Essays in experimental economics

van Veldhuizen, R.R.

Publication date
2013

Chapter 4

Willpower Depletion and Framing Effects[1]

This chapter investigates whether depleting people's cognitive resources (or 'willpower') affects the degree to which they are susceptible to framing effects. Recent research in social psychology and economics has suggested that willpower is a resource that can be temporarily depleted and that a depleted level of willpower is associated with self-control problems in a variety of contexts. In this study, we extend the willpower depletion paradigm to framing effects and argue that willpower depletion should increase framing effects. To test this idea we ran an experiment in which we depleted participants' willpower and subsequently had them take part in a series of framing tasks, including a framed prisoner's dilemma, an attraction effect task, a compromise effect task and an anchoring task. However, we found no evidence that framing effects were indeed more prevalent in willpower depleted participants than in controls.

4.1 Introduction

Every day people are subjected to temptations that require self-control to be resisted. A recent literature in social psychology has emphasized that exercising self-control requires willpower and that people's willpower should be regarded as

[1]This chapter is based on De Haan and Van Veldhuizen (2012). We would like to thank Gary Charness, Joep Sonnemans and other colleagues from the Center of Research in Experimental Economics and Political Decision-making (CREED) and the University of Amsterdam for their helpful comments. Financial support from the University of Amsterdam Research Priority Area in Behavioral Economics is gratefully acknowledged.

a limited resource (see for example Muraven and Baumeister, 2000; Baumeister et al., 2008; or see Hagger et al., 2010 or Bucciol, Houser, and Piovesan, 2010). In particular, this literature argues that exercising self-control can temporarily 'deplete' the willpower resource. Once someone's willpower has been temporarily depleted, this will make him or her be more likely to yield to subsequent temptations.

The resource (or willpower) depletion literature has recently started to get a foothold in economics.[2] Most commonly, willpower depletion has been related to self-control, procrastination and affiliated concepts. Recent theoretical models that have explicitly incorporated willpower depletion include Ozdenoren, Salant, and Silverman (2011); Ali (2011) and Fudenberg and Levine (2012). The empirical literature is slightly sparser, although Bucciol, Houser, and Piovesan (2011a,b) find that a willpower depleting activity reduces subsequent productivity in both children and adults. Another related paper is Burger, Charness, and Lynham (2011), who investigate the effect of willpower depletion on procrastination and find that depleted subjects are more likely to postpone filling out a questionnaire for a day.

In this paper we extend the economic literature on willpower depletion to another aspect of individual decision making. In particular, we use a laboratory experiment to investigate if depleting people's willpower increases their susceptibility to *framing effects*. Framing effects (or context effects) occur when people's decisions are affected by seemingly irrelevant aspects of decision problems (see e.g. Tversky and Kahneman, 1981; Levin, Schneider, and Gaeth, 1998; Druckman, 2001; De Martino et al., 2006). A popular explanation of framing effects argues that they are the result of the use of simplified decision rules, heuristics or 'system 1' processes that selectively process only a limited number of (possibly irrelevant) details of the decision problem (Tversky and Kahneman, 1974, 1981). System 1 processes can be overruled by higher level or 'system 2' processes, but this requires willpower (Pocheptsova et al., 2009). As a consequence, willpower depletion should increase susceptibility to framing effects.

In the experiment, we use a five-minute version of the well-known 'Stroop' task (Stroop, 1935) to deplete the willpower of half our participants. The Stroop task is followed by one of five secondary tasks. These tasks are an 'attraction' effect

[2] Alternative terms used for resource depletion and willpower depletion are ego depletion, cognitive depletion and psychological depletion, amongst others.

task, a ‘compromise’ effect task, an ‘anchoring’ effect task, a framed prisoner’s dilemma and a cognitive task. The first four tasks test for different types of framing effects, whereas the cognitive task allows us to test for differences in cognitive performance. This sequence of a Stroop task followed by a secondary task is repeated five times, such that participants go through all five secondary tasks exactly once.

By comparing framing effects in all four framing tasks between depleted and non-depleted participants, we can investigate if depleted individuals are indeed more susceptible to framing effects. The experiment also allows us to address at least three additional questions of interest. Firstly, the cognitive task allows us to investigate the effect of willpower depletion on cognitive performance. Secondly, the repeated nature of our set-up allows us to trace the effects of depletion over successive tasks. Thirdly, making participants go through four different framing tasks allows us to investigate if participants who are susceptible to one framing effect are more susceptible to other framing effects as well.

All in all, however, we find little evidence that depleted participants are more susceptible to framing. Framing effects appear in the attraction, compromise and anchoring tasks but do not differ between depleted and non-depleted participants. Furthermore, we do not find a depletion effect on the performance of participants in the cognitive task or differential willpower depletion effects over successive tasks; we also find little evidence of between-task correlations in framing effects. In many ways, these results are quite striking since the depleting effects of the Stroop task have been documented in many studies and especially since the effect of depletion on the attraction effect and compromise effect has already been documented by Pocheptsova et al. (2009).[3] This suggests it may be important from a methodological perspective to critically examine the conditions under which willpower depletion is likely to have an effect on performance and under which it is not.

The structure of the rest of the paper is as follows. In section 2 we present the design, where we describe each task in greater detail. In section 3 we describe the research questions, after which we provide the results in section 4 and conclude in

[3]Pocheptsova et al. (2009) describe a series of willpower depletion experiments, two of which examine the compromise effect and one of which examines the attraction effect. To deplete participants they use different methods in different experiments, including the Stroop task for one of the compromise effect experiments. We will say more about the differences and similarities between our study and Pocheptsova et al. (2009) in the following sections.

section 5.

4.2 Experimental Design

The goal of this study was to investigate the effect of willpower depletion on participants' susceptibility to framing effects. We therefore conducted an experiment where subjects first participated in a task that either depleted or did not deplete their cognitive resources or willpower. Upon completing this task participants then went through a secondary task in which framing effects had previously been shown to play a role. In the remainder of this section, we first discuss the depletion task and then discuss all five secondary tasks individually, followed by a short description of the experimental procedure.

4.2.1 Depletion Task

For the depletion task, we used the Stroop (1935) task. Originally created to test psychological interference theory, the Stroop task has since been used in many applications (see MacLeod, 1991 for an overview of the first 55 years of applications). Importantly, the Stroop test has also been used as a way to deplete participants' willpower in several studies (see for example Webb and Sheeran, 2003; Pocheptsova et al., 2009; Burger, Charness, and Lynham, 2011). We used a computerized version of the Stroop task, in which participants were asked to indicate the font color a color name was printed in. There were five possible font colors and color names: blue, red, yellow, orange and purple. We adopted two different versions (or treatments) of the Stroop task to experimentally vary the level of willpower depletion in participants. For the control treatment, the font colors were always identical to the color names. For example, the word 'blue' would always be printed in blue letters. For the depletion treatment the font color and color name were randomly matched, so that they were identical in only on average one third of all cases.

The depletion treatment of the Stroop task has been argued to lead to depletion in the following way. When a color word is written in a different font color, our initial impulse is to read the semantic meaning of the word. Naming the font color instead requires us to override our initial tendency to read the color word. In terms of the willpower paradigm, overriding our initial (system 1) tendency to say the color word is a form of self-regulation (system 2) that requires willpower to be undertaken. As a consequence, taking part in the depletion treatment of the Stroop task lowers the amount of remaining willpower that can be used in subsequent tasks. An alternative but closely related way to look at it is that overriding our initial tendencies is a cognitively demanding activity, which leaves fewer cognitive resources for subsequent tasks. Importantly, the control version of the Stroop task requires no self-regulation and fewer cognitive resources, such that control participants should be less depleted than participants in the depletion treatment.

Over the course of the experiment, every participant went through five Stroop tasks; each participant faced the same version (control or depletion) of the Stroop task every time. For each Stroop task, participants faced a random sequence of words drawn by the computer, with randomly matched font colors for the depletion treatment. The color words appeared in the middle of the computer screen; subjects could indicate the color of the word by pressing the corresponding key on the keyboard. After they had given their response, they received a new word after pressing the spacebar and waiting for approximately 0.55 seconds.[4] Each Stroop task lasted precisely five minutes. Incentives were such that participants received one cent for each correct response and were deducted two cents for each incorrect answer. Feedback on the number of correct and incorrect answers was only provided at the end of the experiment.[5]

4.2.2 Secondary Tasks

After every depletion task, participants took part in one of five secondary tasks. Four of these tasks were framing tasks, and one was a cognitive task. The order of the tasks was randomized between participants to minimize potential ordering effects. Each secondary task took six minutes in total; the first two minutes were reserved for an instructions screen, the remaining four minutes were reserved for the task itself.

[4] The time for a new word to appear was uniformly random between 0.3 and 0.8 seconds.

[5] A screenshot of the Stroop task can be found in Appendix C.

Attraction Effect Task

The attraction effect or asymmetric dominance effect (Huber, Payne, and Puto, 1982) occurs in decision problems when adding an asymmetrically dominated alternative to a choice set increases the share of the dominating alternative. For example, Huber, Payne, and Puto (1982) asked subjects to choose between three types of cars that differed on two dimensions. Two of the options were roughly equally attractive, with one option having a better value on one dimension (e.g. price) and the other option having a better value on the other dimension (e.g. fuel efficiency). The third (or 'decoy') option was similar to one of the first two options, but was strictly dominated by this option. In the car example, the decoy option was both less fuel efficient and more expensive than one of the first two options. Huber, Payne, and Puto demonstrated that people are more inclined to choose the alternative that dominates the decoy option. The attraction effect has been replicated in a large number of studies in a variety of contexts.[6]

The attraction effect task in this experiment was based on Herne (1999), who used a lottery task that allows for an incentivized test of the attraction effect. In the attraction effect task, participants had to choose between three binary lotteries which varied along two dimensions: probability of winning and prize. Of the three lotteries, two lotteries (say lotteries x and y) had equal expected value; one lottery had a higher prize and the other lottery a higher probability of winning. The third lottery was a decoy lottery which was dominated on both the probability and the prize dimension by either lottery x or lottery y (and not the other). The attraction effect occurs if participants are more likely to choose the lottery that dominates the decoy. Herne (1999) showed that many people are indeed susceptible to the attraction effect in this task. In total, participants faced five sets of lotteries consisting of three lotteries each; see table 4.12 in Appendix A.[7] We used a between subjects design, so that for half the participants lottery x was the dominating option and for the other half it was lottery y. The outcome of the lotteries was only revealed at the end of the experiment. The expected value of lotteries x and y was equal to 45 cents in all choice menus; the expected value of the decoy lottery was 36 cents.

[6] Ok, Ortoleva, and Riella (2011) mention more than 20 studies documenting the attraction effect in economics, psychology, marketing and political science.

[7] The lotteries are identical to those used by Herne (1999) except that all prizes were multiplied by 1.5.

Compromise Effect Task

The compromise effect (Simonson, 1989; Simonson and Tversky, 1992) occurs in decision problems when a given alternative is chosen more often when it is presented as the middle alternative compared to when it is an extreme alternative. For example, in Simonson and Tversky, subjects had to make a hypothetical choice between three of four calculator batteries varying in expected life (in hours) and the probability of corrosion. They found that subjects were more likely to pick the battery with the second lowest probability of corrosion if this battery was the middle option, i.e. if the choice set included the battery with the lowest probability of corrosion as well. The compromise effect has been supported empirically by a range of studies (see e.g. Bernatzi and Thaler, 2002; Busemeyer et al., 2007; Müller, Kroll, and Vogt, 2010, 2011).

The compromise effect task we used in this experiment was also based on Herne (1999). Similar to the attraction effect task, participants had to choose between three lotteries varying in two dimensions. Two lotteries (say x and y) were identical for all participants, whereas the third (or decoy) lottery differed between decision frames. In one frame, a decoy lottery was included that made lottery x the middle option, whereas in the other frame the decoy lottery that was included made lottery y the compromise. Herne (1999) provides evidence that suggests that some people may be susceptible to the compromise effect in this task. As with the attraction effect task, participants faced five different sets of lotteries and we used a between subjects design. Table 4.13 in the appendix shows the lotteries we used in the experiment.[8] The outcome of the lotteries was only revealed at the end of the experiment. For all lotteries (including the decoys), the expected value was 45 cents.

Anchoring Task

The anchoring effect task was based on Tversky and Kahneman (1974), who asked their subjects to estimate various statistics stated in percentages, including for example the percentage of African countries in the United Nations. They then randomly drew a number between 0 and 100 by spinning a wheel of fortune in their subjects' presence and asked their subjects to specify if the statistic to

[8] As with the attraction effect task, the lotteries are identical to those used by Herne (1999) except that all prizes were multiplied by 1.5.

be estimated was higher or lower than the randomly drawn number. Subjects were then also asked to give a precise estimate of the statistic. Even though participants knew that the randomly drawn number provided no information on the true value of the statistic, participants who were subjected to a larger number (or 'anchor') on average still provided a higher answer to the second question. Since Tversky and Kahneman, the anchoring effect has been consistently replicated both in a similar setting and in several generalizations (see e.g. Epley and Gilovich, 2001; Mussweiler and Strack, 2001; Ariely, Loewenstein, and Prelec, 2003; or see Furnham and Boo, 2011 for a recent literature review).

The anchoring set-up we used in this experiment was very similar to that of Tversky and Kahneman (1974). Participants had to answer five rounds of two questions. In each round, the first question asked participants if a particular statistic was larger or smaller than a randomly drawn integer between 0 and 1000, the second question then asked them to estimate the true value of this statistic. An anchoring effect occurred if this estimate was correlated with the randomly drawn number (or anchor). All five statistics were chosen such that all participants were expected to have at least some idea of their values, but would be unlikely to know any value exactly. For example, one of the statistics asked participants for the distance between Paris and the Dutch city of Eindhoven in kilometers.[9] The random integer appeared on screen as if generated by a slot machine to emphasize its randomness. A correct answer to the first question yielded 25 cents. For the estimate of the statistic, participants earned 1 euro minus one cent times the difference between their answer and the true value of the statistic (with a minimum of zero). No feedback on the correct answers was provided during the experiment, although all answers and earnings were provided at the end.

Prisoner's Dilemma Task

Framing effects have also been shown to occur in prisoner's dilemmas. In the experiment we used a symmetric prisoner's dilemma task, for which we varied the framing in two ways. Firstly, we followed Ross and Ward (1995) in labeling the

[9]The other statistics were: the highest measured top speed of the fastest road car in the world in kilometers per hour, the number of years ago the Dutch city of Alkmaar received city rights, the height of the tallest building on earth in meters and the number of inhabitants of Vatican City.

prisoner's dilemma as either a 'community game' or a 'banker game'.[10] Ross and Ward found that subjects were more likely to choose the cooperative option when the game was presented as a community game.

Secondly, we used the decomposed prisoner's dilemma framing of Pruitt (1967). Prisoner's dilemmas are typically presented to subjects by stating the payoffs for both players that correspond to each *outcome*. Pruitt instead presented prisoner's dilemmas to his subjects by stating the payoffs as a function of a subject's *choice*. Two prisoner's dilemmas can have identical payoffs for outcomes but different payoffs for choices. Prisoner's dilemmas 1a and 1b of table 4.1 -taken from Pruitt (1967)- give an example of two such prisoner's dilemmas. Pruitt showed that the percentage of cooperative choices varies as a function of how the choices were represented, even if the corresponding outcomes were identical.

As with all secondary tasks, the prisoner's dilemma task started with two minutes of instructions time. During the instructions, subjects were told that they would be anonymously matched to another participant in the experiment for the prisoner's dilemma only, and that they would only learn the choices of the other participant at the end of the experiment. During the instructions, the banker-vs-community frame was already visible in the page heading; half the participants were assigned to each label. The banker or community heading persisted into the four minutes of decision time. Participants played the three prisoner's dilemmas displayed in table 4.1. The sequence was either 1a/2/1b or 1b/2/1a. Prisoner's dilemmas 1a and 1b were the two different representations of the same prisoner's dilemma representing the (decomposition) framing of Pruitt, allowing us to compare cooperation rates within subjects. Prisoners dilemma 2 was different from the other two in terms of payoffs and was placed in between to make the similarity between dilemmas 1a and 1b less obvious.

Cognitive Task

Finally, we included a cognitive task to check for differences in cognitive functioning between depleted and control participants. The task itself consisted of adding three two-digit integers, which has been used in several studies before (e.g. Sloof and Van Praag, 2010). The three numbers were presented vertically to make it

[10] Ross and Ward (1995) used 'Wall Street game' rather than 'banker game'. Since 'Wall Street' does not translate into Dutch very easily, we elected to use a word with similar connotations ('banker') instead.

Table 4.1: Payoffs in the Prisoner's Dilemma Task

	Prisoner's Dilemma 1a		Prisoner's Dilemma 1b		Prisoner's Dilemma 2	
Payoffs	You	Other	You	Other	You	Other
Cooperate	+60	+60	+120	+0	+40	+0
Defect	+120	-60	+180	-120	+100	-100

Notes. This table gives the payoffs in cents for the prisoner's dilemma task. The sequence of prisoner's dilemmas was 1a/2/1b for half the participants and 1b/2/1a for the other half. Note that prisoner's dilemmas 1a and 1b have identical payoffs for every outcome.

easier for subjects to do the calculations. Participants could do as many such calculations as they liked and were capable of doing within the allotted four minutes. Every correct answer was worth 10 cents and every incorrect answer cost a participant 10 cents. No feedback was given during the experiment about whether a given answer was correct; subjects only learned the number of correct and incorrect answers at the end of the experiment.

4.2.3 Procedure

The experiment was computerized using PHP/MySQL software and conducted at the CREED laboratory of the University of Amsterdam. The experiment started with an initial set of instructions that explained the Stroop task. As part of the instructions, subjects had to answer two check-up questions and had to perform 10 practice trials of the Stroop task to familiarize themselves with the Stroop interface. The instructions also contained some information about the general structure of the experiment. Specifically, participants were told that the Stroop task would be repeated five times and that each repetition would be followed by a different task. They were also told that the other tasks would be explained later and that they would have five minutes for every Stroop task and six minutes for each other task. Apart from the practice trials for the Stroop task, all instructions were identical regardless of whether subjects were in the depletion or control treatment. A copy of the initial instructions was provided to participants as well, so that they could refer back to the instructions over the course of the experiment if necessary.[11]

The first Stroop task started after all participants in a session had finished the initial instructions. After the five minutes for the first Stroop task had expired,

[11] A translation of all instructions can be found in Appendix B.

participants moved on to their first secondary task. Each participant received a short set of instructions on screen; they had exactly two minutes to read these instructions and four minutes to complete the task. During all secondary tasks, the remaining time was displayed on a clock on the bottom right of the screen. After the fifth and final secondary task had been completed, the computer matched the choices of two participants for the prisoner's dilemma task and calculated the outcome of the lotteries for the attraction and compromise tasks. Participants then received a detailed overview of their results including their earnings separately for each task. After having reviewed their earnings, they were asked to fill out a short questionnaire that contained background questions, five questions per task asking if they thought the task had been interesting, difficult, tiresome, fun or boring as well as several general questions related to participants' willpower and emotional states.[12] After filling out the questionnaire, participants received their earnings and could leave the laboratory. All in all, the experiment lasted for approximately an hour and a half including the instructions, questionnaire and payment.

4.3 Research Questions

In this study, we seek to investigate the effects of willpower depletion on participants' susceptibility to framing effects. In particular, we argue that depleted participants are more susceptible to framing effects than control participants. In this section, we go into what this hypothesis implies for the four framing tasks. We also discuss three additional research questions related to the cognitive task, changes in depletion effects over successive tasks and correlations between different framing effects.

We expect depleted participants to be more susceptible to framing effects in the *attraction effect* task. Pocheptsova et al. (2009) argue that the attraction effect is the result of a simple and largely automated (system 1) heuristic that can potentially be overwritten by a more thoughtful (system 2) decision process.

[12] These questions asked them how frustrated, happy, satisfied, irritated, content, disappointed and tired they felt. We also asked them how hungry they were, how often they practiced sport, drank alcohol, smoked and how long they had slept the previous night and we asked them how likely they would be to study, smoke, do something annoying or fill out important forms immediately after the experiment.

However, overriding the simple heuristic requires willpower or cognitive effort. To the extent that depleted participants have less willpower available, they should then be less likely to override the simple heuristic and thus be more susceptible to the attraction effect. Pocheptsova et al. (2009) depleted participants using an attention regulation task wherein participants in the depletion treatment were asked not to look at phrases that appeared at the bottom of a video screen and control participants received no such instructions. They then asked participants to make a hypothetical choice between three apartments to investigate the attraction effect. They find a strong attraction effect for depleted participants, but no attraction effect for controls. In line with these results, we expect depleted participants to be more susceptible to the attraction effect in this study as well.

For the *compromise effect*, the argument is slightly more subtle. Simonson and Tversky (1992) argue that the compromise effect is the result of an extended notion of loss aversion. In compromise effect tasks with three options, the middle option has a small advantage (on one dimension) and a small disadvantage (on the other dimension) relative to the extreme options, whereas both extreme options have a large advantage and a large disadvantage with respect to the other extreme option. If disadvantages loom larger than advantages, this should make the middle option most attractive. If reference dependence and loss aversion are lower level (system 1) decision processes that can be overruled by exercising willpower or cognitive effort (as argued by Kahneman, 2003), we expect depleted participants to be more susceptible to the compromise effect. We get a similar prediction if we regard “choosing the middle option” as a simple heuristic (as argued by for example Bettman, Luce, and Payne, 1998) that can be overruled by a more sophisticated decision process. Thus, we expect depleted participants to be more susceptible to the compromise effect.

It must be noted, however, that Pocheptsova et al. (2009) found exactly the opposite effect. They investigated the effect of depletion on the compromise effect using two separate experiments. In one experiment they depleted participants using the video task described above, in the other experiment they used a 40 trial non-incentivized version of the Stroop task. They found that the compromise effect is actually *weaker* in depleted participants. They argue that this makes sense if the compromise effect is the result of a higher level decision strategy that requires significant cognitive resources to be used. We favor the idea that the compromise effect is based on lower level processes, but acknowledge that the opposite argument can also be made.

We expect depleted participants to be more susceptible to the *anchoring effect* as well. As first suggested by Tversky and Kahneman (1974), an anchoring effect appears when participants use the anchoring-and-adjustment heuristic but adjust insufficiently. People who use this heuristic make estimates by starting from an initial value that is subsequently adjusted to yield the final estimate. When this adjustment is insufficient, different starting values may lead to different final estimates. We expect depleted participants to have fewer cognitive resources or less willpower available to override the anchoring-and-adjustment heuristic; as a consequence depleted participants should be more susceptible to the anchoring effect.

For the *prisoner's dilemma* task we also expect depleted participants to be more susceptible to both framing effects. Calculating the payoffs associated with the four possible outcomes based on the payoffs associated with choices takes cognitive effort, which depleted subjects may not be willing or able to provide. On the other hand, basing a choice on the community-versus-banker title or on the payoffs associated with *choices* (as opposed to *outcomes*) does not require any computation. Thus we predict that both framing effects will appear more strongly in depleted participants than in controls.

We can exploit the structure of this experiment to investigate at least three more questions of interest. Using the *cognitive task*, it is possible for us to investigate the effect of willpower depletion on cognitive performance. If cognitive resources are in fact depleted in the Stroop task as well, depletion should also lead to reduced cognitive performance.

The repeated nature of the experiment allows us to investigate if and how the effect of depletion on framing effects changes over successive tasks. Repeatedly administering the Stroop task can have several effects. For example, it is possible that participants in the depletion treatment get more strongly depleted over every repetition of the Stroop task, leading to a large difference with controls in later rounds. Conversely, if control participants are also being slightly depleted by every Stroop task and depleted participants are fully depleted in the first Stroop task, the difference between depleted participants and controls could also grow smaller in later tasks.

Finally, we can use the fact that participants took part in multiple framing tasks to examine if participants who display a framing effect on one framing tasks are also more susceptible to a framing effect on another framing task. We can also investigate if such correlations between tasks are larger for depleted or control participants.

4.4 Results

We ran six sessions of the experiment in November 2010, in which a total of 104 subjects participated. Half of these subjects were assigned to the depletion treatment, the other half were controls. Sixty-five percent of participants were male, with an average age of 21.17 years and the majority (68%) studying economics and/or business. Mann-Whitney tests on social demographic variables show no significant difference in characteristics between the two treatment groups.

In the remainder of the section we first present the results of the Stroop task as a manipulation check. We subsequently present the results separately for each of the four framing tasks, where we both investigate if the framing effect occurred at all and if it occurred to a different degree for depleted participants. We then look at the cognitive task, followed by an overview of how the size of the depletion effects develop over time and a look at correlations between tasks. For all statistical tests, each participant is treated as an independent observation.

4.4.1 Stroop Task

As a manipulation check it is useful to first compare the results of depleted and control subjects on the Stroop task. Recall that participants in the depletion treatment did the incongruent version of the Stroop task, which was more difficult than the control version. Thus, we expected depleted participants to perform worse on the Stroop task overall. Table 4.2 and figure 4.1 show that this is indeed the case; depleted participants earned less on the Stroop task overall than control participants. The second and third row of table 4.1 show that this difference is driven exclusively by depleted participants having a slower response time; the number of mistakes does not differ between treatments.[13] Moreover, according

[13]Interestingly, participants in our study responded almost four times as quickly as participants in Pocheptsova et al. (2009). This effect could be due to learning, increased incentives and

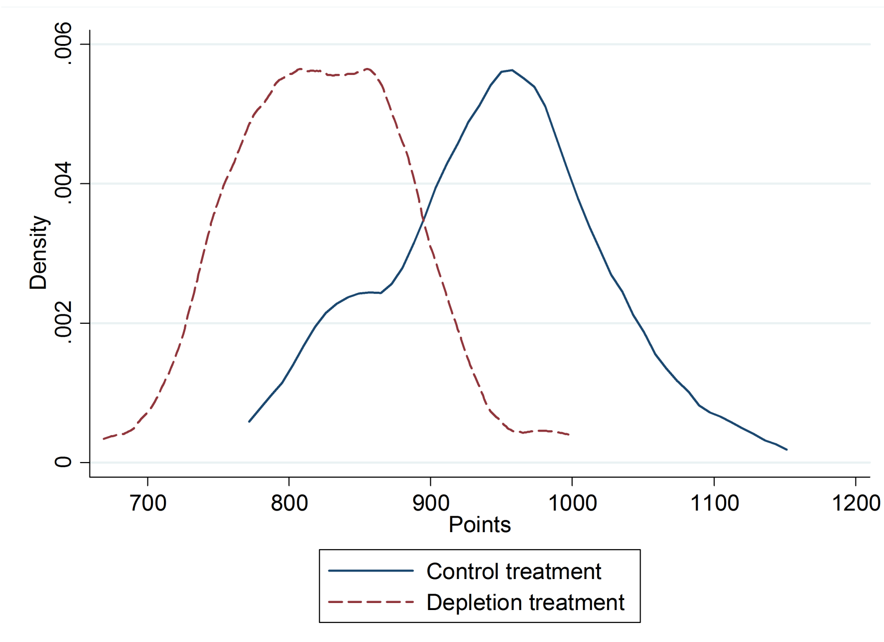

Figure 4.1: The Distribution of Earnings in the Stroop Task (in Cents)

Notes. The figure plots the smoothed density (Epanechnikov kernel) of the earnings of participants over all five Stroop tasks separately for each treatment. The bandwidth was equal to 25.2 for the depletion treatment and 28.5 for the control treatment.

Table 4.2: Stroop Task Statistics

	Control	Depletion	Difference	P-value
Average total Stroop task earnings	944.06 (73.50)	824.19 (61.69)	119.87	$p<0.0001$
Average total number of mistakes	13.13 (12.21)	13.52 (7.83)	0.39	$p=0.2917$
Average reaction time in seconds	0.747 (0.081)	0.960 (0.107)	0.213	$p<0.0001$

Notes. The table gives the total earnings, the number of mistakes and the average reaction time in the Stroop task separately for both treatments; the bracketed numbers are standard deviations. P-values are calculated using Mann-Whitney tests.

to the questionnaire participants in the depletion treatment found the Stroop task more exhausting, more interesting, more fun, less boring (at the 5% level or better) and more difficult (at the 10%) level than participants in the control treatment. [14] All in all, this is in line with previous studies and suggests that the Stroop manipulation was successful.

4.4.2 Attraction Effect Task

For the attraction task, participants went through five decision problems, in each of which they had to choose between three lotteries. One of these lotteries was always the decoy lottery, which had both a higher risk and lower earnings than one of the other lotteries. The attraction effect predicts that participants are more likely to choose the lottery that dominates the decoy lottery. To investigate if this was the case, we compare the fraction of times the dominating lottery was chosen over all five decision problems to the fraction of times the non-dominating lottery was chosen. In line with the attraction effect and Herne (1999), table 4.3 shows that the dominating lottery was picked more often than the non-dominating lottery; this difference is significant for both the control treatment and the whole sample.

software differences or all three.

[14]In terms of the questionnaire variables, we also found a positive correlation between Stroop income and the answer to the question “in general as a person I would describe myself as easily influenced” (p=.003, Pearson correlation). Students of economics performed better on the Stroop task in the control treatment (p=.03, Mann-Whitney) but not in the depletion treatment (p=.3, Mann-Whitney). Males were better than females at the Stroop task in the control treatment (p=.046, Pearson correlation) but not in the depletion treatment (p=.68). Participants who performed well on the Stroop task in the depletion treatment were less likely to report they were going to study after the experiment (p=.048), more likely to exercise (p=.045), and were heavier smokers (p=.023).

Table 4.3: Attraction Effect Task Statistics

	Decoy	Dominating	Non-Dominating	P-value (Attraction Effect)
Control subjects	1.5%	56.9%	41.5%	.002
Depleted subjects	1.9%	53.1%	45.0%	.126
All subjects	1.7%	55.0%	43.3%	.001

Notes. The table gives the percentage of lottery sets participants chose the dominating, non-dominating and decoy lotteries respectively; this uses the results from all five lotteries. To compute the p-values, we take the fraction of dominating choices and non-dominating choices for every individual and use a Wilcoxon test to check if they are equal. Thus, the number of independent observations is 52 per treatment.

However, the attraction effect does not differ between treatments (Mann-Whitney, p-value=0.2994). This finding stands in stark contrast to Pocheptsova et al. (2009), who found that depleted participants were 37 percentage points more likely to choose the dominant option than control participants. Here, depleted participants are actually 4 percentage points *less* likely to choose the dominant lottery, with a 95% confidence interval of [-11,3] percentage points. All in all, we find evidence of an attraction effect, but this effect does not differ between depleted participants and controls.

4.4.3 Compromise Effect Task

For the compromise effect task, participants also went through five decision problems with three lotteries, similar to the attraction effect task. We used a between-subjects design where the two main lotteries (say x and y) were always the same and the third lottery (or ‘decoy’) differed between participants. The decoy was either strictly riskier or strictly safer than both other lotteries, but had the same expected value. The compromise effect predicts that participants are more likely to choose option x or y if it is the middle option. To investigate if this was the case, we compared the percentage of trials the ‘compromise’ lottery was chosen and compared this percentage to the percentage of trials the other main (non-decoy) lottery was chosen. Table 4.4 shows that participants were more likely to choose the non-compromise lottery than the compromise lottery. Thus, we find a reverse compromise effect, which is significant (with a p-value of .002 for the whole sample). The compromise option was also chosen significantly less often than the decoy (Wilcoxon, p-value<0.001), whereas the difference between the percentage of decoy choices and non-compromise choices is not significant.

Table 4.4: Compromise Effect Task Statistics

	Decoy	Compromise	Non-Compromise	P-value (Comp. Effect)
Control subjects	42.7%	23.5%	33.9%	.012
Depleted subjects	38.5%	24.2%	37.3%	.045
All subjects	40.6%	23.9%	35.6%	.002

Notes. The table gives the percentage of trials participants chose the compromise, non-compromise and decoy lotteries respectively; this uses the results from all five lotteries. To compute the P-values, we take the fraction of compromise choices and non-compromise choices for every individual and use a Wilcoxon test to check if they are equal. Thus, the number of independent observations is 52 per treatment.

Table 4.5: Compromise Effect Task Statistics by Riskiness

	Compromise	Most risky choice	Least risky choice
Control subjects	23.5%	27.7%	48.9%
Depleted subjects	24.2%	25.0%	50.8%
All subjects	23.9%	26.4%	49.8%

Notes. The table gives the percentage of trials participants chose the compromise, most risky and least risky lotteries respectively; this uses the results from all five lotteries.

The finding of a reverse compromise effect is quite surprising, since it is contrary to the results of the studies we previously mentioned including Herne (1999), who used the same lottery task.[15] However, this apparent puzzle can be explained by risk attitudes. Table 4.5 shows that the least risky lottery was chosen in nearly half of the decision problems, whereas the middle option and the most risky option were both chosen approximately 25% of the time.[16] The least risky lottery was chosen significantly more often than either alternative (Wilcoxon, p-value<.0001), which suggests that at least some choices were driven by risk aversion and that risk aversion is more common than risk seeking, which is intuitive.[17]

[15]Herne (1999), however, used a within-subjects design that allows her to look only at the small number of subjects (at most 12 and sometimes as few as 2) who (a) never chose the decoy and (b) did not choose the same lottery regardless of the decoy. She reports a compromise effect (at the 10% level or better) among these subjects even if there are only 2 observations (table 4.4, choice set e, p=.0784).

[16]We also observe a preference for the least risky lottery in the attraction effect task, where the least risky lottery is chosen 62.7% of the time and the riskiest option is chosen 35.6% of the time.

[17]Interestingly, we find no gender differences in the percentage of trials the risk averse option was chosen for either the attraction effect task or the compromise effect task. This stands in contrast with a recent review of Croson and Gneezy (2009), who concluded that most studies suggest that women are on average more risk averse.

The finding that many participants tended to choose the least risky lottery does not change the fact that several participants still picked the compromise. Thus, in principle it would still be possible for the compromise to be chosen more often in the depletion treatment by subjects who were not influenced by risk attitudes. However, this does not seem to be the case (Wilcoxon, p-value=0.9868), which contrasts with Pocheptsova et al. (2009) who found an effect in two separate experiments.[18] Note, however, that Pocheptsova et al. (2009) used a hypothetical decision task in which risk attitudes do not matter. To sum up, we find a reverse compromise effect which is driven by risk aversion and does not differ between treatments.

4.4.4 Anchoring Effect Task

For the anchoring effect task, participants had to answer five rounds of questions. In each round, the first question asked participants if a particular statistic was larger or smaller than a randomly drawn number, the second question then asked them to estimate the true value of this statistic. An anchoring effect occurred if the estimate of the true value was correlated with the randomly drawn number (or anchor). We used OLS to investigate whether the value of the anchor affected the answer to the second question. In the regression, we also controlled for the true answer and included a dummy for depleted participants. Moreover, we interacted both the true answer and the anchor with the dummy to test if anchoring effects differed between depleted and control participants.

Table 4.6 displays the results of this regression. Two things are apparent. Firstly, there is strong evidence of an anchoring effect (p=.000); increasing the anchor by 100 increases the estimate of the statistic by 28.[19] Secondly, there is no evidence that the size of the anchoring effect differed between depleted and control participants. If anything, the anchoring effect was slightly smaller (.202 instead of .280) among depleted participants; however this difference is not statistically significant (p=.342).

[18] Additionally, there is no difference in the number of risk averse choices between depleted and control participants.

[19] The anchoring effect is significant at the 1% level for three statistics taken individually; these are the city rights, the car speed and the Vatican city statistic. The anchoring effect is significant at the 10% level for the tallest building statistic, and not significant (p=.186) for the distance between Eindhoven and Paris.

Table 4.6: Anchoring Effect Task Estimates

	Dependent variable: estimate of the statistic		
	Coefficient	Std Error	P-value
Anchor	0.28	0.06	0.000
True answer	0.20	0.07	0.003
Depleted	−52.50	59.87	0.380
Depleted×Anchor	−0.08	0.08	0.342
Depleted×True answer	0.11	0.09	0.240
Constant	243.25	42.28	0.000

Notes. This table gives the results of a linear regression of the estimate of the true value of the statistic (i.e. the answer to the second question in the anchoring task) on a constant, a treatment dummy ('Depleted'), the random anchor, the true answer and interaction terms between the treatment dummy and the other two variables. The regression uses the data from all five statistics in the anchoring task. Standard errors are clustered at the participant level; thus there are 104 independent observations.

Table 4.7: Prisoner's Dilemma Task Statistics

% Cooperation	PD 1a	PD 1b	PD 2
	Control Participants		
Community label (n=26)	30.8%	46.2%	19.2%
Banker label (n=26)	46.2%	50.0%	26.9%
	Depleted Participants		
Community label (n=33)	54.5%	42.4%	30.3%
Banker label (n=18)	33.3%	50.0%	16.7%

Notes. This table gives the percentage of cooperative choices for the three prisoner's dilemma tasks (PDs) separately for both treatments. Prisoner's dilemmas 1a and 1b were either the first or third prisoner's dilemma participants received, the second prisoner's dilemma was always the same for all participants, see table 1 for more details.

4.4.5 Prisoner's Dilemma Task

In the prisoner's dilemma task, participants had to choose between a cooperative and a non-cooperative choice in three separate prisoner's dilemmas. Recall that there are two framing effects that could occur. First, there was a between-subject difference in labeling: half the participants had the game presented to them with the label 'community game' and the other half were told they were playing the 'banker' game. Second, there was a within-subject varying in framing à la Pruitt (1967). In particular, all participants were presented with the same prisoner's dilemma in two different representations as in table 4.1.

Table 4.7 gives the cooperation rates for the two composition frames (prisoner's dilemmas 1a and 1b) and the two labels separately for depleted and control participants.[20] Within either depleted or control participants, label framing effects are observed if cooperation rates differ between rows and composition framing effects are observed if cooperation rates differ between the first two columns. Overall, prisoner's dilemma 1a leads to slightly higher cooperation rates than prisoner's dilemma b in three of four cases. Similarly, the banker frame leads to slightly higher cooperation rates than the community frame in three of four cases.[21] However, none of the differences in cooperation rates are statistically significant either overall or within specific subgroups based on willpower depletion, label and/or composition frame.

To investigate the difference in framing effects between depleted and control participants, we ran a set of probit regressions where we regressed the indicator for a cooperative choice on a depletion dummy, a framing dummy and the interaction between framing and depletion. For the label framing, we ran this regression separately for all three prisoner's dilemmas as well as for all three combined. Similarly, for the composition framing we ran this regression for prisoner's dilemmas 1a and 1b separately as well as the combined data. Table 4.8 reports the results for the only regression that yielded a statistically significant treatment effect. The results suggest that changing the label from 'community' to 'banker' in prisoner's dilemma 1a has a positive effect on the share of social choices in the control treatment and a net negative effect in the depletion treatment. This treatment-framing interaction effect is significant at the 10% level. However, the effect is small and the main framing effect is significant for neither depleted nor control participants. Moreover, there was no significant depletion effect in the other 6 specifications we tested. Thus, there is little evidence that willpower depleted participants are differently affected by framing effects than controls.[22]

[20]For one participant in the banker/depleted group, the prisoner's dilemma task did not set up correctly due to a software error, so her results were dropped from the analysis of the prisoner's dilemma task.

[21]Note that participants were statistically significantly less cooperative in prisoner's dilemma 2, but this is the result of different payoffs rather than a framing effect.

[22]Interestingly, depleted participants were on average more cooperative than controls in this specification. However, this effect did not appear in any of the other 6 specifications we tested and was also not robust to removing the interaction term from the regression.

Table 4.8: Prisoner's Dilemma 1a Probit Estimates

	Dependent variable: cooperative choice (1=yes)		
	Coefficient	Std Error	P-value
Community frame	0.16	0.14	0.26
Depleted	0.56	0.23	0.01
Community frame×Depleted	−0.37	0.20	0.07

Notes. This table gives the results of a probit regression of a dummy for a cooperative choice on a constant, a framing dummy, a treatment dummy ('Depleted') and the interaction between the treatment and framing dummies. The reported results are marginal effects. The regression uses the data from prisoner's dilemma 1a.

Table 4.9: Cognitive Task Statistics

	Control	Depletion	Difference	P-value
Cognitive task earnings	193.65 (72.84)	195.00 (73.26)	1.35	p=0.9637
Number of exercises completed	22.52 (7.01)	22.00 (7.67)	0.52	p=0.6231
Number of mistakes	1.58 (1.76)	1.25 (1.41)	0.33	p=0.4437

Notes. The table gives the earnings, the number of completed exercises and the number of mistakes in the cognitive task separately for both treatments; the bracketed numbers are standard deviations. P-values are calculated using Mann-Whitney tests.

4.4.6 Cognitive Task

Examining the cognitive task allows us to investigate the effect of the Stroop task on subsequent cognitive performance. Recall that in this task participants had to solve addition problems consisting of three two-digit numbers. Table 4.9 shows the average earnings in cents from the cognitive task per treatment. The difference between the performance of willpower depleted and control participants is not significant. The number of completed problems and the number of mistakes also did not differ between treatments.

4.4.7 Depletion over Successive Tasks

Our design also allows us to investigate how the effects of willpower depletion on framing develop over successive repetitions of the Stroop task (or rounds). Before turning to the secondary tasks, we first examine if performance on the Stroop task also changed over successive repetitions. Table 4.10 shows that participants do worse in the first Stroop task in both treatments and that performance in general increases slightly for each repetition of the Stroop task. Importantly, the difference in performance between treatments stays roughly constant and is

Table 4.10: Stroop Task Earnings per Round

Order	Control	Depletion	Difference	Mann-Whitney P-value
1	177.33 (16.37)	148.92 (13.16)	28.41	p<0.0001
2	188.73 (15.92)	164.83 (14.26)	23.90	p<0.0001
3	190.56 (17.23)	168.73 (14.02)	21.83	p<0.0001
4	193.23 (14.90)	170.40 (13.60)	22.83	p<0.0001
5	194.21 (16.31)	171.31 (14.56)	22.90	p<0.0001

Notes. The table gives the earnings for the Stroop task per treatment separately for each successive repetition of the Stroop task (or round); the bracketed numbers are standard deviations.

significant in all repetitions of the Stroop task (p<0.0001).

To trace framing effects over successive rounds, we need to use an individual-specific measure of the size of the framing effect for each task. For the anchoring task, we use the coefficient for the anchor variable in the regression of table 4.6 estimated at the individual level.[23] For the compromise and attraction task we use the fraction of times the compromise and dominating options were chosen respectively. Finally, for the prisoner's dilemma we include an dummy for whether a participant changed her choice between prisoner's dilemmas 1a and 1b. We also include performance on the cognitive task (in euros) to check for differences in cognitive performance over different rounds. We then take the difference between depleted and control participants on these respective measures to check if depletion had a different effect in different rounds.

Figure 4.2 plots the difference in the size of the framing effects between depleted and control participants. Overall, no clear time patterns can be discerned from the picture. Moreover, none of the depletion effects is significant in any round for any task (Mann-Whitney, p-value=0.101 for the cognitive task in round 1). However, not finding any differences between rounds is perhaps not surprising given the lack of a depletion effect overall.

4.4.8 Correlations between Tasks

Finally, our data also allow us to investigate if participants' behavior is correlated between each of the six experimental tasks. To investigate this question, we use

[23] To be precise, we estimate a regression of the estimate of the true value of the statistic on a constant, the random anchor and the true answer. The depletion variable dummy not vary at the individual level and is therefore omitted.

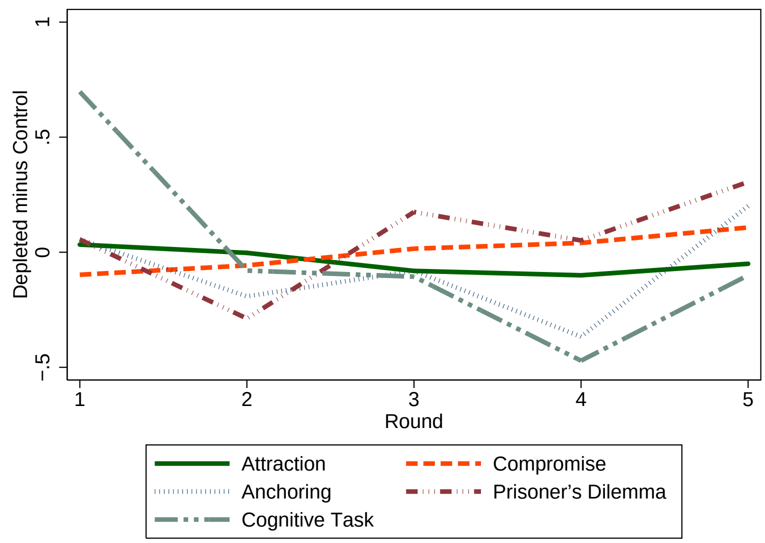

Figure 4.2: Depletion Effects over Successive Rounds

Notes. This figure gives the effects of willpower depletion on behavior in successive rounds of the game. A positive value indicates that depleted participants were more susceptible to the framing effect in this round (for framing tasks) or were more successful (for the cognitive task).

Table 4.11: Correlations between Tasks

	Overall		Depleted		Control	
	corr.	p-value	corr.	p-value	corr.	p-value
Stroop/Cognitive	0.126	0.201	0.388	0.002	0.002	0.990
Stroop/Prisoner's Dilemma	0.149	0.132	0.427	0.004	0.086	0.545

Notes. This table gives the Pearson correlation coefficients between performance in the Stroop task and the cognitive task and between performance in the Stroop task and a dummy that indicates whether a participant change their choice between prisoner's dilemmas 1a and 1b.

the same measures mentioned in the previous section to measure framing effects in the framing tasks. We correlate these measures at the individual level with each other as well as overall income on the cognitive task and Stroop tasks.

Table 4.12 displays the correlation coefficients for which we found a significant effect at the 5% level or better among either the whole sample or only the depleted or control participants.[24] Most intuitive is the positive correlation between performance on the Stroop task and performance on the cognitive task. Interestingly, this correlation is present only among depleted participants. This suggests that the depletion version of the Stroop task is also a cognitively demanding task (whereas the control task is not) so that subjects who excel at cognitive tasks perform well both in the summation task and in the depletion version of the Stroop task. There is also a correlation between the Stroop task and the prisoner's dilemma. Depleted participants who performed better on the Stroop task were also more likely to change their choice between prisoner's dilemmas 1a and 1b. Both these effects are significant at the 1% level. However, of all 45 correlations we computed, only two are significant at the 5% level or better. Thus, overall we find little evidence that behavior is correlated between the six experimental tasks.

4.5 Discussion

In this study we investigated the effect of willpower depletion on participants' susceptibility to framing effects. Framing effects appear in the attraction, compromise and anchoring tasks but do not differ between depleted and non-depleted participants. Furthermore we do not find a depletion effect on the performance of participants in the cognitive task or differential depletion effects over time. In many ways, these results are quite striking since the depleting effects of the Stroop task have been documented in many studies and in particular since the effect of depletion on the attraction effect and compromise effect has already been documented by Pocheptsova et al. (2009).

Our failure to find differences in framing effects between depleted and control participants could be due to our implementation of the Stroop task. In particular, it is possible that the Stroop task did not manage to more strongly deplete the

[24] In total, we computed 45 correlations. All results are identical if we use Spearman correlations or OLS instead. For the Prisoner's dilemma variable (which is a binary variable), the results are identical if we use a t-test, Mann-Whitney test or probit regression as well.

willpower of participants in the depletion treatment. In this light it is useful to compare our implementation of the Stroop task with previous studies. Most strikingly, to our knowledge this is the first study that uses an incentivized version of the Stroop task to deplete participants' willpower. Incentivizing the Stroop task may have increased the motivation of control participants to do well. The congruent Stroop task is less interesting than the incongruent version and doing well on the congruent Stroop task requires participants to maintain focus and concentration. Hence, it is possible that control participants were also depleted by the Stroop task.[25]

Our failure to find differences in framing effects could also be because of the secondary tasks used in the experiment. However, it is important to note that we found a framing effect in the anchoring, attraction and compromise tasks that was significant at the 1% level or better. Moreover, in all four secondary framing tasks there was sufficient scope for framing effects to vary between treatments. However, a crucial difference with most previous studies (such as Pocheptsova et al., 2009; Burger, Charness, and Lynham, 2011; Muraven and Baumeister, 2000) is that participants in our study were also incentivized to do well on the secondary task. Comparing our results to the literature, one possibility is that depletion effects disappear when participants are incentivized to do well on a secondary task.

If our finding that willpower depletion does not increase framing effects is caused by our implementation of the Stroop task, this suggests that the willpower depletion effects of the Stroop task are not robust to adding incentives, which would add an important insight to the willpower depletion methodology. If instead willpower depletion effects disappear when incentives are added to secondary tasks, this could indicate that willpower depletion is only relevant in contexts where the stakes are low or nonexistent. To the extent that economic behavior is typically subject to real incentives, willpower depletion may then not be as relevant to economic behavior as suggested by previous studies. We leave it to future research to investigate under what conditions the effect of willpower depletion on framing is or is not likely to appear.

[25] Additionally, the average number of trials for a single Stroop task (180) larger than the number of trials participants went through in Pocheptsova et al. (2009) (40); however, it is not considered to be particularly high in the literature. For example, Burger, Charness, and Lynham (2011) use 250 trials. Since we repeated the Stroop task five times, the total number of trials in our experiment is larger; however, note that we also found no treatment effect looking only at the results of the first secondary task for any of the four framing tasks.

Appendix 4.A Lotteries of the Compromise and Attraction Effect Tasks

Table 4.12: Lotteries of the Attraction Effect Task

		Probability of Winning (%)			Prize		
lottery	framing	lottery 1	lottery 2	lottery 3	lottery 1	lottery 2	lottery 3
1	1	40	80	35	112	56	103
1	2	40	80	75	112	56	48
2	1	65	30	70	56	150	64
2	2	25	30	70	144	150	64
3	1	40	75	80	112	48	56
3	2	40	35	80	112	103	56
4	1	70	45	50	64	80	90
4	2	70	65	50	64	56	90
5	1	30	25	60	150	144	75
5	2	30	50	60	150	72	75

Notes: This table gives the ten lotteries used in the attraction effect task. Every participants went through five lotteries; half the participants went through the lotteries corresponding to framing 1, the other half went through the lotteries of framing 2. The probability of winning for each lottery is expressed in percentages, the prize is expressed in cents. Participants who did not win the prize got zero instead.

Table 4.13: Lotteries of the Compromise Effect Task

		Probability of Winning (%)			Prize		
lottery	framing	lottery 1	lottery 2	lottery 3	lottery 1	lottery 2	lottery 3
1	1	40	80	30	112	56	150
1	2	40	80	90	112	56	50
2	1	30	70	80	150	64	56
2	2	30	70	20	150	64	225
3	1	50	30	20	90	150	225
3	2	50	30	60	90	150	75
4	1	90	80	70	50	56	64
4	2	60	80	70	75	56	64
5	1	50	30	40	90	150	112
5	2	50	60	40	90	75	112

Notes: This table gives the ten lotteries used in the compromise effect task. Every participants went through five lotteries; half the participants went through the lotteries corresponding to framing 1, the other half went through the lotteries of framing 2. The probability of winning for each lottery is expressed in percentages, the prize is expressed in cents. Participants who did not win the prize got zero instead.

Appendix 4.B Experimental Instructions[26]

This section contains the experimental instructions, separately for each task. The headings used in this section are different from the ones used in the experiment. The initial instructions were simply called 'instructions' in the experiment; they explained the structure of the experiment as well as the Stroop task. All secondary tasks had the heading "Instructions for Round 2/4/6/8/10", with the exception of the prisoner's dilemma, which had the heading "Instructions for Round 2/4/6/8/10: Community/Banker game".

4.B.1 Initial Instructions

Welcome to the CREED laboratory. Today's experiment will consist of ten rounds; every round will last between 5 and 6 minutes. Your earnings will be determined by the results of your choices in this experiment. This experiment makes use of a point system; 100 points translate to 1 euro at the end of the experiment. The number of points you earn in any round will only become known to you at the end of the experiment, so you will not receive any feedback about your earnings during the rounds.

On the next page you will receive the instructions for the first round of the experiment. The instructions for the other rounds will be given immediately preceding those rounds. Using the navigation tool at the top of the screen you can go back to any of the previous pages of the instructions. You will also receive a printed version of these instructions.

Round 1: Name the color

During the first round of the experiment a sequence of words will be shown on the screen. These words will be printed in the colors yellow, blue, purple, orange or red. Your task is to indicate the color the word is printed in. Each correct color will earn you 1 point, while each incorrect color will cost you 2 points. In other words, the more colors you name correctly, the more points you earn. This task will last for 5 minutes, after which you will automatically continue to next task.

[26]The original instructions (in Dutch) are available on request.

You can indicate the color of your choice using the keyboard. The relevant keys are g (for yellow), r (red), p (purple), o (orange) and b (blue).[27] The key-color combinations will also be visible at the bottom of the screen throughout the task. Be aware: if you press any other key than the key corresponding to the correct color, this will be counted as an incorrect assignment. This also holds for keys that do not refer to any color. On the next page you will have the opportunity to practice the task for 10 rounds with no payoff consequences.

4.B.2 Compromise/Attraction Effect Task

In this task the procedure is as follows. You are going to make a choice between 3 lotteries. These lotteries will vary in the amount of points you can win and the chance that you will win this amount of points. Below an example of a lottery choice menu is displayed. Each lottery will have only 2 possible outcomes. The lengths of the green and yellow parts of the rectangle symbolize the chances of either a green or yellow outcome. The exact probability and the amount of points you will win in this event are also printed in both the green and yellow parts of the rectangle.

In total you will make 5 lottery choices; each time you make a choice between 3 lotteries. After the instruction time has run out you will be automatically directed to the first lotteries; in other words there are no practice rounds for this task. This task itself will last for 4 minutes; During the task you can see how much time is left to answer on the clock in the lower right corner of the screen.

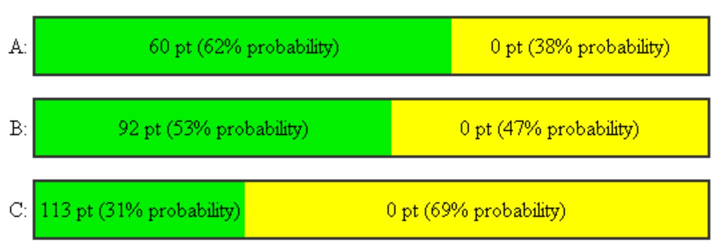

Figure 4.3: Example Used for the Compromise and Attraction Effect Tasks

[27]The Dutch word for yellow is 'geel', hence the key for yellow was 'g'.

4.B.3 Anchoring Effect Task

In this task the procedure is as follows: First a random number between 0 and 1000 is drawn. Then you will be asked to answer two questions. The first question is a yes/no question, in which a correct answer will earn you 25 points. To answer the second question you will have to enter a number; you will earn an amount of points depending on how close your answer is to the correct answer. Answering exactly correctly will earn you 100 points but each step you are removed from the answer will cost you 1 point, up to earning a minimum of 0 points in this task. For example if the difference between your answer and the correct answer is 36, then you will earn 64 points in this task.

In this task there will be a total of 5 question rounds, each containing 2 questions. After the instruction time has run out you will be automatically directed to the first question round; in other words there are no practice rounds for this task. This task itself will last for 4 minutes; In case you fail to give answers to some questions in time you will earn 0 points for these questions. During the task you can see how much time is left to answer on the clock in the lower right corner of the screen.

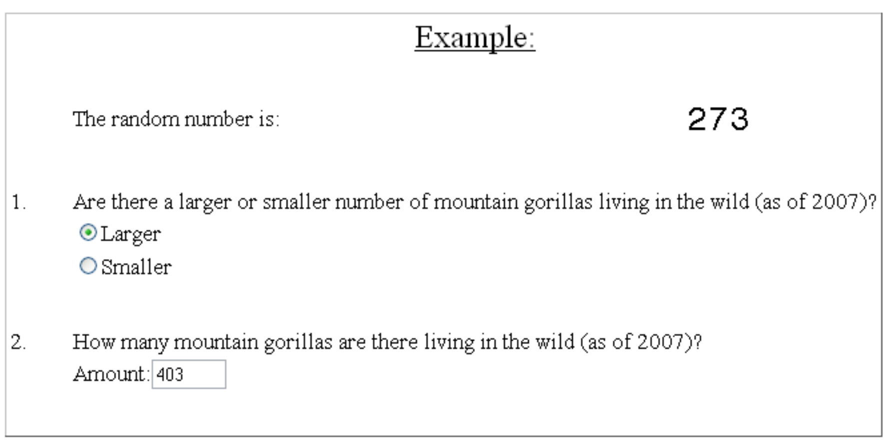

Figure 4.4: Example Used for the Anchoring Effect Task

4.B.4 Prisoner's Dilemma Task

For this task you will be randomly matched to another participant. The identity of the participant you will be matched with will stay hidden to you and similarly your identity will remain unknown to the other participant. The rest of the procedure is as follows: You will be presented with several situations where you have to choose between one of two options. A choice for a particular option will have consequences both for the amount of points you will earn for this task as well as the point earnings of the matched participant. The other participant faces the exact same choice task with the same consequences for you and the other participant. Your point total will be calculated based on your choices and the other participant's choices. Your earnings for this round will be made known to you at the end of this experiment.

In total there will be 3 choice situations with 2 possible choices each. After the instruction time has run out you will be automatically directed to the first choice situation, so there is no practice time for this task. This task itself will last for 4 minutes; in case you fail to make a choice in time you and the participant you are matched with will earn 0 points for said choice situation. During the task you can see how much time is left to answer on the clock in the lower right corner of the screen.

Example of a decision situation:

	Points:	You	The Other
Option 1:		+50	+50
Option 2:		+60	-40

Figure 4.5: Example Used for the Prisoner's Dilemma Task

4.B.5 Cognitive Task

For this task the procedure will be as follows: You will be shown 3 two-digit numbers on the screen. Your task is to calculate the sum of these three numbers. For each correct answer you will receive 10 points, while for each incorrect answer 10 points will be deducted. After the instruction time has run out you will be automatically directed to the first problem set, so there is no practice for this task. This task itself will last for 4 minutes. During this time you can do as many calculations as you want. During the task you can see how much time is left to answer on the clock in the lower right corner of the screen.

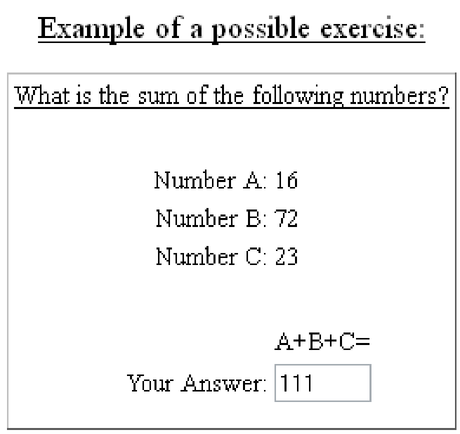

Figure 4.6: Example Used for the Cognitive Task

Appendix 4.C Screenshot of the Stroop Task

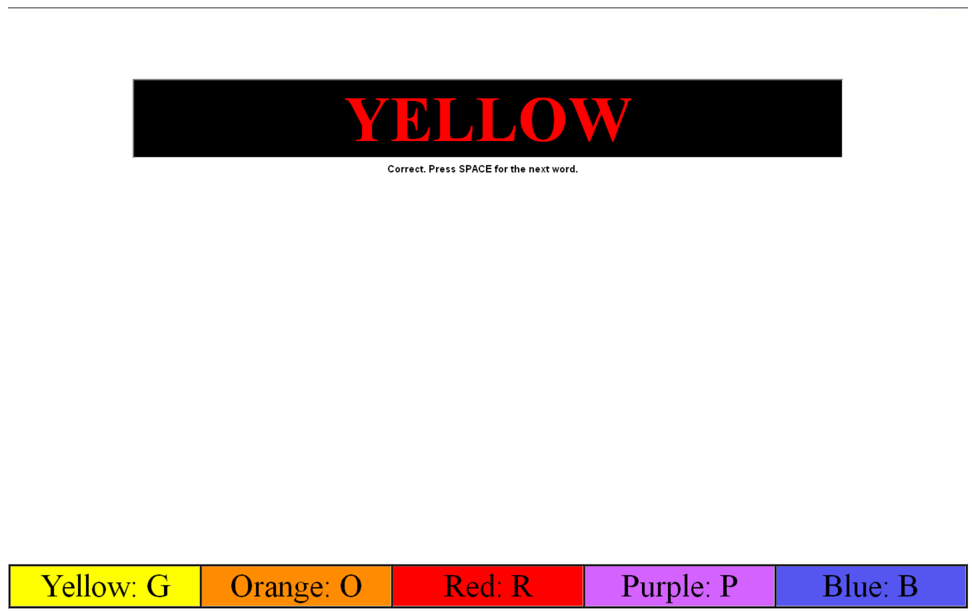

Figure 4.7: Screenshot of the Stroop Task

Notes: This figure is a screenshot of the Stroop task we used in the experiment, translated from Dutch to English. The capital letters next to the colors at the bottom are the keys used to indicate a given color in the experiment; a G was used for yellow since the Dutch word for yellow is ‘Geel’.