability. Women turn

out to be more risk averse than men, making them less willing to take on competition. Men, on the other hand, depend their choice more on the gender of their opponent, and not so much on risk. Moreover, Groβe and Riener (2010) relate to these papers, as they find that gender differences in competition entry rely mostly on gender specific task stereotypes. Leaving behind the general verbal and mathematical tasks, the authors also find that with a physical task – throwing balls into a bucket – men also choose more competition than women. Lastly, Markowsky and Beblo (2022) have done a meta-analysis of all existing literature regarding gender differences in competitiveness, and they conclude that men are 13 percentage points more likely to choose competition than women. They also find that the gender gap is biggest for math tasks, and smallest for non-students and verbal tasks.

Another aspect that will be investigated in this paper, is time pressure. Men and women differ in their perception of time pressure (Mattingly & Sayer, 2006). Considering gender differences in competition under time pressure, Shurchkov (2012) finds that with math tasks under hight time pressure, women are much worse performers than men. However, with verbal tasks under low time pressure, women are substantially better performers and their willingness to compete also rises. Research by Cahlíková, Cingl, and Levely (2020) contains a lot of factors that coincide with this research as well. However, the researchers use psychological stress, to see how this potentially changes the behaviour of their participants. They show that both men and women are less competitive in a stressful setting. For men, they conclude that they lower their competitiveness, since their stress level influences their competitive preferences. Still, men are more competitive in both a stressful setting and a non-stressful setting.

Another feature that is of importance in this research, is the gender difference in competitive behaviour for multiple task difficulties. For children, it is found that in different settings, girls are to show more consistency in their competitive behaviour between easy and difficult tasks, whereas the behaviour of boys might change (Lewis, Alessandri, & Sullivan, 1992). From another angle, when considering spatial orientation, Coluccia and Louse (2004) find that with difficult tasks, there is a gender difference in orientation skills, whereas for easy tasks this is not the case. This gender difference often means that men have on average better orientation skills than women. Looking at it the other way, given the choice of easy and difficult questions, women prefer more easy questions than men (Slade & Rush, 1991). When investigating the gender difference in competition and keeping the difficulty of the task in mind, research by Hoyer et al. (2020) shows that with easy questions, the gender gap of choosing competition is bigger than when the questions are more difficult. They find that this is the result of men

changing their competitive behaviour, as the female preference remains the same for different task difficulties. For easy questions, as much as 41% of the men choose competition, whereas for difficult questions this is only between 26% and 27%. For women, the difference between easy and difficult is only 1.8%.

Following this literature, there are four hypotheses that will be investigated through this research. They are as follows:

H1: *'Men perform better under time pressure than women.'*

H2: '*Men are more likely to choose a competing scheme than women, under time pressure.'*

H3: *'Men are more likely to choose a competing scheme than women, no matter the difficulty of the task.'*

H4: *'Participants are most likely to choose a competing scheme when the task is easy and time pressure is low, and they are least likely to choose a competing scheme when the task is difficult and time pressure is high.'*

3.3 Report structure

The hypotheses will be answered using data collected through an online survey experiment. Specifically, it will be investigated how male and female participants act differently under time pressure and with different task difficulties. In the remainder of this paper, it will firstly be explained what the experiment looks like and how the data is collected. Secondly, the methods that will be used to help analyse the results will be explained, separating the models for each hypothesis. Finally, the results of these methods will be presented, followed by a fitting conclusion and some corresponding implications.

4. DATA AND METHODOLOGY

4.1 Data and experimental design

The data that is used in this paper is gathered through an online survey experiment, conducted via Qualtrics. The set-up of the experiment is based on the experimental design as described in the paper from Niederle and Vesterlund (2007). The experiment consists of three rounds. In each round, participants must answer 25 math questions. However, the circumstances in each round will be rather different.

In the first round, participants will be participating in a piece-rate task. They answer 25 math questions and receive one point per correct answer. In the second round, participants will be taking part in a competition task. They again answer 25 math questions, but this time they are competing against a random other participant of the experiment. The number of points received depends on winning or losing. If the participant has more correct answers than their opponent, they will receive two points per correct answer. If they have less correct answers than their opponent, they will receive zero points.

In the third and final round, participants can choose to either participate in the piece-rate task or the competition task again. In this round, the division of points will be the same as previously described. The only difference is the determination of the winner in the competition task. If the participant chooses the competition task, their number of correct answers in the third round, will be compared to the number of correct answers of their opponent, from the second round. To incentivise participants to choose the option – competition or not – that they think is most successful, all participants have a chance to win their number of points from a certain round paid out in euros. One random participant will be selected to win, but all others are incentivised to think thoroughly about their choice.

The experiment will consist of four treatment groups. The distribution of the participants amongst these groups will be random, by using a randomiser tool that is provided in Qualtrics. The groups differ in the height of an imposed time limit and the difficulty of the math questions asked. The time limit can either be one minute, to ensure high time pressure, or two minutes, to ensure low time pressure. The difficulty of the math questions can either be easy questions or difficult questions. The easy questions are determined by using the multiplication tables, creating questions such as 3×4 or $12 : 2$. The difficult questions are formed using multiplication tables as well, only this time higher numbers are used. This gives, for example,

12 × 8 or *128* : *16*. All these questions, together with the entire survey, are shown in Appendix A. However, before answering the questions, individuals are divided into one of the treatment groups as presented in Table 1.

Table 1 – Treatment group division

	TIME PRESSURE	**DIFFICULTY**
Group 1	Low time pressure	Easy questions
Group 2	High time pressure	Easy questions
Group 3	Low time pressure	Difficult questions
Group 4	High time pressure	Difficult questions

Note: The table shows the treatment groups as used in the experiment. The second column shows whether the participants are presented a low time pressure or high time pressure, and the third column shows whether participants have to answer easy or difficult math questions.

After the participants have answered all the math questions, they are asked some final demographic questions. Specifically, they are asked for their age, gender, highest completed education, occupation, and marital status. Gender is of considerable importance in this research; therefore this must be asked as well. The other demographics are used to perform a randomisation check, to be sure that participants are randomly divided between all the treatment groups.

4.2 Descriptive statistics

The data collection for this study ran from the 21st of April 2022 until the 9th of June 2022. The total sample consists of 165 participants. In order to make analysis easier, the decision is made to consider only those participants who have reported to be either male or female. Also, since there are only two participants who report to be non-binary or third gender, this would not result in a well-founded conclusion. Therefore, deleting these observations does not change any conclusion.

One other respondent is also deleted from the sample, as this respondent is just 11 years old. Firstly, in the Netherlands it is not allowed to work or earn your own money when you are under 13 years old (Ministerie van Algemene Zaken, 2022). Next to this, since this respondent might not fully understand the risk of choosing competition or not, it is decided that this respondent will be excluded from the sample. The descriptive statistics of the remaining sample of 162 observations are shown in Table 2.

Table 2 – Descriptive statistics

	MEAN	STD. DEV.	MIN	MAX
Gender	0.593	0.493	0	1
Age	27.426	10.495	15	68
Education				
Primary school	0.006	0.079	0	1
High school	0.179	0.385	0	1
Vocational education	0.086	0.282	0	1
Bachelor's degree (HBO/WO)	0.556	0.498	0	1
Master's degree (HBO/WO)	0.160	0.368	0	1
PhD	0.012	0.111	0	1
Occupation				
Student	0.710	0.455	0	1
Employed	0.253	0.436	0	1
Unemployed	0.012	0.111	0	1
Retired	0.012	0.111	0	1
Other	0.012	0.111	0	1
Marital status				
Single	0.475	0.501	0	1
In a relationship	0.420	0.495	0	1
Married	0.105	0.307	0	1
Divorced	0	0	0	0
Widowed	0	0	0	0

Note: The table shows the descriptive statistics of 162 observations. The first column shows the variable, the second column shows the mean value of the variable. The third column shows the standard deviation of the variable. In column 4 and 5 the minimum and maximum value of variables is presented, respectively.

The gathered dataset consists of 162 observations. Of the sample, 59.3% of the respondents is female, which corresponds with the value of 0.593 in Table 2. This means that the other 40.7% of the respondents is male. The average age of the respondents is 27.4 years old, the youngest respondent being 15 years old, and the oldest respondent being 68 years old.

For all other demographics, the decimal number reported in Table 2 represents the fraction of the sample that falls within this category. Therefore, the majority of the sample (55.5%) reports to have a bachelor's degree as their highest completed education. Furthermore, 17.9% of the sample reports to have completed high school, followed by 16.0% of the sample who have a master's degree. Additionally, there are a lot of students (71.0%) and employed respondents

(25.3%) in the sample. Finally, almost the entire sample either reports being single or being in a relationship, 47.5% and 42.0% respectively.

4.3 Methodology

To start the analysis, it will be analysed whether the participants are randomly divided between the four treatment groups. This will be done by investigating if the demographics are the same on average for each group. There are two types of variables in the dataset. First, age is the only continuous variable in the dataset. To do a randomisation check for a continuous variable, a Kruskal-Wallis test needs to be performed. The other variables are either binary or categorical. However, since the categories within one demographic variable are considered separately, they are all analysed as binary variables. These variables can have the value of 0 or 1, meaning a participant either belongs in the category (value of 1) or does not belong in the category (value of 0). To determine randomisation for a binary variable, the Chi-squared test should be performed.

4.3.1 Investigating gender differences

Subsequently, gender differences in overall performance will be investigated. This is done to investigate the first hypothesis, that states that men are better performers than women under time pressure. This will be done by using a Mann-Whitney U test, to compare the mean number of correct answers of men and women, to see if they are statistically different under time pressure. To do this, only observations with high time pressure are considered, as only these observations are of interest.

For hypotheses 2 and 3, the data analysis consists of two parts. The first part investigates whether there are any gender differences in choosing competition. This will be done by performing several Fisher exact tests, which can determine if there is a difference in the proportion of men and women that chooses competition for a certain condition. These conditions can be low time pressure, high time pressure, easy questions, or difficult questions. So, for example, the test will see if there is a difference in the proportion of men and women that chooses competition, considering only easy questions.

The null hypothesis for a Fisher exact test is that there is no difference between the proportion of men and women that chooses competition, for a certain condition. The corresponding p-value of the Fisher exact test can help determine whether this hypothesis should be rejected or not.

After this is determined, an Ordinary Least Squares (OLS) regression is performed, to determine the size of the possible effect. Therefore, the main statistical method of this paper focuses on OLS regression models. They will focus on the effect of gender on the choice of competition. Accordingly, the dependent variable is choice of competition and the independent variable of the model is gender. The analysis of these OLS models can help answer the second and third hypothesis. They both state that men are more likely to choose a competing scheme than women. However, the second hypothesis states that this is true under time pressure and the third hypothesis states that this is true for both task difficulties. Consequently, both aspects have to be investigated.

To test the second hypothesis, two OLS models will be considered. The first model focuses only on participants under low time pressure and the second model focuses only on participants under high time pressure. The models are as follows:

$$Y_{LOWPRESSURE} = \alpha + \beta * Gender + \varepsilon \quad (1)$$

$$Y_{HIGHPRESSURE} = \alpha + \beta * Gender + \varepsilon \quad (2)$$

To test the third hypothesis, two more OLS models will be considered. The first model focuses only on participants with easy questions and the second model focuses only on participants with difficult questions. The models are as follows:

$$Y_{EASYQUESTIONS} = \alpha + \beta * Gender + \varepsilon \quad (3)$$

$$Y_{DIFFICULTQUESTIONS} = \alpha + \beta * Gender + \varepsilon \quad (4)$$

For all four models, Y stands for the choice of competition. In equations 1 and 2, it stands for the choice of competition for participants with low time pressure and those with high time pressure, respectively. In equations 3 and 4, it stands for the choice of competition for participants that have to answer easy questions and those that have to answer difficult questions, respectively. The alpha (α) at the beginning of the equations equals the constant, which is the value of Y when all independent variables equal zero. In all four equations, the coefficient of interest follows. This coefficient is multiplied by the gender variable, of which the value either equals 0 if the participant is male, or 1 if the participant is female. Together, this determines how the probability that someone chooses competition changes. Lastly, the epsilon (ε) stands for the error term that remains.

To gain a more precise estimate of the gender coefficient, all four models will be performed once more, only this time with added control variables. The general model looks the same, only

with the control variables age, highest level of completed education, employment status and marital status added to it.

4.3.2 Interaction between time pressure and task difficulty

As an additional analysis, the interaction between time pressure and task difficulty will be investigated. So far, only the separate effects of these two have been investigated, but to draw a proper conclusion, the interaction effect is something that is also of importance. This interaction effect will be calculated by using an OLS regression model. This time, the only variables that are interesting is the dependent variable, choice of competition, and the interaction between time pressure and task difficulty as independent variable. The model that follows from this is as follows:

$$Y = \alpha + \beta_1 * \text{Time pressure} + \beta_2 * \text{Difficulty} + \beta_3 * \text{Time pressure} * \text{Difficulty} + \varepsilon \quad (5)$$

Here, Y again stands for the choice of competition. Because an interaction effect is added, the model now consists of three different coefficients. In the model, time pressure and difficulty are both binary variables, meaning they either have the value of 0 or 1. Time pressure takes the value of 0 if there is a low time pressure, and the value of 1 if there is a high time pressure. Also, difficulty takes the value of 0 if the questions are easy, and the value of 1 if the questions are difficult.

The coefficients β_1 and β_2 should be multiplied with time pressure and difficulty, respectively, to determine the change in probability to choose competition. However, this model also contains an interaction term. Using an interaction term means that the effect of time pressure depends on the difficulty of the task, and vice versa. Otherwise, the interaction term measures how the effect of one variable changes, with a one unit change in the other variable. Therefore, the single coefficients of the variables do not have a meaning on their own. β_3 stands for the interaction term, but the total effect of both time pressure and difficulty will be discussed in the results section.

4.3.3 Additional research on age

As multiple scientific papers have proven (Flory, Gneezy, Leonard, & List, 2018; Mayr, Wozniak, Davidson, Kuhns, & Harbaugh, 2012), competitive preferences can change with age, resulting in a changing gender gap as well. To investigate if this is also the case with the gathered data for this research, an analysis will be done looking at different age groups and their preference for competition. As Flory et al. present evidence that females are most likely to

change their competitive preferences, the analysis will focus both on multiple age groups, as well as gender differences.

Similar to the paper from Flory et al. (2018), this paper will focus on the age groups below and above 50 years old. The age of 50 is seen as a threshold for women to have at least once experienced symptoms of menopause. Considering the sign of the effect of ages above 50 on competitive preferences, existing research is not conclusive. On the one hand, it is stated that the hormone oestrogen represses feelings of competition (Wozniak, Harbaugh & Mayr, 2014). The levels of this hormone decrease with menopause, meaning that competitive feelings could increase. On the other hand, the hormone cortisol is said to positively influence competitiveness (Buser, Dreber, & Mollerstrom, 2017), and this hormone actually increases during menopause (Woods, Mitchell, & Smith-DiJulio, 2009) which would naturally result in more competitive feelings. Therefore, it cannot be stated with certainty which hormonal effect will dominate, but this will be investigated in the final analysis of this paper.

To analyse the effect within this sample – partly the same as Flory et al. (2018) have done – two OLS regression models will be performed. The first model to be performed contains two independent variables; gender and whether a female is over the age of 49. This way it can first be determined whether it is the case that the competitive preferences of women change when they pass the age of 50. The dependent variable of this model is still the choice of competition, and the complete model is as follows:

$$Y = \alpha + \beta_1 * \text{Gender} + \beta_2 * \text{Female over 49} + \varepsilon \quad (6)$$

The same as in other models, Y stands for the choice of competition. The betas represent the coefficients that belong to the independent variables. The alpha (α) stands for the constant term, and epsilon (ε) stands for the remaining error term.

Now, to investigate possible differences between men and women over time, another OLS regression model will be performed. This model is as follows:

$$Y = \alpha + \beta_1 * \text{Female over 49} + \beta_2 * \text{Male under 49} + \beta_3 * \text{Male over 49} + \varepsilon \quad (7)$$

In this model, Y represents the choice of competition, which is the dependent variable. The independent variables are whether a woman is over the age of 49, whether a man is under the age of 49 and whether a man is over the age of 49. Within this model, the alpha (α) represents the constant term, which in this case is the value for a female participant under the age of 49 – as all other variables equal zero. The epsilon (ε) represents the error term that remains.

5. RESULTS

5.1 Randomisation check

Firstly, the results of the randomisation check – to determine if respondents are randomly divided between all four treatment groups – will be discussed. The results of the randomisation check are presented in Table 3. In the table, the considered variables are shown, as well as the method used to determine the randomisation and the p-value. This p-value shows whether or not the hypothesis that there are no differences in the characteristics of respondents in the groups can be rejected.

Table 3 – Randomisation check results

	METHOD	**P-VALUE**
Age	Kruskal-Wallis test	0.793
Gender	Chi-squared test	0.154
Education	Chi-squared test	0.761
Employment	Chi-squared test	0.368
Marital status	Chi-squared test	0.751

Note: This table shows the results of a randomisation check to determine if respondents are randomly divided over the four treatment groups. In the second column, it is presented which method is used to determine this randomisation. The third column shows the p-value of the performed test.

In column 3 of Table 3, the p-values of the performed tests are given. As all these values exceed the significance level of 0.05, it can be concluded that none of the hypotheses can be rejected. Therefore, the randomisation has worked, as there appears to be no difference in the characteristics of the respondents between the separate groups.

5.2 Hypotheses testing

5.2.1 Hypothesis 1

The first hypothesis focuses on the difference in overall performance between men and women. The hypothesis states that under time pressure, men are better performers than women. The result of the Mann-Whitney U test to investigate this possible difference consists of a p-value. The found p-value is 0.2840. This value is higher than the significance level (0.05), meaning that the hypothesis of the Mann-Whitney U test that there are no differences between men and women cannot be rejected. This means that statistically speaking, there are no differences between the number of correct answers of men and women.

However, the performed test also presents the probability that the number of correct answers is bigger for men than for women. This probability is 0.571. Overall, there are no statistical differences in the number of correct answers between men and women, however, there is evidence that men actually perform better than women under time pressure. Therefore, the first hypothesis that men are better performers than women under time pressure cannot be rejected.

5.2.2 Hypothesis 2

Now, the results of the methods used to test the second hypothesis will be discussed. This hypothesis states that men are more likely to choose a competing scheme than women, under time pressure. First, two Fisher exact tests are conducted, to analyse if there is any difference in performance between men and women. The first test determines if, for low time pressure, there is a difference in the proportion of men and women that chooses competition. The two-tailed p-value found, is 0.005, which means that the null hypothesis should be rejected. There appears to be a significant difference in choice of competition between men and women, for low time pressure. Now, for the difficult questions, another Fisher exact test is performed. The two-tailed p-value that is found with this test, is 0.162, which means that the null hypothesis cannot be rejected. There is not enough evidence to conclude that there is a difference between genders in choosing competition, under high time pressure.

Furthermore, the size of the effect of gender on the probability to choose competition is investigated through multiple OLS regression models. The results of this regression are shown in Table 4. Columns 1 and 3 show the results for low and high time pressure, with only gender as the independent variable. In columns 2 and 4, the overall model is the same, only this time the coefficient of gender can be interpreted more precisely, as several control variables are added to the model.

Table 4 – OLS regression results

	LOW TIME PRESSURE	**LOW TIME PRESSURE**	**HIGH TIME PRESSURE**	**HIGH TIME PRESSURE**
	(1)	(2)	(3)	(4)
Gender (female)	-0.324*** (0.097)	-0.277*** (0.096)	-0.161 (0.109)	-0.129 (0.113)
Age		-0.022*** (0.008)		0.003 (0.009)
Education				
Primary school		Omitted variable		Omitted variable
High school		0		0.305*

				(0.174)
Vocational education (MBO)		-0.310 (0.251)		0.375 (0.257)
Bachelor's degree (HBO/WO)		-0.197 (0.126)		0.805*** (0.147)
Master's degree (HBO/WO)		0.016 (0.166)		0.741*** (0.141)
PhD		-0.149 (0.225)		-0.016 (0.202)
Employment				
Student		Omitted variable		Omitted variable
Employed		0.191 (0.134)		-0.190 (0.211)
Unemployed		0		0.513** (0.231)
Retired		0		-0.501 (0.565)
Other		0.421** (0.165)		-0.426 (0.294)
Marital status				
Single		Omitted variable		Omitted variable
In a relationship		-0.144 (0.109)		0.085 (0.135)
Married		0.228 (0.278)		0.062 (0.246)
Constant	0.833*** (0.069)	1.506*** (0.185)	0.722*** (0.076)	-0.009 (0.216)
R^2	0.1023	0.2698	0.0280	0.3211
Observations	85	85	77	77

Note: This table shows the results of an OLS regression with the choice of competition as the dependent variable, and gender as independent variable. In columns 1 and 3, the coefficients of gender are presented for a simple model with only gender as independent variable for low and high time pressure, respectively. In columns 2 and 4, the coefficients of gender are presented for more complicated models with gender as independent variable and age, education, employment status and marital status as added control variables. Robust standard errors are shown between brackets. Stars are used to indicate significance levels. (* p<0.10, ** p<0.05, *** p<0.01)

Within columns 1 and 3 of Table 4, it shows that for both simple models, the probability to choose competition is lower for women than for men. This can be seen by looking at the negative coefficients for female participants. For low time pressure, the probability of choosing competition decreases with 32.4 percentage points (found by multiplying the coefficient by 100%) when a respondent is female, ceteris paribus. This result is statistically significant at a 1% significance level. For the model with high time pressure, the probability of choosing competition decreases when a respondent is female, however, this result is not statistically significant. Therefore, an interpretation of this coefficient is not dependable, and no conclusive conclusion can be drawn.

For the models with added control variables, of which the results are presented in columns 2 and 4, the coefficient of gender is still the only coefficient of interest. Therefore, no other effects will be discussed. For low time pressure, the coefficients for high school, unemployed and retired are zero, as there are no observations that fall in these categories. The results show that with low time pressure, the effect of gender on the probability of choosing competition is still negative. In this model, being female decreases the probability with 27.7 percentage points, ceteris paribus. This result is also statistically significant at a 5% significance level. Next, considering high time pressure, it is found that the association between gender and choice of competition is also still negative. Yet, this result is also insignificant. Therefore, to interpret this coefficient would be unreliable.

It can be stated with certainty that men are more likely to choose a competing scheme under low time pressure than women, but it is uncertain how this likelihood will be under high time pressure. Therefore, there is not enough evidence that men are indeed more likely to choose a competing scheme than women under time pressure, which means that the second hypothesis cannot be rejected.

5.2.3 Hypothesis 3

The methods to test the third hypothesis are similar to those to test the second hypothesis, as the hypotheses are very similar. The only difference is that the effect of gender on choice of competition will be investigated, considering the distinction between easy and difficult questions. The third hypothesis specifically states that men are more likely to choose a competing scheme than women, no matter the difficulty of the task. First, to investigate this difference, two Fisher exact tests are performed.

The first test considers the difference in the proportion of men and women that chooses competition, with easy questions. The two-tailed p-value that is found, is 0.068. This means that the null hypothesis cannot be rejected, therefore, there seems to be no statistical difference between the proportion of men and women that chooses competition with easy questions. The second test looks at gender differences in choosing competition for difficult questions. The two-tailed p-value for this test is 0.004, meaning that the null hypothesis should be rejected. There appears to be a significant difference between men and women in choosing competition, for difficult questions.

Next, to determine the size of the effect of gender and the probability to choose competition, multiple OLS regressions are performed. The results of these regressions are shown in Table 5. Columns 1 and 3 consider a model with the choice of competition as the dependent variable, and only gender as independent variable. Columns 2 and 4 consider the same general model, just with several added control variables, to be more precise when interpreting the gender coefficient.

Table 5 – OLS regression results

	EASY QUESTIONS	**EASY QUESTIONS**	**DIFFICULT QUESTIONS**	**DIFFICULT QUESTIONS**
	(1)	(2)	(3)	(4)
Gender (female)	-0.194**	-0.155	-0.331***	-0.397***
	(0.091)	(0.096)	(0.107)	(0.107)
Age		0.001		-0.009
		(0.008)		(0.008)
Education				
Primary school		Omitted variable		Omitted variable
High school		0		0.280*
				(0.163)
Vocational education (MBO)		0.248*		-0.244
		(0.138)		(0.223)
Bachelor's degree (HBO/WO)		0.028		0.471***
		(0.143)		(0.126)
Master's degree (HBO/WO)		0.257*		0.452***
		(0.140)		(0.157)
PhD		-0.707***		-0.400*
		(0.222)		(0.230)
Employment				
Student		Omitted variable		Omitted variable
Employed		0.028		0.213
		(0.134)		(0.148)
Unemployed		-0.951***		-0.261
		(0.094)		(0.165)
Retired		0.032		0
		(0.502)		
Other		0.049		0.049
		(0.094)		(0.276)
Marital status				
Single		Omitted variable		Omitted variable
In a relationship		0.051		-0.082
		(0.105)		(0.123)
Married		-0.307		-0.054
		(0.250)		(0.249)
Constant	0.867***	0.783***	0.694***	0.625***
	(0.063)	(0.200)	(0.078)	(0.203)
R^2	0.0456	0.2043	0.1084	0.3658
Observations	82	82	80	80

Note: This table shows the results of an OLS regression with the choice of competition as the dependent variable, and gender as independent variable. In columns 1 and 3, the coefficients of gender are presented for a simple model with only gender as independent variable for easy and difficult questions, respectively. In columns 2 and 4, the coefficients of gender are presented for more complicated models with gender as independent variable and age, education, employment status and marital status as added control variables. Robust standard errors are shown between brackets. Stars are used to indicate significance levels. (* p<0.10, ** p<0.05, *** p<0.01)

In the first place, the results of the simple model consisting of only gender as independent variable will be analysed, to get a partial understanding of the effect of gender. Columns 1 and 3 of Table 5 show that for the easy questions, when a respondent is female, the probability of choosing competition decreases with 19.4 percentage points, ceteris paribus. This result is statistically significant at a 5% significance level. For the difficult questions, it can be seen that when a respondent is female, the probability of choosing competition decreases with 33.1 percentage points, ceteris paribus. This result is statistically significant at a 1% significance level. The results also show that the gender gap in competitiveness is bigger for difficult questions than for easy questions, as with difficult questions there is a bigger difference in the probability to choose competition for men and women.

Because of the addition of multiple control variables, the interpretation of the gender coefficient, can be done more precisely. These results are shown in columns 2 and 4 of Table 5. This time, the coefficient of high school for easy questions and the coefficient of retired for difficult questions is zero, as no observations fall within these categories. The results show that for the easy questions, the effect of gender on choice of competition remains negative. However, in this model, the coefficient is not statistically significant. The coefficient cannot be interpreted reliably. As for the difficult questions, this effect also remains negative. It even becomes a little more negative, showing a decrease of 39.7 percentage points when a participant is female, ceteris paribus. This result is statistically significant at a 1% significance level.

These results are in line with the third hypothesis. Even though not all results are statistically significant, there is enough evidence that for both task difficulties, men are more likely to choose a competing scheme than women. Therefore, the third hypothesis cannot be rejected**.**

5.2.4 Hypothesis 4

As for the final hypothesis, it is stated that participants are most likely to choose a competing scheme when the task is easy and the time pressure is low, and least likely to choose a competing scheme when the task is difficult and time pressure is high. This is tested by performing an OLS

regression with choice of competition as the dependent variable, and an interaction term between time pressure and task difficulty as independent variable. The results of this regression are presented in Table 6.

Table 6 – OLS regression results

	Probability of choosing competition
Time pressure (high)	0.023 (0.098)
Task difficulty (difficult)	-0.233** (0.104)
Time pressure (high) × *Task difficulty (difficult)*	-0.002 (0.150)
Constant	0.733*** (0.067)
R^2	0.0580
Observations	162

Note: This table shows the results of an OLS regression with the choice of competition as the dependent variable, and an interaction term between time pressure and task difficulty as independent variable. In the second column, the coefficients are shown that represent the change in probability of choosing competition. Robust standard errors are shown between brackets. Stars are used to indicate significance levels. (* p<0.10, ** p<0.05, *** p<0.01)

To determine the effect of time pressure and task difficulty, it is useful to fill in the found coefficients in the regression equation and analyse the coefficients like this. This way, the following formula is formed:

$$Y = 0.733 + 0.023 * \text{Time pressure} – 0.233 * \text{Difficulty} – 0.002 * \text{Time pressure} * \text{Difficulty} \quad (8)$$

It is found that the effect of time pressure on choice of competition is 0.023 – 0.002 * difficulty. This means that when the task is easy (difficulty = 0), the effect of time pressure on choice of competition is 2.3 percentage points, and when the task is difficult (difficulty = 1), the effect of time pressure on choice of competition is 2.1 percentage points, ceteris paribus. However, since the results are not statistically significant at a 5% significance level, it cannot be concluded that time pressure has a significant effect on the choice of competition.

For task difficulty, the same analysis can be done. The effect of task difficulty on choice of competition is -0.233 – 0.002 * time pressure. This means that when the time pressure is low (time pressure = 0), the effect of task difficulty on choice of competition is -23.3 percentage points, and when time pressure is high (time pressure = 1), the effect of task difficulty on choice of competition is -23.5 percentage points, ceteris paribus. However, since these results are not

all statistically significant at a 5% significance level, it cannot be concluded that task difficulty has a significant effect on the choice of competition.

Now that the separate effects are discussed, it can be determined when the probability of choosing competition is highest and lowest. As the hypothesis states, the highest probability should be found when considering a low time pressure and easy questions, and the lowest probability should be found when considering a high time pressure and difficult questions. Again, the model with filled in coefficients is used, to determine the probability of choosing competition. All the possible outcomes of this model are presented in Table 7.

Table 7 – Possible outcomes

TIME PRESSURE	**TASK DIFFICULTY**	***PROBABILITY OF CHOOSING COMPETITION***
Low time pressure	Easy questions	0.733
High time pressure	Easy questions	0.756
Low time pressure	Difficult questions	0.500
High time pressure	Difficult questions	0.503

Note: This table presents all possible outcomes in the OLS regression model with choice of competition as the dependent variable, and an interaction term between time pressure and task difficulty as independent variable. In the first two columns, it is shown which time pressure and which task difficulty is considered. The probability is found by filling in values for time pressure and task difficulty (for time pressure: low time pressure = 0, high time pressure = 1, for task difficulty: easy questions = 0, difficult questions = 1) into regression equation (6).

Table 7 shows the probability of choosing competition in all possible outcomes of the model. The exact way these numbers are found, are shown in Appendix B. Results show that the highest probability to choose competition is when the time pressure is high, and the questions are easy. On the other hand, the lowest probability is found when considering low time pressure and difficult questions. Both of these findings are not in line with the hypothesis, and therefore, the final hypothesis is rejected.

5.3 Changing competitive preferences

This final analysis focuses on the potential changing competitive preferences when an individual reaches the age of 50. The results of the OLS regression models conducted to investigate this change, are shown in Table 8. Column 1 of the table shows the results for the simple model, with just gender and whether a woman is over the age of 49 as independent variables, to determine if the predicted change in female competitive behaviour is true. In column 2, the results of the second OLS regression model are presented. This model has choice

of competition as dependent variables, and other variables showing whether a woman is over the age of 49, whether a man is under the age of 49 and whether a man is over the age of 49, as independent variables. These variables are used to determine if women know a bigger change in competitive behaviour than men, or vice versa.

Table 8 – OLS regression results

	(1)	**(2)**
Gender (female)	-0.239*** (0.074)	
Female over 49	-0.033 (0.213)	-0.033 (0.213)
Male under 49		0.294*** (0.073)
Male over 49		-0.158 (0.181)
Constant	0.773*** (0.052)	0.533*** (0.053)
R^2	0.0605	0.0987
Observations	162	162

Note: This table shows the results of an OLS regression with the choice of competition as dependent variable. In column 1, the independent variables are gender and whether a woman is over 49 years old. In column 2, the independent variables are whether a woman is over 49 years old, a man is under 49 years old and a man is over 49 years old. Robust standard errors are shown between brackets. Stars are used to indicate significance levels. (* p<0.10, ** p<0.05, *** p<0.01)

When looking at column 1 of Table 8, it is found that women over the age of 49 are less likely to choose competition than women under the age of 49. This would mean that when women pass the age of 50, they indeed become less competitive. However, the coefficient for women over the age of 49 is not statistically significant, meaning that the coefficient cannot be interpreted and there is not enough evidence to draw a fitting conclusion.

The results in column 2 of Table 8 present some more information regarding changing competitive preferences with age. First, it is found that men under the age of 49 are most likely to compete, as the coefficient for this variable has the highest positive impact on the probability to choose competition. Also, this result is statistically significant at a 1% level. Next, the results show that men over the age of 49 are less likely to choose competition than women over the age of 49. However, both of these coefficients are not statistically significant, so again, no conclusive answer can be given. Finally, considering only the men, it shows that men also

become less competitive when they are older. Although, considering the insignificant coefficients, this conclusion is inconclusive.

All in all, it appears as if women indeed become less competitive when they pass the age of 49 years old. This is the same for men, who also become less competitive when they are over the age of 49 years old. However, there are no significant results to confirm these findings from other literature. Next to this, it is found that when both over the age of 49, men are less likely to choose competition than women, although this cannot be statistically proven.

6. DISCUSSION

Understanding the topic of gender differences is becoming increasingly important every year. The goal of this research is to become more aware of possible gender differences in competitive behaviour. Considering both time pressure and task difficulty, this research sees how different circumstances change performance as well as competitive choices. The main results of this paper focused on gender differences in performance and competitive behaviour. Considering the difference in performance – which was investigated in the first hypothesis – it was found that there are no statistical differences between men and women. This means that their scores were the same, on average. This supports existing literature from Niederle and Vesterlund (2007), as they state that men and women perform equally well.

As for the choice of competition, some more aspects are investigated. First, gender differences in choice of competition are investigated, only addressing low and high time pressure. Results show that for low time pressure, men are more likely to choose a competing scheme than women, which supports the findings of Shurchkov (2012). However, it is inconclusive what the gender gap looks like under high time pressure. The gender gap is expected to close up with high time pressure (Shurchkov, 2012), and this does correspond to the findings of this study, only results are not significant. As the second hypothesis states that men are more likely to choose a competing scheme under time pressure, the hypothesis cannot be rejected, as there is not enough evidence to do so.

Second, the difference in choice of competition is again investigated, only this time considering the difference between easy and difficult questions. The third hypothesis states that men are more likely to choose a competing scheme than women, no matter the task difficulty. Results show that for both easy and difficult questions, men indeed choose the competing scheme more than women. This is in line with the third hypothesis, and therefore, this is not rejected. However, it is also found that the difference between men and women increases, when going from easy to difficult questions. This is not in line with literature from Hoyer et al. (2020) as they state that the gender gap changes in the opposite direction. A potential explanation for this is found in the discussed paper from Lewis et al. (1992). They observe that the competitive behaviour of men might change between easy and difficult tasks. Therefore, the difference in the gender gap can be traced back to male competitive behaviour.

The final hypothesis ties the previous hypotheses together, as it looks into the combined effect of both time pressure and task difficulty. Specifically, the hypothesis states that when questions

are easy and time pressure is low, the probability of choosing competition is highest. Also, when questions are difficult and time pressure is high, it is expected that the probability of choosing competition is lowest. Results show that this is both not the case. The highest probability is found for high time pressure and easy questions and the lowest probability is found for low time pressure and difficult questions. This is not in line with previously discussed literature, but it can be partly explained by using behavioural economic theory. Burson, Larrick, and Soll (2005) explain the concept of the *hard-easy* effect. This is a bias that makes people feel better than average for easy tasks, but makes them feel worse than average for hard tasks. Next to this, Yang, Thompson, and Bland (2012) present evidence that for easy tasks, time pressure increases self-confidence, whereas for difficult tasks, time pressure decreases self-confidence. Adding these theories together, this supports the findings of this study.

6.1 Limitations

As for limitations of this study, there are a few that should be mentioned. Choosing to examine only gender differences in performance under time pressure, has been one of them. The main focus of this paper lies with gender differences in competition, both under time pressure and with different task difficulties. To test the first hypothesis, no distinction was made between performance with easy and difficult questions. There might be a difference in the gender gap of performing under time pressure, when looking at those conditions separately.

After doing some more analysis, this can indeed be confirmed. Results from this analysis can be found in Appendix C. For easy questions, there are no statistical differences between men and women. However, for difficult questions, there are very significant differences. Also, there is a high probability that men have a higher score than women. These results would be more in line with existing literature, as that states that men are indeed better performers under time pressure than women. It is very probable that because the easy questions are more basic knowledge, there are no differences here, whereas when individuals have to think thoroughly, men can outperform women.

Finally, next to this methodological limitation, there is also a limitation in the gathered sample size. When looking at gender, there are 66 male and 96 female participants in the study. To gain a power – a representation of how true the effect is – that is sufficient to make conclusive conclusions about certain effects, both groups should contain about 100 participants. Therefore, the current sample size is too small to be able to measure the true effect.

7. CONCLUSION

The research question of this paper is: "*What are the gender differences in performance and competitive behaviour when performing under time pressure and with multiple task difficulties?*". Multiple hypotheses were tested to answer this question and to determine these gender differences. First, no significant gender differences were found when considering performance under time pressure. This means that when both given a task under time pressure, the performance will be equally good. Looking back at the headline regarding the domination of male managers, the stress that comes with the job, should not be the reason for this domination, as it was found that men perform equally good as women.

Second, gender differences in competitive behaviour were investigated. It was found that for low time pressure, men are more likely to choose competition than women, but for high time pressure these gender differences are inconclusive. This means that when there is a situation of low time pressure, men are more willing to enter competition, meaning they are more self-confident than women in this case. However, for high time pressure, no such conclusion can be drawn.

Next, again gender differences in competitive behaviour were investigated, but this time for both easy questions and difficult questions. It turned out that for both task difficulties, men are more likely to choose competition than women. Relating this back to actual tasks on for example the work floor, this means that for both easy and difficult tasks, men are more willing to compete with others.

Finally, when looking at both time pressure and task difficulty, it was found that the highest probability to choose competition is with high time pressure and easy questions and the lowest probability is with low time pressure and difficult questions. This can be explained by the *hard-easy* effect, and the fact that with easy tasks, individuals tend to be more self-confident under time pressure, whereas with difficult tasks, they tend to be less self-confident under time pressure.

Overall, this research shows that there are indeed multiple gender differences when considering performance and competitive behaviour. Understanding all these differences and relating them back to real surroundings can help form the best environment for every individual and potentially close the gender gap piece by piece.

7.1 Practical implications

Since there are no gender differences in overall performance, men and women perform equally well when doing a certain task. This helps shine new light on the domination of male managers. Since men and women perform equal under time pressure, the stress that this job entails, should not be the reason for this domination. Therefore, to address this issue and help women become more on top of the business world, other possible explanations should be searched to explain and solve this domination.

Furthermore, the apparent higher competitiveness of men is also something to take into account in the working environment. When considering both easy and difficult tasks, men think they can outperform others, because of their relatively high competitiveness. For easy tasks, this might not have the biggest consequences, as easy tasks are usually relatively unimportant compared to difficult tasks. However, being more competitive than others might have consequences when it considers difficult tasks. For example, getting a promotion when a difficult task is completed successfully. Women might shy away from this competition, not trying to get the promotion at all, whereas men feel more confident to compete and end up getting promoted. To close this gender gap, firms should be made aware of this difference, given them the possibility to handle, for example, this path to promotion in a different manner, resulting in fairer and more equal environments.

8. REFERENCES

Burson, K. A., Larrick, R. P., & Soll, J. B. (2005). Social Comparison and Confidence: When Thinking You're Better than Average Predicts Overconfidence. *Ross School of Business Paper*, (1016).

Buser, T., Dreber, A., & Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental Economics*, *20*(2), 506-530.

Cahlíková, J., Cingl, L., & Levely, I. (2020). How stress affects performance and competitiveness across gender. *Management Science*, *66*(8), 3295-3310.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of environmental psychology*, *24*(3), 329-340.

Cox, J. (2021b). *Why women are more burned out than men*. BBC Worklife. Retrieved from https://www.bbc.com/worklife/article/20210928-why-women-are-more-burned-out-than-men

Datta Gupta, N., Poulsen, A., & Villeval, M. C. (2005). Male and female competitive behavior-experimental evidence.

Delfgaauw, J., Dur, R., Sol, J., & Verbeke, W. (2013). Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics*, *31*(2), 305-326.

Ding, C. S., Song, K. I. M., & Richardson, L. I. (2006). Do mathematical gender differences continue? A longitudinal study of gender difference and excellence in mathematics performance in the US. *Educational Studies*, *40*(3), 279-295.

Flory, J. A., Gneezy, U., Leonard, K. L., & List, J. A. (2018). Gender, age, and competition: A disappearing gap?. *Journal of Economic Behavior & Organization*, *150*, 256-276.

Große, N. D., & Riener, G. (2010). *Explaining gender differences in competitiveness: Gender-task stereotypes* (No. 2010, 017). Jena Economic Research Papers.

Haynes, S. G., & Feinleib, M. (1980). Women, work and coronary heart disease: prospective findings from the Framingham heart study. *American Journal of Public Health*, *70*(2), 133-141.

Herrero, S. G., Saldaña, M. Á. M., Rodriguez, J. G., & Ritzel, D. O. (2012). Influence of task demands on occupational stress: gender differences. *Journal of Safety Research*, *43*(5-6), 365-374.

Hoyer, B., van Huizen, T., Keijzer, L., Rezaei, S., Rosenkranz, S., & Westbrock, B. (2020). Gender, competitiveness, and task difficulty: Evidence from the field. *Labour Economics*, *64*, 101815.

Jericho, G. (2017). *Men still dominate managerial positions. The excuses are running out*. The Guardian. Retrieved from https://www.theguardian.com/business/grogonomics/2017/aug/29/men-still-dominate-managerial-positions-the-excuses-are-running-out

Lewis, M., Alessandri, S. M., & Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child development*, *63*(3), 630-638.

Markowsky, E., & Beblo, M. (2022). When do we observe a gender gap in competition entry? A meta-analysis of the experimental literature. *Journal of Economic Behavior & Organization*, *198*, 139-163.

Mattingly, M. J., & Sayer, L. C. (2006). Under pressure: Gender differences in the relationship between free time and feeling rushed. *Journal of marriage and family*, *68*(1), 205-221.

Matud, M. P. (2004). Gender differences in stress and coping styles. *Personality and individual differences*, *37*(7), 1401-1415.

Maume, D. J. (2006). Gender differences in restricting work efforts because of family responsibilities. *Journal of marriage and Family*, *68*(4), 859-869.

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.

Ministerie van Algemene Zaken. (2022). *Wat voor werk mag ik doen als ik 13 of 14 jaar ben?* Rijksoverheid.nl. Retrieved from https://www.rijksoverheid.nl/onderwerpen/bijbaanvakantiewerk-en-stage-door-jongeren/vraag-en-antwoord/wat-voor-werk-13-of-14jaar#:%7E:text=Als%20je%2013%20of%2014%20jaar%20bent%2C%20mag%20je20helpen,lichte%20werkzaamheden%20in%20de%20landbouw.

Nelson, D. L., & Quick, J. C. (1985). Professional women: Are distress and disease inevitable?. *Academy of Management Review*, *10*(2), 206-218.

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*(1), 601-630.

Petersen, J. (2018). Gender difference in verbal performance: A meta-analysis of United States state performance assessments. *Educational Psychology Review*, *30*(4), 1269-1281.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189-1213.

Slade, L. A., & Rush, M. C. (1991). Achievement motivation and the dynamics of task difficulty choices. *Journal of Personality and Social Psychology*, *60*(1), 165.

Timmins, B. B. (2021a). *Pay gap between men and women fails to improve*. BBC News. Retrieved from https://www.bbc.com/news/business-58786739

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: a meta-analysis. *Psychonomic bulletin & review*, *18*(2), 267-277.

Wikipedia. (2022). *Mental rotation*. Retrieved from https://en.wikipedia.org/wiki/Mental_rotation

Woods, N. F., Mitchell, E. S., & Smith-DiJulio, K. (2009). Cortisol levels during the menopausal transition and early postmenopause: observations from the Seattle Midlife Women's Health Study. *Menopause (New York, NY)*, *16*(4), 708.

Wozniak, D., Harbaugh, W. T., & Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, *32*(1), 161-198.

Yang, H., Thompson, C., & Bland, M. (2012). The effect of clinical experience, judgment task difficulty and time pressure on nurses' confidence calibration in a high fidelity clinical simulation. *BMC medical informatics and decision making*, *12*(1), 1-9.

9. APPENDIX

9.1 Appendix A

Survey text **(math questions will be added at the end of Appendix A)**

- **Introduction:**

"Welcome to this survey, and thank you in advance for filling it in. This survey is a part of my master's thesis in Behavioral Economics. If you are interested to know what I am investigating for my thesis, or if you have any other questions, you can always contact me via 507499kk@student.eur.nl.

In this survey, you will take part in an experiment that consists of three rounds. In every round, you will be asked to answer 25 math questions within a specific time. Please read the instructions of each round very carefully.

PAY ATTENTION – It is very important to get as many points as you can. The number of points you receive for one correct answer differs between rounds. For one randomly chosen participant of this survey, the received points in one random round will be paid out in euros. For a chance to win, you must leave your email address at the end of the survey.

After answering all math questions, you will be asked some demographic questions (age, gender, etc.). Any personal information that is gathered, will be handled safely and deleted after the research is done. By continuing, you consent to participating in the research study as shown above."

- **Explanation round 1:**

Low time pressure / high time pressure: (time display will differ)

"**Welcome to round 1.**

In this round you will be asked to answer 25 math questions, within 2 minutes / within 1 minute.

All questions will be shown on one page, you just have to click the box the answer them yourself.

When you click on the arrow to go to the next page, round 1 will start. **Good luck!**"

- **Explanation round 2:**

Low time pressure / high time pressure: (time display will differ)

"**Welcome to round 2.**

In this round you will be asked to answer 25 math questions, in a competition against one of the other participants of this survey. You have 2 minutes / 1 minute to answer the questions.

If you give more correct answers than the other participant, you will receive 2 points per correct answer. If you do not give more correct answers, you will receive 0 points. In case of a tie (gelijkspel), the winner will be determined by a coin flip.

Your points will be compared after filling in the survey, so you will not know how many correct answers the other participant has given. When you go to the next page, your 2 minutes / 1 minute will start. **Good luck!**"

- **Explanation round 3:**

Low time pressure / high time pressure: (time display will differ)

"**Welcome to round 3 - the final round.**

In this round you have a choice to partake in either round 1, or round 2 again.

When you choose round 1, you again have to answer 25 math questions. You will receive 1 point per correct answer.

When you choose round 2, you will also answer 25 math questions, again competing to the number of correct answers from the (same) other participant. Your correct number of answers from this round will be compared to their correct number of answers from round 2. You will receive 2 points per correct answer if you have more correct answers, but you will receive 0 points, if you have less correct answers. In case of a tie (gelijkspel), the winner will be determined by a coin flip.

In both cases you will have 2 minutes / 1 minute to answer the questions.

Below, you can select the round in which you want to participate again.

I want to participate again in:

- Round 1 – no competition, 1 point per correct answer
- Round 2 – competition, 2 points per correct answer if you win, 0 otherwise"

If round 1:

"You have chosen to participate again in round 1.

Once you click the arrow to go to the next page, your 2 minutes / 1 minute will start. **Good luck!**"

If round 2:

"You have chosen to participate again in round 2.

If you give more correct answers than the other participant (from round 2), you will receive 2 points per correct answer. If you give less correct answers, you receive 0 points. In case of a tie (gelijkspel), the winner will be determined by a coin flip.

Your points will be compared after filling in the survey, so you will not know how many correct answers the other participant has.

Once you click the arrow to go to the next page, your 2 minutes / 1 minute will start. **Good luck!**"

- **Math questions**:

Table 9 – Math questions

EASY QUESTIONS			**DIFFICULT QUESTIONS**		
Round 1	*Round 2*	*Round 3*	*Round 1*	*Round 2*	*Round 3*
3 x 4 = 12	15 : 5 = 3	6 x 1 = 6	12 x 8 = 96	90 : 15 = 6	11 x 5 = 55
14 : 2 = 7	8 x 8 = 64	8 x 2 = 16	56 : 4 = 14	5 x 14 = 70	18 x 3 = 54
6 : 2 = 3	30 : 5 = 6	30 : 6 = 5	13 x 5 = 65	9 x 11 = 99	68 : 4 = 17
36 : 6 = 6	42 : 7 = 6	10 x 10 = 100	9 x 19 = 171	64 : 4 = 16	14 x 8 = 112
5 x 1 = 5	3 x 5 = 15	4 : 2 = 2	84 : 12 = 7	4 x 19 = 76	96 : 8 = 12
8 x 10 = 80	4 x 2 = 8	72 : 8 = 9	14 x 7 = 98	20 x 9 = 180	180 : 9 = 20

28 : 7 = 4	45 : 5 = 9	24 : 3 = 8	91 : 7 = 13	48 : 12 = 4	7 x 12 = 84
4 x 8 = 32	9 x 6 = 54	9 x 3 = 27	126 : 9 = 14	54 : 3 = 18	13 x 4 = 52
6 x 6 = 36	28 : 7 = 4	5 x 8 = 40	2 x 13 = 26	120 : 20 = 6	38 : 19 = 2
10 x 2 = 20	2 x 10 = 20	6 x 2 = 12	60 : 15 = 4	11 x 8 = 88	6 x 11 = 66
12 : 3 = 4	8 x 1 = 8	7 x 5 = 35	7 x 18 = 126	17 x 4 = 68	84 : 6 = 14
16 : 4 = 4	63 : 7 = 9	9 : 1 = 9	12 x 6 = 72	108 : 18 = 6	8 x 15 = 120
7 x 3 = 21	3 x 3 = 9	6 x 10 = 60	99 : 11 = 9	3 x 20 = 60	133 : 19 = 7
2 x 7 = 14	3 x 6 = 18	8 x 7 = 56	11 x 11 = 121	114 : 6 = 19	72 : 4 = 18
40 : 10 = 4	8 : 2 = 4	7 x 4 = 28	39 : 13 = 3	6 x 12 = 72	7 x 14 = 98
21 : 3 = 7	35 : 7 = 5	2 x 1 = 2	9 x 13 = 117	14 x 3 = 42	10 x 16 = 160
48 : 6 = 8	50 : 5 = 10	30 : 3 = 10	18 x 6 = 108	135 : 9 = 15	108 : 12 = 9
4 x 7 = 28	4 x 9 = 36	18 : 9 = 2	11 x 7 = 77	5 x 15 = 75	2 x 19 = 38
9 x 5 = 45	12 : 2 = 6	9 x 4 = 36	48 : 3 = 16	80 : 16 = 5	51 : 3 = 17
81 : 9 = 9	2 : 2 = 1	63 : 9 = 7	34 : 17 = 2	152 : 8 = 19	112 : 7 = 16
2 x 3 = 6	7 x 9 = 63	27 : 9 = 3	5 x 12 = 60	6 x 13 = 78	7 x 15 = 105
16 : 2 = 8	10 x 7 = 70	4 x 5 = 20	119 : 7 = 17	52 : 13 = 4	5 x 13 = 65
5 x 5 = 25	49 : 7 = 7	7 x 7 = 49	78 : 6 = 13	7 x 20 = 140	14 x 2 = 28
7 x 6 = 42	5 x 4 = 20	7 : 7 = 1	13 x 7 = 91	44 : 4 = 11	144 : 18 = 8
27 : 3 = 9	10 : 2 = 5	72 : 9 = 8	42 : 14 = 3	45 : 3 = 15	18 x 5 = 90

Note: This table contains all math questions that were asked in the online survey experiment. The first three columns show the easy math questions, per round in which they were asked. The last three columns show the difficult questions, per round in which they were asked. For both difficulties, round 1 is the piece-rate task, round 2 is the competition task, and in round 3 participants are given a choice between the two.

9.2 Appendix B

The computation of the probabilities to choose competition, using regression equation (8).

Y = 0.733 + 0.023 * Time pressure – 0.233 * Difficulty – 0.002 * Time pressure * Difficulty

Possible values of time pressure:	Possible values of difficulty:
0 = low time pressure	0 = easy questions
1 = high time pressure	1 = difficult questions

Low time pressure, easy questions:

Y = 0.733 + 0.023 * Time pressure – 0.233 * Difficulty – 0.002 * Time pressure * Difficulty

Y = 0.733 + 0.023 * 0 – 0.233 * 0 – 0.002 * 0 * 0 = 0.733

High time pressure, easy questions:

Y = 0.733 + 0.023 * Time pressure – 0.233 * Difficulty – 0.002 * Time pressure * Difficulty

Y = 0.733 + 0.023 * 1 – 0.233 * 0 – 0.002 * 1 * 0 = 0.756

Low time pressure, difficult questions:

Y = 0.733 + 0.023 * Time pressure – 0.233 * Difficulty – 0.002 * Time pressure * Difficulty

Y = 0.733 + 0.023 * 0 – 0.233 * 1 – 0.002 * 0 * 1 = 0.500

High time pressure, difficult questions:

Y = 0.733 + 0.023 * Time pressure – 0.233 * Difficulty – 0.002 * Time pressure * Difficulty

Y = 0.733 + 0.023 * 1 – 0.233 * 1 – 0.002 * 1 * 1 = 0.503

9.3 Appendix C

Investigating possible gender differences in performance, considering easy questions and difficult questions separately.

To determine this possible difference, two Mann-Whitney U test will be performed, to see if there is a difference in the number of correct answers for men and women. The hypotheses for these tests are as follows:

H_0: The mean number of correct answers of men and women are the same.
H_1: The mean number of correct answers of men and women are not the same.

Next, the found p-values will be presented, together with a short conclusion about the test.

For easy questions:

P-value = 0.7463.

This means that the null hypothesis cannot be rejected. The mean number of correct answers appears not to be different for men and women, with easy questions.

For difficult questions:

P-value = 0.0175.

This means that the null hypothesis should be rejected. The mean number of correct answers appears to be different for men and women, with difficult questions.

The probability that the number of correct answers is higher for men than for women, is 0.718.