CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Bachelor thesis

2014/2015 **Edita Ďurovčíková**

CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Edita Ďurovčíková

Comparison of progress in information processing after a term of studies of mathematical and non-mathematical programs

Bachelor thesis

Prague 2015

Author: Edita Ďurovčíková

Supervisor: PhDr. Julie Chytilová, Ph.D.

Academic Year: 2014/2015

Bibliographic note

ĎUROVČÍKOVÁ, Edita. *Z Comparison of progress in information processing after a term of studies of mathematical and non-mathematical programs*. 37 p. Bachelor thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor: PhDr. Julie Chytilová, Ph.D.

Abstract

If we follow the rules of neoclassical microeconomics, the way the information is told to individuals should not matter. But the research has been showing that this does matter. In my bachelor thesis, I firstly introduce the term behavioral economics and describe some of the effects which influence information processing of individuals. In the second part of my thesis I present the main part of my bachelor thesis – detailed description of my research where I try to answer the main question: Is there a significant difference in progress in information processing between mathematicians and non-mathematicians after a term of studies? I also study the initial difference between mathematicians and non-mathematicians and correlation of their progress and other important factors like gender, GPA and an interest in the studied program. I observe a significant difference in progress in 2 out of 7 studied effects, however, correlation with other variables is very limited by a short-time studied period.

Keywords

behavioral economics, information processing, framing effect, research, econometrics

Abstrakt

Podľa neoklasickej mikroekonómie by nemalo záležať na spôsobe, akým je informácia podaná jednotlivcom. Výskumy však ukazujú opak. V mojej bakalárskej práci najprv predstavujem pojem behaviorálna ekonómia and popisujem niektoré z efektov, ktoré ovplyvňujú spracovávanie informácií. V druhej časti sa venujem najdôležitejšej časti mojej práce – detailnému popisu môjho výskumu, v ktorom sa pokúšam zodpovedať hlavnú otázku: Pozorujeme rozdiel v pokroku v spracovávaní informácií medzi matematikmi a nematematikmi po semestri štúdia? Študujem tiež počiatočný rozdiel medzi matematikmi a nematematikmi a koreláciu progresu s ostatnými dôležitými faktormi, akými sú napríklad pohlavie, študijný priemer a záujem o študovaný odbor. Viditeľný rozdiel v progrese pozorujem u dvoch zo siedmych skúmaných efektov, avšak korelácia s ostatnými premennými je veľmi limitovaná krátkym študovaným časovým úsekom.

Kľúčové slová

behaviorálna ekonómia, spracovávanie informácií, framing efekt, výskum, ekonometria

Range of thesis: 56 274 symbols

Declaration of Authorship

I do solemnly declare that I have written the presented research thesis by myself without undue help from a second person others and without using such tools other than that specified.

Prague, May 7, 2015 Edita Ďurovčíková

Acknowledgments

I would like to express my gratitude to my supervisor, PhDr. Julie Chytilová, Ph.D., and to Mgr. Barbara Pertold-Gebická M.A., Ph.D. for helping me with the analysis. I would also like to thank students who were willing to respond my questionnaire.

Institute of Economic Studies
Bachelor thesis proposal

V mikroekonómií sa často hovorí o preferenciách, na základe ktorých spotrebiteľ nachádza pre seba najoptimálnejšie riešenie. Podľa tejto teórie by teda spôsob, akým je informácia podaná, nemal hrať úlohu. Výskumy však ukazujú, že spôsob podania informácie zohráva v jej spracovaní veľkú rolu, keďže totožná informácia podaná rôznymi spôsobmi vyvoláva u ľudí odlišné, niekedy až protichodné reakcie. Hovoríme o tzv. framingu.

Vo svojej práci by som sa rada venovala pozorovaniu zmeny v spracovávaní informácií po semestri štúdia matematického odboru a porovnala ho s rozdielom v spracovávaní informácií po semestri štúdia vyslovene nematematického odboru.

Výskum bude prebiehať v dvoch častiach. Na začiatku školského roka 2014/2015 budú „matematici” a „nematematici“ požiadaní o odpoveď na niekoľko otázok, ktorými sa inšpirujem napríklad u Daniela Kahnemana, aby som zistila, aké boli ich schopnosti prijímania a spracovávania informácií pred začatím štúdia. Druhé kolo otázok im položím po absolvovaní prvého semestra, aby som videla, kam a či vôbec sa ich spracovávanie informácií posunulo. Výsledky budú následne spracované a vyhodnotené.

Medzi otázky, ktorými sa budem zaoberať, patria:

- Existujú už na začiatku štúdia značné rozdiely v spracovávaní informácií u ľudí, ktorí si zvolili matematický odbor, a u ľudí, ktorí sa rozhodli pre odbor nematematický?
- Je vôbec možné, aby v tak krátkom období nastal viditeľný posun v spracovávaní informácií.
- Ak áno, potom koreluje tento posun so známkami študentov a časom, ktorý venujú štúdiu?
- Nastane u študentov matematiky výraznejší posun než u ich kolegov z nematematických vied?
- Uvidíme väčší posun u osôb mužského alebo ženského pohlavia?

- Ovplyvní tento posun u jednotlivcov fakt, či ich štúdium úprimne baví a v danom odbore sa našli, alebo sú tam takpovediac „nasilu?“

Predpokladaná štruktúra práce:

1. Predstavenie témy
2. Úvod do spracovávania informácií, oboznámenie pojmov ako framing a behaviorálna ekonómia
3. Detailný popis podstaty a priebehu výskumu
4. Vyhodnotenie zozbieraných informácií

Literatúra:

Kahneman, Daniel. 2012. *Myšlení – rychlé a pomalé (podla anglického originálu Thinking, Fast and Slow preložila Eva Nevrlá).* Brno: 2012. ISBN 978-80-87270-42-4

Wilkinson, Nick. 2008. *An Introduction to Behavioral Economics.* Houndmills, Basingstoke:2008. ISBN 978-0-230-53259-5

Ariely, Dan. 2008. *Predictably Irrational: The Hidden Forces That Shape Our Decisions.* HarperCollins:2008. ISBN 978-0-06-135323-9

Ariely, Dan. 2008. *The Upside of Irrationality: The Unexpected Benefits of Defying Logic at Work and at Home*. HarperCollins:2008. ISBN 978-0-06-199503-3

Baláž, Vladimír. 2009. *Riziko a neistota: Úvod do behaviorálnej ekonómie a financií.* Bratislava:2009. ISBN 978-802-2410-823.

Contents

List of tables

Introduction

In classical microeconomics we can often hear about preferences. It is preferences which lead individuals to make the optimal decision – the decision which maximizes their utility. In classical microeconomics we consider these decision-takers rational. A rational decision-taker follows four main rules:

1. The decision-taker sets out all the feasible alternatives, rejecting any which are not feasible
2. He takes into account whatever information is readily available, or worth collecting, to assess the consequences of choosing each of the alternatives
3. In the light of their consequences he ranks the alternatives in order of preference, where this ordering satisfies certain assumptions of completeness and consistency
4. He chooses the alternative highest in this ordering, i.e. he chooses the alternative with the consequences he prefers over all other available to him

(Gravelle, Rees, 1992, p.6)

Under this theory, the way the information is told to individuals should not matter. But the research has been showing that this does matter. If one tells the same information about a topic differently it can lead to a change in opinion of an individual about the topic. Sometimes it can even lead to a total opposite opinion. How is this possible? Do the students of mathematical programs tend to use more logical and statistical thinking while information processing? And does this ability get better after a term of their studies? I will try to find the answers to these question in my research.

In the first part of my bachelor thesis I focus on a theoretical background of my research. I start from the basics by introducing behavioural economics. We will have a look at the history of this branch of economics and see some main differences between behavioural and neoclassical economics.

In the second part I describe some of the effects which influence information processing. All the effects used in research and their meaning in the research are described in details.

In the third part I start focusing on the main part - the practical part of my bachelor thesis. I introduce the studied sample and the questionnaire given to students in sample. I describe the variables I used and the process I followed to answer the questions in my research:

- Is there a significant difference in the progress in information processing between mathematicians and non-mathematicians?
- Has there already been a significant difference in information processing between mathematicians and non-mathematicians at the beginning of their studies?
- Is it even possible to see a significant progress in information processing after only a term of studies?
- If yes, is this progress correlated with the GPA of a student?
- Do we observe a larger progress in the group of men or women?
- Is the progress correlated with their affection for the studied program? (Students were asked: Honestly, do you enjoy the program that you study?)

The fourth and the last part of my bachelor thesis is dedicated to a comprehensive description of the research results supported by econometrics analyses. All information given us from the research are summarized in the conclusion.

1. Behavioral economics

"Behavioral economics increases the explanatory power of economics by providing it with more realistic psychological foundations" (Camerer, Loewenstein, Rabin, 2004, p.3)

The first chapter of my bachelor thesis is intended to an introduction of a young branch of economics – behavioral economics, its meaning and history.

1.1 Explanation of the term

The classical economy works with the standard economic model (SEM). In the centre of SEM stands a rational decision maker. The nature of this rational decision maker has already been described in the introduction, but to repeat it briefly, we can say, that he always follows his consistent preferences to find the optimal solution – the solution which maximizes his utility.

This model however suffers from many restrictions and it fails to explain some of the individuals´ actions. In many cases our decision maker is not rational because he cannot help avoiding the influence of other factors surrounding him.

Let us have a look at some of them: He for example predicts the upcoming result differently when he is under risk, he sees the possibility to win more realistic if the winning is described as "more beautiful" to him, he tends to overestimate the peculiar events when they make a big impression on him and many other things…

Behavioral economics tries to avoid these problems by implementing insights from psychology to the field of economics however this does not mean that it rejects the neoclassical economics based on the theory of utility maximization. It only takes into account also psychological parameters while modelling so that these models can be used in a much wider range of situations.

1.2 History

If we took a look forty years back the term *behavioral economics* would not exist however the psychological approach in economics is not new at all.

Everybody that has something to do with economics must have heard about the concept of the "invisible hand" and the author of this concept – Adam Smith. Many of these people however have not heard about the book, *The Theory of Moral Sentiments*, where he presented his ideas, which are identical with the loss aversion theory.

Also Jeremy Benthan, the author of utility concept, wrote a lot about the psychological approach in uptake of utility. Francis Edgeworth introduced the simple model of social utility where the decision maker is no longer *homo economicus* (a decision maker who satisfies the four assumptions of SEM described in the introduction) but he can also show signs of altruism.

After naming these neoclassical economists it is even bigger paradox that rejection of psychological approach in economics started somehow simultaneously with the beginning of the era of neoclassical economics. It was probably caused by the willing of economists to make their discipline an exact, natural science. On the other hand, psychology back than was in its youth and was not considered very scientific.

Implementing psychology back to economics started in the second half of the twentieth century, when the first articles about the importance of psychology in economics were published (Katona, Liebenstein, Scitovsky, Simon). Another big move forward was the worldwide acceptance of the theory that individuals do not make decisions according to the expected value but according to the expected utility.

Finally during the seventies the behavioral economics, as we know it in present, started being formed, mostly thanks to psychologists Ward Edwards, Duncan Luce, Amos Tversky and Daniel Kahneman. The milestone was the release of important articles written by Tversky and Kahneman. In 1974 it was *Science*, where they described the deviation of individuals´ probability judgements from statistical principles. In 1979 *Prospect theory: Decision making under risk* was released and denied the theory of expected utility (Bernoulli´s theory). Instead of that they implemented a new theory, which says that individuals do not make decisions according to expected utility but according to expected change in utility.

In 1986 a conference at the University of Chicago took place, where a large number of social scientist presented their papers and articles focused on behavioral economics.

"In 1997 a special issue of the Quarterly Journal of Economics was devoted to behavioral economics." (Camerer, Loewenstein, Rabin, 2004, p.7)

Since then there have been many lines of research which use behavioral economics. First the normative assumptions of the model are identified, second the anomalies in the model which break normative assumptions are being looked for (subjects´ confusion, transaction costs…). And then new theories which could explain these anomalies are tried to be created and then tested.

2. Studied effects

The second part of my bachelor thesis is dedicated to a detailed description of seven effects which can influence information processing of an individual and which were implemented on students from my research sample.

2.1 Framing effect

First of all I would like to introduce one of the most important phenomena in behavioral economics – the framing effect. "Framing effects are particularly important since they account for a high incidence of preference reversal." (Wilkinson, 2008, p.53) Preference reversal represent the situation when individual chooses the first option when the information is told (framed) one way and the second option when the information is told in a different way.

In my research I paid attention to the kind of framing effects which appear while solving the problem known as "Asian disease problem". Let us have a closer look at this problem:

"Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that exact scientific estimate of the consequences of the programs are as follows:

Version 1:

Program A: 200 people will be saved.

Program B: There is a 1/3 probability that 600 people will be saved and 2/3 probability that no one will be saved.

Version 2:

Program A: 400 people will die.

Program B: There is a 1/3 probability that nobody will die and 2/3 probability that 600 people will die."

(Kahneman, 2012, p.393)

The answers in both versions express exactly the same outcome. Research has been showing that in the case of Version 1 people tend to prefer Program A to Program B (72%) and in the case of Version 2 it is opposite – they prefer Program B to program A (78%).

To explain this we use the Prospect theory developed in 1992 by Tversky and Kahneman. The Prospect theory tries to model the behaviour of people under uncertainty and it shows that people tend to avoid risk in positively framed situations and, on the other hand, they rather take risk when the situation is framed negatively.

People evaluate gains and loses differently and obviously the fear of loss of the same amount exceeds the pleasure of gaining this amount. Implementing this knowledge into the "Asian disease problem" we now see, that people tend to risk more in Version 2 because they simply feel like they have nothing to lose.

Questions from research studying framing effect:

1. There are 600 people with a serious illness. Choose between using cure A and cure B.

A: Exactly 200 people will survive.

B: With probability 2/5 everybody will survive, with probability 3/5 nobody will survive.

2. There are 1200 people with a serious illness. Choose between cure A and cure B.

A: Exactly 400 people will survive.

B: With probability 2/5 everybody will survive, with probability 3/5 nobody will survive.

3. There are 1800 people with a serious illness. Choose between cure C and cure D.

C: Exactly 600 people will die.

D: With probability 2/5 everybody will die, with probability 3/5 nobody will die.

4. There are 600 people with a serious illness. Choose between cure C and cure D.

C: Exactly 200 people will die.

D: With probability 2/5 everybody will die, with probability 3/5 nobody will die.

To create the four questions above, which are suitable for my research, I modified the “Asian paradox problem” Firstly, I changed the probability values so that the options A and B (or C and D) would not have the same expected value. In each of the four cases people choose between an answer with a higher expected value and an answer which follows the Prospect theory.

Secondly, some scientists criticized the “Asian paradox problem” because the answer “200 people will be saved.” can be understood in two ways. It can mean that exactly 200 people will be saved and also it can mean that 200 will be surely saved but there can be more. I treated this problem by adding word “exactly” to my answers and made them undoubtful.

2.2 Representativeness heuristics and Conjunction fallacy

The Representativeness heuristics refers to situations when people tend to ignore probability and statistics while thinking.

“*Tom W. is of high intelligence, although lacking in true creativity. He has a need for order and clarity, and for neat and tidy systems in which every detail finds its appropriate place. His willing is rather dull and mechanical, occasionally enlivened by somewhat corny puns and by flashes of imagination of the sci-fi type. He has a strong drive for competence. He seems to feel little sympathy for other people and does not enjoy interacting with others. Self-centred, he nonetheless has a deep moral sense*.” (Kahneman, 2012, p.159)

Following this description, people were asked to order 9 study programs , business administration, computer science, engineering, humanities and education, law, library science, medicine, physical and life sciences, social science and social work, according to the probability that Tom W. studies these programs. The description of Tom´s personality makes us see him as a typical student of computer science. On the other hand, one would hardly say that he is a student of law or social science.

We have to take into account the fact that the study containing this question was done by Kahneman and Tversky in 1973. In seventies there was much smaller number of computer science students than law students. If we then hold the basic principles of probability, we can definitely say, that the probability that Tom W. is a law student is much higher than the probability that he is a computer science student. However, about 95% of respondents considered Tom to be more likely a student of a computer science than humanities and education.

The conjunction fallacy problem also works with probability principles. This fallacy occurs when people consider a situation with more specific conditions more probable than a situation with only one condition. It can be explained on the “Linda experiment” also done by Kahneman and Tversky.

"*Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?*

1. Linda is a bank teller.

2. Linda is a bank teller and is active in the feminist movement"

(Kahneman, 2012, p.169)

Again, the description of Linda leads us to an opinion that the second option is the right one. It was also the answer of about 80% of respondents from the research done by Kahneman and Tverski.

Let us implement a simple probabilistic rule – probability of a subset can never exceed the probability of a set. Now it is clear that the first option is the more probable.

These two types of failures in statistical thinking of an individual inspired me to create another two questions of my research:

1. Anne is creative, she likes singing, dancing and enjoys going to theatre. Her mother used to act in amateur theatre. Is there a higher probability that Anne studies law or that Anne studies acting?

2. Peter is 35 years old. He likes travelling and adrenaline sports. Is there a higher probability that Peter is a teacher at school or a teacher at school who climbs mountains in his free time?

2.3 Quantity expression

In this part I would like to introduce a method which is very easy to implement and widely used in advertisements and campaigns as a perfect way to manipulate people.

People respond much more sensitively to information where amount is expressed in real numbers than in percentage. To give a simple explanation of this phenomenon we use the following example. Let us imagine a lawyer who defends his client and uses the following information:

"*a) There is a 0,1% chance that the DNA match is false while testing.*

b) The false DNA match while testing happens in 1 of 1000 cases."

(Kahneman, 2012, p.354)

Which of the two sentences would make a bigger impression? It would be probably the second one. When people hear the first sentence the first thing that comes to their mind is simply a number. On the other hand, when they hear the second sentence they can very clearly imagine the one poor innocent person sitting in the jail just because of the mistake in DNA testing.

By using exact numbers our minds become automatically emotional, sympathetic, protective and unable to control only the numerical information that was given to us in the sentence.

Questions from research studying the effect of quantity expression:

"1. There is an illness which kills 1286 of 10000 patients. How serious do you consider the illness? (1 – the least serious, 10 – the most serious)

2. There is an illness which kills 13,2% patients. How serious do you consider the illness? (1 – the least serious, 10 – the most serious)

3. Patients similar to Mr. Jones are estimated to have a 10% chance of committing an act of violence to others. Would you set him free?"

(Kahneman, 2012, p. 353-354)

In Kahneman´s research there were 41% right answered questions among statistically trained people and 21% among the others.

4. Of every 200 patients similar to Mr. Brown, 20 are estimated to commit an act of violence to others. Would you set him free?

2.4 Overweighting the rare events

While making rational opinions people are very limited by the information they take into account the most. This is equivalent to the information they heard about or read about a lot or a situation they experienced.

This explains a lot about, for example, the fear of travelling by plane. It is statistically proved that flying is one of the safest ways how to travel, however, media makes us see it much more dangerous. When there is a car accident, media hardly inform us about it because it is probably the most common accident that happens quite often. On the other hand, plane accidents are the perfect way to catch reader´s eye.

Plane accidents cause death of a large number of innocent people at once, they are often followed by many unproved theories created by media about who caused the accident and why he did it, and they happen very rarely. It is their rarity which makes us experience a very intensive emotions every time they happen.

Another example of overweighting peculiar events is a danger of terrorist attacks. In the book *Thinking Fast and Slow* Daniel Kahneman describes his experience in Israel.

It was known that within about three years there were 23 suicidal attacks in busses. According to a number of people in Israel who use public transport, this was a very negligible number. Kahneman noticed immediately that local people had seen the danger bigger that it was. They avoided bus transport and when they had to use the bus after all they were nervously and suspiciously checking other people and their luggage in the bus. There was another thing that Kahneman noticed and that surprised him much more. Not even Kahneman himself was able to feel fully comfortable while travelling by bus. Calling himself *the one to know the truly probability of danger*, he could not help feeling anxiety everytime he stopped at the traffic lights with a bus next to him, and relief when the green light appeared on the traffic lights giving him permission to move on with his vehicle.

This example shows that one is never able to avoid overweighting rare situations. Kahneman did not avoid "dangerous" situations because he feared of his life.

He did it primary because when he got into this “dangerous” situation, he started thinking automatically about terrorist attacks, bombs, death… Leaving these thoughts behind simply increased his utility.

There is also positive overweighting. Lottery is a perfect example. The advertisement showing us plenty of happy winning faces makes us overestimate our chance to win. We do not have a good feeling only when we win. We get a great feeling immediately after buying the lottery ticket. Buying the lottery ticket then increases our immediate utility.

Kahneman´s story inspired me to create another two questions of my research:

1. Between December 2001 and September 2004 there were 23 suicidal attacks in busses in Israel which caused 236 deaths. Based on this information, would you prefer travelling to job by taxy to travelling by bus even if it is much more expensive?

2. During last summer there were three plane accidents. There are about 10 000 flights daily in this season. Based on this information, would you prefer travelling for holiday by bus to travelling by plane even if it takes much longer? (You do not suffer from fear of flying or sickness during flight.)

2.5 Denominator neglect

“1. Red ball means winning. Choose bowl from which you are going to draw.

A: 10 balls, 1 red

B: 100 balls, 8 red”

(Kahneman, 2012, p.352)

The probability of winning is in the first case is 10%, in the second case 8%. Choosing the bowl B seems to be really foolish. There are, however, many people who ignore this fact and simply choose the bowl with a higher amount of winning balls. (30-40% in the original research) They pay higher amount of attention to the winning balls then to the rest.

This phenomenon can be explained better if we imagine it visually. There are two bowls standing in front of us – bowl A and bowl B from the example above. The red ball means winning. From which bowl are we going to draw?

In the first bowl we can see very clearly 1 red ball and 9, for example, white balls. In the second bowl we can see 8 red balls and a larger amount of white balls. Exactly. 92 white balls are simply too many to be counted immediately by a human brain. Our eyes can just see a heap of white balls, not an exact number.

In my research I used the question above and I also modified this question by multiplying by two to create another question:

2. Red ball means winning. Choose bowl from which you are going to draw.

A: 20 balls, 2 red

B: 200 balls, 16 red

2.6 Implications

Questions used in research:

"1. Do the first two sentences imply the third one? Every rose is a flower. Some flowers wilt fast. Some roses wilt fast."

(Kahneman, 2012, p.52)

2. Do the first two sentences imply the third one? Every pit bull is a dog. Some dogs have fleas. Some pit bulls have fleas.

I believe that the right answer (No.) does not need any explanation. Everybody who finished high school should solve this without any problem by using basics of mathematical logic. Why is it the so, that people tend to answer positively without any hesitation? Laziness.

Kahneman says that these people have *Lazy controller*. Considering the question easy and unimportant, they have a lack of motivation to solve the problem thoroughly and feel satisfied by answering “Yes.” without any effort.

3. Introduction to the research

In the following chapter I would like to present the most important part of my bachelor thesis – the research. The following pages contain information about theoretical background of my research, research sample, variables used, questions to be answered in research and the research itself.

3.1 Survey questions

All the seven effects mentioned above and also all the questions mentioned above were used to answer the main question of my research: Is there a significant difference in the progress in information processing between *Mathematicians* and *Non-Mathematicians*?

Other questions studied were:

- Has there already been a significant difference in information processing between *Mathematicians* and *Non-Mathematicians* at the beginning of their studies?

- Is it even possible to see a significant progress in information processing after only a term of studies?

- If yes, is this progress correlated with the GPA of a student?

- Do we observe a larger progress in the group of men or women?

- Is the progress correlated with their affection for the studied program? (Students were asked: Honestly, do you enjoy the program that you study?)

3.2 Sample

As it was mentioned before, the research compares progression in information processing of two groups of students. I would like to now introduce these two groups more precisely.

Mathematicians – consists of 36 students, who study their first grade at the Faculty of Mathematics and Physics and who have never studied at any other university before. This group of *Mathematicians* was then randomly divided into two same-sized groups – *Mathematicians1* and *Mathematicians2*. (The reason for this will be explained later.)

Non-mathematicians – consists of 50 students, who study the first grade at the university and who study a program which contains no subjects focused on mathematics. As the group before, they also have not studied at any other university before and as the group before, group of *Non-mathematicians* is also randomly divided into two same-sized groups – *Non-mathematicians1* and *Non-mathematicians2*.

Table nr. 1 - Sample gender distribution 1

Mathematicians				Non-Mathematicians			
Male		**Female**		**Male**		**Female**	
28		8		23		27	
Mathematicians1		Mathematicians2		Non-Mathematicians1		Non-Mathematicians2	
Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
14	4	14	4	11	14	12	13

3.3 Tests

The two following tests were given to the half of the groups introduced above in an opposite order than to the second half. The reason for this is explained in the subchapter 3.4 Data collection.

3.3.1 Test 1

1. There are 600 people with a serious illness. Choose between cure A and cure B.

A: Exactly 200 people will survive.

B: With probability 2/5 everybody will survive, with probability 3/5 nobody will survive.

(Variables *FramingPositive_1, FramingPositive_2*)

2. There are 1800 people with a serious illness. Choose between cure C and cure D.

C: Exactly 600 people will die.

D: With probability 2/5 everybody will die, with probability 3/5 nobody will die.

(Variables *FramingNegative_1, FramingNegative_2*)

3. Anne is creative, she likes singing, dancing and enjoys going to theatre. Her mother used to act in amateur theatre. Is there a higher probability that Anne studies law or that Anne studies acting?

(Variables *StatisticalThinking_1, StatisticalThinking_2*)

4. There is an illness which kills 1286 of 10000 patients. How serious do you consider the illness? (1 – the least serious, 10 – the most serious)

(Variables *QuantityExpression_1, QuantityExpression_2*)

5. Between December 2001 and September 2004 there were 23 suicidal attacks in busses in Israel which caused 236 deaths. Based on this information, would you prefer travelling to job by taxy to travelling by bus even if it is much more expensive?

(Variables *RareEvents_1, RareEvents_2*)

6. Red ball means winning. Choose bowl from which you are going to draw.

A: 10 balls, 1 red

B: 100 balls, 8 red

(Variables *DenominatorNeglect_1, DenominatorNeglect_2*)

7. Do the first two sentences imply the third one? Every rose is a flower. Some flowers wilt fast. Some roses wilt fast.

(Variables *Implications_1, Implications_2*)

8. There is an illness which kills 13,2% patients. How serious do you consider the illness? (1 – the least serious, 10 – the most serious)

(Variables *QuantityExpression_1, QuantityExpression_2*)

3.3.2 Test 2

1. There are 1200 people with a serious illness. Choose between cure A and cure B.

A: Exactly 400 people will survive.

B: With probability 2/5 everybody will survive, with probability 3/5 nobody will survive.

(Variables *FramingPositive_1, FramingPositive_2*)

2. There are 600 people with a serious illness. Choose between cure C and cure D.

C: Exactly 200 people will die.

D: With probability 2/5 everybody will die, with probability 3/5 nobody will die.

(Variables *FramingNegative_1, FramingNegative_2*)

3. Peter is 35 years old. He likes travelling and adrenaline sports. Is there a higher probability that Peter is a teacher at school or a teacher at school who climbs mountains in his free time?

(Variables *StatisticalThinking_1, StatisticalThinking_2*)

4. Patients similar to Mr. Jones are estimated to have a 10% chance of committing an act of violence to others. Would you set him free?

(Variables *QuantityExpression_1, QuantityExpression_2*)

5. During last summer there were three plane accidents. There are about 10 000 flights daily in this season. Based on this information, would you prefer travelling for holiday by bus to travelling by plane even if it takes much longer? (You do not suffer from fear of flying or sickness during flight.)

(Variables *RareEvents_1, RareEvents_2*)

6. Red ball means winning. Choose bowl from which you are going to draw.

A: 20 balls, 2 red

B: 200 balls, 16 red

(Variables *DenominatorNeglect_1, DenominatorNeglect_2*)

7. Do the first two sentences imply the third one? Every pit bull is a dog. Some dogs have fleas. Some pit bulls have fleas.

(Variables *Implications_1, Implications_2*)

8. Of every 200 patients similar to Mr. Brown, 20 are estimated to commit an act of violence to others. Would you set him free?

(Variables *QuantityExpression_1, QuantityExpression_2*)

3.4 Data collection

The data collection consisted of two parts:

The first part of data collection began at the beginning of a school year 2014/2015. The groups *Mathematicians1* and *Non-Mathematicians1* were asked questions from the *Test 1* and the groups *Mathematicians2* and *Non-Mathematicians2* were asked questions from the *Test 2*.

The second part of data collection was done at the beginning of a summer term when all exams were over. Now the tests and the groups were changed. *Test 1* was given to the groups *Mathematicians2* and *Non-Mathematicians2* and *Test 2* was given to the groups *Mathematicians1* and *Non-Mathematicians1*. Students were also asked two more questions:

1. What is your GPA for the first term?

2. Honestly, do you enjoy the program that you study?

The collected data were then processed and studied to answer the questions above.

The reason for dividing of both groups is to avoid observing a progress which would be caused only by the difference in difficulties of *Test 1* and *Test 2*.

3.5 Variables description

In my research I worked with many variables to describe the answers of students:

Math – a binary variable; equals 1 if a student belongs to group *Mathematicians*, equals 0 otherwise

Test – a binary variable; equals 1 if a student belongs to groups *Mathematicians1* and *Non-Mathematicians1*, equals 0 otherwise

FramingPositive_1 – a binary variable; equals 1 if a student chose the answer B to the first question during the first part of the research, equals 0 otherwise

FramingNegative_1 – a binary variable; equals 1 if a student chose the answer C in the second question during the first part of the research, equals 0 otherwise

StatisticalThinking_1 – – a binary variable; equals 1 if the answer of a student to the third question during the first part of the research is statistically right, equals 0 otherwise

QuantityExpression_1 – a binary variable; equals 1 if a student is consistent in answers of the second and the ninth question during the first part of the research, equals 0 otherwise

RareEvents_1 – a binary variable; equals 1 if a student answered positively to the fifth question in during first part of the research, equals 0 otherwise

DenominatorNeglect_1 – a binary variable; equals 1 if a student chooses bowl A in the sixth question during first part of the research, equals 0 otherwise

Implications_1 – a binary variable, equals 1 if a student answers negatively to the seventh question during the first part of the research, equals 0 otherwise

FramingPositive_2, FramingNegative_2, ..., Implications_2 – denotes the same information as the variables above but they are related to the second part of the research

Male – a binary variable; equals 1 if a student is a man, equals 0 if a student is a woman

GPA – denotes the GPA of a student; equals 4 if a student ended his studies

Enjoy – a binary variable; equals 1 if a student enjoys the program that he study, equals 0 otherwise

To simplify the meaning of the variables we can say that every variable equals 1 if the answer was right and equals 0 if the answer was wrong.

4. The research

In the following chapter I would like to interpret the collected data in both parts of my research and to use these data to answer the questions mentioned above.

4.1 Analysis of the first data collected

After collecting data from students at the beginning of their studies I have focused on one important question: Is there a significant difference in information processing between mathematicians and non-mathematicians at the beginning of their studies? I will try to answer this question as thoroughly as possible and also control for other factors like gender and a difference in two tests given to students.

Firstly, I started by using a simple t-test to see if there is a significant difference between *Mathematicians* and *Non-Mathematicians*.

Table nr. 2 – Initial difference Math vs. Non-Math

	Math	Non-math	Difference	t-test	p-value	Observations - M	Observations - N
Framing effect, positive	0.5	0.38	0.12	-1.1038	0.2728	36	50
Framing effect, negative	0.5278	0.24	0.2878	-2.8369	0.0057	36	50
Statistical thinking	0.3611	0.3	0.0611	-0.5909	0.5562	36	50
Quantity expression	0.6944	0.68	0.0144	-0.1407	0.8884	36	50
Rare events	0.9167	0.78	0.1367	-1.7024	0.0924	36	50
Denominator neglect	0.9722	0.7	0.2722	-3.37	0.0011	36	50
Implications	0.4167	0.26	0.1567	-1.5327	0.1291	36	50

To find out if there is a significant difference we have to focus on the column in the table above which represents p-value. P-value is the smallest value at which our hypothesis can be rejected (Hypothesis: No difference between *Mathematicians* and *Non-Mathematicians*).

If the p-value is smaller than 0.1 we reject our hypothesis at 90% level. If the p-value is smaller than 0.05 we reject our hypothesis at 95% level and if the p-value is smaller than 0.01 we reject our hypothesis at 99%.

Therefore we can say that we can see a significant difference in the case of Negative framing effect, Rare events and Denominator neglect. In these three cases, if we look at the means at variables *Math* and *Non-math* we can say that the students who belong to the *Mathematicians* group have answered more correctly than the students from *Non-Mathematicians* group because their coefficient is a larger number.

The coefficients in the case of *Mathematicians* are greater than in the case of *Non-Mathematicians* also in other four studied effects, however, considering the high p-value, we do not have enough evidence to say that there was a significant difference between the two groups in these cases. This is probably caused by our sample size which is not big enough.

Secondly, I would like to have a look at the differences of the two test given to students. These tests were computed the way so that they would be focused on the same seven effects.

Table nr. 3 – Initial difference Test 1 vs. Test 2

	Test 1	Test 2	Difference	t-test	p-value	Observations-T1	Observations-T2
Framing effect, positive	0.3256	0.5349	-0.2093	-1.982	0.0507	43	43
Framing effect, negative	0.2791	0.4419	-0.1628	-1.5765	0.1187	43	43
Statistical thinking	0.186	0.4651	-0.2791	-2.8588	0.0054	43	43
Quantity expression	0.814	0.5581	0.2559	2.6277	0.0102	43	43
Rare events	0.7209	0.9535	-0.2326	-3.0416	0.0031	43	43
Denominator neglect	0.814	0.814	0	0	1	43	43
Implications	0.3023	0.3488	-0.0465	-0.4554	0.65	43	43

Again, considering the p-value, we can see a significant difference between the answering to the test questions in the case of Positive framing effect, Statistical thinking, Quantity expression and Rare events. In these four cases we observe a significant difference between the amounts of correctly answered questions among the students.

This does not have to mean exactly that the test questions were different. For example, if we take a look at the questions studying the Positive framing effect, we can clearly see, that they are identical, only the numbers used are multiplied by 2 in the case of Test2. The two subsamples can simply contain people with different abilities at information processing. That is why we observe significance.

The fact that in the case of some effects we see a significant difference does not have to mean that this difference is caused by the study focused on mathematics. That is why we have to also control for other factors.

If we have a look at the sample description above we see that group *Mathematicians* contains much higher share of men than the group *Non-mathematicians*. This means that if men had answered more correctly than women, we would have probably seen a significant difference between the two studied groups and this difference would not be caused by the influence of mathematics.

To find if there is a significant difference between men and women we again compute t tests and compare means across genders. We get the following table:

Table nr. 4 – Initial difference Male vs. Female

	Male	Female	Difference	t-test	p-value	Observations - M	Observations - F
Framing effect, positive	0.4706	0.3714	0.0992	-0.9062	0.3675	51	35
Framing effect, negative	0.4314	0.2571	0.1743	-1.6605	0.1005	51	35
Statistical thinking	0.3529	0.2857	0.0672	-0.6476	0.519	51	35
Quantity expression	0.7255	0.6286	0.0969	-0.9453	0.3472	51	35
Rare events	0.9608	0.6571	0.3037	-4.0485	0.0001	51	35
Denominator neglect	0.9412	0.6286	0.3126	-3.9365	0.0002	51	35
Implications	0.3529	0.2857	0.0672	-0.6476	0.519	51	35

The table shows us, that in almost all cases we do not have to worry about a difference between men and women. The exceptions are Rare events and Denominator neglect, where we can see a very strong significance. In both of these cases we could have also seen a significant difference between *Mathematicians* and *Non-Mathematicians*, therefore, we cannot make a conclusion about the difference between the studied groups in the case of these two effects at the beginning of their studies.

In another step, I would like to focus on the significant differences in a smaller sub-samples to find out if I observe big differences between these sub-samples.

Firstly, I take a look at the significance separately in groups given the Test1 and groups given the Test2:

Table nr. 5 – Initial difference Math vs. Non-Math only for Test 1

	Math	Non-math	Difference	t-test	p-value	Observations - M	Observations - N
Framing effect, positive	0.4444	0.24	0.2044	-1.4113	0.1657	18	25
Framing effect, negative	0.3889	0.2	0.1889	-1.3599	0.1813	18	25
Statistical thinking	0.2778	0.12	0.1578	-1.3072	0.1984	18	25
Quantity expression	0.7222	0.88	-0.1578	1.3072	0.1984	18	25
Rare events	0.8333	0.64	0.1933	-1.3934	0.171	18	25
Denominator neglect	0.9444	0.72	0.2244	-1.9005	0.0644	18	25
Implications	0,5	0.16	0.34	-2.5121	0.016	18	25

Table nr. 6 – Initial difference Math vs. Non-Math only for Test 2 (43 observations)

	Math	Non-math	Difference	t-test	p-value	Observations - M	Observations - N
Framing effect, positive	0.5556	0.52	0.0356	-0.2253	0.8229	18	25
Framing effect, negative	0.6667	0.28	0.3867	-2.6639	0.011	18	25
Statistical thinking	0.4444	0.48	-0.0356	0.2253	0.8229	18	25
Quantity expression	0.6667	0.48	0.1867	-1.2083	0.2338	18	25
Rare events	1	0.92	0.08	-1,2216	0.2288	18	25
Denominator neglect	1	0.68	0.32	-2.8419	0.007	18	25
Implications	0.3333	0.36	-0.0267	0.1768	0.8605	18	25

In almost all cases *Mathematicians* answered more correctly. The exceptions are Quantity expression in Treatment 1 and Statistical thinking and Implications in Treatment 2. In both treatments we observe a significant difference in answers only in two effects. In the case of Treatment 1 it is Denominator neglect and Implications, in the case of Treatment 2 it is Denominator neglect and Negative framing effect. Low significance can be caused by a not big enough sample.

Secondly, I would like to look whether there are differences between answers across the two tests, among a subsamples of *Mathematicians* and *Non-Mathematicians*:

Table nr. 7 – Initial difference Test 1 vs. Test 2 only for Mathematicians

	Test 1	Test 2	Difference	t-test	p-value	Observations-T1	Observations-T2
Framing effect, positive	0.4444	0.5556	-0.1112	-0.6519	0.5188	18	18
Framing effect, negative	0.3889	0.6667	-0.2778	-1.6889	0.1004	18	18
Statistical thinking	0.2778	0.4444	-0.1666	-1.0272	0.3116	18	18
Quantity expression	0.7222	0.6667	0.0555	0.3523	0.7268	18	18
Rare events	0.8333	1	-0.1667	-1.8439	0.0739	18	18
Denominator neglect	0.9444	1	-0.0556	-1	0.3244	18	18
Implications	0.5	0.3333	0.1667	1	0.3244	18	18

Table nr. 8 – Initial difference Test 1 vs. Test 2 only for Non-Mathematicians

	Test 1	Test 2	Difference	t-test	p-value	Observations-T1	Observations-T2
Framing effect, positive	0.24	0.52	-0.28	-2.087	0.0422	25	25
Framing effect, negative	0.2	0.28	-0.08	-0.6518	0.5177	25	25
Statistical thinking	0.12	0.48	-0.36	-2.9592	0.0048	25	25
Quantity expression	0.88	0.48	0.4	3.288	0.0019	25	25
Rare events	0.64	0.92	-0.28	-2.4879	0.0164	25	25
Denominator neglect	0.72	0.68	0.04	0.3027	0.7635	25	25
Implications	0.16	0.36	-0.2	-1.6222	0.1113	25	25

Focusing on *Mathematicians*, we can see a significant difference between correctly answered questions from Test1 and Test2 only in the case of Rare events. In these case more correct answers were observed in Test2.

Focusing on *Non-Mathematicians*, we can see much more significant differences between correctly answered questions from Test1 and Test2. We can see a significant difference in the case of Positive framing effect, Statistical thinking, Quantity expression and Rare events. This is probably the consequence of high diversity of students who belong to our *Non-Mathematicians* group. The group *Mathematicians* contains only students from the Faculty of Mathematics and Physics, however, *Non-Mathematicians* contain students who fulfil only one condition – no mathematics while studying at the university. Those are students of politics, culture, arts, territorial studies, languages, teaching…

After detailed study of initial correlations we can now run the following regressions where we can study the difference between *Mathematicians* and *Non-Mathematicians* controlling for other important variables. Therefore we run the regression - answer to the question on variables *Math, Test and Male*.

Table nr.9 – Regressions results (86 observations)

		FramingPositive_2	FramingNegative_2	StatisticalThinking_2	QuantityExpression_2	RareEvents_2	DenominatorNeglect_2	Implications_2
Intercept	coefficient	0.4554	0.2804	0.418	0.5004	0.7627	0.5816	0.2749
	p-value	0	0.005	0	0	0	0	0.007
Math	coefficient	0.1003	0.2601	0.0466	-0.0204	0.0467	0.1924	0.151
	p-value	0.378	0.016	0.659	0.846	0.536	0.021	0,17
Test	coefficient	-0.2079	-0.1608	-0.278	0.2584	-0.226	0.0058	-0.0461
	p-value	0.053	0.109	0.006	0.01	0.002	0.939	0.652
Male	coefficient	0.062	0.087	0.0456	0.1097	0.2832	0.2511	0.0177
	p-value	0.587	0.417	0.668	0.3	0	0.003	0.872

The summarized coefficients and the corresponding p-values from the seven models are summarized in the Table nr. 9 above.

Considering the p-value, we see the significant difference between *Mathematicians* and *Non-Mathematicians* only in two cases – Negative framing effect and Denominator neglect.

However, in most of the cases we observe a significant difference between the students who took the *Test 1* first and the students who took the *Test 2* first. This is an unexpected phenomenon because, as it was mentioned before, the tests are almost identical. The significance is probably caused by the fact that our subsamples can contain people with different abilities at information processing. However, the significance also supports the theory that some students may have answered randomly. (I will pay attention to this problem later.)

On the other hand, the significant difference between men and women is not observed in almost any of the cases, where we observe a difference between *Mathemaicians* and *Non-Mathematicians*. Therefore we do not have to worry about the case in which the significant difference between *Mathematicians* and *Non-*

Mathematicians was observed because of the unequal gender distribution in these two groups.

4.2 Analysis of the second data collected

After finishing the second data collection we got the answers from the second part of our test and also the values for the new variables we can work with now – *GPA* and *Enjoy* (honest interest of a student in the studied program). The proportions in the answers in the different groups are summarized in the following tables:

Table nr. 10 – Sample GPA distribution

Mathematicians		Non-Mathematicians	
GPA in Average		**GPA in average**	
2.473		2.0007	
Mathematicians1	Mathematicians2	Non-Mathematicians1	Non-Mathematicians2
GPA in Average	**GPA in Average**	**GPA in Average**	**GPA in Average**
2.4777	2.4683	1.9468	2.0546

Table nr. 11 – Sample distribution of an interest in the studied program

Mathematicians				Non-Mathematicians			
Interested		**Not Interested**		**Interested**		**Not interested**	
26		10		41		9	
Mathematicians1		Mathematicians2		Non-Mathematicians1		Non-Mathematicians2	
Interested	**Not Intr.**	**Intrested**	**Not Intr.**	**Interested**	**Not Intr.**	**Interested**	**Not Intr.**
13	5	13	5	19	6	22	3

The main questions we want to answer using our data are: Is there a significant difference in the progress in information processing between Mathematicians and Non-Mathematicians after a term of their studies? And is this progress correlated with gender, GPA after a first term or the true interest of a student in the program that he studies?

While focusing on each of the seven studied effects, our students answered two questions, each with one right and one wrong answer. This gives us four possible combinations of answers. To distinguish these four combinations we create 7 new variables:

FramingPositive_prog – equals 1 if the variable *FramingPositive_1* equals 0 and the variable *FramingPositive_2* equals 1 (progress); equals 2 if the variable *FramingPositive_1* equals 1 and the variable *FramingPositive_2* equals 0 ("negative progress"); equals 3 if the variable *FramingPositive_1* equals 1 and the variable *FramingPositive_2* equals 1; equals 4 if the variable *FramingPositive_1* equals 0 and the variable *FramingPositive_2* equals 0.

All other Variables like *FramingNegative_prog, StatisticalThinking_prog, QuantityExpression_prog, RareEvents_prog, DenominatorNeglect_prog* and *Implications_prog* are created the same way.

For each variable we generate four dummy variables (e.g. *FramingPositive_progress, FramingPositive_reverse, FramingPositive_nochange1, FramingPositive_nochange0*). Now we can study the effects we are interested in.

Table nr. 12 – Progress distribution

Framing effect, positive	mean	std. dev.	observations
0 -> 1	0,0814	0,275	7
1 -> 0	0,1395	0,3485	12
1 -> 1	0,2907	0,4567	25
0 -> 0	0,4884	0,5028	42

Framing effect, negative	mean	std. dev.	observations
0 -> 1	0,1628	0,3713	14
1 -> 0	0,1628	0,3713	14
1 -> 1	0,1977	0,4006	17
0 -> 0	0,4767	0,5024	41

Statistical thinking	mean	std. dev.	observations
0 -> 1	0	-	0
1 -> 0	0	-	0
1 -> 1	0,3256	0,4713	28
0 -> 0	0,6744	0,4713	58

Quantity expression	mean	std. dev.	observations
0 -> 1	0,2558	0,4389	22
1 -> 0	0,0698	0,2562	6
1 -> 1	0,6163	0,4891	53
0 -> 0	0,0581	0,2354	5

Rare events	mean	std. dev.	observations
0 -> 1	0,093	0,2922	8
1 -> 0	0,1279	0,3359	11
1 -> 1	0,7093	0,4567	61
0 -> 0	0,0698	0,2562	6

Denominator neglect	mean	std. dev.	observations
0 -> 1	0,0698	0,2562	6
1 -> 0	0,0465	0,2118	4
1 -> 1	0,7674	0,4249	66
0 -> 0	0,1163	0,3224	10

Implications			
	mean	std. dev.	observations
0 -> 1	0,2326	0,4249	20
1 -> 0	0,1047	0,3079	9
1 -> 1	0,2209	0,4173	19
0 -> 0	0,4419	0,4995	38

The tables above can be interpreted as the proportions of the four combinations of answers to the seven test questions. Let us have a look at the Implications table at the first number in the first row – 0,2326. This means that approximately 23% of students answered the seventh question at the beginning of their studies wrong and after the first term of studies right.

What else do the previous tables tell us? In all seven cases we can see that majority of students did not change their answers among time. One term of studies seems to be too short to make a large progress in information processing. However, not all the students remained consistent in their answers. We observe a progress but, unexpectedly we observe also a negative progress.

Firstly, let us focus on Statistical thinking – the only effect where we cannot observe any change between answers to the two questions during the first term of studies. Therefore we also cannot see any progress in processing this effect.

What about the other six studied effects? To find out if there is a significant difference we focus on the variable *FramingPositive_progress* (*FramingNegative, StatisticalThinking…*) which represents the progress and compute t tests controlling for the *Math* variable.

Table nr. 13 – Progress Mathematicians vs. Non-Mathematicians

	Math	Non-math	Difference	t-test	p-value	Observations - M	Observations - N
Framing effect, positive	0.1111	0.1613	-0.0502	0.4751	0.6369	18	31
Framing effect, negative	0.0588	0.3421	-0.2833	2.2939	0.0258	17	38
Statistical thinking	0	0	0	-	-	23	35
Quantity expression	1	0.6875	0.3125	-2.1517	0.0413	11	16
Rare events	1	0.4545	0.5455	-1.7566	0.1044	3	11
Denominator neglect	1	0.3333	0.6667	-	-	1	15
Implications	0.4286	0.2973	0.1313	-1.0022	0.3206	21	37

Unluckily, in the case of Denominator neglect, we cannot observe the t-test and p-value because of a small sample. (There is only one student from *Mathematicians* who answered firstly wrong and secondly right.)

To confirm significance, we again need the p-value to be lower than 0,1. This can be observed in two effects – Negative framing effect and Quantity expression. In both cases we observe the p-value even lower than 0,05. This implies significance at 95% level.

Let us now interpret the coefficients. In the case of Negative framing effect, 5,88% of *Mathematicians* who answered wrong in the first part of the research now answered right and 34,21% of *Non-Mathematicians* who firstly answered wrong answered right now. We can then observe a greater progress in the group of *Non-Mathematicians.* In the case of Quantity expression we observe the opposite. There are 68,75% of *Non-Mathematicians* who made progress but in the group of *Mathematicians*, everyone who previously answered wrong now answered right.

Again, we have to pay attention to the unequal distribution of other factors that we control between the two studied groups. The significant progress between *Mathematicians* and *Non-Mathematicians* which we observed at the two effects does not have to be automatically correlated with the studied program. If we take a look at the three proportion tables above we see that *Mathematicians* contain much larger amount of men. On the other hand, *Non-mathematicians* have better GPA in average and also contain more people, who enjoy the program that they study. The question that comes up is: Can the progress in information processing be correlated rather with this disproportion in the studied groups?

Table nr. 14 – Progress Male vs. Female

	Male	Female	Difference	t-test	p-value	Observations - M	Observations - F
Framing effect, positive	0.07407	0.2273	-0.15323	1.5296	0.1328	27	22
Framing effect, negative	0.1724	0.3462	-0.1738	1.4793	0.145	29	26
Statistical thinking	0	0	0	-	-	33	25
Quantity expression	0.8571	0.7692	0.0879	-0.5691	0.5744	14	13
Rare events	0,5	0.5833	-0.0833	0.2045	0.8414	2	12
Denominator neglect	0.3333	0.3846	-0.0513	0.1548	0.8792	3	13
Implications	0.3636	0.32	0.0436	-0.3406	0.7347	33	25

Looking at the p-values of the two effects we are interested in – Negative framing effect and Quantity expression, we see, that we do not have to be afraid of a possibility, that the progress would be correlated with gender. The p-value is too high. Looking at the p-values of all the studied effects we can say the same. We do not observe any significant difference between men and women in the progress at information processing after a term of studies.

Table nr. 15 – Progress Enjoy vs. Not Enjoy

	Enjoy	Not Enjoy	Difference	t-test	p-value	Observatons - E	Observations - NE
Framing effect, positive	0.1143	0.2143	-0.1	0.8925	0.3767	35	14
Framing effect, negative	0.25	0.2667	-0.0167	0.1241	0.9017	40	15
Statistical thinking	0	0	0	-	-	47	11
Quantity expression	0.8571	0.6667	0.1904	-1.0412	0.3078	21	6
Rare events	0.5455	0.6667	-0.1212	0.3499	0.7325	11	3
Denominator neglect	0.4167	0.25	0.1667	-0.5641	0.5816	12	4
Implications	0.3333	0.3846	-0.0513	0.337	0.7374	45	13

The p-values in the Table nr. 15 are again too large to give us permission to make any conclusions. We do not have enough evidence to reject the hypothesis about no correlation between the progress in information processing and interest in the studied program.

Table nr. 16 – Correlation of progress and GPA (86 observations)

	GPA	t-test	p-value	Observations
Framing effect, positive	-0.000006	-0.38	0.701	49
Framing effect, negative	-0.00001	-0.57	0.57	55
Statistical thinking	0.00001	1.1	0.277	58
Quantity expression	0.000002	0.07	0.943	27
Rare events	-0.000007	-0.41	0.68	14
Denominator neglect	-0.000005	-0.35	0.724	16
Implications	-0.00002	-0.71	0.478	58

The Table nr. 16 above shows us correlation between the progress and the GPA of a student using t test. In all seven cases we observe very small coefficients with very small t-tests and very high p-values. Therefore we observe no correlation between the progress and GPA.

Table nr. 17 – Progress Test 1 vs. Test 2

	Test 1	Test 2	Difference	t-test	p-value	Observations-T1	Observations-T2
Framing effect, positive	0.1724	0,1	0.0724	-0.7009	0.4868	29	20
Framing effect, negative	0.2903	0.2083	0.082	-0.6825	0.4979	31	24
Statistical thinking	0	0	0	-	-	35	23
Quantity expression	0.875	0.7895	0.0855	-0.5052	0.6178	8	19
Rare events	0.6667	0	0.6667	-1.8516	0.0888	12	2
Denominator neglect	0.375	0.375	0	0	1	8	8
Implications	0.5	0.1786	0.3214	-2.6868	0.0095	30	28

The split of the studied groups into two half-groups where each group were given the two tests in a different order has already been explained before. We wanted to avoid the possibility, that the observed progress could by correlated with the order of the test and the possibility that one test was easier than the second one. To answer the questions we look at the table above.

The p-values at the two effects where we observed progress (Negative framing effect and Denominator neglect) are too high (0,4979 and 1) to observe any significant difference in progress according to the order of the tests given.

On the previous pages I focused fully on the progress, however, after collecting the data we get to a problem we have to solve. We did not observe only positive progress among the answers, we observed also negative progress.

From the analysis above we concluded that there was a significant difference in progress between Mathematicians and Non-Mathematicians after a term of studies in two effects – Negative framing effect and Quantity expression. However, what if we also observed a significant negative progress while studying these two effects? We could no longer say that students made a progress. The answer to this problem is in the Table nr. 18 below:

Table nr. 18 – Negative progress Mathematicians vs. Non-Mathematicians

	Math	Non-math	Difference	t-test	p-value	Observations - M	Observations - N
Framing effect, positive	0.2222	0.4211	-0.1989	1.2852	0.2071	18	19
Framing effect, negative	0.5263	0.3333	0.193	-1.0358	0.3088	19	12
Statistical thinking	0	0	0	-	-	13	15
Quantity expression	0.04	0.1471	-0.1071	1.3422	0.1849	25	34
Rare events	0.1818	0.1282	0.0536	-0.6229	0.5353	33	39
Denominator neglect	0.0286	0.0857	-0.0571	1.0228	0.31	35	35
Implications	0.2	0.4615	-0.2615	1.4831	0.1501	15	13

Focusing on the two effects with visible progress (Negative framing effect and Quantity expression) and also on all other studied effects, we see that the p-values are not small enough to observe a significant difference in negative progress between *Mathematicians* and *Non-Mathematicians*.

This, however, still does not give us totally valuable information. We have to deal with an effect known as the ceiling effect. If *Mathematicians* have answered more correctly already the first set of questions then there are less *Mathematicians* who have answered wrong and could make a progress answering the second set of questions. N*on-Mathematicians* then have an advantage – a larger set of students who could have made progress.

How to deal with this? Let us have a look at the Table nr. 19 below:

Table nr. 19 – Distribution of the correct answers across the time

	1st p. of research			2nd part of research			At the end	
	Math	Non-math	Significant?	Math	Non-math	Significant?	Math	Non - Math
Framing effect, positive	18 (50%)	19 (38%)	No	2 (6%)	5 (10%)	No	16 (44%)	16 (32%)
Framing effect, negative	19 (53%)	12 (24%)	Yes	1 (3%)	13 (26%)	Yes	10 (28%)	21 (42%)
Statistical thinking	13 (36%)	15 (30%)	No	0 (0%)	0 (0%)	-	13 (36%)	15 (30%)
Quantity expression	25 (69%)	34 (68%)	No	11 (31%)	11 (22%)	Yes	35 (97%)	40 (80%)
Rare events	33 (92%)	39 (78%)	Yes	3 (8%)	5 (10%)	No	30 (83%)	39 (78%)
Denominator neglect	35 (97%)	35 (70%)	Yes	1(3%)	5 (10 %)	-	35 (97%)	37 (74%)
Implications	15 (42%)	13 (26%)	No	9 (25%)	11 (22%)	No	21 (58%)	18 (36%)

The table shows us the distribution of right answered questions across the time. Firstly, we see the amount of *Mathematicians* and *Non-Mathematicians* who answered the test questions right at the beginning of the research. Secondly, there is the amount of *Mathematicians* and *Non-Mathematicians* who did not answer the test questions right at the first part of the research but answered right at the second part of the research. And finally, there is a final amount of *Mathematicians* and *Non-Mathematicians* who answered right at the second part of the research (= 1st column + 2nd column – amount of students who answered firstly right and secondly wrong (Negative progress)).

Now, I will focus separately on each of the effects, where I observed significance at least once while the data collection:

Negative framing effect – Unlike the Positive framing effect, the Negative framing effect showed us very significant differences in both parts of data collection.

Firstly, we observed a higher amount of right answers among *Mathematicians.* Then, however, *Mathematicians* made a very small progress and a large negative progress. On the other hand, *Non-Mathematicians* made significantly larger progress than *Mathematicians*. In this case we do not have to be afraid of the previously mentioned ceiling effect. After the first data collection, there were still 47% of Mathematicians who did not answer right and could have made progress at the end of the term and only 3% of them did. Their progress was then not limited by the low amount of students who could have made progress for sure. We can say, that in this case we observe a significant difference in progress between our two groups.

Quantity expression – After the first data collection there was almost the same percentage of the right answered questions among *Mathematicians* and *Non-Mathematicians*. The both groups then had the same possibility to make a progress and we did not have to be afraid of the ceiling effect anymore. After the second data collection we observed significantly larger progress in the group of *Mathematicians* and *Mathematicians* also made smaller negative progress. We can then observe a significant difference in progress between our two groups.

Denominator neglect – We observed a significant difference between the correctly answered question after the first data collection. Unluckily, after the second data collection, we did not have enough students in the sample and we cannot say anything about the significance. However, if we compare the distribution of right answered questions at the beginning of the research and at the end of the research we observe almost similar percentages and therefore we do not observe significant difference in progress between our studied groups.

Let us now take the six studied effects at which we observed different answers among time (except for Statistical thinking) to run the six regressions of new answers to our six studied effects on our variables – *Math, Test, Male, Enjoy, GPA* and the old answers.

Table nr. 20 – Final regressions

		FramingPositive_2	FramingNegative_2	QuantityExpression_2	RareEvents_2	DenominatorNeglect_2	Implications_2
Intercept	coefficient	0.3295	0.3695	0.7572	0.1577	0.3639	0.0733
	p-value	0.173	0.162	0	0.518	0.048	0.796
Math	coefficient	0.0646	-0.2486	0.1556	-0.0065	0.0761	0.1933
	p-value	0.504	0.026	0.045	0.942	0.292	0.086
Old answer	coefficient	0.5657	0.3818	0.1014	0.331	0.4965	0.3303
	p-value	0	0.001	0.209	0.014	0	0.005
Test	coefficient	0.1267	-0.0684	-0.0907	0.126	-0.0904	0.1662
	p-value	0.171	0.5	0.229	0.164	0.169	0.111
Male	coefficient	-0.0329	-0.0451	0.0568	0.0226	0.0732	-0.0778
	p-value	0.735	0.674	0.465	0.818	0.985	0.484
GPA	coefficient	0.00002	-0.00002	0.000005	-0.000015	-0.0000003	-0.00002
	p-value	0.364	0.512	0.792	0.749	0.985	0.504
Enjoy	coefficient	-0.1147	-0.1643	0.0434	-0.0324	0.0468	-0.0029
	p-value	0.298	0.179	0.615	0.749	0.551	0.981

The answers to questions in the second part of our research – our dependent variables are negatively correlated with variable *Math* in the case of Negative framing effect and positively correlated in the case of Quantity expression and implications. This mean that a larger progress was observed in the answers of *Mathematicians* at the Negative framing effect and Implications and a smaller progress at the Quantity expression.

The correlation between new and old answers is very high at almost all effects except for Quantity expression. This means that students were highly consistent in the answers among the two tests and did not make a big progress probably because of short time between the first and the second test (one semester).

No correlation is observed between the dependent variables and the order of the tests given. Therefore we do not have to worry, that the observed progress in information processing could be correlated with a different difficulty of given tests.

Focusing on the rest of the variables – *Male, Enjoy* and *GPA*, we observe no significant correlation with the explanatory variable.

4.3 The problem of randomly answered questions

In the analysis we observed phenomenon like negative progress or a significant difference in answering almost identical questions. Therefore it is necessary to take into account the possibility that students from our sample answered randomly.

We use a t test to test the hypothesis that the students answered randomly (i.e. the distribution of the right answers among all the questions is exactly 50%). The *Table nr. 21* shows us the corresponding results.

Table nr. 21 – T-test for randomness

	Data collection	
	1	2
Framing effect, positive	0.1974	0.0168
Framing effect, negative	0.0089	0.0089
Statistical thinking	0.0009	0.0009
Quantity expression	0.0004	0
Rare events	0	0
Denominator neglect	0	0
Implications	0.0009	0.3915

Luckily, in almost all questions we can reject the hypothesis about random answering because of a very low p-value. The only exception is a question studying the Positive framing effect in the first part of the data collection and a question focused on Implications in the second part of the data collection.

Conclusion

Is there a significant difference between students of mathematical programs and students of nonmathematical programs in the progress in information processing after a term of studies? This was the main question I wanted to answer in my bachelor thesis.

Firstly, I introduced the term behavioral economics and the seven effects, which I chose to study in my research focused on information processing of students – Framing effect, Statistical thinking (Conjunction fallacy and Representativeness heuristics), Quantity expression, Overweighting of rare events, Denominator neglect and Implications (focused on mathematical logic). Secondly, I described the background of my research – the process of data collection, the tests I implemented – *Test 1* and *Test 2*, the variables I used to interpret the outcomes of my research and the two samples of students called *Mathematicians* and *Non-Mathematicians* (divided randomly into another subsamples – *Mathematicians1*, *Mathematicians2*, *Non-Mathematicians1* and *Non-Mathematicians2*).

The data collection consisted of two parts. The first part took place at the beginning of school year 2014/2015. The *Mathematicians1* and *Non-Mathematicians1* were given the *Test 1* and the *Mathematicians2* and *Non-Mathematicians2* were given the *Test 2*. In the second part of data collection, at the beginning of the second term of studies, I gave the *Test 1* to *Mathematicians2* and *Non-Mathematicians2* and the *Test 2* to *Mathematicians1* and *Non-Mathematicians1*. This was done to avoid the possibility of observing a progress because of a different difficulty of the two tests.

Then I processed the results and tried to find answers to the questions I stated at the beginning of the research:

Is there a significant difference in progress in information processing between *Mathematicians* and *Non-Mathematicians* after a term of studies? - From the analysis where we took into account also the problem of the negative progress, the ceiling effect and the randomly answered question we can conclude that a significant difference was observed at 2 out of 7 effects – Negative framing effect and Quantity expression.

Has there already been a significant difference in information processing between *Mathematicians* and *Non-Mathematicians* at the beginning of their studies? – Yes, we observed a significant difference at 3 out of 7 effects – Negative framing effect, Denominator neglect and Rare events. In all three cases there were more correctly answered questions among Mathematicians.

Is this progress correlated with the GPA of a student, gender or the affection of a student for the studied program? – We did not observe any significant correlation between these three variables and the progress.

The results that came of the research gave us answers to the questions that we stated at the beginning of the research. We observed some significant differences between the studied samples at the beginning and we also observed a significant difference in the progress in answering among the samples. However, after processing the collected data we must conclude that one term of studies is a too short range of time to see any correlation between the progress and variables like gender or GPA and to observe more significant results.

Bibliography

1. KAHNEMAN, Daniel. *Myšlení - rychlé a pomalé*. Vyd. 1. Překlad Eva Nevrlá. V Brně: Jan Melvil, 2012, 542 s. Pod povrchem. ISBN 9788087270424.

2. KAHNEMAN, Daniel, Paul SLOVIC a Amos TVERSKY. *Judgment under uncertainty: heuristics and biases*. New York: Cambridge University Press, c1982, xiii, 555 p. ISBN 0521284147.

3. CAMERER, Colin, George LOEWENSTEIN a Matthew RABIN. *Advances in behavioral economics*. Princeton, N.J.: Princeton University Press, 2004, xxvi, 740 p. ISBN 0691116814-.

4. WILKINSON, Nick. *An introduction to behavioral economics*. New York: Palgrave Macmillan, c2008, xxi, 511 p. ISBN 9780230532595.

5. TALEB, Nassim. *The black swan: the impact of the highly improbable*. London: Allen Lane, c2007, xxviii, 366 s. ISBN 9780713999952.

6. BLACK, John, Nigar HASHIMZADE a Gareth D MYLES. *A dictionary of economics*. 4th ed. /. Oxford: Oxford University Press, 2012, 1 v. ISBN 0199696322.

7. ARIELY, Dan. *Predictably irrational: the hidden forces that shape our decisions*. 1st ed. New York, NY: Harper, c2008, xxii, 280 p. ISBN 006135323x.

8. ARIELY, Dan. *The upside of irrationality: the unexpected benefits of defying logic at work and at home*. New York: Harper, 2010, xi, 334 p. ISBN 9780061995033.

9. BALÁŽ, Vladimír. *Riziko a neistota: úvod do behaviorálnej ekonómie a financií*. 1. vyd. Bratislava: Veda, 2009, 451 s. ISBN 9788022410823.

10. GRAVELLE, Hugh, REES, Ray. *Microeconomics*. 3rd edition. Gosport: Ashford Colour Press, 2004, 738 p. ISBN 978-0-582-40487-8.