

**UNIVERSITAT  
JAUME·I**

# **GENDER DIFFERENCES IN THE IMPROVEMENT OF ACADEMIC PERFORMANCE BY MEANS OF ECONOMIC INCENTIVES**

**Author:** Aarón Alejandro Mas Rivero

al346030@uji.es

**Tutor:** Gerardo Vicente Sabater Grande

**Degree in Economics**

**Academic course 2019/2020**

# INDEX

1. ABSTRACT.....	1
2. INTRODUCTION .....	1
3. LITERATURE REVIEW .....	2
4. EXPERIMENTAL DESIGN .....	8
5. DATA ANALYSIS.....	12
6. ANALYSIS OF THE SAMPLE.....	17
7. REFERENCES .....	22
8. ANNEXES.....	24

## 1. ABSTRACT

The aim of this paper is to determine whether or not there is a difference in how economic incentives affect academic performance. Also, to focus on checking if there is a possible significant difference in terms of gender, that is, if there is a real difference between men and women as a result of the way in which they are economically encouraged in the academic field. Statistics are used in order to confirm if the result is focused on a goal, in this case, the final result of the subject, or if it is the result of a rank tournament where it is also focused on analysing competitiveness between students and between genders. To do that, diverse statistics are used to examine the significance of those results and also with support of a voluntary and non-incentive control group so as to observe and compare the results with a neutral sample.

## 2. INTRODUCTION

The vast majority of students have lack of motivation towards studying. They do not think about how the future would be when they finish their studies, they just focus on the present. Ethics, analysis and explanation will be involved with results in order to know how economic incentives affect students. Diverse elements will be taken into account to discover what encourages them to change their attitude to try their best and to accomplish academic success. Taking into account the previous analyses of other specialists in the matter, we will be able to determine how it will affect the economic influence, both psychological and through the results of the tests, on one way, we have a previous analysis of the students, a series of samples that allow us to observe how these incentives have come to affect the students, on the other way, we must estimate, test, analyse and contrast these effects in order to obtain results that give us significance, in order to be able to say and demonstrate the differences that exist or not in the sample that we will analyse below. Results will be developed to know which is the best way to encourage them. This way to motivate students will lead them on the right track in their present as well as in their future. The way it affects them will be determined, based on a sample of two incentive treatments and a control group which will be useful to actually measure the results. This allows us to recognise the polemic case of the differentiation of the results and the actions carried out to different gender people, which has always been a subject to take into account due to the controversy that usually generates this issue.

### 3. LITERATURE REVIEW

In this study some key criteria had been followed. One of them follows the iconic and confronted investigation line of giving an economic incentive to the students so as to improve their grades all course long as in Herranz-Zarzoso, N., & Sabater-Grande, G. (2018). They obtained results about diverse motivation methods for students, such as effectiveness of economic incentives depending on their absolute and relative academic performances by using a piece-rate and a rank-order payment mechanisms and classifying students in different categories. As their study was read, it was thought to be an inspiring source to this study due to its detailed explanation.

Other studies to be taken into consideration, even if they are ethical or not, are the ones that give economic incentives to students to promote the boost of their abilities and to ameliorate their academic achievements. Authors, institutions and even nations call into question that giving students economic support encourages students to increase their academic performance. However, these days some grants are offered to them, such as the ones of the Universitat Jaume I, the Department for Education of the Valencian Community or the Spanish Ministry of Education. But to get one of those scholarships, the student has to get good grades and to have certain household incomes. These institutions can pay them only for their education or can give them some money too depending on their situation. So, these study grants affect their behaviour and force them to get a minimum grade to pass the cut-off mark required in order to receive financial support.

Incentives in both lines of investigation can be seen as similar even if their methodologies are slightly different. This field of knowledge is very controversial in academic and scientific areas because some people think that is unethical to encourage them with money to improve their grades because they believe that external factors are the ones that encourage them to give the best of themselves. But other researchers such as professor Jere R. Behrman have achieved to improve academic performance on certain groups of students by means of those motivation techniques with help of relevant institutions for the purpose of proving a theory based on monetary effectiveness and human development.

It is demonstrated that an impetuous stimulus stirs up a purpose on the students, such as to motivate their efficiency by monetising it. This stimulus gets them more interested in their subjects and changes the way of seeing them. Teaching methodologies have always been called into question. The ones based on past study practices as in Spain and even the newest ones of Nordic countries such as Finland, that is based on a demanding education to teachers, as it is explained in Enkvist, I. (2010).

Several times has been discussed what to do in order to improve education and to increase the efficiency of the students. Some lines of research defend that good quality facilities and better teachers assure a better performance as well as better study techniques. Nieto, D. A. (2005) explains how diverse and innovative measures related to reading comprehension and the influence of a metacomprehension (closely related to brain waves and the better study development) allow to affect positively in the students' educative and cognitive development and help them to develop facets that they could not obtain with conventional education. The purpose of these measures is to put an end to indolence and lack of interest by the students by applying innovative study techniques or developing new aspects in education that by means of experiments, demonstrate the possibilities of these original measures.

Besides, the development of this experiment casts serious doubts on whether to consider it ethical or not. Researchers and authors such as Peter Fredriksson (1997) or Richard J. Murnane (1993) have made some research to prove that it is unethical.

As it is detailed in Herranz-Zarzoso, N., & Sabater-Grande, G. (2018), "From the pioneering work of Latham and Locke (1979) to the present, there has been an extensive body of empirical literature testing goal-setting theory in controlled environments. However, when it comes to self-chosen goals, the number of empirical contributions is limited and not all of them use incentive compatibility." These slight details make us contemplate the ethics of those incentives and reminds us that there is not that much research about this matter. That suggests that the lack of information might be cause and effect, because researchers are not keen on giving incentives.

To talk about how academic incentives have influence in progress of academic performance is necessary to talk about motivation. Keller (1983) provides a definition: "Motivation refers to the choices people make as to what experiences or goals they will approach or avoid, and the degree of effort they will exert in that respect."

It is possible to motivate the students economically in different ways, but we can put them into two groups: intrinsic motivation or extrinsic motivation. Elements of distinction as personality or in this case the area of study we are focusing on, academic results, can determine the dominant type of motivation. Regardless of the type of motivation, the importance of it can be reinforced in Clark et al. (2016) and Goerg & Kube (2012).

Extrinsic motivation can be defined in this context as a motivation arisen from outside the individual that drives a person (a student in this case) to make an activity or work due to an external factor influence, such as money. In this type of motivation what creates an effect on the subject is external or comes from outside, even if it is caused by a good or

bad stimulus, a reward or a punishment. These extrinsic motivation factors can be determined by the type of reward obtained by the execution of an activity. Examples of this are the salary obtained after workdays or an academic incentive if certain grades are obtained when passing a subject, in a matter related to the scope of economic incentives, another example of extrinsic motivation, which leads directly to more psychological factor, already entering scientific matter, caused by our own body thanks to an external factor, is the encouragement of crowds in sports, such as soccer, rugby or football. The most important point is, the support of the crowd generates in people a continuous generation of hormones, which include, for example, dopamine, the well-known hormone of happiness, or adrenaline, which provides that extra strength and determination to people, which causes them to be in a state of continuous emphasis, a greater predisposition to physical and mental effort, all this caused by the moral support of other people; The point I want to get to is, extrinsic motivation is not just a physical factor, like in the previous case commented, as the money was in the last example, if it is not a type of motivation that is born from a spectrum external to ourselves, we can extrapolate a lot of examples and multitude of types of extrinsic motivation at all.

On the other hand, intrinsic motivation is the one that involves personal rewards and comes from within. That is to do an activity for its own sake. The motivation to do an activity arises from the inside and the performance of that action or activity because you enjoy doing it and because of the personal satisfaction you get it is like fulfilling a wish. So, if a person does not succeed it does not see it as failure. It is not a disaster because the psychological need of performing that activity has been achieved and if it is done wrong, it is possible to learn from its mistakes. Some examples are the happiness or fulfilment received when passing an exam or getting our dream job. That means motivation accomplished by putting material rewards aside and just feeling fulfilled and valued. Those distinctions between intrinsic and extrinsic motivation can also be seen and exemplified in Iriarte (2007).

Analysis of motivation related to academic incentives and self-managed achievements can be seen in Clark et al. (2016) and Georg & Kube (2012).

Clark et al. (2016) with great samples of university students made two experiments. "One experiment asked treated students to set goals for performance in the course; the other asked treated students to set goals for a particular task (completing online practice exams)". It was found that performance-based goals had not meaningful impact on students' performance and task-based goals had exceptional positive effects on course performance. Extrinsic motivation in those experiments was inconclusive and not effective

because students participating in the experiment were not paid even if they accomplished goals.

Money incentives relevance was analysed in Georg & Kube (2012). In their experiment people were hired to rearrange a library. Their contract was incentive because it combined self-chosen goals (personal work goals) and a piece-rate payment. It was found that personal goals had a positive effect even if monetary incentives fall back. That involves a bigger intrinsic motivation because workers were more motivated if they chose their goals than if they reached pre-specified goals.

Other studies to consider are the ones that have not taken into account monetary incentives. Falk & Knell (2004) compare social models in whose people choose their own references in respect of the goals they want to reach, their target or what they want to improve. They seek their ideal archetype to be and those reference standards predict that people tend to compare to similar others so as to increase their own abilities.

Hirshleifer (2015) is a big source of inspiration to this current study because in his experiment students' capital build-up is tested. By means of economic input, it is observed how incentives influence students and put them in critical situations. Money tends to affect directly in students' effort and is intricately connected to productivity and the results of the experiment. Those results were checked, as we could also check, by student's performance on tests to evaluate their effect. In this case, it is noted how economic incentive is more effective and encourages students to get better marks than students that do not receive incentives.

The exploratory analysis of this experiment bestows us a key criterion that is attention. It was noticed that students that were given that incentive paid more attention and that could be the cause of their better performance. That reward could not be a direct reason, but it could allow the development of other attributes, for instance, attention, whose raise is proven. Economic input could be the perfect complement with the aim to incentivize other features that allow to enhance academic performance.

In several fields of economy and politics resort to economic incentives so as to boost inversions in labour force and effort instead of work. Examples of this enforcement can be seen in Gneezy, Meier & Rey-Biel (2011) and in Lemieux, Macleod & Parent (2009). This research allows to contend that a greater effort supplemented by economic incentives leads to a performance upgrade. Frequent rewards, even if they are not substantial, allow to enhance efficiency. Therefore, extrinsic motivation abovementioned has an effect on the output.

Study sources related to gender differentiation have also taken into account.

Mehreen, Salman & Zafar (2015) attempted to determine the impact of monetary incentives in students by having gender and its performance in mind at the university level. In their study it is made an analysis of samples from three groups that attend to first year of maths and from other three groups that take their fourth year of accounting. It is detected that student's performance is affected by grade incentives for all six groups of both courses. Gender has a significative importance in math student's performance but lacks of it in accounting students and similar results were drown by Severiens & Ten Dam (1994) that found out that men are good at profound information processing and women are good at methodical study style. These results rely on the type of subject that is thought because there are courses in which profound and phycological mind aspects are used, for instance, in logarithms and mathematics or other more focused on logic, as accountancy. Interaction between gender incentives and academic performance was not found to be clear, so it is not possible to confirm that gender understandably affects academic background.

Furthermore, the study conducted by Zaidi et al (2006) was conducted to analyse the influence of economic incentives and individual performance reliant on the gender (male, female) of the students at the graduate level. A two-way analysis of variance was conducted. It was performed on a sample of three groups that were taking their first-year of core mathematics course and another three groups that were taking a fourth-year compulsory accounting course.

In this case, economic incentives affect significantly students' performance. Gender associated data show that affects significantly the performance of maths students but not the performance of accounting students. Studies on this matter such as Schmeck et al. (1977) support that significant gender differences were not detected even if that experiment was not as incidental as the Zaidi one.

Interaction on how monetary incentives and gender affects academic performance has no remarkable impact in none of those experiments. Hence, to this current study draws from the premise that the aforementioned factor will be irrelevant.

Additionally, the first part of the study conducted by Iturra et al. (2012) goes beyond. They establish minimum differences between genders, but not only the ones related with the academic field, but also related with other elements, like the social sphere. On their study those elements interact as a triangle, having three starting points and taking into account three interactions: social sphere, gender and academic performance. Only two of them are not correlated (like the last two abovementioned), but when the social factor appears, the approval of the others creates interactions related to academic scope.



In their study those results are interpreted as a consequence of influence of gender roles on expectation and differential adaptation of men and women. It was expected that men were competitive so abilities were not useful to improve their performance but helped in upgrading their interpersonal relationships. Efficiency on women is linked by the quality of interpersonal relations, so their orientation towards cooperation would strengthen their relationships and contribute to enhance academic performance.

Outcomes of those experiments can be isolated and compared with other studies with a similar connection between genders such as Severiens & Ten Dam (1994). Said results, particularly display that male students are better at deep processing of information and female students are better at methodical studies. Those slight differences create certain doubts. Gender dependent differences among students exist? or gender has no clear influence?

Furthermore, as Jacob (2002) states, most of the differences between genders appear to be determined by non-cognitive aspects, that is to say that are related on how those subjects interact with the others. Therefore, it cannot be taken as an alternative option because personal relationships between subjects are difficult to measure properly without falling into assumptions and possibly, into inappropriate methods.

This study is closely associated with various experimental fields of economy and its literature is based in ethics and consequences of financial incentives on academic performance. Economists that keep more defined to experimental areas are convinced that greater incentives lead to bigger effort and efficiency on students' performance. However, psychologists confirm that incentives focused on improving performance on mathematic, algorithmic or repetitive tasks are less effective, or counter-productive on tasks that require more creativity, concentration or intuition.

Learning usually has been classified as a method in which extrinsic motivation predominates. Is seen as something that will contribute to a future profit that by making efforts at the present, will generate external benefits: a better salary, a nice job or a relaxed life. As some authors stand by, to ameliorate academic performance, economic incentives would be favourable but not entering in fact ethics that in the presence of an extrinsic motivation finding a result find effective those motivators.

In this literature review it was seen how testers see a matter. In a way, some focus on practicality to society of a controlled motivation on students. And others find it unrealistic and frown upon it because creating motivation through economic matters could modify the educative base that stand up for every day.

## 4. EXPERIMENTAL DESIGN

### THE AREA OF DESIGN

All the data and evidence about how experimental design will be are based on various studies and are distant from giving us a clear view of the results obtained from this experiment. A wide number of variants and results were explained and taken into account as the base when analysing the sample, so as to be able to explain in a realistic and compelling way, with the intention of finding whether or not determining relationships exist. And all of that, to know if affirm or reject the hypothesis that will be provided later.

At first, as the title suggests, the experimental design is based on how economic incentives affect students' academic performance. The contribution of data is based on the experiment conducted by Herranz-Zaroso, N., & Sabater-Grande, G. (2018), that was focused on the students who studied in the 2017/2018 academic year the subject of Introduction to Microeconomics at Universitat Jaume I, where they were offered the possibility of taking part in an incentive program related to the performance they had with that specific subject. Those students were invited to participate in the middle of the course, where they were informed that the ones who accepted the invitation would be randomly assigned to a treatment group. The treatment groups were:

- Treatment group T0: the participants would not receive any monetary incentive for their participation in the experiment, we only take his final mark to analyse the others treatment groups. This is the control group, which will be used as the basis for determining the significance of the independent samples.
- Treatment group T1: the participants would receive a payment according to their academic performance, that is, the final grade of the course. This is a way of encouraging the student as it has been done in other studies such as in Gneezy, Meier & Rey-Biel (2011) or in Lemieux, Macleod & Parent (2009). While these studies were focused on the work environment, in our case we will observe how economic incentives affect when they are directed to specific goals, more precisely, passing the course, and the higher grade you get, the more economic reward you will get.
- Treatment group T2: the participants would receive an economic incentive by a rank-order tournament, with a payment mechanism according to their performance in order to reward students. This method is focused on encouraging competitiveness between students. This is a variant to be taken into account, since in several studies it has been observed how gender difference can be focused in

the psychological field, where women concentrate their efforts as a result of rational feelings, while men are certainly more competitive. Therefore, this treatment can help to clarify if the results obtained by other experimenters are true and are also observed in our case.

With that stated, students were informed about the group they were assigned and the conditions of their corresponding group, so that they could choose a grade objective for the final exam. Besides, they were told that as they were volunteers, in other words, they have not been forced to participate in the experiment. From the outset, this experimental design rules out possible effects related to the willingness to participate, since they all know which group they belong to and the conditions that this represents. This variant is a key point that must be taken into account in the whole experiment and specially in the results, since it is a necessary condition to obtain more significant results and to have more control in them.

The starting point is that each of the students makes a bet or an approximation according to the personal objective they want to achieve by the end of the academic year. In this way, different types of students can be identified and divided into different categories, depending on their ambition (from more to less) or their self-control degree. Although we do not take it as a variable, it can be useful to establish a psychological profile. To carry out the experiment, Herranz-Zarzoso & Sabater-Grande use a similar strategy to that of Jackson (2010), in which those students who get the grade they set or have a degree of approximation between the bet they made in relation to the grade they obtained, would be rewarded as a payment mechanism.

In addition to this, a grade obtained in a test taken during that course will be given to students at mid-term. This allows them to get an approximation of how they are doing in the course, as well as to compare the effort they made to the grade they have obtained. This is a factor to take into consideration, since those who have had an excess of confidence or ambition, can see how those expectations have been reduced, and have the possibility of betting downwards to go for the monetary bonus.

This strategy has to do with a key factor. Knowing the difference that exists at that moment between students with and without monetary incentives, to find if, at first, it has had a positive effect on them or not.

In terms of the sample, we will principally analyse the effect of economic incentives on students, both in terms of the final grade of the course and whether the effect of belonging to one treatment group or another has any effect on gender, all through the analysis of the sample.

The variables that we are going to emphasize the most are:

- Gender: When the student is male, it will take the value 1 and when the student is female, it will take the value 2. This variable can only have two options.
- Treatment group: When the student has been included in the treatment group 0 (T0), the control group, it will take the value 3. When the student has been included in the treatment group 1 (T1), the one focused on results according to their final mark in the subject, it will take the value 1. And when the student has been included in the treatment group 2 (T2), the rank tournament group, it will take the value 2.
- Exam result: It will take a value between 0 and 10 in regards to the grade obtained by the student at the end of the course. This is the grade of the final exam.

As a result of this sample of students in which the abovementioned variables are included, the effect of incentives will be seen. Firstly, by the effect that each treatment has had on the final average mark, considering all the students' profiles divided into different samples. That is to see if regardless of gender there is any evidence that differentiates one treatment group from another. Secondly, the effect on students, taking into account how has affected each treatment group and each gender to emphasize whether or not there is an actual difference between genders. All of that with a general analysis of the sample, so as to see if there is or not a clear difference on how economic incentives affect each gender of students.

The preliminary hypothesis about the results of the experiment is that it will be possible to observe how the effect on treatment groups is significant regards to the final average grade of students, so it affects positively their final grades. Besides, regarding to the gender section, it will be seen that there are no significant differences in the final average mark to draw conclusions, that is, that final average grade of both genders will not be markedly different.

#### First hypothesis

H0 = There is no significant difference between the treatment groups and the final average grade of the students, so the financial incentive does not have an effect on the final mark, the academic performance.

H1= There is a clear difference between the treatment groups and the final average grade of the student, that is, financial incentives have a positive impact on the final mark, the academic performance.

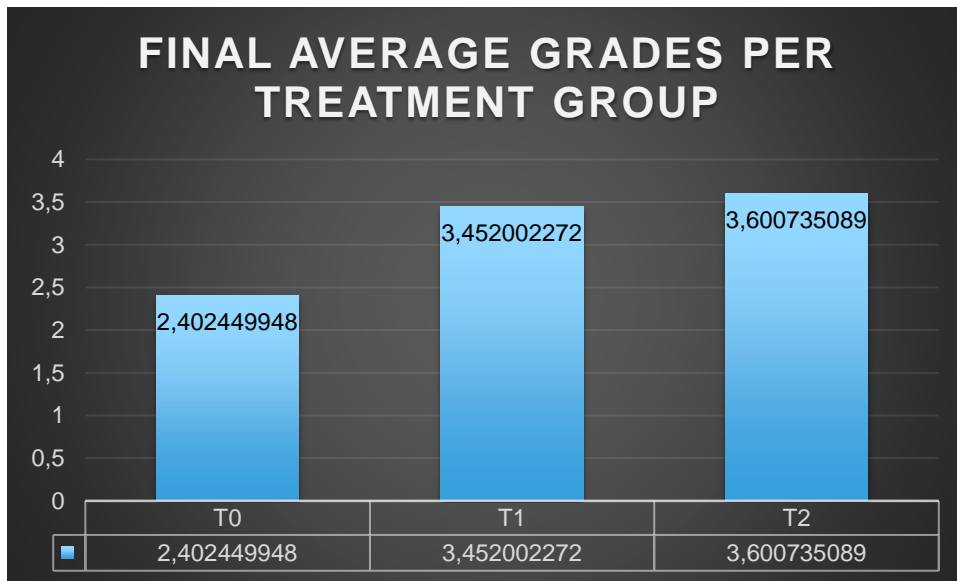
Second hypothesis.

H0 = There is no significant difference between the gender of the student and the final average mark. Therefore, there is no correlation between the economic incentive and different gender influence on the final mark, the academic performance.

H1= There is a substantial difference between the student's gender and the final average grade, hence, there is a relationship between economic incentive and its different gender influence on the final mark, the academic performance.

## 5. DATA ANALYSIS

An analysis of how different treatment groups and the gender of the students have affected their final average grades will be carried out.



*Figure 1 Final average grades per treatment group.*

This first graphic is focused on how the experiment has affected the students' final average grades, regardless of gender, centring attention on the treatment group, men and women T0, men and women T1 and men and women T2.

As it can be seen in figure 1, on the first place and with a final average grade of 3.6, we have the students who have been economically motivated through a rank-order tournament. These are the ones who have been tested in the T2 incentive group so as to encourage them to study. This can be explained by the human reason to competitiveness, in this case, there is more than a psychological profile of wanting to get the highest grade, there is the incentive to want to get a higher grade than your classmates, which are seen as academic rivals.

Then, students which belong to the T1, and therefore have been given financial incentives according the final grade obtained, have scored a final average grade of 3.45 that is quite similar to the one obtained by the treatment group 2.

Lastly, we have the students of the control group or T0, who have not been given any incentive. These ones have a lower average grade of 2.4 that is below the grades of the other two groups, with nearly an average of one point less in relation to the groups T1 and T2.

As a result, it will also be taken into account the pass/fail ratio for each group so as to continue with the analysis of the sample.

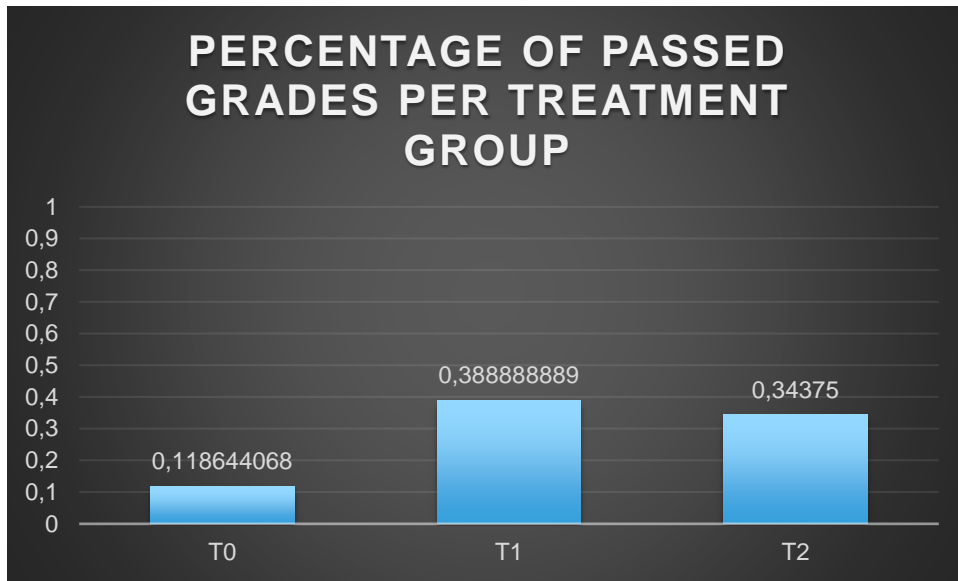


Figure 2 Percentage of passed grades per treatment group.

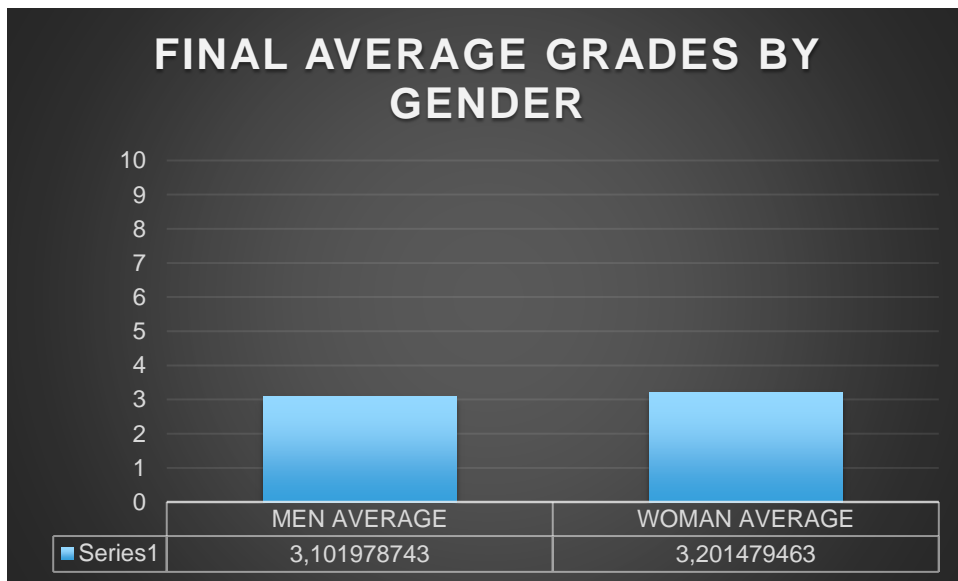
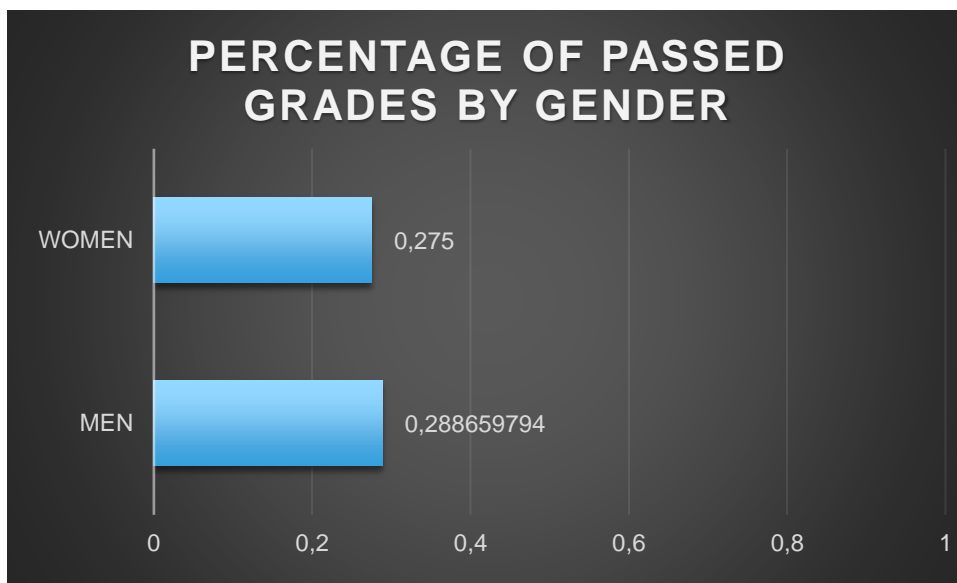


Figure 3 Final average grades by gender.

As it is seen in the figure 2, there is a large difference between the percentage of passed grades in the treatment groups 1 and 2 compared to the control group. This can be clearly explained by the emotional and rewarding factor that economic incentives have on the students. The motivation that each student has can be divided in two types: intrinsic motivation if it is focused on the goals and self-improvement or extrinsic motivation if those economic incentives are the ones that generate the need to pass the subject.

Looking more closely at gender, as it can be seen in figure 3, the average passed mark is similar for both genders, at around 3.15 average points for them. It is appreciated that females obtained 0.1 better results than males, but this is not seen as a very remarkable difference. Later on, attention will be drawn to statistical tests that allow to state if there is a significant difference in gender or not.

Finally, the passed rate has a slight contrast with the average grades in whose female gender was slightly above than the male, whereas in the passing rates, male has 1% ahead than female, as it can be seen in figure 4 reference to the passed grades rate by gender.



*Figure 4 Percentage of passed grades by gender.*

It can be seen in the figure 4 with regards to the pass/fail ratio that 29% of men have been able to pass the subject compared to 28% of women. As it happened previously, these results do not give a clear indication to be able to draw some conclusions.

This could give us some clues about the relationship between these two samples and if there really are differences between how economic incentives affect students' performance in connection with male and female genders.

In addition, attention will be centred on the analysis of the sample by making a deeper contrast between the different groups, taking into account each gender and treatment group, as we can see on the figure 5.

In this case, in a more general vision, differences in the average grades can be seen. Women in the treatment group 1 have a higher average grade (3.76), and were focused on the economic incentive by the value of their final grade, while for men, the average



grade is higher in the treatment group 2 (4.13), the one in which students are encouraged by a ranking tournament. The other groups show a decreasing direction, even the control group, having the lowest average grade for women of 2.77 and of 2.03 for men.

These data are related to various studies discussed above, such as the one of Iturra et al. (2012) which presented different emotional environments between genders. One of them was focused on emotional activities that at first, female gender had an advantage. And the other one more involved in the emotional competitive environment, in which the final grade would be the goal and where competitive measures would be established among the students to obtain the highest amount of money.

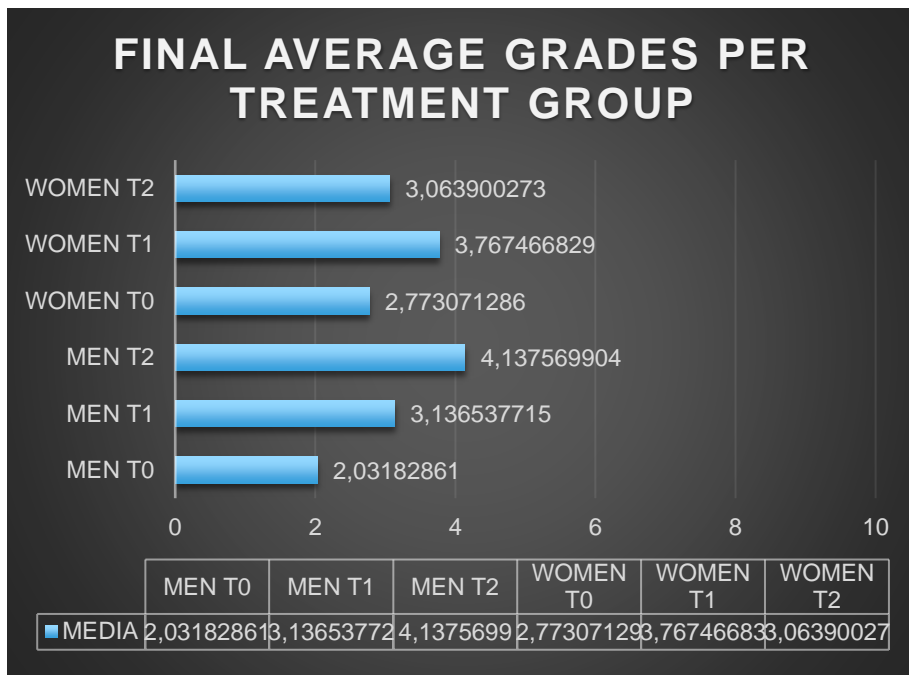


Figure 5 Final average grades per treatment group.

In addition to this, checking the table 1, we almost have various statistics, maximum marks per treatment group, minimum marks of treatment group and standart deviation, that have been useful and contributed to the analysis are:

	<b>H.T0</b>	<b>H.T1</b>	<b>H.T2</b>	<b>M.T0</b>	<b>M.T1</b>	<b>M.T2</b>
<b>Maximum</b>	7.42	10	9.64	8.28	10	8.55
<b>Minimum</b>	0.35	0.27	0.71	0.54	1.07	0.55
<b>Standard deviation</b>	1.589	2.686	2.466	2.131	2.609	2.011

*Table 1 Useful statistics.*

In the chart number 1 several useful statistics can be observed to check initially visual differences in the treatment groups. It is seen that the most noticeable peaks in the in-centived groups, as well as the standard deviations of the samples. The minimums are similar in the treatment groups.

## 6. ANALYSIS OF THE SAMPLE

With regard to the analysis of the sample, some very interesting results have been found. On the one hand, when analysing the sample, the first to consider was the treatment of the groups, differentiated by gender, to examine the results compared between them. That was to see if there were differences between the students' average marks of those groups by taking into account the gender of the students. There is a sample with the following features in terms of sample size.

<b>GROUPS</b>	<b>SAMPLE</b>
Men T0	34
Men T1	31
Men T2	32
Women T0	25
Women T1	23
Women T2	32
TOTAL OF WOMEN	97
TOTAL OF MEN	80

On the other hand, It is noticed that there are differences in the total number of men and women that participated, so it is expected to the test conducted on women to be more rigorous than the one performed on men. When comparing ratios between samples, this will be considered for the final result.

It will be conducted a normality test so as to investigate whether or not these samples follow a regular distribution by means of the Shapiro test. This test enables to contrast the normality samples in which departing from the null hypothesis that the sample is distributed as normal. Also, the alternative hypothesis would exclude the possibility of taking the sample as normal, by starting from a 95% reliability interval.

SHAPIRO TEST	P-VALUE
MEN T0	0.04226
MEN T1	0.1954
MEN T2	0.408
WOMEN T0	0.3974
WOMEN T1	0.05626
WOMEN T2	0.2694

As it is displayed in the chart, once the Shapiro test or the normality test for samples has been conducted, different results have been obtained. The samples of men T1, men T2, women T0 and women T2 present a normal distribution, and the p-value does not show the opposite. There is not enough statistical evidence to determine the non-normality of the sample in those cases.

Conversely, the samples of men T0 and women T1, due to their p-value and a 95% confidence interval, expose the conclusion that there is enough statistical evidence to confirm that they do not follow a normal distribution.

Because of the above mentioned that there is not a normal distribution in all the samples, the Mann-Whitney test will be conducted. This is a non-parametric test applied to two independent samples to help us determine the differences in average grades between treatment groups. It will first start by distinguishing within the same gender and then compares the treatment groups with different gender. This test can be used because the characteristics of the samples are appropriate for it.

### GROUP AND GENDER SAMPLING TABLE

<b>MANN-WITHNEY TEST</b>		<b>P-VALUE</b>
MEN T0	MEN T1	0.01352
MEN T0	MEN T2	8.808e-05
MEN T1	MEN T2	0.2033
WOMEN T0	WOMEN T1	0.2312
WOMEN T0	WOMEN T2	0.6995
WOMEN T1	WOMEN T2	0.2823
MEN T0	WOMEN T0	0.08999
MEN T1	WOMEN T1	0.6682
MEN T2	WOMEN T2	0.04116

All the tests can be seen thoroughly in the annexes and it can be noted that the samples that show a significant difference between samples is the comparison between men. To be more precise, the comparison between the control treatment, group 0 and the treatment groups 1 and 2 is the one that makes the difference. This means that with a 95% of confidence interval, significant differences between the control group and the economic incentives groups have been found in men samples. Nonetheless, there is no substantial difference in the ratio between the treatment groups 1 and 2, despite the fact that the average grade of the group 2 is higher but not significantly, and therefore, it is not possible to determine statistically that difference.

Besides, with regard to women, it has not been able to find a substantial difference in their final average grades, but there is the possibility mentioned above that due to its smaller sample size, the test is more demanding. In the sample of men, when there was found a similar disparity in data, it was detected a significant and determining difference in order to obtain some conclusive results. But in the case of the samples of women, due to the size of the samples, it is not possible to draw the conclusion that economic incentives have a positive effect on the academic performance of students.

In relation to the comparisons of treatment groups in terms of gender, it is determined that with a 95% of confidence interval, there is a significant difference on how pertaining to the treatment group 2 affects men and women. In this case, it would be convenient to mention the differences previously commented in the literature review section about the effect and differences in competitiveness between men and women. At first, the study presents a point in favour of men, as it was stated that men have a greater instinct for competition than women. However, women seem to be more understanding and tend to work better in a team. This result states the study by Iturra et al. (2012) by which different emotional areas between genders were presented.

Once it was compared how the treatment groups affect the students, both dependent and independent of their gender. An analysis was performed through a normality test or a Shapiro test and the results obtained were a p-value of 0.0004938 in men and of 0.005535 in women, as it can be seen in the annex. Concretely, as a consideration of the results obtained, it was discarded the null hypothesis of normality of the samples, at a general level and in behalf of the results procured, it is concluded that there is statistical evidence that the samples do not follow a typical pattern.

From this perspective, a look at the results of the tests conducted will be taken so as to analyse if average marks among genders have clear differences. As it was analysed in detail before between treatment groups, now the analysis will be centred on the samples in general, regardless of the type of treatment to which students belong, only taking into account their gender so as to compare the results with the hypothesis set at the beginning of this paper. Following a Mann-Whitney test to compare ratios that are not from the same population, a p-value of 0.825 was obtained.

As an outcome of the results obtained and with 95% of reliability, it can be stated that there is not exactly a significant difference in terms of gender, as far as the average grade of the students is considered. This was assumed from the graphical analysis which gave us a minimum difference in grade averages. In a broader scope, it is observed that there are differences in the matter of gender regarding of how economic incentives affect academic performance, as they are very effective in men but not in women.

## CONCLUSION

As a result of the study, it has been possible to determine several factors that have been useful in clarifying the questions set before the experiment.

First of all, substantial differences were found due to the effect of economic incentives on students. This result was expected because in terms of psychology matter, it is known the mental and psychological effect of money, that affects positively to the motivation of people. It is clearly appreciated in the labour scope, where the production and the time of the worker is transformed into a monetary reward. In this context, the worker is exchanging time for money. Money generates a positive impact on people's motivation and effort, just as it does in this experiment. In the case studied, differences between the groups that are economically inactivated and those that are not can be observed. So, it can be said that the results obtained are positive, because in a more general scope it has been able to determine as positive and real the hypothesis that has been set out.

Then, from a more pessimistic point of view, significant results were only found in male gender, as far as the effect of economic incentives to improve academic performance were concerned. With regard to the female gender, relevant differences were not found, so it can be determined that the effect of economic incentives to improve their academic performance is not significant, in other words, it does not have a decisive effect on their final grades.

To sum up, the best result obtained is a positive impact on men. On the one hand, economic incentives have had a favourable effect on their final average marks as it has been proven through the different texts conducted in the study. On the other hand, with regards to women, the conclusion reached is that none of the treatment groups has obtained a significant result in the final average mark of studies, so it is possible to affirm that there is no substantial prove that economic incentives affect female students.

## 7. REFERENCES

- Brown, Reva Berman (2006): *Doing your dissertation in business and management: the reality of researching and writing*, SAGE (cuarta edición), London.
- Clark, Damon & Gill, David & Prowse, Victoria & Rush, Mark. (2016). Using Goals to Motivate College Students: Theory and Evidence from Field Experiments in *SSRN Electronic Journal*.
- Enkvist, I. (2010). El éxito educativo finlandés in Bordón. *Revista de pedagogía*, vol. 62, no.3, pp. 49-67.
- Falk, Armin and Knell, Markus. (2004) Choosing the Joneses: Endogenous Goals and Reference Standards. *Scandinavian Journal of Economics*, vol. 106, no. 3, pp. 417-435.
- Fredriksson, P. (1997). Economic incentives and the demand for higher education. *Scandinavian Journal of Economics*, vol. 99, no. 1, pp. 129-142.
- Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*. Pp. 191-210.
- Georg, Sebastian J., Kube, Sebastian. (2012). Goals That Work—Goals, Monetary Incentives, and Workers' Performance. *MPI Collective Goods Preprint 2012/19*.
- Herranz-Zarzoso, N., & Sabater-Grande, G. (2018). Monetary incentives and self-chosen goals in academic performance: An experimental study in *International Review of Economics Education*, vol. 27, pp. 34-44.
- Hirshleifer, S. (2015). Incentives for effort or outputs? A field experiment to improve student performance. Unpublished manuscript. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.
- Iriarte, M. (2007). Motivación "intrínseca" y "extrínseca". Retrieved from: [https://www.in-press.com/wpcontent/uploads/2011/05/13Motivacion\\_Intrinseca\\_Extrinseca.pdf](https://www.in-press.com/wpcontent/uploads/2011/05/13Motivacion_Intrinseca_Extrinseca.pdf).
- Iturra, G. O., Astete, E. P., & Jara, M. O. (2012). Habilidades sociales y rendimiento académico: Una mirada desde el género. *Acta colombiana de psicología*, vol. 15, no. 2, pp. 21-28.
- Jacob, B. (2002). Where the boys aren't: non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, vol. 21, pp. 589 – 598.
- Keller, J.M. (1983). Motivational design of instruction. *Instructional Design Theories and Models: An Overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum.



- Latham, Gary P., Locke, Edwin A., (1979). Goal setting—a motivational technique that works. *Organizational Dynamics*. Vol. 8, no. 2, pp.68–80.
- Lemieux, T., Macleod, W. B., & Parent, D. (2009). Performance pay and wage inequality. *Quarterly Journal of Economics*. Vol. 124, no. 1, pp. 1-49.
- Murnane, R. J. (1993). Economic incentives to improve teaching. *Teachers in Developing Countries*, vol. 137.
- Nieto, D. A. (2005) Efectos del resumen de la mejora de la metacompreñión, de la comprenñión lectora y del rendimiento acadmico in *Revista de educacin*, no. 337, pp. 281-294.
- Schmeck, R.R., Ribich, F.D., & Ramanaiah, N. (1977). Development of a self-report inventory for assessing individual differences in learning processes. *Applied Psychological Measurement*, vol. 1, pp. 413-431.
- Severiens, S.E., y Ten Dam, G.T. (1994). Gender differences in learning styles: a narrative review and quantitative meta-analysis. *Higher Education*, vol. 27, pp. 487-501.
- Zaidi, S.A.R (2016) The Impact of Gender Difference Motivating Assessing and Grading Student's Performance.
- Beall, A.E., & Sternberg, R.J. (1993). The psychology of gender. New York, The Guilford Press.
- Garca, F. C. (2000). Diferencias de gnero en estrategias y estilos de aprendizaje. *Psicothema*, vol. 12, no. 3, pp. 360-367.
- Grimes, S.K. (1995). Targeting academic programs to student diversity utilizing learning styles and learning-study strategies. *Journal of College Student Development*, vol. 36, no. 5, pp. 422-430.
- Nnez, J.C., Gonzlez-Pienda, J.A., Garca-Rodrguez, M.S., Gonzlez Pumariega, S., y Garca, S.I (1994). Estrategias de aprendizaje en estudiantes de 10 a 14 aos y su relacin con los procesos de atribucin causal, el autoconcepto y las metas de estudio. *Revista Galega de Psicopedagoga*, vol. 7, no. 10-11, pp. 219-242.
- Williams, Joseph M. (1990), "Ten Principles for Writing Clearly," in *Style: Toward Clarity and Grace*, University of Chicago Press, Chicago.

## 8. ANNEXES

### Normality test of independent samples

`shapiro.test(Hombrest0$`NOTA EXAMEN10`)`

Shapiro-Wilk normality test

data: Hombrest0\$`NOTA EXAMEN10`

W = 0.93446, p-value = 0.04226

`shapiro.test(hombrest1$`NOTA EXAMEN10`)`

Shapiro-Wilk normality test

data: hombrest1\$`NOTA EXAMEN10`

W = 0.95355, p-value = 0.1954

`shapiro.test(hombrest2$`NOTA EXAMEN10`)`

Shapiro-Wilk normality test

data: hombrest2\$`NOTA EXAMEN10`

W = 0.96646, p-value = 0.408

`shapiro.test(mujerest0$`NOTA EXAMEN10`)`

Shapiro-Wilk normality test

data: mujerest0\$`NOTA EXAMEN10`

W = 0.95913, p-value = 0.3974

`shapiro.test(mujerest1$`NOTA EXAMEN10`)`

Shapiro-Wilk normality test

data: mujerest1\$`NOTA EXAMEN10`

W = 0.91656, p-value = 0.05626

`shapiro.test(mujerest2$`NOTA EXAMEN10`)`

Shapiro-Wilk normality test

data: mujerest2\$`NOTA EXAMEN10`

W = 0.95969, p-value = 0.2694

## Mann-Whitney test for independent samples

Comparison between general sample of men and women

```
wilcox.test(hombresgeneral$`NOTA EXAMEN10`,mujeresgeneral$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: hombresgeneral\$`NOTA EXAMEN10` and mujeresgeneral\$`NOTA EXAMEN10`

W = 3804.5, p-value = 0.825

alternative hypothesis: true location shift is not equal to 0

Comparison between independent sample processing groups

```
> wilcox.test(Hombrest0$`NOTA EXAMEN10`,hombrest1$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: Hombrest0\$`NOTA EXAMEN10` and hombrest1\$`NOTA EXAMEN10`

W = 338.5, p-value = 0.01352

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(Hombrest0$`NOTA EXAMEN10`,hombrest2$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: Hombrest0\$`NOTA EXAMEN10` and hombrest2\$`NOTA EXAMEN10`

W = 238, p-value = 8.808e-05

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(hombrest1$`NOTA EXAMEN10`,hombrest2$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: hombrest1\$`NOTA EXAMEN10` and hombrest2\$`NOTA EXAMEN10`

W = 403, p-value = 0.2033

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(mujerest0$`NOTA EXAMEN10`,mujerest1$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: mujerest0\$`NOTA EXAMEN10` and mujerest1\$`NOTA EXAMEN10`

W = 229, p-value = 0.2312

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(mujerest0$`NOTA EXAMEN10`,mujerest2$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: mujerest0\$`NOTA EXAMEN10` and mujerest2\$`NOTA EXAMEN10`

W = 375.5, p-value = 0.6995

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(mujerest1$`NOTA EXAMEN10`,mujerest2$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: mujerest1\$`NOTA EXAMEN10` and mujerest2\$`NOTA EXAMEN10`

W = 431.5, p-value = 0.2823

alternative hypothesis: true location shift is not equal to 0

### Comparison between equal treatment groups and different gender of independent samples

```
> wilcox.test(Hombrest0$`NOTA EXAMEN10`,mujerest0$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: Hombrest0\$`NOTA EXAMEN10` and mujerest0\$`NOTA EXAMEN10`

W = 314, p-value = 0.08999

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(hombrest1$`NOTA EXAMEN10`,mujerest1$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: hombrest1\$`NOTA EXAMEN10` and mujerest1\$`NOTA EXAMEN10`

W = 331.5, p-value = 0.6682

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(hombrest2$`NOTA EXAMEN10`,mujerest2$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: hombrest2\$`NOTA EXAMEN10` and mujerest2\$`NOTA EXAMEN10`

W = 664.5, p-value = 0.04116

alternative hypothesis: true location shift is not equal to 0

```
> wilcox.test(hombresgeneral$`NOTA EXAMEN10`,mujeresgeneral$`NOTA EXAMEN10`)
```

Wilcoxon rank sum test with continuity correction

data: hombresgeneral\$`NOTA EXAMEN10` and mujeresgeneral\$`NOTA EXAMEN10`

W = 3804.5, p-value = 0.825

alternative hypothesis: true location shift is not equal to 0