ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Appetite, rationality, and honesty?

A classroom experiment to measure the impact of appetite on rationality and dishonesty

April 30 - 2020

Master thesis of:
Wouter ten Brink
351004
Economics & Business:
Behavioural Economics specialisation

Supervisor:
Aurelien Baillon

Abstract

Classic economic models generally depend on the assumption that a decision maker is purely rational, and therefore should decide what is in their own best interest. However, for choices in real life, decisions often do not follow what rational models would predict. In this study, a classroom experiment was conducted to explore whether it is possible to find a relation between a stimulus designed to increase the level of appetite and the likelihood for respondents to lie in a setup where dishonesty would increase their expected reward. Over four sessions, the participants were assigned to either a treatment or control group, where the treatment group had an alternative set of instructions, providing them with the prospect of food as a reward as opposed to the control group, who also had the prospect of a reward but for whom the reward was not specified. To incentivise the commitment of the participants, a reward was randomly awarded to two respondents after the sessions based on the participant's answers. The results of this experiment did not show any significant effect supporting the hypothesis.

Contents

List of Tables

List of figures

1 Introduction

Behavioural economics has often been lauded for its capability to predict real behaviour where standard economic theory falls short. Essentially it tries to model the behaviour of people rather than the behaviour of purely rational decisionmakers, "homo economicus" or, as Richard Thaler would put it, Econs (Thaler, 2015). However, Tversky and Kahneman introduced heuristics and biases that caused systematic inconsistencies in how economic theory would expect people to behave (Tversky & Kahneman, 1974). The two-system theory, as published by Kahneman (2011), further explores these inconsistencies and, in a way, provides an explanation as to why standard economic theory can so often be proven to be a poor predictor of real behaviour, even though it always tries to predict what real behaviour should be like. Essentially, people make decisions using their somewhat instinctive sub-conscious decision-making systems, but they also make some decisions using a more rational-like approach.

This paper aims to explore whether a different physiological state, increased appetite, could also be a predictor for different behavior. Specifically, when it comes to decisions that have a social or moral context, such as decisions involving dishonesty. One could argue that fully rational, Spock-like, Econs would not care about the social or moral implications of lying, if it would mean that their own well-being improves because of it. Mazar, Amir & Ariely (2008) argue however, that acting in a way that opposes one's social constructs and self-image in itself can be a negative internal incentive. Essentially, if acting dishonest makes you feel bad, are you really maximising your own well-being by ignoring this factor?

A classroom experiment was done to explore whether a seemingly irrelevant factor, level of appetite, affected either the likelihood to be dishonest or the level of rationality. Triggering appetite or hunger could act as an exogenous shock in the level of rationality and the degree of honesty. The level of rationality was measured with an adaptation of Afriat's method for determining the level of economic efficiency. The tendency to lie was estimated by allowing the subjects to hand in self-stated test scores which could not be linked on an individual level with their actual test results. Finally, two different frames were used in the experiment set-up in an attempt to enhance the level of appetite for half of the subjects. The statistics did not show any significant relation between the explored factors.

2 Literature review

2.1 Dishonesty

Cressey's fraud triangle

Cressey (1986; 1953) investigated the motivations of embezzlers leading up to an act of fraud. By interviewing a sample of convicted fraudsters in both state and federal US prisons, he found that there were commonalities in the explanations of their motivations leading up to fraud. Cressey describes that the enabling factors for committing fraud can be categorised into three common elements, or risk factors:

> *"1) The feeling that a personal financial problem is unsharable.*
>
> *2) The knowledge of how to solve the problem in secret, by violating a position of financial trust.*
>
> *3) The ability to find a formula which describes the act of embezzling in words which do not conflict with the image of oneself as a trusted person."*
>
> (Cressey, 1986, p. 199)

The first factor is described as a financial obligation that the violator feels is not socially sanctionable. It could be that the violator has acquired debts due to spending on products or services which are illegal or otherwise socially unacceptable. In which case the violator would not want his social environment to know that he has these financial obligations.

The second factor implies that the violator is in a position of responsibility that can be abused to solve own their financial problem, which can be done in such a way that other involved parties will not find out. In other words, they have the opportunity to solve their problem by violating their position of trust without other people finding out.

The third factor explains that people generally do not feel that their acts have implications on their perception of their own morality. In the interviews, the violators generally seemed to feel that, although their actions were not socially acceptable, they did not imply that they were untrustworthy or immoral in general. The violators were able to verbalize their actions in a way that it did not have to conflict with their overall self-image of being a good person.

Cressey's three-factor model is one of the more prevalent theories describing fraud and is commonly referred to as Cressey's fraud triangle. The idea is, that someone only engages in an act of fraud if all the three factors are present (Kasem & Higson, 2012).

Although Cressey originally defined pressure as a non-shareable financial issue, Kassem and Higson (2012) propose that pressure or motivation to commit fraud might not derive financial pressure exclusively. Rather, sources of pressure could either be financial or non-financial and could be classified as either personal pressure, professional or career related pressure, or external pressure. Whereas the opportunity aspect of the fraud triangle would imply the perceived opportunity of the potential violator to commit fraud without being caught. The final part of the fraud triangle is also commonly described as rationalization, which involves the process of explaining the action in a way that does not have implications on one's moral self-image (Kasem & Higson, 2012).

Dishonesty and its incentives

While Cressey's research views fraud as an isolated problem with specific motivations. Mazar, Amir and Ariely (2008) approach the concept of fraud as a general issue of dishonesty and theorise that the rewards and punishments that people experience when engaging in dishonest acts are similar for most cases of dishonesty. There is generally a reward to be gained, which can be material, social or career related, which is weighed against the prospect of punishment if a dishonest decisionmaker gets exposed. Besides these external incentives, people also weigh how much an action follows their own concept of fairness, going against their internal moral standards could result in experiencing a sense of discomfort. According to this theory, a decision is made based on the balance of the positive and negative incentives, similar to how a rational individual would weigh utility values when evaluating a decision under uncertainty (Mazar, Amir, & Ariely, 2008).

The impact of social considerations also becomes evident in altruism experiments, such as dictator games, which show that people have a sense of fairness towards others when it comes to distributing rewards (Kahneman, Knetsch, & Thaler, 1986). Subsequently, people are willing to even lower their own payoff if it will increase the payoff of others in comparison. Moreover, people appear to generally have an intrinsic preference to not deceive others, even if deception would increase their own material payoff (Gneezy, 2005; Hurkens & Kartik, 2009). In fact, as Erat & Gneezy (2012) found, people even prefer not to

deceive in a scenario where deception would result in an increased financial reward for all participants. The intrinsic punishment mechanism of dishonesty can be identified through experiments involving neuroimaging methods, which have shown that complying with social standards in itself triggers a reward mechanism, which would imply that the lack of social standard compliance also means that someone pays an opportunity cost of not triggering this reward mechanism (Quervain, et al., 2004; Rilling, et al., 2002).

2.2 Rationality

Predictive value of rational models

Capturing incentives as predictive variables is generally a difficult task. Theoretically, if it is assumed that people behave based on purely rational self-oriented motivations, behaviour could be predicted trough expected utility theory. For example, in case of tax fraud, people would consider fraud purely based on the financial aspect, which is thoughtfully weighed against the risk of being caught and a factor for non-social behavior (Allingham & Sandmo, 1972).

However, it is often found that people behave irrationally, even in situations where they are fully informed and prospective payoffs are predictable and quantifiable. People tend to assign subjectively weighted values to risk in their decision-making (Tversky & Kahneman, 1974; 1992; Quiggin, 1982).

Moreover, people can change their preferences based on irrational biases and heuristics which results in them making suboptimal decisions. People's preferences can change based on the way a problem is presented to them, or even by being presented with information which should be irrelevant for their decision-making process (Tversky & Kahneman, 1974).

In classic economic theory, a decisionmaker's motivations have commonly been described based on the idea that he is fully rational and acts in his own best interest (Neumann & Morgenstern, 1944). However, numerous examples can be found where decisionmakers make choices that do not seem to be in their own best interest. Economic theories that dismiss the assumption that decisionmakers are purely rational utility maximisers have been finding more support.

Although rationality is a well-known concept, the definition that is used in academic publications can vary. Rationality is often defined as the manner in which decisionmakers make decisions that are aligned with their own preferences. However, Meglino and Korsgaard (2004) distinguish two types of rationality:

1) *Instrumental rationality*, which reflects behaviour that is in line with the optimal result when purely looking at the decisionmaker's own rewards.

2) *Epistemic rationality*, which reflects the ability to find the optimal outcome when external consequences are also considered.

They propose that people can be differentiated in how much they value the well-being of others involved in their decisions compared to their own. More self-interested people will generally assign more weight to the value of their own rewards compared to the value of the rewards of others. While people that are more other-oriented assign relatively more weight to the well-being of others. Coincidentally, people that have a lower level of other orientation also tend to show a higher level of instrumental rationality when it comes to decision making.

Irrational heuristics and biases cause decisionmakers to behave in ways other than what would be expected from someone who behaves rationally (Tversky & Kahneman, 1974). The heuristics and biases however are situational and do not necessarily apply to all situations. Kahneman (2011) provides an explanation for this: when faced with a question, people can either use an intuitive way of thinking, or they can use a more computational, rational-like approach. The intuitive thinking process is described as system 1 and the computational approach as system 2. While system 2 generally causes choices to be approached more deliberately, it also requires more cognitive effort. System 1 is able to simplify problems through associations that come from prior experiences in order to reduce cognitive effort and speed up the decision-making process. Decisions involving a system 1 thinking process are expected to be more prone to heuristics and biases, as situations can be incorrectly recognised to be similar to a previously experienced example, or some information can be incorrectly assumed to be relevant. Decisions that are made using system 2 are generally more rational, and more likely to be in line with what normative rational models, such as utility functions, could predict. Kahneman's theory of system 1 and system 2 helps to explain

why utility functions do not always actually predict actual behaviour, but also how examples can be found of situations where they do.

Revealed preferences

The theory of revealed preference suggests that the choices that a decisionmaker makes are the best indicator for their actual preference (Samuelson, 1938). Using Afriat's cyclical consistency theorem (Afriat, 1967), Varian (1982) proved that data for which the General Axiom of Revealed Preference (GARP)[1] holds, can be modelled as a well-behaved utility maximising function. More specifically, if A and B represent bundles of that goods a utility maximising agent can choose between, and both are affordable, it can be said that his choice can be plotted as a well-behaving utility function if the following statement holds true:

If choice A is made with the existence of the alternative B, it must mean that this consumer does not strictly directly prefer option B over option A.[2]

Although Afriat's theorem, strictly speaking, only allows for behaviour to be classified into two categories: (1) behaviour that satisfies GARP and (2) behaviour that does not satisfy GARP, some methods have been introduced in which theorem can be "partially" true, which

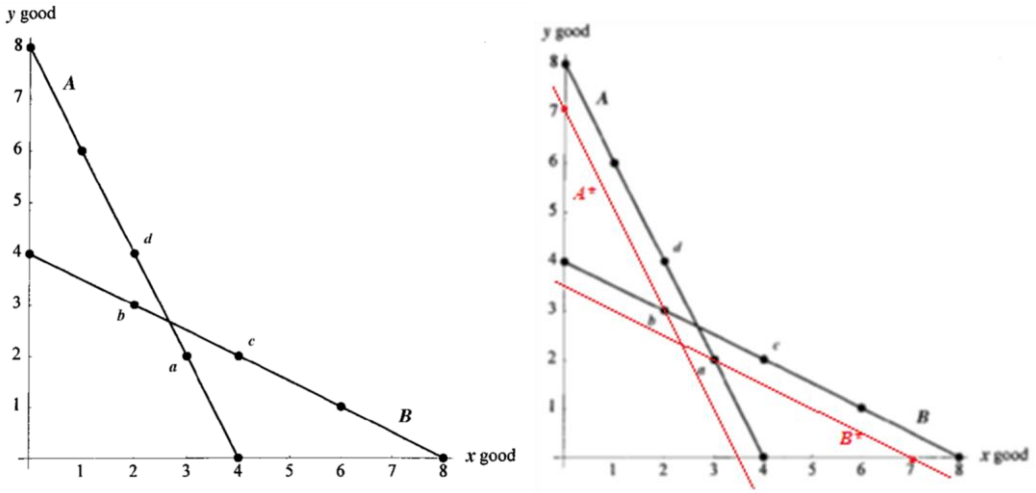


Figure 1: Illustration of Afriat's Efficiency Index (Harbaugh, Krause, & Berry, 2001)

allows for the extent to which GARP holds to be used as an indicator for rationality. One

[1] Throughout the rest of this paper, GARP means: General Axiom of Revealed Preference

[2] The exact definition from Varian is: "A set of data satisfies the: (…) Generalized Axiom of Revealed Preference (GARP) if x^iRx^j implies not $x^jP^Ox^i$" (Varian, 1982, p. 947). In this example, x^iRx^j represents a revealed preference of good x^i over good x^j and $x^jP^Ox^i$ represents a strictly direct preference of x^j over x^i.

such methods is introduced by Varian (1990) as the Critical Cost Efficiency Index, which is sometimes referred to as the Afriat Efficiency Index (AEI)[3].

The Afriat Efficiency Index is based on the amount of GARP errors in a set of observations, as well as the magnitude of the GARP violations. The AEI can be described as the highest possible value of factor e_x, with $0<e<1$, by which GARP violating budget restrictions have to be multiplied in the dataset, in order to create a new dataset that would not have any GARP violations. The figure from Harbaugh, Krause and Berry (2001) below illustrates how the Afriat Efficiency Index can be determined:

Imagine the image above shows the choices of the consumer on two budget restrictions. For budget restriction A, the consumer chose bundle *a*, while for budget restriction B, he chose bundle *b*. In the situation of budget restriction A, the consumer does not prefer any of the other possible bundles over bundle *a*. Therefore, through Afriat's theorem, it must mean that, for a subject following GARP, any given alternative bundle *a** on budget restriction A should at most be of equal value to bundle *a*. In other words, if this were the choice, the following could be concluded: $U(a)\geq U(a^*)$. Assuming strong monotonicity, this also means that $U(a)>U(b)$, since *a* was chosen with the existence of alternative bundle *d*(4,2), and *d* has at least as much of one product as bundle b(3,2), and more of the other. However, if the same reasoning is followed for budget line B, looking at bundle c, it means that $U(b)>U(a)$. Since it cannot be true that $U(a)>U(b)$ and $U(b)>U(a)$, this combination of choices violates GARP.

However, if either budget line would have crossed the other in bundle *a* or *b* respectively, the consumer would not have been violating GARP, since the new relation could be written as $U(a)\geq U(b)\geq U(a)$, which means that $U(a)=U(b)$ is a valid. The factor by which either of the budget lines must be multiplied to cross the other in the GARP violating choice, respectively creating A* or B*, is what would be the Afriat Efficiency Index in this example. Meaning these two choices would have an Afriat Efficiency Index of 8/7=0.875.

It is worth noting that in this example, the budget restrictions had a one-to-one inverse relation, therefore calculating factor $e_{B\rightarrow B^*}$ for moving B to B* would be the same as $e_{A\rightarrow A^*}$ for moving A to A*. The AEI is defined as the highest possible factor e_x, so that if either factor

[3] Throughout the rest of this paper AEI means: Afriat Efficiency Index.

$e_{A \to A^*}$ or $e_{B \to B^*}$ would be higher, that would be the one that determines the AEI. Moreover, the AEI is based on the factor by which all GARP violating budget restrictions must be multiplied, in order to create a dataset without GARP violations. Meaning that if another budget restriction, C, was included in the dataset, which had to be moved by $e_{C \to C^*}$, to remove a GARP violation, while $e_{C \to C^*} < e_{A \to A^*}$, the Afriat Efficiency Index of the entire dataset would be AEI=$e_{C \to C^*}$.

The Afriat Efficiency Index can be taken as an indication of the level of rationality of an individual. The rationale for this is based on two assumptions. The first assumption is that a perfectly rational individual will follow GARP. After all, if GARP is not satisfied there is an opportunity for a money-pump. Secondly, it is assumed that people that are not fully rational, can vary in their magnitude of irrationality, based on how much the outcome of their choice differs from the utility maximising outcome.

In their study on economic rationality amongst kids, Harbaugh, Kruse and Berry (2001) designed a test to calculate the Afriat Efficiency Index test by using choice lists of bundles of two desirable goods. The idea behind the choice lists, is that any bundle option in a given list should also be represented in another choice list, either directly or indirectly. This way, when a preference is given for a bundle in one list, it can be checked whether this preference holds in another list where the same options are presented.

For example, table 1 shows two list of bundles that are both desirable for the subject. If GARP holds, it should mean that the subject does not prefer the other options in a given list over his chosen bundle. This implies that the subject is either indifferent between his choice and the other options or he strictly prefers his chosen bundle over the others.

List A		
Selectable Alternatives	**Distribution of goods x and y**	
	Qx	Qy
1	0	8
2	1	6
3	2	4
4	3	2
5	4	0

List B		
Selectable Alternatives	**Distribution of goods x and y**	
	Qx	Qy
1	0	5
2	1	4
3	2	3
4	3	2
5	4	1
6	5	0

Table 1: Example of two choice sets

In this example, option 4 in list B has the same distribution of goods as option 4 in list A. It can also be said that the options 1, 2 and 3 from list B are strictly inferior to options 1, 2 and

3 from list A respectively, since they contain the same amount of product X, but less of product Y. Therefore, it can be inferred that a subject following GARP will not choose options 1, 2 and 3 in list B, if he has chosen option 4 in list A. In other words, if a subject chooses option 4 in list A and either 1, 2 or 3 in list B, he would be violating GARP.

2.3 Visceral states

George Loewenstein (1996) defined visceral states as follows:

"The defining characteristics of visceral factors are, first, a direct hedonic impact, and second, an influence on the relative desirability of different goods and actions."

(Loewenstein, 1996, p. 273)

In other words, a visceral factor is a factor that makes you feel more, or less comfortable and as a result changes your preferences for certain goods and actions. One of the most recognizable examples of the effect that a visceral state has on people's preferences, is the example of someone who feels hungry when he is doing groceries. It is found that people tend to buy products that have higher caloric values when they feel hungry compared to when they do not (Nederkoorn, Guerrieri, Havermans, Roefs, & Jansen, 2009). In other words, the desirability of these high-calorie foods increased due to the state of feeling hungry.

While the effect of visceral states on preferences for products related to the physiological state make sense, as in the case of how being hungry increases the preference for food, some less obvious relations between visceral states and the desirability of goods or actions have also been found. As Briers et al. (2006) have shown, the state of feeling hungry can negatively influence how much money people were willing to donate to charity, or to others in general in a dictator game setting. They also found that priming people to a state of desiring money, subsequentially increased their desire of food, indicating a positive correlation between money and food.

Tuk, Trampe and Warlop (2011) found that people with higher sense of urination urgency, also had a higher sense of restraint when choosing between an immediate reward with a low payoff, versus a delayed reward with a higher payoff. They propose a theory that visceral states associated with inhibition, such as delaying the act of urinating, also increase self-

control in unrelated tasks, which they call the inhibitory spill over effect. Tuk, Zhang and Sweldens (2015) tested this theory and found a spill over effect of inhibition for tasks that were done simultaneously. However, they also found that inhibition is limited and depletable, and that the amount of self-control that someone has declines when faced with a series of inhibition related tasks. Fenn et al. (2015) explored the relation between deception and inhibition. The act of deception by hiding the truth appears to be similar to an act involving inhibition. Suggesting that people who are more convincing deceivers are also able exert more self-control.

2.4 Hypothesis

Assumptions

The difference between rational and actual behaviour raises one of the main issues that this paper explores, and it raises several questions: How rational are people? Why is it that people can make perfectly rational decisions in one situation and make an irrational decision in the next? Furthermore, if people make an irrational decision, how can the level of rationality be deterred?

In this light the following assumptions are made for the analysis:

- A perfectly rational individual will maximise his own utility and act accordingly.
- The revealed preference of the participants is in line with their true preference when they answer the questions.
- More violations of GARP indicate a lower level of rationality of the participant.

Hypothesis development

The aim of this paper is to explore whether someone's likelihood to cheat can be affected by a seemingly unrelated visceral state, appetite. Furthermore, the experiment also explores whether rationality can be considered as a mediating factor between appetite and dishonesty in the form of an incentive-based classroom experiment.

As discussed in the literature review, there is a link between instrumental rationality and other orientation. More specifically, it seems that there is a positive relation between rationality and the difference in the valuation of someone's own payoff compared to those

of others. This would imply that, if someone can improve their own financial reward by being dishonest (i.e. by lying about the artist of a painting in an art-sale) more rational subjects can be expected to be more dishonest. After all, the negative effect of decreasing the payoff of other parties (i.e. the buyer) is related to other orientation.

The hypotheses focus on the relation between increased appetite as a visceral state and its effects on rationality and the likelihood to be dishonest:

> *H1: On average, when participants are required to do a task that allows for dishonesty without the risk of detection, and where acting dishonestly will yield a bigger financial reward, participants will act dishonestly.*
>
> *H2: A group of participants that receives a treatment through which they are primed to think about food, due to the prospect of food as a reward for their participation, will have a lower average rationality score compared a group of participants that has the prospect of an unknown reward for their participation.*
>
> *H3: A group of participants that receives a treatment through which they are primed to think about food, due to the prospect of food as a reward for their participation, will display a higher level of dishonesty on average if this would increase their expected financial payoff, when compared a group of participants that has the prospect of an unknown reward for their participation.*

3 Methods

3.1 Experiment Design

Treatments

An experiment was conducted to test the hypotheses. The experiment was designed so that it would provide an indicator for rationality, as well as a task that can determine the likelihood for the participants to lie, while controlling for the level of appetite.

The subjects were divided between a control group and a treatment group. All the subjects knew beforehand that they would receive the same fixed reward for their participation, but the reward was only specified for the treatment group, and not for the control group. The instruction sheet for the control group only stated that the subjects would be rewarded for

their participation in the experiment, without further specification. By contrast, the instructions for the treatment group specified the reward by stating that the subjects would receive a Mars, Snickers or Kit-Kat chocolate bar as a reward for their participation.

The treatment was done so that the subjects in the treatment group would anticipate receiving food at the end of the experiment, which was expected to increase their level of appetite. The instructions for the treatment group also displayed pictures of the chocolate bars to further enhance the potential appetite increasing effect (Blechert et al., 2014).

To gather the required variables, a two-part questionnaire was conducted amongst the participants. The questionnaires were separated in two versions, version A including the control group instructions and version B including the treatment group instructions. Apart from the difference in instructions and the version indicator, the questionnaires were identical. In order to retain the anonymity of the participants, the questionnaires did not include any means of identification, other than being either of the two versions, this was underlined in the instructions that the participants received.

Measures: Part 1

For the first part of the questionnaire, the participants had to make a series of decisions, similar to the study by Harbaugh, Krause and Berry (2001). The subjects were faced with eleven choice sets with three to seven options of a distribution of two goods. Every choice set included the same two goods, a number of tokens which could be used in a coffee machine on the university campus, and a number of tokens which were exchangeable for beer, wine or soda in the university bar. For each choice set, the trade-off costs between the two products were kept constant, so that the options could be transformed into classic, linear budget restrictions in the analysis. To incentivise the subjects to give their actual preference, the subjects were told beforehand that, for the session they were in, two subjects would be chosen to come to the experimenter and be rewarded the actual amount of tokens from their answer on one of the choice sets.

The choice sets for part 1 were based on the budget restrictions from the paper of Harbaugh et al. (2001). It should be mentioned that, although the study originally was meant to have eleven choice sets, one of the questions has been excluded from the final dataset, as it was erroneously formatted, causing the trade-off values between the two goods to follow a non-

linear pattern. Having less choice sets means there are fewer violations of GARP possible per subject, which means the overall variance of the indicator for rationality would be lower.

However, considering that the study of Harbaugh et al. (2001) reported an average of 2.0 violations at the group with the lowest amount of GARP violations, GARP violations were still expected to be found in the answers to the remaining choice sets.

Choice sets

Check one of the "preferred option" boxes

1) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select one)*
9	0	
6	1	
3	2	
0	3	

Figure 2: First choice set

	List number																			
	1		2		3		4		5		6		7		8		9		10	
Distribution of goods	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy	Qx	Qy
	9	0	8	0	6	0	6	0	6	0	5	0	4	0	4	0	3	0	2	0
	6	1	6	1	5	1	4	1	3	1	4	1	3	2	3	1	2	3	1	3
	3	2	4	2	4	2	2	2	0	2	3	2	2	4	2	2	1	6	0	6
	0	3	2	3	3	3	0	3			2	3	1	6	1	3	0	9		
			0	4	2	4					1	4	0	8	0	4				
					1	5					0	5								
					0	6														

Table 2: Bundles per choice set

Measures: Part 2

The second part of the questionnaire is based on the study by Nazar, Amir and Ariely (2008). The questions were taken from an online cognitive reasoning test and were intended to be challenging but doable. However, the questions had to be done within the timeframe of three minutes, so that it would be highly unlikely for any participant to be able to complete all the questions. The participants then received an answer key with the correct answers to all the questions and were asked to write down the amount of questions that they answered correctly. In each session the participants were told beforehand that two of them would be

picked at random, and would be paid an amount in euros, which was half of the number they stated on the answer sheet.

This part is designed so that the participants were incentivised to be dishonest. Writing down a higher number on the answer sheet would result in potentially receiving a bigger financial reward. Since the question handouts and the answer key were separated and had no identification marker, there was also no way for the experimenter or for the other participants to know if one individual lied about whether the number on the answer key correctly reflected the amount of questions he had correctly answered.

3.2 Implementation

The experiment was conducted as a classroom experiment. Four different tutorial groups of students who were following a bachelor level marketing course participated. For each group, the experiment was done at the end of the tutorial and participation was voluntary, the students were however told beforehand that they would be rewarded for participating. As a result, the experiment was conducted in four separate sessions. Although the groups followed the same course, the groups were somewhat different in background, two of the groups consisted out of first year students from the International Bachelor Economics and Business, one of the groups followed an IT/Economic Bachelor programme in Dutch and the last group followed a Dutch Bachelor Programme for Fiscal Economics.

To control for the difference between the tutorial groups and other confounding factors such as time and date, the treatment groups were randomly assigned within each session.

For each session, the subjects were told not to converse with each other, they were however allowed to ask questions to the experimenter regarding the instructions if something was unclear. In the sessions, the subjects first received a page with the general instructions which they were asked to read, which also included the treatment condition. After confirming that the subjects read and understood the instructions, the experimenter handed out the first handouts. The handouts consisted of another instruction page and the cognitive reasoning questions. These instructions, which were identical for both treatment groups, described the tasks, the time limit, and the rules for receiving the extra reward. Again, upon confirming that the subjects read and understood the instructions, the experimenter

allowed the subjects to proceed with reading and answering the questionnaire within the time limit of three minutes. After the three minutes were over, the subjects were told to stop, and were handed a page with the answer key, which also included an area where the subjects could fill in the number of correct answers. Upon confirmation that all subjects were done writing the number on the answer key, two subjects were randomly selected and asked to show the number that they wrote on the answer key to the experimenter.

For the second part, the subjects were again given a handout comprising of a page of instructions followed by the choice lists of budget restrictions. The instructions, again being identical for both the treatment and the control group, also described the tasks and rules for receiving the extra reward. For this part, the subjects had no time limit, and they were free to proceed with answering the questions at their convenience. Upon confirming that the subjects were done answering the questions, two subjects were again picked at random and asked to show their choice list to the experimenter. The participants were then informed that the experiment had concluded and were rewarded with a chocolate bar of their choice.

In order to maintain a degree of anonymity towards the other participants, the answers of the extra reward winners for both parts of the experiments were only shown to the experimenter. These rewards were handed out after the experiment in closed envelopes at a later date.

3.3 Analysis

Determining rationality conflicts

For the analysis, the preference relation for the budget restrictions were entered in a table (see Table 3 below). For each choice list the bundles of goods that could be chosen are indicated with a 2. Any combination of products that was included in the other choice sets, that were strictly worse than the options in the question were assigned a 1. This way, for any choice list it could be said that the chosen combination of products was weakly preferred over the options with a preference indicator of 2 and, under the assumption of strong monotonicity, strictly preferred over the product combinations that were indicated with a 1.

Preference indicators per distrubition of goods										
Choice list / Bundle	1	2	3	4	5	6	7	8	9	10
Qx=0 Qy=2	1	1	1	1	2	1	1	1	1	1
Qx=0 Qy=3	2	1	1	2	0	1	1	1	1	1
Qx=0 Qy=4	0	2	1	0	0	1	2	2	1	1
Qx=0 Qy=5	0	0	1	0	0	2	0	0	1	1
Qx=0 Qy=6	0	0	2	0	0	0	0	0	1	2
Qx=0 Qy=8	0	0	0	0	0	0	0	0	1	0
Qx=0 Qy=9	0	0	0	0	0	0	0	0	2	0
Qx=1 Qy=3	0	1	1	0	0	1	1	2	1	2
Qx=1 Qy=4	0	0	1	0	0	2	0	0	1	0
Qx=1 Qy=5	0	0	2	0	0	0	0	0	1	0
Qx=1 Qy=6	0	0	0	0	0	0	2	0	2	0
Qx=2 Qy=0	1	1	1	1	1	1	1	1	1	2
Qx=2 Qy=2	1	1	1	2	0	1	1	2	1	0
Qx=2 Qy=3	0	2	1	0	0	2	2	0	2	0
Qx=2 Qy=4	0	0	2	0	0	0	0	0	0	0
Qx=3 Qy=0	1	1	1	1	1	1	1	1	2	0
Qx=3 Qy=1	1	1	1	1	2	1	1	2	0	0
Qx=3 Qy=2	2	1	1	0	0	2	1	0	0	0
Qx=3 Qy=3	0	0	2	0	0	0	0	0	0	0
Qx=4 Qy=0	1	1	1	1	1	1	1	2	0	0
Qx=4 Qy=1	1	1	1	2	0	2	1	0	0	0
Qx=4 Qy=2	0	2	2	0	0	0	2	0	0	0
Qx=5 Qy=0	1	1	1	1	1	2	1	0	0	0
Qx=5 Qy=1	1	1	1	0	0	0	1	0	0	0
Qx=6 Qy=0	1	1	2	2	2	0	1	0	0	0
Qx=6 Qy=1	2	2	0	0	0	0	2	0	0	0
Qx=8 Qy=0	1	2	0	0	0	0	2	0	0	0
Qx=9 Qy=0	2	0	0	0	0	0	0	0	0	0
Interpretation: 0=Bundle of goods was not present in this choice list, 1=A strictly better bundle of goods was included in the choice list, 2=Bundle of goods was included in the choice list										

Table 3: Overview of included bundles

The rationality conflicts were determined using a cross-referencing function in MS Excel. For every subject, the chosen product combination together with all alternatives are referenced against the other choice lists. If any combination of the chosen answer and the alternative answers were, directly or indirectly, included in another choice list and the subject chose the alternative bundle in the other situation. It was measured as a conflict in rationality. The total number of conflicts in rationality is used as an indication of how irrational the respondent was at the time he answered the questionnaire.

Statistics

To test the hypotheses, the following questions should be answered:

- Whether there is a significant difference between real and stated answers for the entire sample.
- Whether there is a significant difference between the level of rationality between the treatment and control group.
- Whether there is a significant difference between the number of real and stated answers between the treatment and control group.

The answers for the questions could not be traced back to a participant in order to guarantee a degree of anonymity. Which means that the results of the two parts of the experiments cannot be matched on an individual level, only on a session level. Therefore, the data only allowed for comparison of results between groups of different treatments and between groups of different sessions.

For choosing the right method for comparison between the treatment groups, it is first tested whether the answers are normally distributed using the Shapiro-Wilk's test. This was used to determine whether the comparison could be done using a parametric, or a non-parametric test. As a parametric test a standard t-test could be done, while a Mann-Whitney U test could be used for cases where the results were not normally distributed. These tests determine whether the means of two samples are significantly different. Based on results from similar experiments (Mazar, Amir, & Ariely, 2008) it was expected that the means of the stated answers are higher than the means of the real answers for all groups. Moreover, in this experiment setup the expected external reward increased when a respondent reports a higher number of correctly answered questions. Meaning, the respondents were externally incentivised to lie and overstate the amount of questions they correctly answered.

Finally, two regression analyses are done. The first regression analysis will test whether there is a linear relation, taking the amount of correct answers as dependent variable and two dummy variables as predicting variables. The first dummy indicates whether the amount of correct answers is real or stated, while the second dummy indicates whether the answers come from a participant of the treatment or the control group. The second regression will include the same variables as the first one, while also adding another variable: the average number of violations of the treatment group within the session that the answer came from.

4 Results

4.1 Terms and definitions:

Table 4 shows the terms and definitions used in the results (4) and discussion (5) sections:

Terms	**Definition**
Full sample	All the participants that took part in the experiment.
Control group	All participants that received the set of the instructions and questions that did not include the "candy bar prime" (version A).
Treatment group	All participants that received the set of the instructions and questions that included the "candy bar prime" (version B).
Budget restrictions	The first set of questions in the experiment, which consists of the choice lists of different distributions of bundles of goods.
Logical reasoning test	The second set of questions in the experiment, consisting of the logical reasoning test questions themselves and the answer sheet on which the participants stated the amount of correctly answered questions
Session	A session in which the experiment was conducted. In total, the experiment was conducted over four sessions.
Rationality violation/conflict in rationality	A combination of bundles in the participants' answers for the budget restrictions, for which a respondent did not satisfy GARP/Afriat's cyclical consistency theorem. Method can be read in section 3.3.
Real answers	The answers that the participants gave to the second questionnaire.
Number of real correct answers	The number of answers that the participants gave to the second questionnaire, that were the correct answers.
Stated correct answers	The number of questions that the participants reported to have answered correctly, for the second questionnaire.
Treatment group (dummy)	A dummy variable indicating to which treatment group an observation belonged, 0=control, 1=treatment.
Real/stated answers (dummy)	A dummy variable indicating whether an observed number of correct answers came from the real answers or from the participant's stated answers, 0=real, 1=stated.
Average number of violations in treatment & session	The average number of rationality violations of the respondents that were in the same treatment group and in the same session.

Table 4: Terms and definitions for result and analysis section

4.2 Budget restrictions

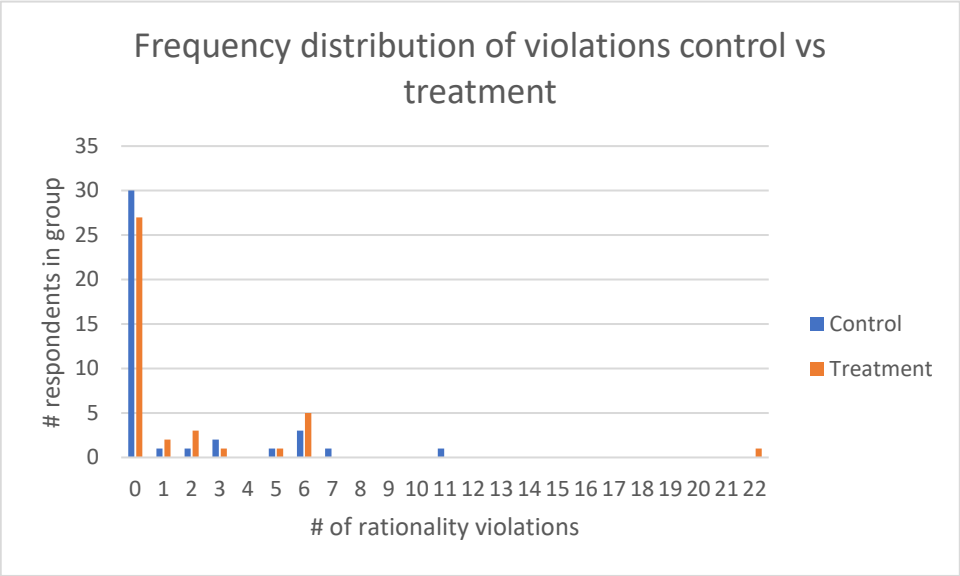


Figure 3: Distribution of rationality violations in the control and treatment groups

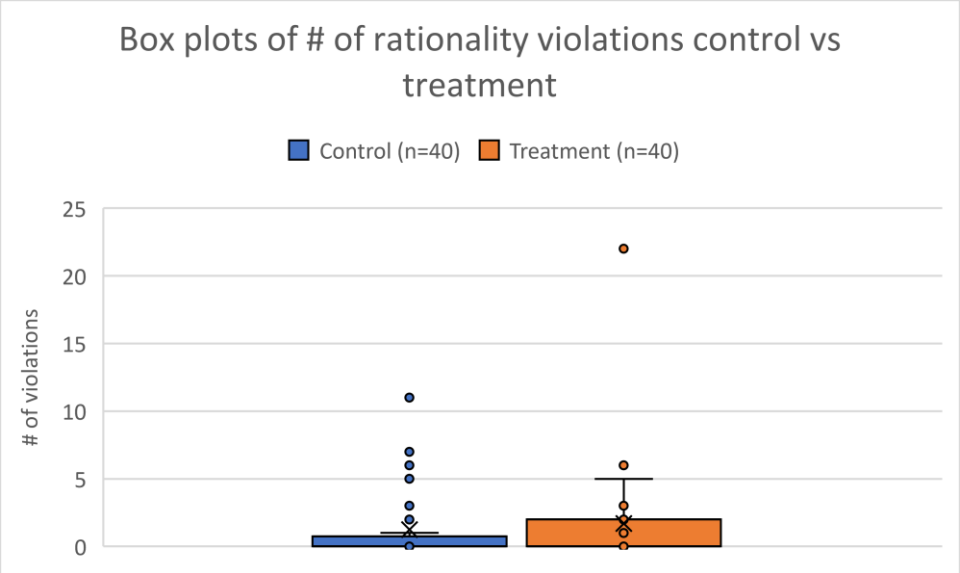


Figure 4: Box plots of rationality violations in the control and treatment groups

The distribution of violations between the control and treatment group, as well as the box plot, suggest that there could be an effect of the treatment on the amount of conflicts. More specifically, the treatment group appears to have more conflicts on average than the control group and appears to have more variance.

Shapiro-Wilk's test		
	A (N=40)	B (N=40)
W	0.56	0.50
P	<0.01**	<0.01**

Table 5: SW test - rationality violations

Mann-Whitney U test	
Z	-0.60
p (1-tailed)	0.27

Table 6: MW-U test - rationality violations

The Shapiro Wilk's test confirms that, for $\alpha<0.05$, the test values are significant for both samples, therefore the null hypothesis is rejected, and it cannot be assumed that the samples are normally distributed. To determine whether the two samples have significantly different distributions, a Mann-Whitney U test is used. The P-value of neither test is significant for $\alpha<0.05$, hence, the test results do not show a statistically significant difference between the two distributions.

4.3 Logical reasoning test

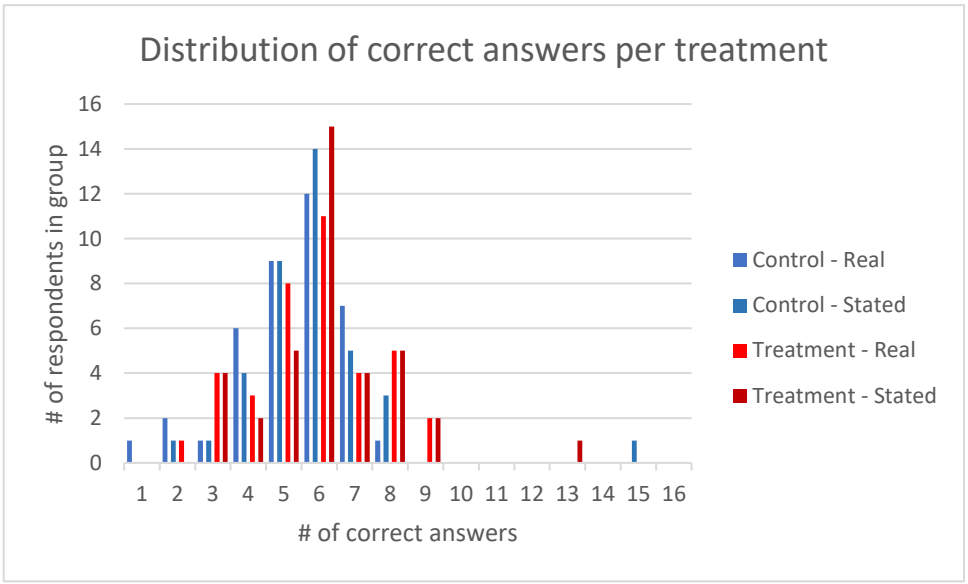


Figure 5: Distribution of number of stated and real correct answers for the control and treatment group

Descriptive Statistics						
	Control		Treatment		Full sample	
	Real correct answers	Stated correct answers	Real correct answers	Stated correct answers	Real correct answers	Stated correct answers
Mean	5.28	5.89	5.45	5.85	5.37	5.87
D-Mean		0.61		0.40		0.50
Median	6	6	6	6	6	6

Table 7: Descriptive Statistics

Shapiro-Wilk's tests						
	Treatment		Control		Full sample	
	Real correct answers	Stated correct answers	Real correct answers	Stated correct answers	Real correct answers	Stated correct answers
W	0.92	0.77	0.93	0.92	0.94	0.87
p	<0.01**	<0.01**	0.02**	<0.01**	<0.01**	<0.01**

Table 8: Shapiro-Wilk's tests for normal distribution of number of correct answers

Mann-Whitney U test			
	Real answers treatment vs. control	Stated answers treatment vs. Control	Full sample real vs stated correct answers
Z	0.55	0.50	-1.27
p (1-tailed)	0.29	0.31	0.10

Table 9: MW-U test for difference of real and stated correct answers

The means in table 7 show a higher average of number of stated correct answers than real correct answers, moreover, there appears to be a difference between the answers the control group than the answers of the treatment group, which could indicate an effect of the treatment on either the likelihood for people to lie, or the magnitude of the "overstatement" of liars. To test whether these differences are statistically significant, the difference is first tested between real correct answers and stated correct answers, to see if there is a significant overstatement on the test performance. The difference is also compared between the distribution of real answers between the treatment and control groups, as well as the distribution of stated answers between the two groups. If a significant difference could be found between either group, it could be an indication of a treatment effect.

For all samples, the Shapiro-Wilk's test shows that the tests are not normally distributed, therefore the required assumptions for comparing the treatment and control group with a t-test do not apply, nor for the comparison between the real and stated correct answers of the full sample. Instead, Mann-Whitney U tests are used to compare whether there is a statistically significant difference in either the real or the stated answers between the treatment and control group. The Mann-Whitney U tests did not show a significant result for any of the comparisons.

4.4 Linear regressions

Coefficients: base regression				
	B	Std. Error	t	Sig.
Constant	5.34	0.28	19.15	0.00**
Treatment Group (dummy)	0.06	0.32	0.19	0.85
Real/Stated answers (dummy)	0.50	0.32	1.57	0.12

Dependent Variable: # correct answers – R^2=0.016

Table 10: Base regression

Table 10 shows the base regression model, only including the treatment group dummy and the real/stated answer dummy as predicting variables. The dummy variables do not have a significant effect on the number of correct answers. Moreover, the R-squared value of 0.016 indicates that the dummy variables only explain a very small portion of the variance of the number of correct answers.

Coefficients: regression including average violations				
	B	Std. Error	t	Sig.
Constant	5.23	0.33	16.01	0.00**
Treatment/Control Group	0.07	0.32	0.22	0.83
Real/Stated answers	0.46	0.33	1.40	0.16
Average number of violations in treatment & session	0.09	0.13	0.64	0.52

Dependent Variable: # correct answers - R^2=0.019

Table 11: Regression including average violations

For the second model, as shown in table 11, the variable for average number of violations in treatment & session (as defined in 4.1) is added as predicting variable. Similar to the first regression model, the predicting variables do not have a significant effect on the number of correct answers. Additionally, the R-squared value is only marginally higher than the first regression model, indicating that the average number of violations in treatment & session has a negligible effect on the explanatory value of the model.

5 Discussion

For this paper, the test results were evaluated on a $\alpha=0.05$ significance threshold. Under this condition, there was no significant difference between the number of real correct answers versus the number of stated correct answers overall. This was surprising considering the results of Mazar et al. (2008). These results could indicate that the expected financial reward was too low for the respondents to their moral principles. Additionally, there was no significant difference found between the distributions of answers for the treatment and control groups. The experiment was set-up in a way that only allowed for comparison on a group level in order to maintain a level of anonymity for the participants. One of the limitations of such a setup is that it requires relatively more respondents in order to be able to run statistical tests with sufficient statistical power. In the experiment setup of this paper, it is possible that no statistical difference was found due to the number of respondents being too low. Coincidentally the fact that the regression analyses did not show significant effects for any of the predicting variables, indicates that these variables do not appear to have any significant predictive value for determining the amount of correct answers for a given observation.

This fact that it was set-up as a classroom experiment, in which the experimenter in some cases was also a student assistant or supervisor of the class, meant that there was likely to be an experimenter effect, regardless of the privacy control measures. It is not unlikely that a participant, who is also a student of the experimenter, would be more (or perhaps in some cases, less) hesitant to act in a way that they would expect to have a negative impact on the experimenter. This effect could have been enhanced due to the fact that the experiment setup involved a financial compensation for being dishonest. This could be perceived by the participant as “taking” more money from the experimenter. Since not all sessions had an experimenter who also supervised the participants, it could be interesting to explore the differences between these sessions. However, in the case of this experiment, the samples would have been too small to have a test setup with sufficient statistical power.

A case could be for examining the hypothesis further with an experiment setup that would allow for a test with more statistical power. For example, by using a design that allowed for a comparison on a subject level. Another advantage of such a design would be that the

experiment can distinguish both between the likelihood that someone lies, as well as the magnitude of the lie in case someone would choose to lie.

Regarding the experiment design, another limiting element was that the experiment only compared two treatments amongst the subjects. This was mostly due to considerations regarding the statistical power of the experiment setup. One of the limitations of using only two treatments is that it is hard to isolate what specifically causes an effect between treatments. In essence, a two-treatment setup measures the combined effect of all of the features that are different for a given respondent in the control and treatment group. For this experiment, the difference between the control and the treatment group was that the control group knew that they were getting a reward for participating, but the reward was not specified. whilst the treatment group knew that they were getting food as a reward, which was supported by showing images of the chocolate bars they were getting. This means that in essence, the experiment setup would measure three effects: the effect of knowing what reward you receive versus not knowing the reward, the effect of the reward being food versus the effect of the reward being something else, and the effect of having images of the specified reward included in the description. One of the methods to control for these confounds, and for relating a possible effect to a more specific cause, would be to add a third treatment group. If the third treatment group would receive a reward other than food, which they were informed about beforehand and for which the description would include images of the reward, it would isolate the effect of receiving food versus receiving another specified reward.

Upon examining the data, one interesting finding was that a number of participants decided not to hand in the answer sheet of the cognitive reasoning test after completing the questionnaire. This finding was curious since from a purely financial perspective, players are better of handing in an answer sheet with a random number as opposed to not handing in an answer sheet at all. Moreover, the effort for filling in the answer sheet is minimal, since it only requires the student to write down a number. For the group that did not hand in the answer sheet, the potential reward could have been unclear, or it might not have weighed up to the social cost of being dishonest. Unfortunately, this has implications for interpreting the data, since the research question was partially about relating the answers of the cognitive reasoning test to the estimate for irrationality of the choice sets. Due to the design

of the experiment, it is impossible to filter the budget restriction “answers” that relate to the incomplete cognitive reasoning test. Therefore, in the budget restriction answer batch results are included that should not be related to the batch of answers of the cognitive reasoning test. This caused some impurities in the data which are likely to influence the comparability of the test results.

All things considered, a lot of the complications of the used experiment design could be controlled for by having a setup where participants would digitally participate in the experiment, rather than having them complete a hard-copy questionnaire. Having a digital test setup would allow for participants to be completely anonymous to the experimenter while it would still allow the experimenter to link the results of the rationality test and the cognitive reasoning test answers. However, not having an experimenter present also means that there is less control for confounds during the participation of the experiment. Since this design also considered (an approximation for) the level of appetite of the participants, it was important for the experimenter to be able to confirm that participants were not eating during the experiment. An ideal design for this specific research question should therefore still have the control of a classroom/lab experiment setup, while also having the participants complete the questionnaires digitally. Increasing the sense of the anonymity of the experiment while also allowing for a comparison on a subject level.

References

Afriat, S. N. (1967). The Construction of Utility Functions from Expenditure Data. *International Economic Review, 8*(1), 67-77.

Allingham, M. G., & Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics, 1*, 323-338.

Bechert, J., Meule, A., Busch, N. A., & Ohla, K. (2014). Food-pics: an image database for experimental research on eating and appetite. *Frontiers in Psychology*, Original Research Article.

Cressey, D. R. (1953). *Other people's money; a study of the social psychology of embezzlement.* New York: US: Free Press.

Cressey, D. R. (1986). Why managers commit fraud. *Australian & New Zealand Journal of Criminology, 19*(4), 195-209.

Erat, S., & Gneezy, U. (2012). White Lies. *Management Science, 58*(4), 723-733.

Fenn, E., Blandón-Gitlin, I., Coons, J., Pineda, C., & Echon, R. (2015). The inhibitory spillover effect: Controlling the bladder makes better liars. *Consciousness and Cognition, 37*, 112-122.

Gneezy, U. (2005). Deception: The Role of Consequences. *American Economic Review, 95*(1), 384-394.

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Behavior & Organization, 93*, 293-300.

Harbaugh, W. T., Krause, K., & Berry, T. R. (2001). GARP for Kids: On the Development of Rational Choice Behavior. *The American Economic Review*, 1539-1545.

Hurkens, S., & Kartik, N. (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics, 12*(2), 180-192.

Kahneman, D. (2011). *Thinking Fast & Slow* (Vol. 1). New York: Farrar, Straus and Giroux.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business, 59*(4), S285-S300.

Kasem, R., & Higson, A. (2012). The new fraud triangle model. *Journal of Emerging Trends in Economics and Management Sciences, 3*(3), 191-195.

Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes, 65*(3), 272-292.

Mazar, N., Amir, O., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Working paper series // Federal Reserve Bank of Boston, 06-3*.

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research, 45*(6), 633-644.

Meglino, B. M., & Korsgaard, M. A. (2004). Considering Rational Self-Interest as a Disposition: Organizational Implications of Other Orientation. *Journal of Applied Psychology, 89*(6), 946-959.

Nederkoorn, C., Guerrieri, R., Havermans, R., Roefs, A., & Jansen, A. (2009). The interactive effect of hunger and impulsivity on food intake and purchase in a virtual supermarket. *International Journal of Obesity, 33*, 905-912.

Neumann, J. v., & Morgenstern, O. (1944). Theory of games and economic behavior.

Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fer, E. (2004). The Neural Basis of Altruistic Punishment. *science, 305*(5688), 1254-1258.

Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior & Organization, 3*(4), 323-343.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A Neural Basis for Social Cooperation. *Neuron, 35*, 395-405.

Samuelson, P. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economica, 5*(17), 61-71.

Thaler, R. H. (2015, May 8). *Unless You Are Spock, Irrelevant Things Matter in Economic Behavior*. Retrieved from The New York Times: https://www-nytimes-com.eur.idm.oclc.org/2015/05/10/upshot/unless-you-are-spock-irrelevant-things-matter-in-economic-behavior.html

Tuk, M., Trampe, D., & Warlop, L. (2011). Inhibitory Spillover : Increased Urination Urgency Facilitates Impulse Control in Unrelated Domains. *Psychological Science, 22*(5), 627-633.

Tuk, M., Zhang, K., & Sweldens, S. (2015). The Propagation of Self-Control: Self-Control in One Domain Simultaneously Improves Self-Control in Other Domains. *Journal of Experimental Psychology, 144*(3), 639-654.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science, 185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.

Varian, H. R. (1982). The Nonparametric Approach to Demand Analysis. *Econometrica, 50*(4), 945-973.

Appendix

Appendix 1 - Experiment Setup

To explore the relation between appetite, likelihood to cheat and economic rationality, a classroom experiment will be conducted. Experiments are generally good methods to measure causality between variables, as they allow the experimenter to have a lot of control which can limit the possible confounding factors. One important drawback for conducting an experiment in this setup however, is that the external validity is relatively low. Conclusions can only be drawn about the causality between the variables under the exact conditions of the experiment, whereas natural conditions usually involve many other uncontrollable variables. Since this experiment focuses on finding the causal inference of appetite on likelihood to cheat, the external validity of the experiment is not as important as the internal validity, which is why a lab experiment is the preferred option.

The experiment will be done in four sessions. Each session will be held at the end of a tutorial for first- and second-year bachelor students at the Erasmus University Rotterdam. The sessions are on Wednesday from 11:00 to 12:45, Thursday from 09:00 to 10:45 and 15:00 to 16:45 and on Friday from 09:00 to 10:45. The experiments themselves are expected to be done during the breaks which are held 45 minutes into the tutorial sessions. The tutorials are identical in content and will be supervised by the same teaching assistant; moreover, both sessions are expected to be attended by 30 students.

The level of appetite of an individual likely varies over the day. For instance, the appetite is expected to be lower right after consuming food. As the sessions have a different starting time, one can expect a difference in the average level of appetite between the two groups. To control for the difference possible in appetite levels between the sessions, all sessions will include a control group and treatment group, which can be combined in the analyses to form the total control and treatment groups. Table 1.x shows how the participants of each session are divided in the experiment groups.

Table A1 Experiment groups

Session 1: 30 students	
Control group	Treatment group
15 students	15 students

Session 2: 30 students	
Control group	Treatment group
15 students	15 students

Session 3: 30 students	
Control group	Treatment group
15 students	15 students

Session 4: 30 students	
Control group	Treatment group
15 students	15 students

At the start of the experiment, the experimenter will ask the participants to blindly draw a ticket out of a bowl on the desk. The tickets will have the numbers 1-30, the experimenter will tell the participants that everyone who drew an uneven number is required to sit at one side of the room and that the even numbers are required to sit on the other side.

The experimenter will announce that all participants will receive instructions and ask them to read the instructions. After the announcement, the experimenter will hand out version A of the instructions to the uneven numbers (Appendix A1) and version B to even numbers (Appendix A2). The sheets of the participants will only be marked with an A or B to identify which experiment group they belong to.

The experimenter will wait 3 minutes and then ask the participants if they have finished reading the instructions. If the participants finished reading the instructions, the experimenter will hand out a brochure with a series of cognitive test questions (Appendix X1), and ask the participants to keep this brochure face down until the experimenter tells them they can start answering the questions.

After the experimenter tells the participants to answer the questions, he will set a timer of 3 minutes. When the 3 minutes are over, the experimenter will ask the participants to stop

and turn their brochure face down again. After which the experimenter will collect the red pens.

The experimenter will then hand out the answer sheets of the questions and tell the participants to state the amount of questions they have answered correctly. The experimenter then asks if the participants are finished after 1 minute, and then ask them to put the question lists in a box at the desk of the experimenter.

After the experimenter has received the question lists from all of the participants, the experimenter will announce that part 1 is over and will start handing out the brochure for part 2. The experimenter will announce that there is no time-limit for answering the questions of part 2, and that the participants can read the instructions and answer the questions at their own pace. After 5 minutes, the experimenter will ask whether all of the participants are finished with part 2.

If the participants are finished with part two, the experimenter will announce that part 2 is concluded and that all of the participants who want to be eligible to win one of the variable rewards can write down the number of their ticket on a piece of paper provided by the experimenter.

The experimenter will then randomly pick one of the numbers written on the piece of paper and announce that this numbers is the winners of part 1 and ask the participants with the corresponding tickets to come to the desk, reveal their ticket, and reveal the number of answers that they have reported to have answered correctly on the answer sheet. He will hand out the cash as stated on the instructions of part 1 after the experiment (appendix A1, A2). The experimenter then instructs the students to hand in their answer sheets in the box on the desk.

The experimenter will then again randomly pick one number and ask the participant with the corresponding tickets to come to the experimenter and reveal their tickets. The experimenter then generates a number between 1 and 11 and awards the coin bundles that the participants have chosen in the corresponding choice set.

After the coins are awarded, the experimenter announces that the session is concluded and asks the participants to put the remaining pieces of paper in the box. He hands out a chocolate bar to all of the participants as they return the sheets.

Appendix 2 – Instructions: control group

Version A

Instructions

Thank you for participating in this experiment. This experiment consists out of two parts, part 1 and part 2. Each part is a list of questions which you are required to answer. At the end of the experiment you will receive a certain reward for your participation. You will be informed at the end of the session what your certain reward will be.

On top of this reward, you will also have a chance to be chosen to receive a variable reward, which will be based on the answers of one of the two parts of the experiment. At the end of this session, one participant will be randomly selected for each part and will receive the reward based on their answers. The rewards are described in the instructions. The instructions for part 1 are written on the back of this page. You will receive the instructions for part 2 after part 1 is finished.

Instructions Part 1

In the first part of this experiment you will receive a brochure with 15 multiple choice questions which you have to answer in 3 minutes. You will receive this brochure face down and are only allowed to turn over the brochure and start answering the questions when the experimenter tells you that you are allowed to start. After the three minutes are over, the experimenter will tell you to stop and ask you to turn over the brochure again so that it is face down.

After all of the participants have turned the brochure face down, the experimenter will hand out sheets with the correct answers to the questions. At the bottom of the sheet you are requested to write down how many questions you have answered correctly. This has to be done with a red pen, which will be provided by the experimenter. After you have stated the number of correct answers, the experimenter will ask you to hand in the list of questions. Note that you only hand in the questions and not the answer sheet. The question lists will remain anonymous to the experimenter and the experimenter will not know which participant had which question list.

If you are selected to receive the variable reward for this experiment at the end of this session, you will be asked to show your sheet to the experimenter in after the session and receive €0.50 per claimed correct answer as stated on the sheet. If you are not selected you can hand in your answer sheet, again in a box on the table. In this case, the number that you write on the answer sheet will also remain anonymous to the experimenter.

The experimenter will not know which answer sheet belongs to which question list, only that an answer sheet for version A belongs to a question list for version A. The pages have no distinguishable markings other than an A.

Appendix 3 – Instructions: treatment group

Version B

Instructions

Thank you for participating in this experiment. This experiment consists out of two parts, part 1 and part 2. Each part is a list of questions which you are required to answer. At the end of the experiment you will receive a certain reward for your participation. For the certain reward, you can choose to receive a Snickers, Mars or Kit-Kat bar (as seen below).

On top of this reward, you will also have a chance to be chosen to receive a variable reward, which will be based on the answers of one of the two parts of the experiment. At the end of this session, one participant will be randomly selected for each part and will receive the variable reward after the session, based on their answers. The rewards are described in the instructions. The instructions for part 1 are written on the back of this page. You will receive the instructions for part 2 after part 1 is finished.

Instructions Part 1

In the first part of this experiment you will receive a brochure with 15 multiple choice questions which you have to answer in 3 minutes. You will receive this brochure face down and are only allowed to turn over the brochure and start answering the questions when the experimenter tells you that you are allowed to start. After the three minutes are over, the experimenter will tell you to stop and ask you to turn over the brochure again so that it is face down.

After all of the participants have turned the brochure face down, the experimenter will hand out sheets with the correct answers to the questions. At the bottom of the sheet you are requested to write down how many questions you have answered correctly. This has to be done with a red pen, which will be provided by the experimenter. After you have stated the number of correct answers, the experimenter will ask you to hand in the list of questions. Note that you only hand in the questions and not the answer sheet. The question lists will remain anonymous to the experimenter and the experimenter will not know which participant had which question list.

If you are selected to receive the variable reward for this experiment at the end of this session, you will be asked to show your sheet to the experimenter in after the session and receive €0.50 per claimed correct answer as stated on the sheet. If you are not selected you can hand in your answer sheet, again in a box on the table. In this case, the number that you write on the answer sheet will also remain anonymous to the experimenter.

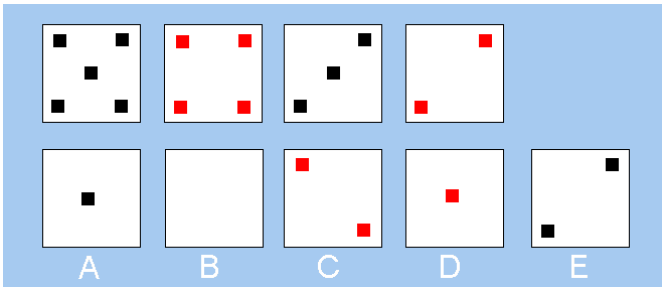
The experimenter will not know which answer sheet belongs to which question list, only that an answer sheet for version B belongs to a question list for version B. The pages have no distinguishable markings other than a B.

Appendix 4 - logical reasoning test questions and answer sheet

Non-verbal reasoning test

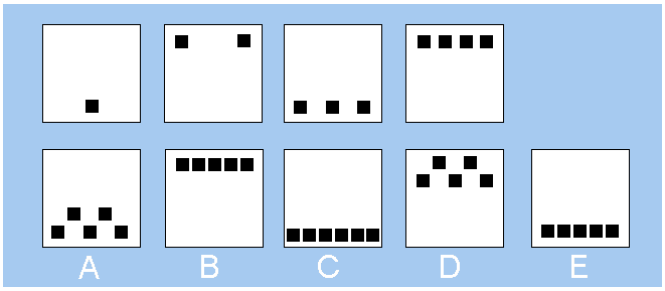
Circle the right answers

1) Which figure is next in the sequence below? *(circle right answer)*



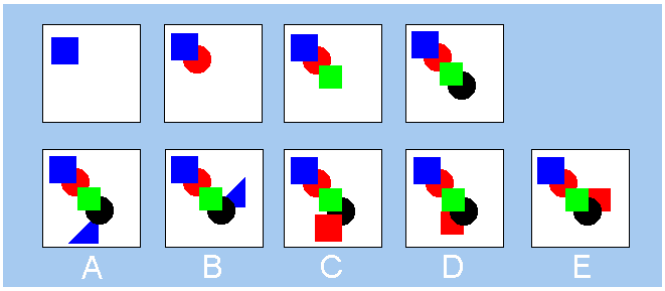
---

2) Which figure is next in the sequence below? *(circle right answer)*



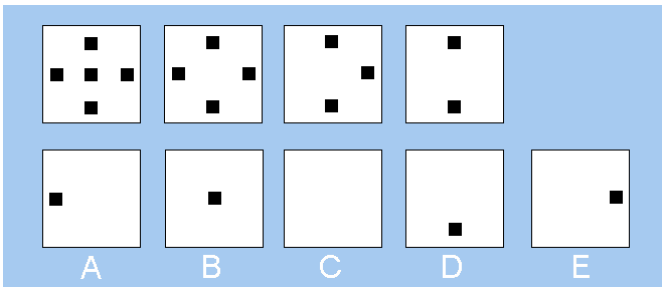
---

3) Which figure is next in the sequence below? *(circle right answer)*



---

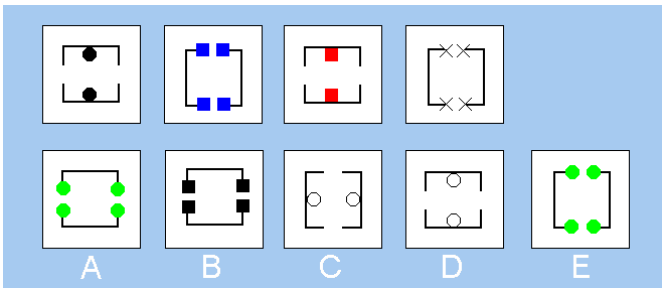
4) Which figure is next in the sequence below? *(circle right answer)*



Non-verbal reasoning test

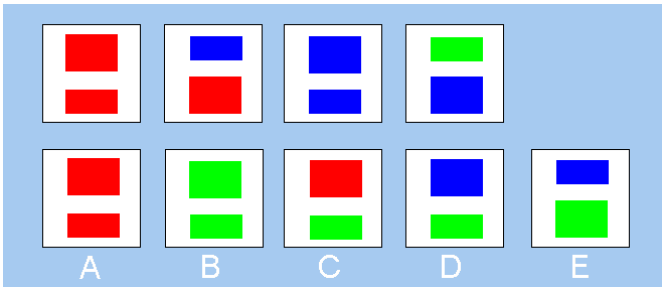
Circle the right answers

5) Which figure is next in the sequence below? *(circle right answer)*



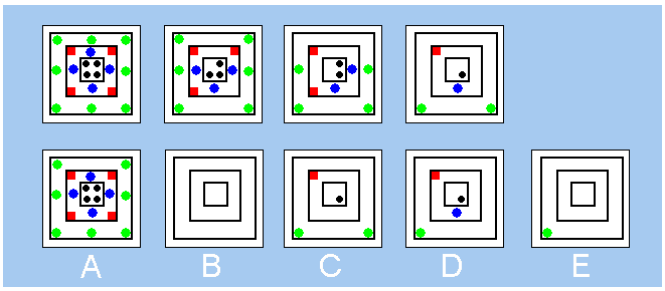
---

6) Which figure is next in the sequence below? *(circle right answer)*



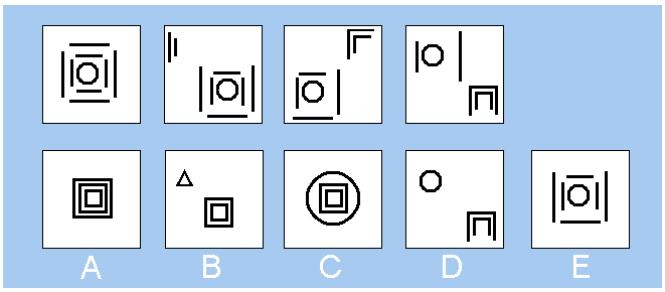
---

7) Which figure is next in the sequence below? *(circle right answer)*



---

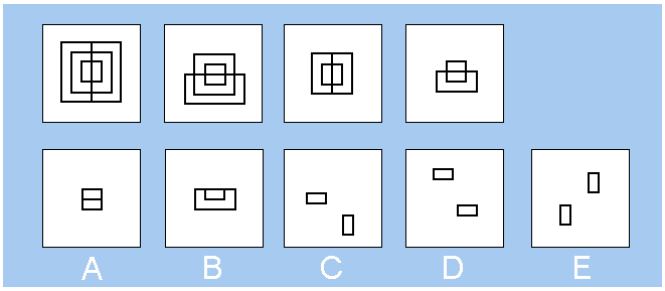
8) Which figure is next in the sequence below? *(circle right answer)*



Non-verbal reasoning test

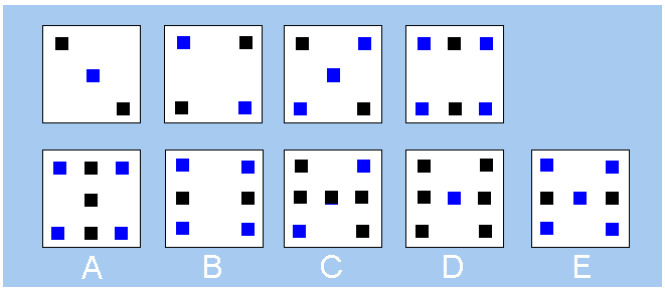
Circle the right answers

9) Which figure is next in the sequence below? *(circle right answer)*



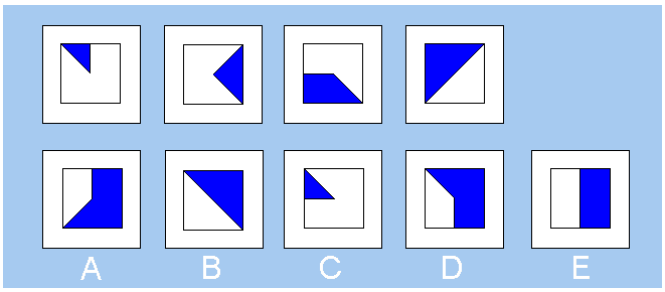
---

10) Which figure is next in the sequence below? *(circle right answer)*



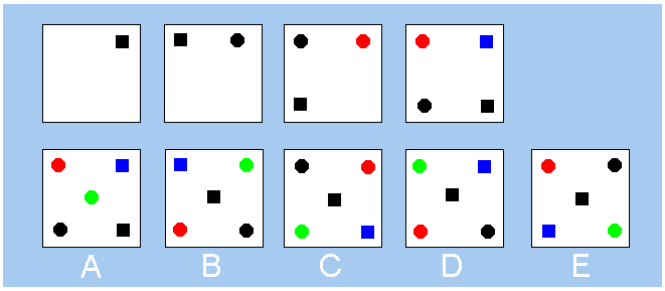
---

11) Which figure is next in the sequence below? *(circle right answer)*



---

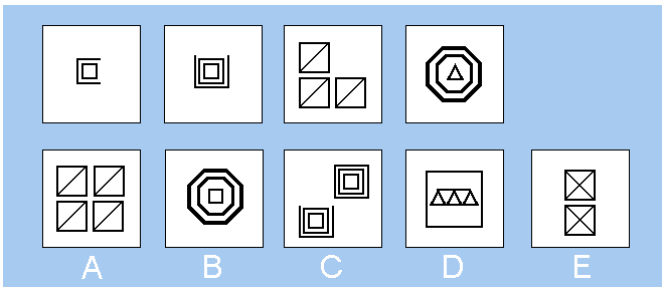
12) Which figure is next in the sequence below? *(circle right answer)*



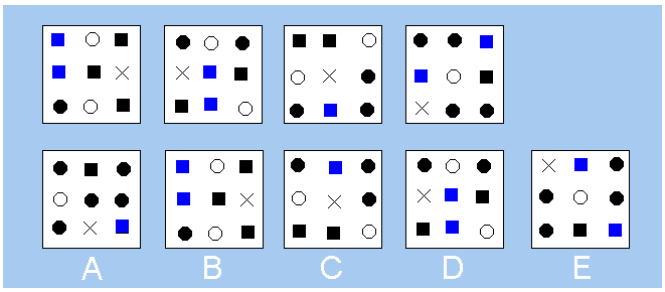
Non-verbal reasoning test

Circle the right answers

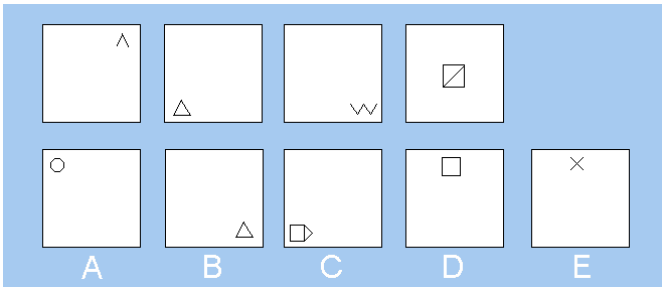
13) Which figure is next in the sequence below? *(circle right answer)*



14) Which figure is next in the sequence below? *(circle right answer)*



15) Which figure is next in the sequence below? *(circle right answer)*



Answer sheet

Please check in the boxes whether you have answered the questions correctly or incorrectly

Non-verbal reasoning test:

Question	Answer
1	A
2	E
3	C
4	D
5	D
6	B
7	B
8	C
9	E
10	E
11	D
12	B
13	C
14	A
15	C

Please fill in the number of answers you have answered correctly below

Number of correct answers: _____

Appendix 5 - Budget restriction questionnaire

On the next pages you will see 11 tables which represent choice sets of coin bundles, for each choice set you have to select which bundle of coins you would prefer. The bundles consist out of a given number of coffee coins and Smitse coins. For each choice set you have to check the "preferred option" box to indicate that this is the bundle of coins that you would prefer in the choice set.

The coffee coins can be used for the coffee machines in the Spar at the Erasmus University Campus, one coin will allow you to get one coffee of your choice.

The Smitse coins can be exchanged at the bar of Café in de Smitse at the Erasmus University Campus, one coin can be exchanged for one beer, or one soda drink.

At the end of the session, one participant will be chosen randomly and be rewarded with the bundle they chose of one of the choice sets. The choice set will also be chosen randomly.

Choice sets

Check <u>one</u> of the "preferred option" boxes

1) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
9	0	
6	1	
3	2	
0	3	

---

2) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
8	0	
6	1	
4	2	
2	3	
0	4	

---

3) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option

		(select <u>one</u>)
6	0	
5	1	
4	2	
3	3	
2	4	
1	5	
0	6	

Choice sets

Check <u>one</u> of the "preferred option" boxes

4) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
6	0	
4	1	
2	2	
0	3	

---

5) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
6	0	
3	1	
0	2	

---

6) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
5	0	

4	1	
3	2	
2	3	
1	4	
0	5	

---

7) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
4	0	
3	2	
2	4	
1	6	
0	8	

Choice sets

Check <u>one</u> of the "preferred option" boxes

8) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
4	0	
3	1	
2	2	
1	3	
0	4	

---

9) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
3	0	
2	3	
1	6	
0	9	

---

10) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option

		(select <u>one</u>)
3	0	
2	2	
1	3	
0	4	

---

11) Out of the options below, which combination of rewards would you prefer?

Spar coffee coins	Smitse coins	Preferred option *(select <u>one</u>)*
2	0	
1	3	
0	6	