The Effect of Fluctuations in Monetary Resources on Dutch Students' Cognitive Function:
Does having less mean that much?

Master's Thesis
Behavioral Economics

Name:	Sten te Vogt
Student number:	427180sv
Supervisor:	Jan Stoop
Date:	14/11/2016

Abstract

Recent research proposed a new perspective on the concept of scarcity by defining it as having less than you feel you need. Scarcity, in any form, leads to less mental resources, known as bandwidth. Evidence has been found that individuals experiencing scarcity, due to fluctuations in income, cognitively underperform their counterparts. Past studies in this new field of science focused on farmers in a third world country and low-income households. This study expands on the existing literature by focusing on Dutch students. An experimental study was conducted comparing cognitive performance before and after the subject population received two very relevant monthly payments. It was found that subjects neither experienced higher feelings of scarcity nor felt more financially pressed as a result of fluctuations in their income. Moreover, no impediment of cognitive function was found. This can be seen as evidence that fluctuations in monetary resources do no universally result in higher perceived scarcity and lower cognitive function. The role of context, environment and uncertainty might be crucial in this new field.

Key words: scarcity, bandwidth, cognitive function, fluid intelligence, monetary resources, payday, students

Table of Contents

List of tables

List of Figures

1. Introduction

"Shantideva, an 8th century Indian Buddhist, vowed to follow a path of service and enlightenment, he made a particular commitment to fulfil the material needs of the poor, who would never otherwise have the mental resources to concentrate on their own enlightenment without distraction" (Spears, 2011, p. 33)

In 2013 Sendhil Mullainathan and Eldar Shafir, an economist and a psychologist respectively, published a book titled "Scarcity". This book was an attempt to describe, what they called, a science in the making. This science is based on the phenomenon of having less than you feel you need (Mullainathan & Shafir, 2013). By defining scarcity as such, the scope of this new field in science is not limited to only financial or material scarcity. According to the authors, there is a common logic underlying all types of scarcity, ranging from time scarcity to obesity, occurring across cultures and societies. The result of feelings of scarcity is that it captures the mind. This causes a scarcity mindset, which affects all aspects of everyday life.

A study by Mani, Mullainathan, Shafir and Zhao (2013) illustrates one of the most remarkable consequences of experiencing (feelings of) scarcity. This study measured the cognitive performance of Indian sugarcane farmers during the annual planting cycle. It was found that the farmers performed significantly worse on cognitive function before harvest as compared to after harvest. The researchers explain this finding by indicating that before harvest the farmers are poor and scarcity captures their mind. Consequently, as the result of poverty in itself, cognitive function is impaired (Mani et al., 2013). To give an indication of the magnitude of the effect of scarcity, in this particular study the IQ of the farmers was diminished by about 10-13 IQ points, which is similar to the effect of missing one night of sleep (Mani et al., 2013).

Vohs (2013) argues that economists favor a theory that adheres to the principle that the more people on earth, the better, as more people imply more ideas to solve the challenges of our time. However, as noticed by Vohs (2013) this theory assumes that all people have a sufficient amount of mental capacity. This assumption is now heavily questioned by the findings of Mani et al. (2013). This new science of scarcity states that feelings of scarcity are not constrained to people living in absolute poverty. Every person can experience scarcity as even rich people can feel that they have less than they need. Hence, it is interesting to

investigate if the findings of Mani et al. (2013) translate to a different environment and culture. Therefore, this study aims to examine if Dutch students are susceptible to feelings of scarcity and if this potentially results in impaired cognitive function.

Similar to the approach of Mani et al. (2013) and Carvalho, Meier and Wang (2016) this study uses a harvest/payday cycle to study differences in cognitive function. The most applicable, cyclical, payments for Dutch students are the monthly payments they receive from the Dienst Uitvoering Onderwijs (hereafter: DUO)[1] and the healthcare allowance payments from the Tax Authority. In a recent study it was found that respectively 93% and 86% of all Dutch students receives a monthly payment from these institutions on the exact same date for everyone (Nibud, 2015). Additionally, these payments are for the majority of students the primary source of income. Therefore, the performance of Dutch higher educated students on a cognitive function test is measured shortly before and after these payments arrive.

The associated research question for this study is:
What is the effect of Dutch students having less monetary resources available (being in a poor state) on their cognitive function as compared to having more monetary resources available (being in a rich state)?

The contribution of this study is to introduce a new field of science to a new context and to a new subject population. To the extent of my knowledge it is the first study in The Netherlands that uses an experimental approach to investigate the scarcity theory as introduced by Mullainathan and Shafir (2013). Section 2 of this paper gives an overview of the current literature on scarcity and cognitive function and presents the hypotheses of this study. Subsequently, section 3 describes the experimental design and other methodological considerations. Thereafter, the results of the experiment are presented in section 4. Lastly, section 5 and 6 conclude this study by discussing its results, limitations and future research directions.

[1] The governmental body that is responsible for student financing.

2. Literature review

To my knowledge this study is the first in the field of scarcity in a Dutch (student population) environment. This section will carefully introduce and describe the concept of scarcity and its characteristics.

2.1 Definitions

2.1.1 Scarcity

In economics, scarcity is a central concept. "Economics is the science which studies human behavior as a relationship between given ends and scarce means which have alternative uses" (Robbins, 1932, p. 16). Scarcity is the discrepancy between human unlimited wants and world's limited resources (Heyne, Boettke & Prychitko, 2010). Economics primarily focuses on material scarcity and regards limited resources mainly as a physical constraint. Standard economic predictions implicitly assume that individuals treat all resources as limited in their decision-making behavior. However, according to Shah, Mullainathan and Shafir (2015) recent studies pose this to be empirically invalid. Individuals do not treat all resources similarly.

Scarcity can be more than physical scarcity. Mullainathan and Shafir (2013) propose a different perspective on the concept of scarcity, namely the feeling of scarcity. This view defines scarcity as "having less than you feel you need" (Mullainathan & Shafir, 2013, p. 4). Crucial in the explanation of this concept is the notion that the feeling of scarcity is purely subjective. Rich individuals can have the feeling that they have less than they want or need to the same extent as poor individuals. The moment individuals *feel* that resources are low compared to their needs a so-called scarcity mindset enters the stage (Shah et al., 2015; Shah, Mullainathan & Shafir, 2012). This mindset alters an individual's decision-making behavior.

Moreover, the feeling of scarcity is not limited to having less money than you feel you need. For instance, Mullainathan and Shafir (2013) argue that there is a resemblance between having too little money and having too little time. They even argue that obesity is a form of scarcity, just as being socially isolated is an example of social scarcity (Mullainathan & Shafir, 2013). In all examples, the underlying mechanism of scarcity is identical. This mechanism will be explained in-depth in section 2.2. As this study focuses on fluctuations in financial

resources, scarcity within the scope of this study refers specifically to the feeling of having less money than you feel you need. This can be classified as monetary scarcity.

2.1.2 Cognitive function and fluid intelligence

To start broadly, cognitive function or cognitive capacity refers to the psychological mechanisms that underlie our memory, perception, attention, language and intelligence (Fuster, 2003). Cognitive capacity together with executive function[2] are the two components of what Mullainathan and Shafir (2013) denote as bandwidth, the mental capacity that is available for use. Our cognitive capacity determines our ability to process and retain information, reason logically and solve problems. Cognitive function is basically at the cornerstone of every decision we make (Burks, Carpenter, Goette & Rustichini, 2009; Dohmen, Falk, Huffman & Sunde, 2010). The main variable of interest in this study is cognitive function and more specifically, fluid intelligence.

Cognitive function can be regarded as an umbrella term and fluid intelligence is one of its concepts. Fluid intelligence is defined as: "the capacity to think logically and solve problems in novel situations, independent of acquired knowledge" (Ferrer, O'Hare & Bunge, 2009, p. 46). It is regarded as being at the foundation of human cognition, in all phases of life, and particularly crucial for cognitive development in early life (Goswami, 1992). Fluid intelligence is associated with success in education and professional life, particularly in challenging and complex environments (Jaeggi, Buschkuehl, Jonides & Perrig, 2008). The current perspective on fluid intelligence is that working memory[3] is its most important determinant (Engle, Tuholski, Laughlin & Conway, 1999).

Although there is evidence for a long term effect of poverty on cognitive function (Lupien, King, Meaney & McEwen, 2001; Shonkoff & Phillips, 2000), this study specifically focuses on short-term fluctuations in cognitive capacity and fluid intelligence. Both concepts are often considered to be constant over time, however evidence shows that this might not be completely true. To give two very distinct examples, fluid intelligence can be increased by training (Jaeggi et al., 2008) and positive affect influences cognitive functioning (Ashby, Isen

[2] Researchers find it difficult to provide an all-encompassing definition of executive function, but it captures skills as self-regulation, response inhibition, planning and organization of behavior (Eslinger, 1996).
[3] Working memory "…refers to the relatively small amount of information that one can hold in mind, attend to, or, technically speaking, maintain in a rapidly accessible state, at one time" (Cowan, 2012, p. 1).

& Turken, 1999). This variability might be seen as striking and is in line with the findings of Mani et al. (2013). Hence, it is worth examining if transitions in financial resources impact students' fluid intelligence.

2.2 Scarcity mindset

The mechanism of scarcity starts by capturing the mind as an individual becomes absorbed when he experiences scarcity. The word "capture" is crucial in understanding the scarcity mindset. Scarcity grabs the attention of an individual and puts the most pressing needs at the top of its mind. The human mind "orients automatically, powerfully, toward unfulfilled needs" (Mullainathan & Shafir, 2013, p. 7). This implies that an individual does not deliberately chooses to think about a need. Scarcity grabs our attention, unconsciously, even if we wish it not to happen (Mullainathan & Shafir, 2013). The mind is now preoccupied. Subsequently, having less than we need does not make us only unhappy, it also changes how we think and how make decisions (Mullainathan & Shafir, 2013).

Experiencing a scarcity mindset has two important, but distinct, consequences. First, having less than you feel you need results in greater focus (Shah et al., 2012). This finding applies to all fields of life. For example, busy individuals who are facing time scarcity are more sensitive to pay greater focus on their tasks when confronted with deadlines (Karau & Kelly, 1992). In addition, it has been found that those who are hungry or thirsty focus more on food- and drink-related prompts (Aarts, Dijksterhuis & Vries, 2001; Radel & Clement-Guillotin, 2012). The participants in these studies that were part of the "hunger" or "thirst" treatment were more likely than participants in the control group to identify food- or drinks related words.

As illustrated by the examples above, a greater focus can be a positive consequence of scarcity. It namely results in a "focus dividend", the benefit of having less resources or time. Again, according to Mullainathan and Shafir (2013) the logic underlying scarcity is the same for money, time and social contact. Shah et al. (2012) argued that scarcity and the focus dividend should be present even when it is artificially created. In several experiments the researchers created scarcity in a laboratory setting. Participants assigned to the "poor" (less resources) condition performed relatively better than participants assigned to the "rich" condition (more resources) as they were automatically forced to focus. These differences in performance were attributed to the effect of scarcity.

Secondly, although the focus dividend can be regarded as a positive consequence, its upside is automatically also its downside. The strength of focusing is the ability to shut out, but the other side of the coin is that this causes people to tunnel at their most pressing needs (Mullainathan & Shafir, 2013). Tunneling is the cognitive equivalent of tunnel vision. When focusing more on a task individuals tunnel and hereby their peripheral vision is diminished (Williams, 1985). By focusing on the most pressing and salient things around, individuals automatically experience tunneling and thereby, neglect other matters. Inhibition is underlying both focus dividend and tunneling. When focusing, all that is outside the tunnel gets cognitively inhibited (Mullainathan & Shafir, 2013).

2.3 Scarcity trap

"Scarcity in one walk of life means we have less attention, less mind, in the rest of life" (Mullainathan & Shafir, 2013, p. 41). The amount of mind available is often called bandwidth, an overarching term that comprises everything from intelligence, working-memory capacity to executive control. Scarcity is believed to tax an individual's bandwidth and thereby, to inhibit our most fundamental capacities (Mullainathan & Shafir, 2013). It should be noted that at any point in time scarcity does not affect or reduce an individual's innate capacity, but only the amount of capacity that is available for instant usage. This implies that scarcity is not limited to physical implications (less time or money) as it also results in having less cognitive resources.

Cognitive function and cognitive skills have been linked to time and risk preferences and overall quality of decision-making. According to Burks et al. (2009) higher cognitive skills favor individual economic success by influencing the choices and preferences of individuals. Consequently, lower cognitive skills imply lower decision-making quality. Moreover, higher (lower) cognitive skills are associated with less (more) risk aversion and lower (higher) impatience (Benjamin, Brown & Shapiro, 2013; Burks et al., 2009; Dohmen et al., 2010). Higher risk aversion and impatience are used by researchers as explanations for why the poor are trapped in poverty (Haushofer & Fehr, 2014; Mosley & Verschoor, 2005).

Individuals that find themselves in such "scarce" environments make worse decisions and this further perpetuates bandwidth and scarcity. Thus, scarcity has its vicious circle, the scarcity trap (Mullainathan & Shafir, 2013). An individual in a scarcity trap is constantly one step behind and "juggling" his activities, moving from one pressing task to another. Only the

most salient task, the "ball" that is about to drop, receives attention. Therefore, (poor) individuals are fighting the most salient problems without solving the underlying causes of their scarcity. A lot of one's bandwidth is allocated to juggling the most pressing needs, leaving fewer resources available for other challenges in everyday life. Having less bandwidth in itself makes it difficult to escape the trap as fewer resources are at hand.

2.4 The effect of a harvest/payday

First of all, it is important to mention that paydays are always known in advance and hence can be fully anticipated upon (Carvalho et al., 2016). Payday (and the associated arrival of money) does not correspond with the reception of new information. Consequently, based on the rational expectations life-cycle or the Permanent Income Hypothesis, for instance consumption should not be influenced by payday (Stephens, 2006). However, Stephens (2003; 2006) found that the consumption of households is extremely sensitive to receiving money on payday. This finding was most pronounced for households for whom social security is the primary source of income. Moreover, it is found that households report to be more financially pressed before payday based on indicators such as the inability to pay rent or bills, eviction due to financial problems and/or being shut-off from telephone service (Caskey, 2010).

In the study of Mani et al. (2013) a harvest cycle instead of a payday cycle was used to measure differences in fluid intelligence and executive control, both of which are components of cognitive function. The sample population in their study consisted of Indian sugarcane farmers for whom the harvest was their primary (annual) source of income. Controlling for nutrition, work effort and stress, it was found that on average farmers performed worse on both fluid intelligence and executive control tasks before harvest compared to after harvest. The researchers explain this finding by stating that the farmers were preoccupied by worries about their monetary resources. This diminished their mental resources (Mani et al., 2013). Moreover, the farmers reported to be more financially pressed before relative to after harvest. This finding was based on some indicators closely related to their environment, such as the rate of pawning, the number of loans outstanding and difficulty in coping with ordinary bills (Mani et al., 2013).

Based on this, Carvalho et al. (2016), studied (among others) the cognitive function of low-income households in the US before and after their payday. In this context, no significant evidence for impaired functioning was found. However, the researchers did not limit themselves to only measuring cognitive function. In addition, a proxy for measuring perceived

scarcity was constructed. It was argued that individuals have to experience higher feelings of scarcity first in order for cognitive function to be affected. The researchers wanted to be able to disentangle this two-stage mechanism (Carvalho et al., 2016). Moreover, the response time on the cognitive tasks was measured. The before-payday group reported to be more preoccupied by scarcity than the after-payday group and response time was found to be higher for the former group. Despite these findings, fluctuations in income did not translate into a difference in cognitive function (Carvalho et al., 2016).

2.5 Hypotheses

Based on the above literature the following is hypothesized for this study.

Hypothesis 1A:

Subjects in the before treatment group will indicate higher feelings of scarcity than subjects in the after treatment group.

Hypothesis 1B:

Subjects in the before treatment group will score higher on a financial pressure measure than subjects in the after treatment group.

Hypothesis 1C:

Subjects in the before treatment group will have a higher response time on the fluid intelligence test than subjects in the after treatment group.

Hypothesis 2:

Subjects in the before treatment group will perform worse on the fluid intelligence test than subjects in the after treatment group.

Hypothesis 3:

The above hypothesized effects are more pronounced for the subgroup of subjects for whom the combined payments are the primary source of income compared to the group for whom it is not.

3. Methodology

This section elaborates on the experimental set-up of this study and describes how the variables and measures of interest were studied.

3.1 The treatment

The treatment of this study comprised of two exogenous shocks on the income of students: the monthly healthcare allowance payment and the monthly DUO loan payment/grant. Formally speaking, these payments cannot be regarded as income shocks as they are fully expected and anticipated and they do not permanently shift income (Carvalho et al., 2016). Students (can) know exactly when both transactions occur and the amount of money they will receive. Coincidentally, both payments are transferred to the students within a few days of each other, usually the 20th and 24th of the month. Hence, both payments were integrated into one externally imposed treatment.

The DUO payment is applicable as part of the treatment as 93% of Dutch students[4] receives this payment every month (Nibud, 2015). This is the only income payment that (almost) all persons in this subject population have in common. The average payment equals €469 which accounts, on average, for 61% of the total monthly income of a Dutch student (Nibud, 2015). Thus, this payment can be classified as the primary source of income for most students. Consequently, students are fairly dependent on it. Additionally, the healthcare allowance payment is received by 86% of the Dutch student population, which makes it the second highest payment that the subject population has in common (Nibud, 2015).

3.2 Experimental design

For the experiment a between-subject design was used, despite the advantages of a within-subject design (primarily more powerful statistical tests and a smaller required sample size) (Charness, Gneezy & Kuhn, 2012). However, its disadvantages are not outweighed and in this particular case a between subject design is more applicable.[5] Subjects can experience learning or training effects in a within-subject design. This is particularly important since it is

[4] Students in this study are both students from universities and universities of applied sciences.
[5] The main drawback of within-subject designs is known as the demand effect or as experimenter demand. Experimenter demand effects are defined as changes in behavior by subjects participating in an experiment due to their own interpretation about what constitutes appropriate behavior (Zizzo, 2010). Subjects either consciously or unconsciously try to guess what the experiment demands from them and change their behavior accordingly (Rosenthal, 1976; White, 1977).

known that performance on an IQ-test can be improved as a result of training (Jaeggi et al., 2008). Consequently, learning confounds the effect that is being measured. Both the treatment and a training effect are believed to work in the same direction (i.e. increase performance). This would rule out the possibility to disentangle both effects and thus, making it difficult to make causal claims. Therefore, this study made use of a between subject design, involving two distinct subject groups.

Due to practical considerations and constraints[6], it was not possible to ex ante randomly assign subjects to either the before or after treatment group (hereafter: before group and after group). This might be seen as a concern, as random assignment is crucial in order to obtain causal estimates (Charness et al., 2012). Comparing students who receive the treatment with students that do not receive the treatment can result in a selection bias. The students that choose to receive the treatment might potentially have a different baseline level of cognitive function or scarcity and thus, bias the results (Gertler, Martinez, Premand, Rawlings & Vermeersch, 2011). Hence, without randomization the effect of the treatment cannot be analyzed by simply comparing average absolute scores of the two groups, as selection bias is not solved. Moreover, comparing the before and after group does not control for confounds due to the passage of time.

To overcome the potential problems of no randomization and confounds this study used a method in which the between-subject design of the experiment is combined with a difference-in-differences (hereafter: DD) approach. Generally, the DD approach uses both before-and-after comparisons, and comparisons between persons who choose or choose not to take a treatment to establish a better estimate of the counterfactual from naturally occurring data. Thus, it combines two counterfeit counterfactuals (Gertler et al., 2011). It must be noted that an assumption has to be made that no major change, except the treatment itself, takes place in the timeframe between the before treatment measurement and after treatment measurement (observed and unobserved factors remain constant).

The design or method of this experiment should not be confused with the classic DD approach - it was merely inspired by this approach. Every subject of the before treatment group (t=0) was asked to complete five easy and five difficult items of the Raven Standard

[6] Upfront it was not known who and how many would participate in the experiment as no database of subjects was available. Moreover, upfront assignment would most likely have limited the sample size of the experiment as the response rate probability was judged to be low.

Progressive Matrices test before the treatment takes place.[7] The difference between both sections was calculated by subtracting the score on the difficult questions (D_0) from the score on the easy questions (E_0). Subsequently, every subject of the after treatment group (t=1) was asked to complete the test after the treatment has taken place and again the difference was calculated (E_1-D_1). Thereafter, the average difference score of the before treatment subjects can be compared with the average difference score of the after treatment subjects. Subjects in both groups had the exact same opportunity to experience a learning effect and for that reason this approach controls for selection bias.

The design of the experiment is visualized in Figure 1 below.

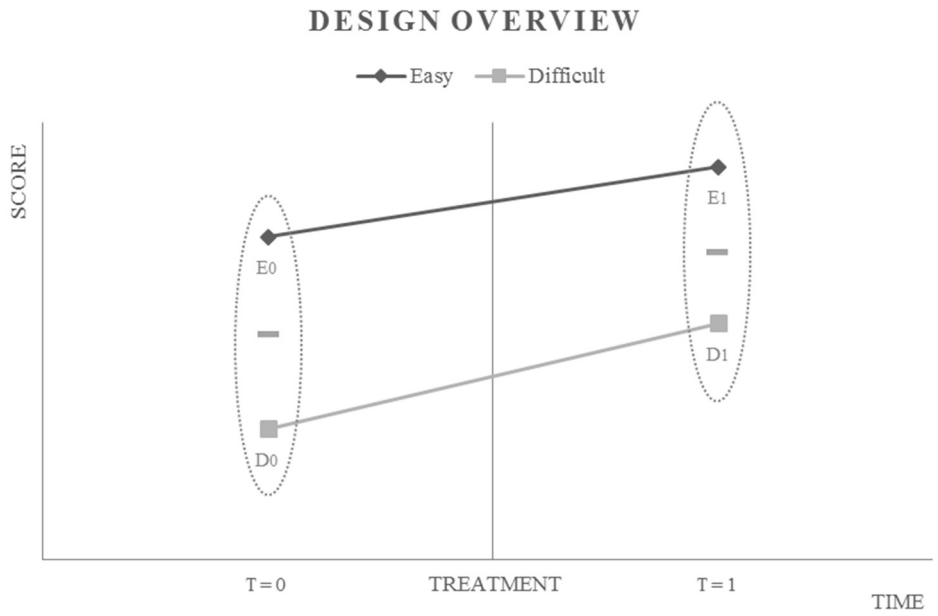


Figure 1. Experimental design overview

As mentioned earlier, the design of this experiment did not allow for comparison of absolute scores. This can be seen as a naturally imposed limit on the interpretation of the variable. Hence, the interpretation of the fluid intelligence variable changed inescapably. Instead of a direct measurement of fluid intelligence (cognitive function), the design only allowed for measuring learning on a fluid intelligence test. Thus, this study tested to what extent subjects' learning ability was impeded as a result of feelings of scarcity. It should be noted, and it can be seen in Table 1, that in this new interpretation learning ability is subjective as two individuals with the same difference score do not need to have the same underlying capability. Consequently, the interpretation of the variable is that it is independent of inherent capability.

[7] See section 3.3 for a description of this test

Table 1. Hypothetical scores of two distinct subjects

Subject	Score on easy	Score on difficult	Absolute score	Difference score
1	5	3	8	2
2	3	1	4	2

3.2.1 Incentives and precepts

Giving incentives to subjects matters as it gives control over the experiment. On the one hand subjects should be rewarded according to their performance and on the other hand subjects should be paid enough in order to exert mental effort (Smith, 1982). Generally, the more subjects are paid, the more effort they are willing to exert.[8] Consequently, their performance will be better leading to more optimal decisions. In this study more optimal decisions should have resulted in a higher absolute score that approached 10 out of 10 questions correct. This automatically implied that the difference score should move towards 0. The closer decisions are to the optimum, the lower the response variance will be, and the more powerful statistical tests are. This relationship is described in the following equations. As a result of a lower variance, the test statistic is higher.

$$V(\beta 1) = \frac{V(\varepsilon)}{N \cdot V(X)} \downarrow \longrightarrow t = \frac{\beta 1}{V(\beta 1)} \uparrow \quad (1)\ (2)$$

Additionally, incentives rule out intrinsic motivation. It is imaginable that in particular this study's higher-educated and/or intellectually curious subjects enjoyed to take up a challenge and to exercise some cognitive tasks. Intrinsic motivation can jeopardize the level of control of an experiment as subjects might have other motivations and objectives to participate. By providing incentives the motivation of subjects can be altered by ruling out intrinsic motivation.

Based on the above, there was a clear need to provide incentives to the subjects in this experiment. However, due to limited resources it was not possible to provide all subjects with both a fixed fee and task-related incentives. Therefore, this study made use of a Binary Lottery Incentive to incentivize the subjects in the experiment (Bardsley, 2010). This mechanism worked as follows: for every correct response on the fluid intelligence test subjects received a lottery ticket. Subsequently, the subjects participated with X lottery tickets in a lottery to win

[8] It should be noted that this effect is diminishing. At a given point, paying more will not result in a higher performance. It might even decrease performance (Ariely, Gneezy, Loewenstein & Mazar, 2009).

a pre-determined prize, which in this case was a voucher worth €50.[9,10] As anonymity was guaranteed to all subjects, subjects had to opt-in to take part in the lottery by filling out their email address and thereby willingly forgoing their anonymity.

The experimental design of this study (partially) complied with the five precepts as formulated by Smith (1982). Both nonsatiation and saliency can be considered to be satisfied as a result of applying the Binary Lottery Incentive. By satisfying these two precepts the study can be regarded as an economic experiment since subjects were being rewarded and not deceived (Smith, 1982). Moreover, the design of the experiment complied with the privacy precept. Subjects were only aware of their own results. Lastly, it was difficult to assess upfront if the dominance precept was satisfied or violated. The fluid intelligence test required a certain level of mental effort and it is questionable if the Binary Lottery Incentive was sufficient in off-setting the subjective costs of participating in the experiment. If the dominance precept was not violated, the experiment can be regarded as a controlled economic experiment (Smith, 1982).

The followings sections elaborate on how the variables of interest are measured.

3.3 Fluid intelligence (cognitive function) – RSPM test

Cognitive function (fluid intelligence) is often interchangeably used with cognitive ability and in this study it was measured by using the Raven Standard Progressive Matrices (hereafter: RSPM) test. The RSPM test is one of the most widely used IQ-tests. Matrices are regarded as an appropriate measure of fluid intelligence (Anastasi & Urbani, 1997). According to most researchers, the RSPM test measures general cognitive ability (Raven, 2000). However, there is no consensus among researchers as others argue that the test measures the concept "eductive ability" which has been introduced by Spearman (1927). In the context of the RSPM test, this implies one's ability to "construct meaning out of confusion" (Raven, 2008, p. 25). In addition, the RSPM test attempts to measure reproductive ability which has been defined as the ability to "absorb, recall, and reproduce information" (Raven, 2000, p. 2).

The original RSPM test consists of 60 items and as can be derived from the name of the test, the difficulty of the items increases. Every item consists of a 3x3 matrix with one of the

[9] X was the total amount of correct answers. X could take values ranging from 0 to 10.
[10] Bol.com voucher.

matrices missing. Respondents are asked to choose the correct response (out of either six or eight options) that completes the matrix (see Figure 2). Firstly, a subject determines what strategy to apply in order to solve the logic of an item. Subsequently the subjects need to memorize the strategy independently from the first sequence in order to apply it for another item (Georgiev, 2008). This study used a reduced form of the RSPM test (ten items[11]) that was also used in studies by Georgiev (2008) and Bilker, Hansen, Brensinger, Richard, Gur and Gur (2012). The classification of an item as either easy or difficult was done on the basis of historical test results (Georgiev, 2008).[12]

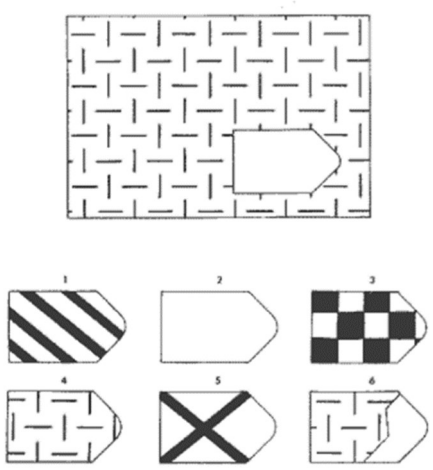


Figure 2. RSPM test example item

3.4 Scarcity measure

Related to the main variable perceived scarcity was measured in this study. This scarcity measure consisted of five questions[13] and was used as a proxy for an individual's level of (monetary) scarcity (Carvalho et al., 2016). As hypothesized in section 2.5 having less money impedes cognitive function because scarcity captures and preoccupies the mind. The scarcity measure was aimed at disentangling this two-stage mechanism. More explicitly, it was interesting to know if an increase in monetary resources would reduce the subjective feeling of scarcity or conversely, if a decrease in monetary resources would increase the scarcity feeling. By measuring this variable, it could be ruled out which part of the two-stage mechanism was or was not impacted, and an informed conclusion could be drawn. These questions had to be answered on a five-point Likert scale.

[11] The full RSPM test would inarguably have violated the dominance precept. See Appendix A for the ten items.
[12] Items with a probability of correctly answering of 0.25 or lower were classified as difficult. Items with a probability ranging from 0.25 to 0.50 were classified as easy. Hereby it was taken into account that the subjects in this study are all higher educated.
[13] See Appendix B1 for the questions.

3.5 Financial pressure measure

The third variable of interest in this study was financial pressure. It has been found that closer to payday both farmers (Mani et al., 2013) and American households (Caskey, 2010) report to have higher financial pressure. As it is questionable if students can be considered to be poor[14], it was interesting to know if and to what extent, students felt financial pressure. Both Mani et al. (2013) and Caskey (2010) measured this variable based on several indicators that specifically mattered for their subject populations. The indicators used in the two cited studies were adapted to make sense for the subject population of this study. Please refer to Appendix B2 for the questions.

3.6 Response time

As argued by Carvalho et al. (2016) if scarcity does impede cognitive function, it can be expected that subjects in the more financially strained group (the before group) would require more time to answer the questions that measure cognitive function as compared to the after group. Similar to the cognitive function variable the response time variable in this study was also measured using the DD approach. For every subject the response time was measured on all ten RSPM test questions. Subsequently, the response time on both the first five (easy) and the last five (difficult) questions was aggregated and subtracted from another. This procedure was repeated for all subjects which resulted in a new variable.

The previous sections all described the elements of the design that were necessary for conducting the experiment. The next section describes the actual data collection procedure.

3.7 Data collection

To gather data this study used an online survey platform.[15] This allowed for collecting data from a large sample while limiting the administration time for both the subjects and for data processing purposes. The subjects were approached during two time slots. The first period of data collection took place on the 18th and 19th of May, directly before (almost) all subjects would receive a healthcare allowance payment on the 20th. The second period of data collection started on the 27th of May and lasted till the 30th of May. This period started a few days after all subjects had received their monthly DUO-payment on the 24th. Moreover, this period

[14] Relative to the subject groups of previous studies one can argue that Dutch students cannot be considered poor. However, in the Dutch society students are relatively poor as most have low or no income and a student debt.
[15] Qualtrics.

commenced a week after the first period in order not to confound the indicators on financial pressure. The procedure and survey was the same for both groups.

During both periods Dutch students who were currently enrolled in a higher education program (at a university or university of applied sciences) were approached to participate in the survey. The subject group was deliberately limited to Dutch students as they all have the opportunity to automatically receive a grant or have the opportunity to apply for a loan.[16] The subjects were approached online and on campus. The communicational message used during the recruitment of subjects was as follows "I am looking for participants to take part in my survey. You will have the opportunity to win a voucher with a value of €50 and the administration time is approximately 10 minutes. Thank you in advance".

The topic of the survey was deliberately not revealed upfront and receiving a payment from the DUO was not a hard requirement for participation. Regarding the latter, it was assumed that the vast majority of the subjects did receive a payment (Nibud, 2015).[17] The introductory paragraph only stated that the survey would start with intelligence test questions and subsequently would proceed with some general questions. These measures were taken to ensure that preceding questions or information did not influence and/or steer subjects towards specific answers or thoughts on following questions (Foddy, 1994). Naturally, the introductory text also clearly explained the tasks of the survey and the incentive mechanism.

In order to make sure that all subjects would fully understand the RSPM test items, a practice question had to be answered correctly (Figure 2). Besides, it was explained that questions would have a time limit of 45 seconds.[18] After the RSPM items, subjects were asked to answer a few questions about their personal finances. These questions were followed by the proxies that measured perceived scarcity and financial pressure. Lastly, subjects needed to provide demographic information.

[16] Theoretically seen, most non-Dutch students are also able to apply for a loan. However, in practice this is not very common. Therefore, I targeted Dutch students to effectively gather as much observations as possible.
[17] If subjects, unexpectedly, did not receive the DUO-payment their responses were removed from the sample. Not revealing the topic of the survey was deemed of more importance.
[18] Based on a pilot this time limit was deemed appropriate - taking into account the differences in item difficulty.

4. Results

In this section the results of the experiment are presented. First descriptive statistics are provided and subsequently the effect of the treatment is analyzed.

4.1 Descriptive statistics

4.1.1 Demographics

In total the survey was completed by 145 subjects (68 before and 77 after treatment). After cleaning the data 111 subjects were left for analysis (51 before and 60 after). Non-students and students not receiving DUO payments were removed from the sample. This harsh cut-off was applied in order to ensure data quality. As mentioned in sections 3.1 and 3.7, based on statistics from Nibud (2015), it was expected that a small percentage of subjects needed to be dropped, but this percentage was higher than expected. The sample was skewed towards men as 74% of all subjects was male. However, as can be seen in Figure 3, this skewness was approximately similar for both the before group (78% male) as well as the after group (69% male). The groups were fairly similar along other dimensions as can be seen in Appendix C1.

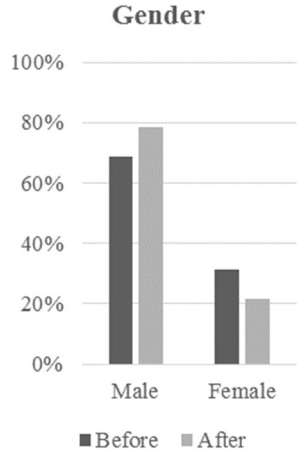


Figure 3. Gender distribution

4.1.2 Payments/Treatment

A further look into the characteristics of the subjects showed that the DUO payment accounts for, on average, 55.4% (52.4% vs. 58.0% for the before and after group respectively) of the monthly income of the students. Within the whole sample, this percentage ranged from 10% to 100% where the latter can be regarded as full dependency on these payments. With regard to the healthcare allowance payment 91.9% of the subjects indicated to monthly receive this payment. As expected the contribution of this payment to the monthly income was found to be lower than the DUO payment. On average, the health allowance payment accounts for

11.4% (before group 12.5% and after group 10.4%). The contribution rate ranged from 5% to 30%.

After merging the two payments into one treatment, consisting of both the DUO and healthcare allowance payments, it was the primary source of income for 71.1% of the subjects.[19] Hereafter this group of subjects will be classified as the "dependent group" as they mainly rely on the treatment's payments when it comes to their personal finances. This group consists of 35 subjects from the before treatment group and 44 subjects from the after treatment group. In the next sections some analyses will specifically focus on this subgroup as it has been hypothesized that the differences between the before and after group are more pronounced for this subgroup.

It should be noted that all the self-reported percentages have to be interpreted with caution as it is unlikely that subjects actually calculated the exact contribution of the payments in their monthly income. Adding up the contribution rate of both the payments revealed that for some subjects the sum was higher than 100% which is impossible. Therefore, the percentages can be regarded as a mere indication of the contribution. Moreover, the percentages might rather be a reflection of how dependent the subjects *feel* on the payments. This perfectly corresponds with the notion that this science of scarcity is based on pure subjectivity and perception.

4.2 Analyses

This experiment has 111 independent observations at the individual level and 1 independent observation at the session level. Experimental data often violates the assumptions of parametric tests as the underlying distribution of a population is generally not known.[20] Generally, in order to compare two samples when having experimental data non-parametric tests are applied.[21] For non-parametric tests far less assumptions are needed. Therefore, these types of tests are applied in the next section.

[19] For those subjects for which both contribution rates added up to more than 100% the rate has been normalized at 100%. This is based on the assumption that the subjects made an attempt to indicate that they are fully dependent on the two payments, however unable to calculate/unaware of the exact division between the two accounts.

[20] See Appendix D for assumptions and tests of violations of parametric tests.

[21] The most important assumption for these tests is that observations in an experiment are independent. Advantageous of non-parametric tests is that they do not require large sample size and variables do not have to be measured in an interval scale only.

4.2.1 Fluid intelligence (cognitive function)

Instead of the absolute scores, the difference scores[22] of the RSPM test[23] were analyzed. As can be seen in Figure 4, not a single subject performed better on the difficult questions since the difference scores range from 0 to 5. The correct interpretation of this variable is that it measured the ability to learn on the individual, subjective, level. It should be noted that scores closer to 0 are regarded as closer to the optimum[24] and thus, associated with higher cognitive function. At the group level (see Figure 5), the average score on the difference variable was 2.745 for the before group and 2.750 for the after group. When only considering the dependent subgroup, the score for the before group is 2.686 versus 2.568 for the after group.

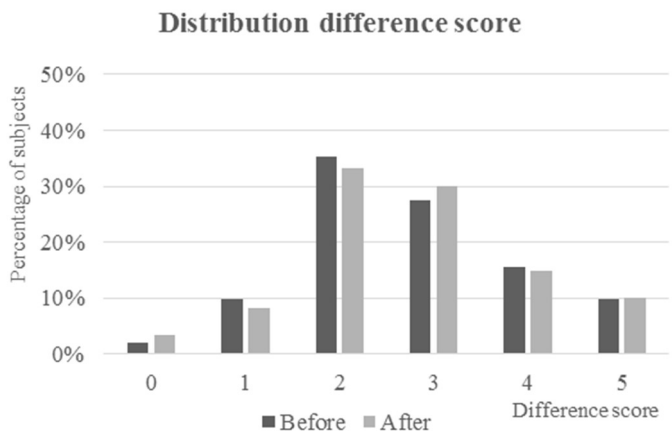


Figure 4. Distribution difference score

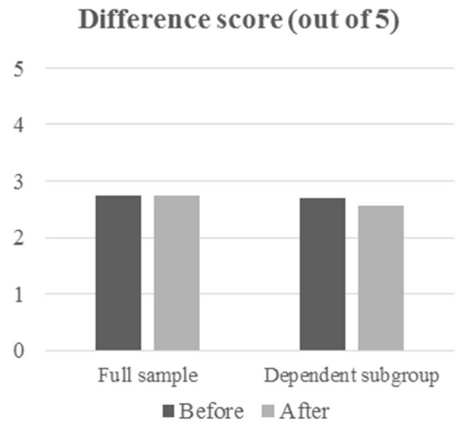


Figure 5. Mean difference score variable

Result 1: No evidence for lower fluid intelligence of the before treatment group compared to the after treatment group is found. Lower monetary resources (before) do not lead to statistically significantly less optimal scores on the difference variable. For the dependent subgroup the score is closer to the optimum and there is more dispersion, although again the difference is not statistically significant. Hence, the findings are not in line with hypothesis 2.

[22] Total of correct answers on the easy questions minus total correct answers on the difficult questions.
[23] Descriptive statistics of the RSPM test can be found in Appendix C2.
[24] The optimum is to have an absolute score of 10 out of 10. This translates to five correct answers on the difficult questions and five correct answers on the easy questions. Consequently, the corresponding optimal difference score is 5-5=0.

Support for Result 1: To analyze this variable, the chi-squared test and the Fisher's exact test are used.[25] The chi-squared test is applicable for data of one sample if the sample population has two or more classes. The test analyses if expected outcomes in a category significantly differ from the observed outcomes in a category. Baseline values for expected outcomes are determined by either a theory or the researcher. By setting the outcomes of one sample as the expected outcomes, this test can also be used to analyze if there is a significant difference between two samples. Following this approach, the chi-squared test is used to analyze if the before and after group differ on the difference variable.

As can be seen and concluded based on the p-values in Table 2A and 2B for both the full sample and the dependent subgroup, the before and after treatment group do not significantly differ. The null hypothesis that there is no difference in the frequency distributions between the two samples cannot be rejected. Thus, it cannot be concluded that those subjects with lower financial resources scored lower on fluid intelligence than their less constrained counterparts.

Table 2A. Chi-squared full sample

Difference score	Before	After	Total
0	1	2	3
1	5	5	10
2	18	20	38
3	14	18	32
4	8	9	17
5	5	6	11
Total	51	60	111

Pearson chi2(5) = 0.3610, p = 0.996

Table 2B. Chi-squared test dependent subgroup

Difference score	Before	After	Total
0	1	2	3
1	4	3	7
2	12	18	30
3	10	13	23
4	4	5	9
5	4	3	7
Total	35	44	79

Pearson chi2(5) = 1.3132, p = 0.934

Alternatively, an additional analysis can be performed by using the Fisher's exact test. The Fisher's exact test is a statistical two sample test suitable for studying if differences between two groups exist. It is used when both samples consist of two classes (2 * 2). The test analyses if the observed outcomes differ from random distribution over the two classes. To be able to use the Fisher's exact test a new variable needed to be constructed as the difference

[25] In fact, in case of a between subject design a Mann-Whitney U test is generally the most applicable to compare two groups. However, in this case this choice is complicated by the data and the statistical technique of the Mann-Whitney U test. The test is based on ranked outcomes and is poor in handling too many 'ties'. A tie implies that two or more independent observations (subjects) have the same score (value) and consequently have the same rank. As in this study the difference variable can only take six values (0, 1, 2, 3, 4 or 5) many ties exist. This is a major drawback and hence, it is not advisable to use this statistical test.

variable had more than two classes. By classifying the subjects as either poor learners or good learners, two classes were constructed. It could be argued that subjects with a score of 0 on the difference variable were good learners whereas those with a score of 5 could be regarded as poor learners.

As the variable scores ranged from 0 to 5, the value 2.5 was the neutral point and therefore it was used as a cut-off point. This yielded the same result as picking the median value which was 3. Those with a difference score of 0, 1 or 2 were classified as good learners whereas subjects with a score of 3, 4 or 5 were classified as poor learners. Following the main hypothesis that feelings of scarcity impede cognitive function, it was hypothesized that the after group should have a relatively higher proportion of good learners as compared to the before group. As can be seen in Table 3A this was not the case for the full sample. Proportionally, there were more improvers in the before group than in the after group. However, this difference in distribution is not statistically significant as the p-value for the Fisher's exact test was 0.49. The null hypothesis that the two samples are evenly distributed over the two classes could not be rejected.

Table 3A. Fisher's exact total sample

Class	Before	After	Total
Poor learners	24	27	51
Good learners	27	33	60
Total	51	60	111

One-sided Fisher's exact: p = 0.49

Table 3B. Fisher's exact dependent subgroup

Class	Before	After	Total
Poor learners	17	23	40
Good learners	18	21	39
Total	35	44	79

One-sided Fisher's exact: p = 0.46

With regard to the more dependent subjects, the proportions were as hypothesized (see Table 3B). However, using the Fisher's exact test this difference is again not statistically significant as the p-value is 0.46. In addition, the chi-squared test was used to analyze this newly constructed variable. However, no significant results were found ($\chi^2 = 0.11$, $p = 0.74$).

Lastly, an OLS regression was used to analyze the effect of the treatment on fluid intelligence. By regressing a dummy variable indicating after treatment on the difference score variable the effect of the treatment could be estimated. The intercept of model I in table 4 equaled the value of the before treatment group. The effect of the treatment equaled the coefficient of the dummy variable. The value of the coefficient was, as expected, exactly the same as the difference between the after treatment score on the difference variable and the

before treatment score (Table 4). The sign of the coefficient was positive which is not in line with the hypothesis 2. However, this variable was again found not to be statistically significant ($t = 0.02$, $p = 0.98$).

In model II and III the simple regression was expanded by adding variables and interaction terms. Regarding model II, the variables *primary_source* and *scarcity_proxy* had an unexpected sign. Based on the hypotheses both coefficients were expected to be positive. Only the variables *female* and *primary_source* are statistically significant. However, with respect to this study the regression of model III was of more importance as the variables were interacted with the dummy variable that indicated after treatment. Again the signs of the coefficients of the interaction terms containing the variables *primary_source* and *scarcity_proxy* were negative. In any case, none of the variables in this regression is statistically significant at any level.

Table 4. Output OLS regression. The dependent variable is *difference score* (*N*=111). The variables *After, Female* and *Primary_source* are dummy variables.

Independent variable	β	Test-statistic	p-value
I:			
(Constant)	2.745098	16.23	0.000
After	.004902	0.02	0.983
II:			
(Constant)	3.277306	7.64	0.000
After	.0691188	0.31	0.757
Female	.7874818	3.10	0.002***
Primary_source	-.4664445	-1.93	0.057**
Scarcity_proxy	-.2124931	-1.27	0.208
III:			
(Constant)	2.672156	4.40	0.000
After	1.181527	1.46	0.148
Female	.5509585	1.57	0.120
Primary_source	-.1875646	-0.53	0.594
Scarcity_proxy	.0133344	0.06	0.955
After * Female	.5039055	0.99	0.324
After * Primary_source	-.435161	-1.29	0.200
After * Scarcity_proxy	-.4786866	-0.98	0.328

Note: * is significant at 0.10, ** is significant at 0.05 and *** is significant at 0.01

4.2.2 Response time

As argued by Carvalho et al. (2016) if scarcity does impede cognitive function, it can be expected that the more financially worried subjects (the before group) needed more time to complete the RSPM test items as compared to the after treatment group. To test this hypothesis, the response time of each subject per item was measured. The overall average response time for all ten questions was 278 seconds (27.76 seconds per item, on average). The fastest subject completed the ten items in 131 seconds and the slowest in 376 seconds. Being fast or slow was not necessarily related to higher absolute scores since the correlation between the two variables is 0.10.[26] Nevertheless, the slowest subject happened to be the subject with the highest score.

***Result 2*:** In absolute terms response time was found to be higher for the before group compared to the after group, both for the full sample as well as for the more dependent subgroup. However, by constructing a new variable it was found that response time does not significantly differ between the before and after group. This does not provide evidence to support the hypothesis that there is a difference in response time as a result of the treatment.

***Support for Result 2*:** Descriptively comparing the groups shows that the before group was slightly slower in answering the items: 279 seconds versus 276 seconds.[27] Regarding the dependent subgroup the absolute response time was further apart (281 seconds vs. 271 seconds). However, similar to the absolute score variable, the absolute response time variable could not be analyzed by comparing the two groups. Therefore, a new variable was constructed that equaled a subject's response time on the easy items subtracted by the response time on the difficult items. Using both a Mann-Whitney U test[28] ($N_1 = 51$, $N_2 = 60$, $p = 0.26$) and a chi-squared test ($\chi^2 = 73.76$, $p = 0.65$) it was found that the two groups do not significantly differ. The same holds for the more dependent subgroup again using the Mann-Whitney U test ($N_1 = 35$, $N_2 = 44$, $p = 0.23$) and a chi-squared test ($\chi^2 = 68.87$, $p = 0.45$).

4.2.3 Scarcity proxy

The scarcity proxy consisted of five questions that had to be answered on a scale from 1 to 5 where the former can be interpreted as not at all preoccupied by thoughts about personal

[26] A correlation of 0.10 should be interpreted as a very weak correlation (Evans, 1996).
[27] See Appendix C2 for histograms on the response time.
[28] A Mann-Whitney U test can be used in this case as there are no ties for the values of the variable. Please refer to Appendix E for the output tables.

financial circumstances and the latter as very preoccupied by these circumstances. The scores per question, both aggregated and subdivided for the two groups, can be found in Appendix F. As can be seen in Figure 6A, and as was hypothesized, the scarcity proxy was higher for the before group (2.16) than for the after group (2.00). The scores for the more dependent subjects (2.10 vs. 2.04) were almost similar to the scores of the full sample as can be seen in Figure 6B. It should be noted that both scores do not necessarily have to indicate a subjective feeling of scarcity as both were fairly low. The scores correspond with "rarely" when it comes to how often subjects indicated to have felt preoccupied, concerned and/or worried.

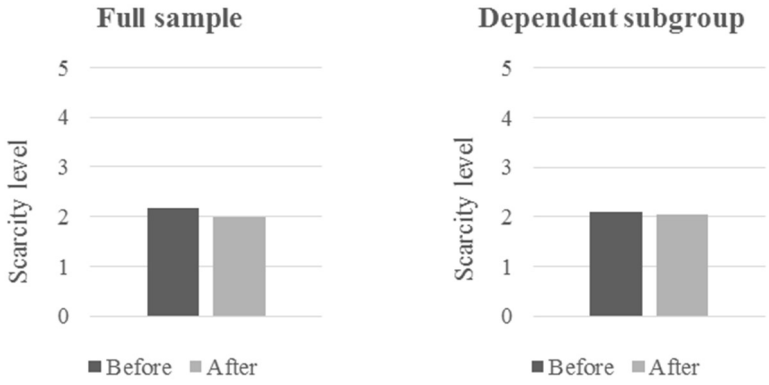


Figure 6A and 6B. Average level of perceived scarcity per group

Result 3: Overall the level of perceived scarcity was fairly low. Nevertheless, this level was higher for the before treatment group, as hypothesized. However, this difference is not (statistically) significant.

Support for Result 3: Using a Mann-Whitney U test, by taking subjects' individual scarcity proxy score as an independent observation, it was found that the before and after group are not statistically different. As the scarcity variable only takes fourteen distinct values, it was questionable how the Mann-Whitney U test dealt with the resulting ties in rank. Therefore, a chi-squared test was used to verify this result. Again, it was found that the difference is not significant ($\chi^2 = 15.08$, $p = 0.30$). Analyzing the dependent subgroups resulted in the same conclusion (Mann-Whitney U test: $N_1 = 51$, $N_2 = 60$, $p = 0.94$; chi-squared: $\chi^2 = 11.94$, $p = 0.53$). In any case, the causal claim that the treatment reduced feelings of scarcity could not have been made since no random assignment to one of the groups was in place.

4.2.4 Financial pressure proxy

This proxy attempted to elicit how financially pressed subjects felt before and after receiving the treatment. In order to measure this three questions are constructed that have to be answered by either yes (1) or no (0). The sum of these three questions is the score for the proxy.

Result 4: Subjects in this study did not feel financially pressed, neither before nor after the treatment.

Support for Result 4: Out of all subjects, 93 subjects (83.7%) answered all three questions with 'no', indicating not to feel financially pressed. This implies that only eighteen subjects indicated to have felt, to a certain extent, pressure. These subjects were almost equally divided between the before and after group (eight vs. ten). As a result of this low number of observations it has no meaning to draw any conclusion about how the two groups differ, other than cautiously concluding that in general the subjects did not feel financially pressed, neither before nor after the treatment. Using a chi-squared test, it was confirmed that there is no significant difference between both groups ($\chi^2 = 0.13$, $p = 0.99$). With regard to the dependent subgroup the same holds ($\chi^2 = 0.54$, $p = 0.91$).

4.3 Power

All results in this study are based on the non-rejection of the null hypotheses. This requires performing power analyses as the risk of a type-II error (false negative) is present. As can be seen in the calculation in Appendix G the power of the tests was extremely low.[29] Hence, this study is heavily underpowered. This implies that the risk of a type-II error is high. The difference between the two samples on the difference score variable was so small that in order for it to be statistically significant, an infeasible amount of observations[30] would have been needed per subject group (based on a significance level of 0.05 and a power level of 0.80). This required sample size is basically larger than the population size of this studies' subjects. Moreover, it should be noted that by infinitely increasing sample size *N* every treatment effect will eventually become significant.

[29] For the power analyses the tool G*Power is used (Faul, Erdfelder, Lang & Buchner, 2007).
[30] Please refer to Appendix G.

5. Discussion

5.1 General discussion

The findings of this experimental study lack evidence for a significant difference between the cognitive functioning of Dutch students in a "poor" or "rich" state. Subjects in the after treatment group are not significantly closer to the optimum for the difference score variable. This finding also applies for the subgroup of subjects whose income is more heavily dependent on the treatment. Fluctuations in monetary resources do not impact the subject population's mental capacity. Likewise, the response time on the fluid intelligence test did not statistically differ between the before and after treatment group. The subjects that were hypothesized to be more preoccupied did not need more time to finish the RSPM test items.

The above might be explained by the fact that the before treatment group was not actually more preoccupied than the after treatment group. In order to measure this a proxy for perception of scarcity was developed. In general, all subjects reported a fairly low level of scarcity. On average, subjects indicated to have been at most occasionally concerned on the day of surveying and this day was indicated to be non-different relative to other days. Although the level of scarcity was slightly higher for the before group, this difference is found to be not statistically significant. Furthermore, the feelings of financial pressure of the student subject population were even lower. Subjects showed neither before nor after the treatment any level of pressure resulting from their personal financial circumstances. Combining these two findings it can be concluded that students' level of scarcity and financial pressure was not affected by fluctuations in their income. It can be that the fluctuations in income were not strong enough to affect perceived scarcity and financial pressure.

It is questionable if the results of this study must be interpreted as evidence for the nonexistence of an impairment of cognitive function as a result of lower monetary resources. As described and tested by Carvalho et al. (2016) individuals should first be captured by higher feelings of scarcity before this phenomenon can impact cognitive function. In other words, if an individual does have less (monetary) resources, but this does not result in higher feelings of scarcity, he or she will subsequently not display lower cognitive performance. Carvalho et al. (2016) found that financially strained individuals reported higher feelings of scarcity before payday as compared to after payday, however this did not translate into lower cognitive

capacity which questioned the scarcity-cognitive function relationship as proposed by Mani et al. (2013).

The findings of this study do not allow for a similar claim. As lower monetary resources did not translate into higher levels of scarcity, it is not completely surprising that no effect on cognitive function was found. Hence, this study does provide evidence for the fact that lower monetary resources not universally have to result in higher levels of scarcity and/or lower cognitive performance. However, the differences between this study and previous studies should not be overlooked. Firstly, Dutch students have rather little in common with US low-income households (Carvalho et al., 2016) and/or Indian farmers (Mani et al., 2013). Moreover, differences in methodological approach might be underlying the distinct findings. Mani et al. (2013) made use of a within-subject design whereas Carvalho et al. (2016) measured cognitive function by the means of an other test than the RSPM test.

Differences in context, circumstances and environment might be of crucial importance when conducting research in this new field of scarcity science. With regard to Mani et al. (2013) and Carvalho et al. (2016), Lichand and Mani (2016) shed some new light on the inconsistency of the results between these studies. According to them the factor risk, instead of anticipated changes in economic circumstances, might be able to explain that the former did find an impairment of cognitive function whereas the latter did not. US households know exactly when and how much income to expect as opposed to the yield of the harvest in case of the Indian farmers. Moreover, the magnitude of a yearly harvest income to Indian farmers is very distinct from a monthly income to (low-income) US households, let alone a monthly income for Dutch students.

In this study paydays were certain and consequently changes in monetary resources perfectly predictable for the students. It is not bold to claim that no risk about the treatment was involved as it is highly unlikely for governmental organizations to default on their obligation. This in combination with intuitively knowing that the majority of students most likely have a safety net (i.e. parents) to fall back on might explain the stability in perceived scarcity and feelings of financial pressure (as far as existent at all among students). Therefore, no effect of having less than you feel you need was found. However, it needs to be stressed that this study does not rule out that scarcity impairs cognitive function. It only provides evidence

to believe that for this particular population fluctuations in monetary resources do not result in an increased scarcity level and an impaired cognitive function.

5.2 Limitations

The main limitation of this study is that the subjects could not be randomly assigned to either the before treatment or after treatment group. Due to a lack of resources and uncertainty with regard to the amount of participants (no database to be used) random assignment was complicated. As a result, the difference-in-differences (DD) approach had to be applied, which inescapably altered the interpretation of the cognitive function variable. Random assignment would have allowed for the comparison of absolute scores whereas the current design only allowed for comparing the difference scores. This reduced the interpretation from measuring fluid intelligence to measuring learning on a fluid intelligence test. Although this is still in accordance with the concepts of fluid intelligence and cognitive function, it is admittedly not similar to what was measured in previous studies.

A second limitation is the lack of statistical power in this study. The consequence of an underpowered study is the risk of failing to reject the null hypothesis when it should be rejected, known as a Type II error. This is a major concern for this study as none of the null hypotheses of the statistical tests was rejected. The differences between the two groups might have been statistically significant for one or more variables, however due to the lack of power this could not be detected. For future research this must be taken into account for the design of experiments. Increasing the power of a study can be done by increasing the sample size, reducing the error variance or increasing the treatment level variance. A within subject design might be recommended as that reduces the variance for all that is unobserved, however this inevitably requires to deal with a learning effect and possibly experimenter demand.

Furthermore, in order to limit administration time for the subjects it was chosen to use timed questions. This might have led to adverse effects as subjects could have felt more pressed. However, the experimental design controlled for this since the condition was similar for both groups. As mentioned in section 3.2.1 it was upfront questioned if the dominance precept as described by Smith (1982) would hold. Most likely this precept was violated as qualitatively several subjects indicated to have struggled with the difficulty of the questions. This implies that a considerable amount of mental effort had to be exerted. The experiment was made incentive compatible by applying the Binary Lottery Incentive, however the probability

of winning the lottery was small and subjects might have estimated that. In the same line of thought it is questionable if intrinsic motivation was ruled out by the incentive mechanism.

Examining a new field of research is highly interesting, but it also implies that there is no abundance of knowledge to build on. With regard to perceived scarcity and financial pressure no universally accepted measures existed. As a result, the proxies used in this study were based on limited empirical evidence, which might have biased effectively measuring these constructs. Moreover, this study was only conducted in one month which does not control for a seasonal or calendar effect. Ideally, the effect of the treatment is measured in a longitudinal study to rule out that particular months influence the baseline values for the variables. Additionally, detailed information about the personal finances of the subjects was missing. Although knowing that the treatment applied to all subjects, it cannot be ruled out that some subjects were surveyed as being in a poor state, while in fact they were in their rich state. Therefore, for future research in this field it is advisable to have more insight in individual's financial situations to individually tailor the experimental design to subjects, as has been done by Carvalho et al. (2016) and Spears (2011).

6. Conclusion

This study introduced a new science of scarcity to a (Dutch) student context and subject population. In previous studies it was found that fluctuations in income can affect cognitive function depending on the context and environment. Implications of these findings are that for example the existing view on poverty must be altered. Poverty does not always have to be the result of bad decisions and behavior, but can merely be the result of a scarcity mindset that captures each and every person that lives in such circumstances. This leads to a, difficult to escape from, vicious circle. However, this study does not confirm these previous findings. It was found that perceived scarcity did not vary with fluctuations in income and as a result cognitive function was not impacted.

Based on the findings of this study it can be concluded that (Dutch) students are not captured by feelings of scarcity or feelings of financial pressure. Moreover, it can be concluded that changes in monetary resources do not impact cognitive function for this particular subject population. This does not imply that for every population in and across countries cognitive

function will not be impaired as a result of having less. The science of scarcity as proposed by Mullainathan & Shafir (2013) needs further research to accept or reject its propositions. First, better understanding of the concepts is required. Future research has the challenging but highly interesting and important task to examine what factors determine (monetary) scarcity to influence cognitive function.

Researchers should for example investigate the role of economic situation, risk or uncertainty with regard to future payments, frequency of payments and culture on the relationship between fluctuations in monetary resources, scarcity and cognitive function. In order to effectively conduct research in this immature field of science standardized measures need to be developed. Measuring scarcity, mental bandwidth and financial pressure using universally accepted constructs allows for comparison between different settings and studies. Once a stronger case is built for the existence of this new field, research should focus on how public authorities and organizations can better take into account fluctuations of individuals' income, and thus cognitive capacity, in their policies.

7. Reference list

Aarts, H., Dijksterhuis, A., & Vries, P. (2001). On the psychology of drinking: Being thirsty and perceptually ready. *British Journal of Psychology*, *92*(4), 631-642.

Anastasi A, Urbina S (1997). *Psychological Testing* (Prentice–Hall, Englewood Cliffs, NJ), 7th Ed.

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies, 76*(2), 451-469.

Ashby, F. G., & Isen, A. M. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological review*, *106*(3), 529.

Bardsley, N. (2010). *Experimental economics: Rethinking the rules*. Princeton University Press.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, *11*(6), 1231-1255

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's Standard Progressive Matrices Test. *Assessment*, 1073191112446655.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, *106*(19), 7745-7750.

Carvalho, L. S., Meier, S., & Wang, S. W. (2016). Poverty and economic decision-making: Evidence from changes in financial resources at payday. *The American Economic Review*, *106*(2), 260-284.

Caskey, J. (2010). Payday lending: New research and the big question.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*(1), 1-8.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, Lawrence Erlbaum Associates Inc.

Cowan, N. (2012). Working memory capacity. Psychology press.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, *100*(3), 1238-1260.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309.

Eslinger, P. J. (1996). Conceptualizing, describing, and measuring components of executive function: A summary.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods, 39(2), 175-191.

Ferrer, E., O'Hare, E. D., & Bunge, S. A. (2009). Fluid reasoning and the developing brain. *Frontiers in neuroscience*, *3*, 3.

Foddy, W. (1994). *Constructing questions for interviews and questionnaires: Theory and practice in social research*. Cambridge university press.

Fuster, J. M. (2003). *Cortex and mind: Unifying cognition*. Oxford university press.

Georgiev, N. (2008). Item analysis of C, D and E Series from Raven's Standard Progressive Matrices with item response theory two-parameter logistic model. *Europe's Journal of Psychology*, *4*(3).

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2011). *Impact evaluation in practice*. World Bank Publications.

Goswami, U. (1992). *Analogical reasoning in children*. Psychology Press.

Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. Science, 344(6186), 862-867.

Heyne, P. T., Boettke, P. J., & Prychitko, D. L. (2010). *The economic way of thinking*. Pearson Education International.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*(19), 6829-6833.

Karau, S. J., & Kelly, J. R. (1992). The effects of time scarcity and time abundance on group performance quality and interaction process. *Journal of experimental social psychology*, *28*(6), 542-571.

Lichand, G., & Mani, A. Cognitive Droughts.

List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, *14*(4), 439-457.

Lupien, S. J., King, S., Meaney, M. J., & McEwen, B. S. (2001). Can poverty get under your skin? Basal cortisol levels and cognitive function in children from low and high socioeconomic status. *Development and psychopathology*, *13*(03), 653-676.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. science, 341(6149), 976-980.

Mosley, P., & Verschoor, A. (2005). Risk attitudes and the 'vicious circle of poverty'. *The European journal of development research*, *17*(1), 59-88.

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. Macmillan.

Nibud (2015). Studentenonderzoek 2015 (Achtergrondstudie bij Handreiking Student & Financiën).

Radel, R., & Clément-Guillotin, C. (2012). Evidence of motivational influences in early visual perception hunger modulates conscious access. *Psychological science*, *23*(3), 232-234.

Raven, J. (2000). The Raven's progressive matrices: change and stability over culture and time. *Cognitive psychology*, *41*(1), 1-48.

Raven, J. (2008). The Raven progressive matrices tests: their theoretical basis and measurement model. *Uses and abuses of Intelligence. Studies advancing Spearman and Raven's quest for non-arbitrary metrics*, 17-68.

Robbins, L. (1932). The nature and significance of economic science. *The Philosophy of Economics: An Anthology*, *1*, 73-99.

Rosenthal, R. (1976). Experimenter effects in behavioral research. (Enlarged ed).

Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. *Science*, *338*(6107), 682-685.

Shah, A. K., Shafir, E., & Mullainathan, S. (2015). Scarcity frames value. *Psychological Science*, 0956797614563958.

Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academies Press

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review*, *72*(5), 923-955.

Spearman, C. (1927). The abilities of man.

Spears, D. (2011). Economic Decision-Making in Poverty Depletes Behavioral Control. *The BE Journal of Economic Analysis & Policy*, *11*(1), 1-44.

Stephens, M. (2003). " 3rd of tha month": Do Social Security recipients smooth consumption between checks?. *The American Economic Review*, *93*(1), 406-422.

Stephens, M. (2006). Paycheque receipt and the timing of consumption. *The Economic Journal*, *116*(513), 680-701.

Vohs, K. D. (2013). The poor's poor mental power. *Science*, *341*(6149), 969-970.

White, R. A. (1977). The Influence of experimenter motivation, attitudes, and methods of handling subjects on Psi test results. *Handbook of parapsychology*, 273-301.

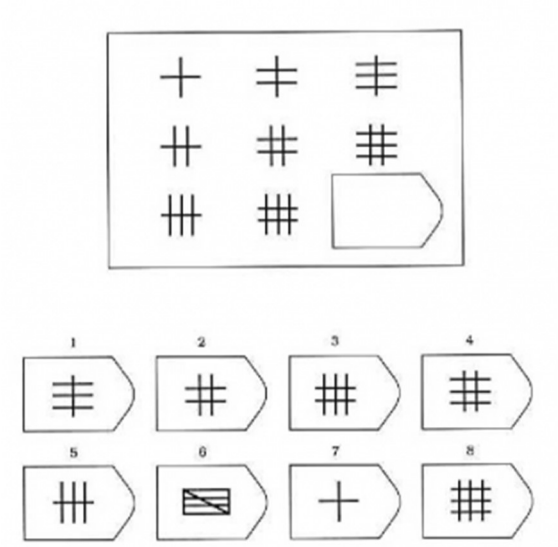
Williams, L. J. (1985). Tunnel vision induced by a foveal load manipulation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *27*(2), 221-227.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75-98.

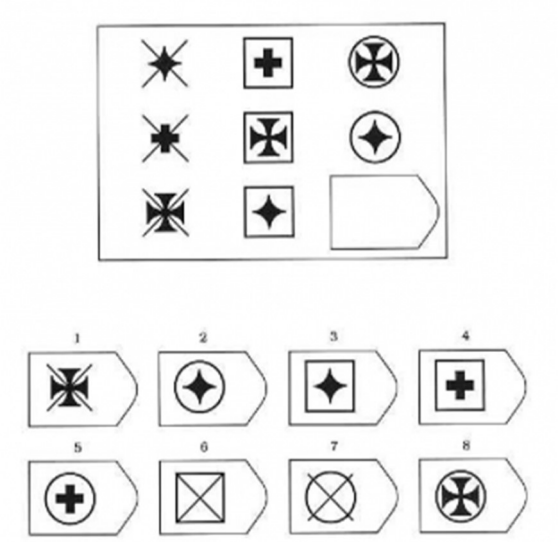
Appendix

Appendix A. RSPM test items

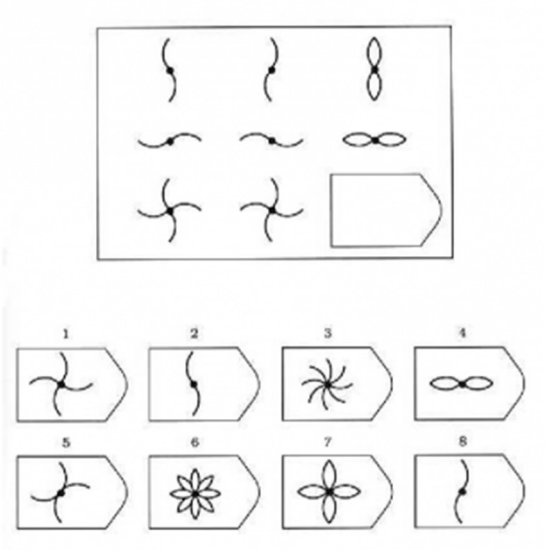
Q1 - Easy
Correct answer: 8



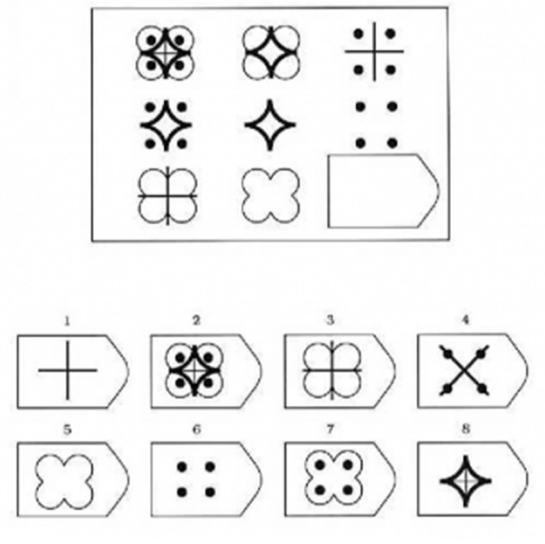
Q2 – Easy
Correct answer: 5



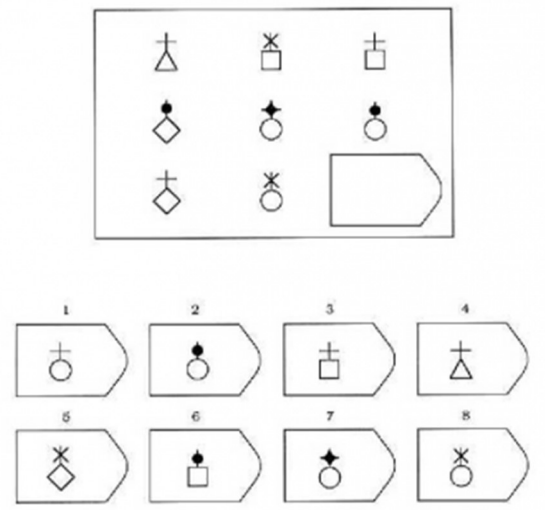
Q3 - Easy
Correct answer: 7



Q4 – Easy
Correct answer: 1

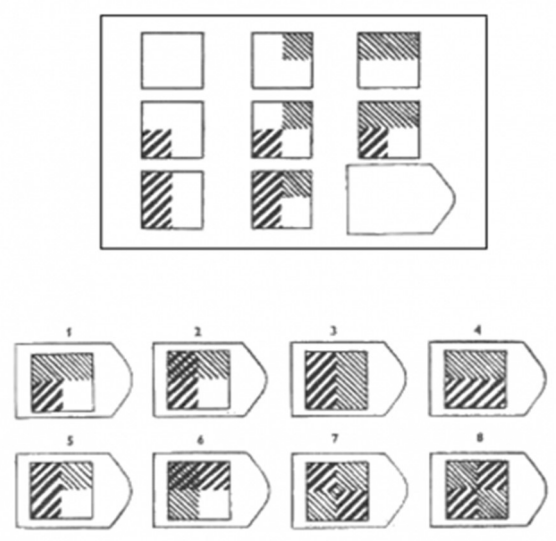


Q5 - Easy
Correct answer: 1



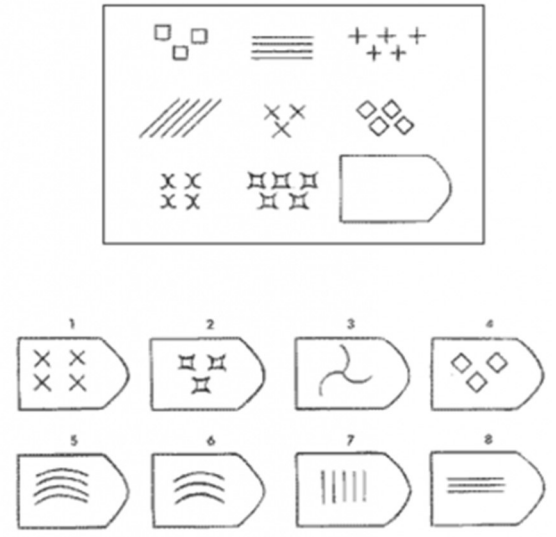
Q6 – Difficult

Correct answer: 2



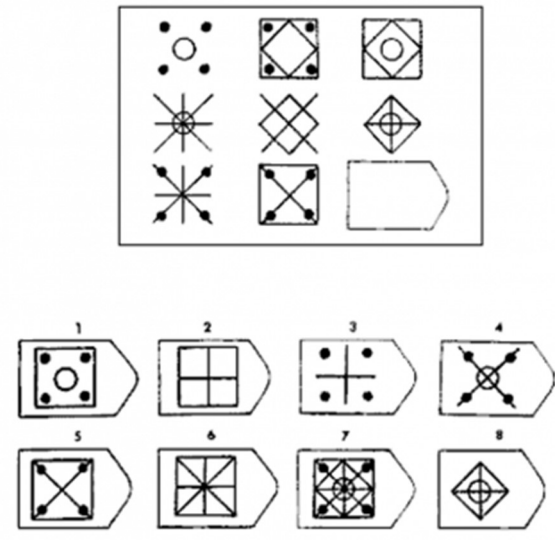
Q7 – Difficult

Correct answer: 6



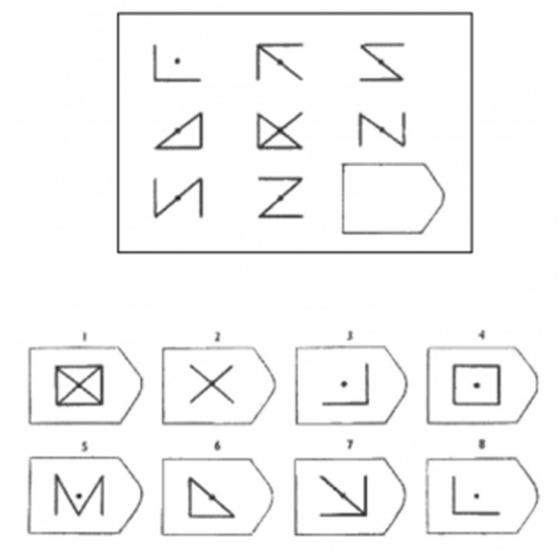
Q8 – Difficult

Correct answer: 2



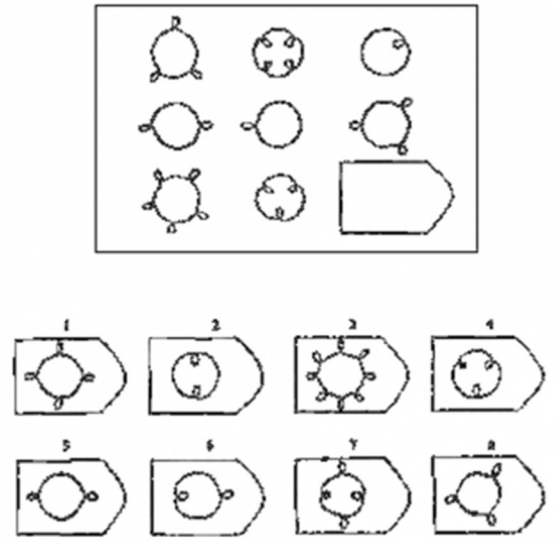
Q9 - Difficult

Correct answer: 4



Q10 – Difficult

Correct answer: 5



Appendix B1. Scarcity proxy measure

"In the last 24 hours, how often:

1) …were you troubled about coping with ordinary bills?

2) …did you worry about having enough money to make ends meet?

3) …did you think about future expenses, some of which may be unexpected?'

4) …were you preoccupied with thoughts about your personal finances?"

These questions had to be answered on a five-point Likert scale ranging from "never" to "very often".

5) “I am interested in understanding if people’s concerns about having enough money to make ends meet change over the month. Relative to other days this month, how concerned were you in the last 24 hours about having less money than you need to make ends meet?”

This question had to be answered on a five-point Likert scale ranging from “much less concerned” to “much more concerned”.

Appendix B2. Financial pressure measure

1) Did you in any kind of way borrow money within the last week (except from DUO)?
2) Did you let someone pay something for you because you were unable to pay yourself within the last week?
3) Did you request to postpone any type of payment within the last week?

All three indicators had to be answered on a dichotomous scale (either yes or no).

Appendix C1. Demographics continued

The average age of the full sample was 22.3 years old. The two groups were similar on this characteristic as both have the exact same average age. The sample consisted of both university of applied sciences (HBO) and university students[31]. The before group appeared to have more HBO students (23%) as compared to the after group (13%). This difference mainly coincided with a lower number of Bachelor students in the before group as compared to the after group. Master students accounted for approximately the same percentage in both groups (47% vs. 48%).

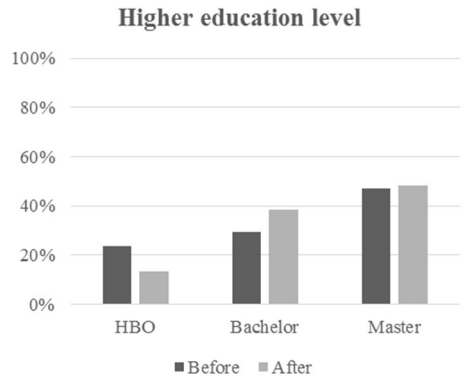


Figure 7. Education level

[31] Both can be regarded as higher educated students. For university students a distinction can be made between Bachelor and Master students.

Appendix C2. Descriptive statistics RSPM test

On average, all subjects from the two groups combined, the subjects answered 4.98 items (out of 10) correctly. The lowest individual score was 2 and the highest score 10 (see Figure 8A for the distribution of scores). Zooming in on the individual questions the (apparent) easiest question was answered correctly by 92.7% and the most difficult question was answered correctly by 8.1%. From the data (Figure 8B) it can be seen that the scores on the questions classified as easy (Q1-5) were better as compared to the difficult questions (Q6-10). As subjects were not randomly assigned to either the before or after treatment group, it was not possible to statistically compare the average score of the before group with the score of the after group. However, descriptively, the scores of the two groups were almost identical (4.980 for the before group versus 4.983 for the after group).

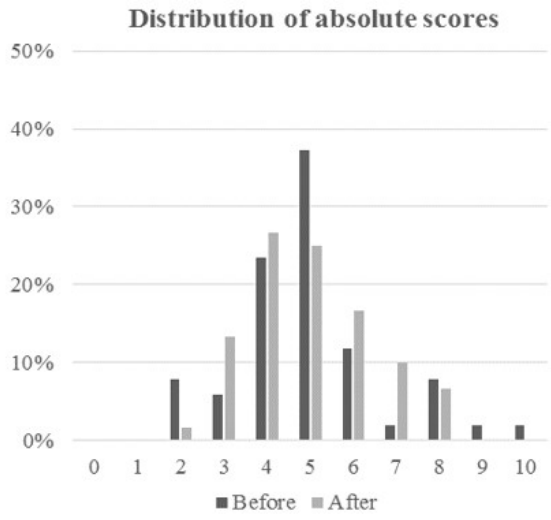


Figure 8A: Absolute scores RSPM test

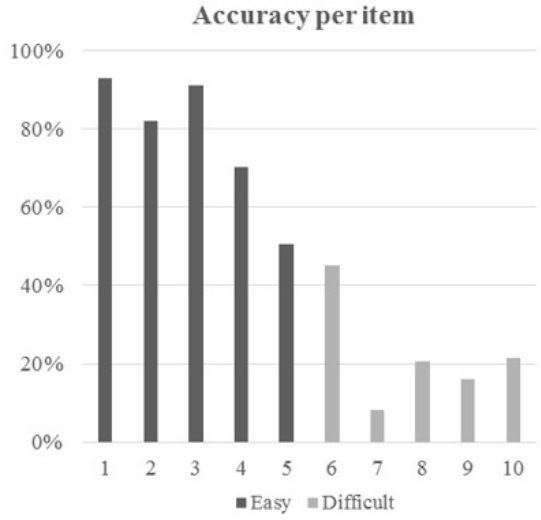


Figure 8B: Accuracy per RSPM

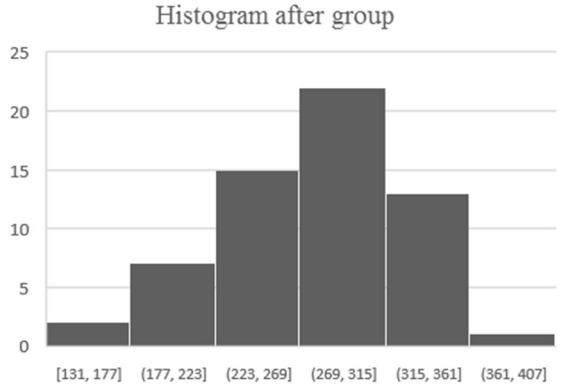


Figure 9A: Histogram response time before group

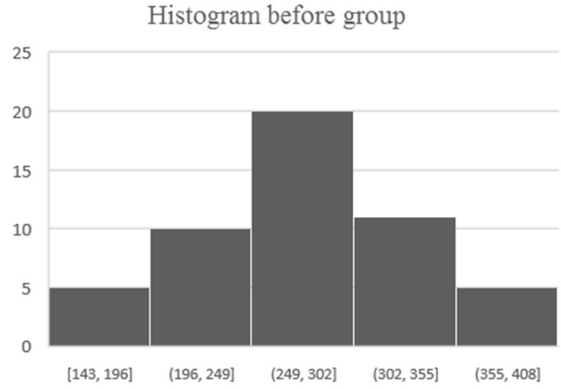


Figure 9B: Histogram response time after group

Appendix D1. Assumptions parametric tests

1. The observations are independent. The value/selection of one observation must not influence the value/selection of another observation.
2. The observations must be drawn from a normally distributed population
3. In case two groups are analyzed, they must have the same variance.
4. Variables must be measured in an 'interval scale', in order to interpret results.

Appendix D2. Normality of distribution

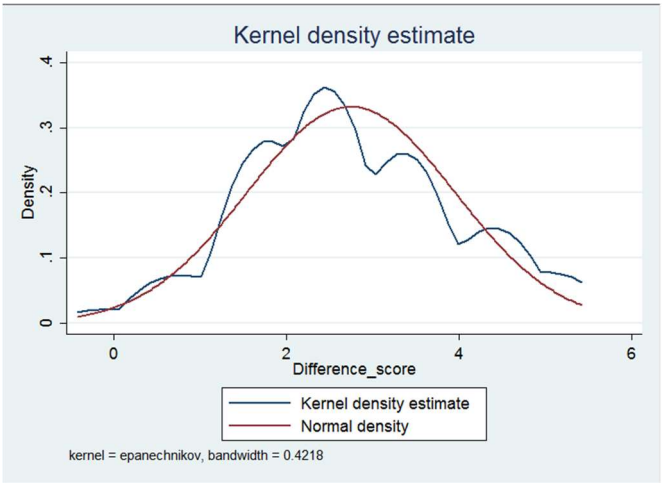


Figure 10: k-density graph difference score variable

Based on the k-density graph there were no major deviations from normality. The Shapiro-Wilk test for normality confirmed this as the null hypothesis of a normal distribution could not be rejected (p = 0.85). Thus, the assumption of normality has not been violated.

Table 5: Test of normality

SW test statistic	N	Z	p-value
0.99297	111	-1.018	0.8456

Appendix D3. Homogeneity of variances

Table 6: Test of homogeneity of variances

Levene statistic	Df1	Df2	p-value
0.00003415	1	109	0.9953

Levene's test is a test that is used to determine if *k* samples have the same variance. It tests the assumption of homogeneity of variances. The null hypothesis is that the variances of samples

are equal. Hence, based on Levene's test the assumption of homogeneous variances could not be rejected (p = 0.995).

Appendix E. Output Mann-Whitney U test - Response time

Table 7A: Mann-Whitney U test – dependent variable: Response time – full sample

Ranks				
	After	N	Mean Rank	Sum of Ranks
Response time	0	51	2664	2856
	1	60	3552	3360
	Total	111		

Test Statistics	
Mann-Whitney U	28554.24
Z	-1.136
p-value	0.2556

Table 7B: Mann-Whitney U test – dependent variable: Response time – dependent subgroup

Ranks				
	After	N	Mean Rank	Sum of Ranks
Response time	0	35	1277.5	1400
	1	44	1882.5	1760
	Total	79		

Test Statistics	
Mann-Whitney U	10266.67
Z	-1.75
p-value	0.2266

Appendix F. Scarcity measure scores

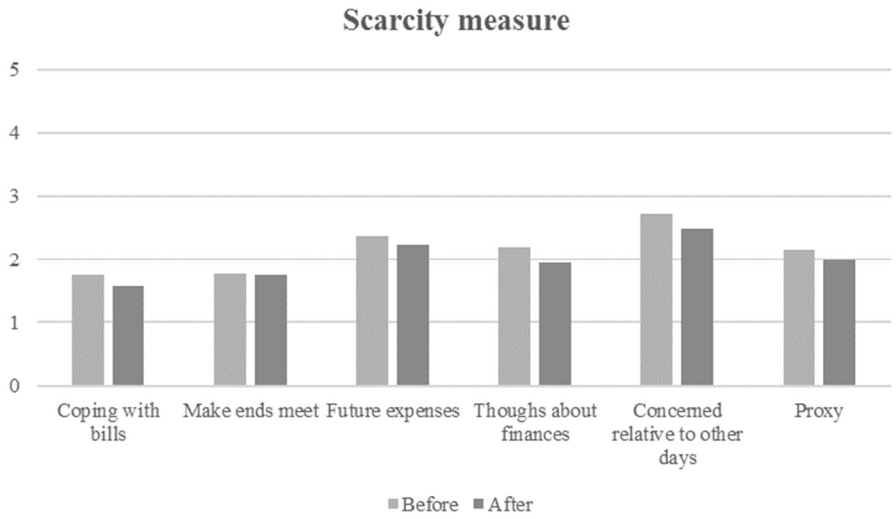


Figure 11. Scarcity measure scores per construct and aggregated

Appendix G - Power calculation

The power calculations below are based on the assumption (rule of thumb) that a student t-test is used to analyze the variables. This was actually not the case in this study and hence, the calculations are approximations.

Post-hoc power analysis

For the power analysis the tool G*power is used. Input parameters for the post-hoc power analysis are the effect size index, the significance level α and sample sizes of the two groups. The effect size index *d* can be calculated as follows:

$$d = \frac{\mu 1 - \mu 2}{\sigma} = \frac{2.745 - .750}{1.20} = \text{-}0.004 \qquad (3)$$

This results in the G*power output as shown in Figure 12. The statistical tests had almost no power to detect a difference between the two groups (0.05).

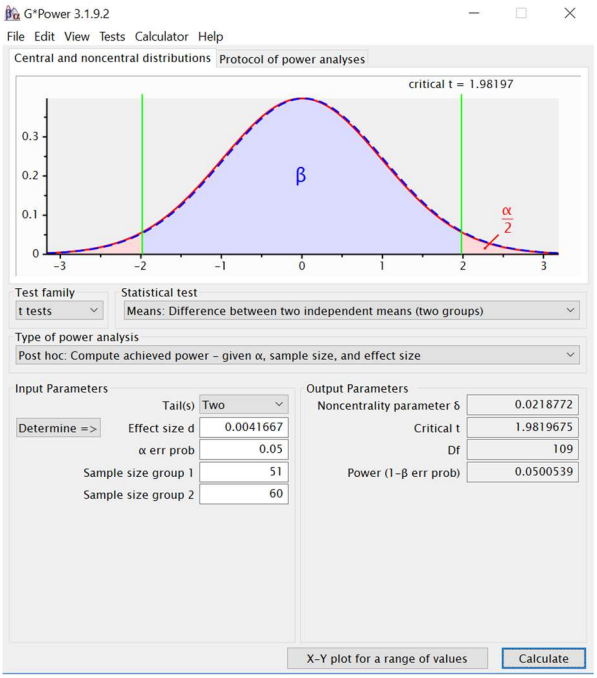


Figure 12. G*Power post-hoc power analysis

A priori power analysis

Using an a priori power analysis it can be determined what the optimal sample size of an experiment is. It can also be used to determine what the sample size of a study should have been to detect a statistical difference. The standard values for an a priori power analysis are a significance level of $\alpha = 0.05$ and a power level of $1\text{-}\beta = 0.80$. Again using G*power and the above calculated effect size index the optimal sample size is enormously high (Figure 13).

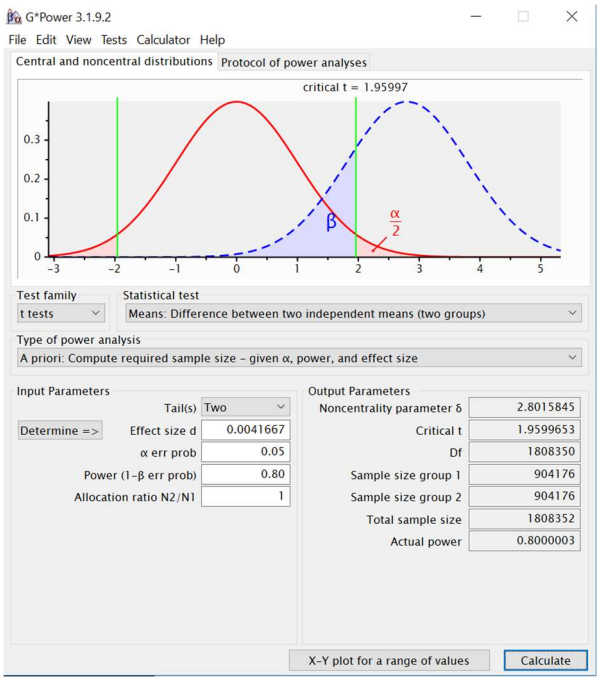


Figure 13. A priori power calculation