WZB

Wissenschaftszentrum Berlin
für Sozialforschung

Thomas de Haan
Roel van Veldhuizen

Willpower Depletion and Framing Effects

Discussion Paper

SP II 2013–206

May 2013

Wissenschaftszentrum Berlin für Sozialforschung (WZB)

Research Area
Markets and Choice

Research Unit
Market Behavior

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Thomas de Haan, Roel van Veldhuizen
Willpower Depletion and Framing Effects

Affiliation of the authors:

Thomas de Haan
Friedrich Schiller Universität Jena

Roel van Veldhuizen
WZB Berlin Social Research Center

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Abstract

Willpower Depletion and Framing Effects

by Thomas de Haan and Roel van Veldhuizen*

We investigate 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 we designed two experiments in which we depleted participants' willpower and subsequently had them take part in a series of tasks, including a framed prisoner's dilemma, an attraction effect task, a compromise effect task, and an anchoring task. However, we find no evidence that framing effects were indeed more prevalent in willpower-depleted participants than in controls.

Keywords: *willpower, ego depletion, framing, willpower depletion, experiment, behavioral economics*

JEL classification: D81, C91

---

* E-mail: thomas.de-haan@uni-jena.de, roel.vanveldhuizen@wzb.eu.

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 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 by exercising self-control, this will make him or her more likely to yield to subsequent temptations.

The resource (or willpower) depletion literature has recently started to get a foothold in economics.[3] 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 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.

If willpower depletion leads to stronger framing effects, this has immediate

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

implications for behavior on markets. Framing effects can be (and have been) used by sellers to influence consumers, for example, by using an anchoring effect to increase consumers' willingness to pay for a product (e.g., Wansink, Painter, and Ittersum, 2001; Simonson and Drolet, 2004) or encouraging them to buy a more expensive product by strategically changing consumers' choice sets (Simonson, 1999; Ariely, 2008). Hence if willpower depletion increases consumers' susceptibility to framing effects, sellers may want to approach buyers when buyers are likely to be low in willpower. By contrast, consumers may wish to move important decisions to moments of the day (e.g., early in the morning or after a short break) when they are likely to have plenty of willpower reserves, since this will weaken any attempt by sellers to use framing to influence their purchases.

To test whether willpower depletion increases susceptibility to framing effects, we run an experiment in which we follow several previous studies (see e.g., Burger, Charness, and Lynham, 2011, or Pocheptsova et al., 2009) in using 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 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 particular sequence, with a Stroop task followed by a secondary task, is repeated five times, such that participants go through all five secondary tasks exactly once.

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. 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).[4]

These results prompted us to run a second experiment in which we administered a more traditional, non-incentivized version of the Stroop task nearly identical to the version used in Pocheptsova et al. (2009). The results of ex-

[4]Pocheptsova et al. (2009) describe five willpower depletion experiments, two of which examine the compromise effect and two of which examine 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.

periment two are remarkably similar to experiment one: framing effects appear in the attraction, compromise, and anchoring tasks as well as the prisoner's dilemma task, but none of these framing effects differ between depleted participants and controls. Thus, even when we deliberately designed the Stroop task to be as close as possible to the previous literature we find no depletion effect.

Our results show that from a methodological perspective it is important to more critically examine the conditions under which willpower depletion is likely to have an effect on economic behavior and under which it is not. In particular, taken together with the results of Pocheptsova et al. (2009), our results suggest that willpower depletion effects may disappear in incentivized environments, suggesting that depletion effects are unlikely to be relevant in economic decision problems.

The structure of the rest of the paper is as follows. In section 2 we present the design of the first experiment, where we describe each task in greater detail. In section 3 we describe our research questions, after which we provide the results in section 4 and discuss the results of the first experiment in section 5. Section 6 presents the design of the second experiment, the results are presented in section 7. Section 8 concludes.

2. Experimental Design of Experiment One

The goal of this study was to investigate the effect of willpower depletion on participants' susceptibility to framing effects. We therefore conducted two experiments 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. A unique feature of our design is that we repeated this sequence of a depletion task followed by a framing task five times. This allows us to investigate if repeatedly exposing participants to the depletion task changed the size of framing effects. In the remainder of this section, we first discuss the depletion task and then we discuss all five secondary tasks individually, followed by a short description of the experimental procedure.

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; Mead et al., 2009; 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 approximately 33 percent 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 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.

During 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 pressed a key, they received a new word after pressing the space bar and waiting for 0.55 seconds on average.[5] Each Stroop task lasted precisely five minutes. Incentives were such that participants received one cent for each correct response and were deducted two

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

cents for each incorrect answer. Feedback on the total number of correct and wrong answers was only provided at the end of the experiment.A screen shot of the Stroop task can be found in Appendix D.

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. For experiment one, the order of the tasks was randomized between participants to minimize potential order 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.

2.2.1. 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 in two dimensions. Two of the options were roughly equally attractive, with one option having a better value in one dimension (e.g., price) and the other option having a better value in the other dimension (e.g., fuel efficiency). The third (or 'decoy') option was similar to one of the first two options, but was asymmetrically 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 and this 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 in both the probability and the prize dimension either by lottery x or lottery y (and not the other). The attraction effect occurs if participants are more likely to

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

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 21 in Appendix B.[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.

2.2.2. 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 set-up used in this experiment was also based on Herne (1999). Similar to the attraction effect task, participants had to choose between three lotteries that varied 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 22 in the appendix shows the lotteries we used in the experiment. 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.

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

2.2.3. Anchoring Task

Another secondary task was the anchoring task, which was based on Tversky and Kahneman (1974). Tversky and Kahneman 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 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 similar settings 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 nearly all participants should have had at least some idea of their values, but would have been unlikely to know the precise answer. For example, one of the statistics asked participants for the distance between Paris and the Dutch city of Eindhoven in kilometers.[8] 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.

[8]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 the Vatican City.

2.2.4. 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 prisoner's dilemma as either a 'community game' or a 'banker game.'[9] 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 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 made by 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 respectively. The banker or community heading persisted into the four minutes of decision time. Participants played the three prisoner's dilemmas displayed in table 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. Prisoner's dilemma 2 was different from the other two in terms of payoffs and was placed in between to make the similarity of dilemmas 1a and 1b less obvious.

2.2.5. Cognitive Task

Finally, we included a cognitive task to check for differences in cognitive functioning between depleted and control participants. The task itself consisted

[9] 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').

Table 1: Payoffs in the Prisoner's Dilemma Task

Payoffs	Prisoner's Dilemma 1a		Prisoner's Dilemma 1b		Prisoner's Dilemma 2	
	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.

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 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.

2.3. Procedure for Experiment One

Experiment One 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 explaining the Stroop task. As part of the instructions, subjects answered two check-up questions and performed 10 practice trials 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 each Stroop task and six minutes for every 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 during the course of the experiment if necessary. A translation of the instructions can be found in Appendix C.

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, 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 at 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.[10] 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.

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. Finally, we 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. 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) investigated this hypothesis in two separate experiments. In the first, they 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

[10]These questions about participants' emotional state 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.

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. In the second experiment, they depleted participants by having them write essays without using the letters 'n' or 'a' and then asked participants to choose between gift cards for two popular retailers. In both cases, they found 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 the 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) find 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 find 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 *sophisticated* decision heuristic that requires the use of significant cognitive resources. We favor the idea that the compromise effect is based on lower level processes, but acknowledge that the opposite argument could also be made.

In terms of the *anchoring effect*, we expect depleted participants to be more susceptible to the framing effect than control participants. 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 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 framing 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 control participants.

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.

We can exploit the repeated nature of the experiment 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 become more strongly depleted over every repetition of the Stroop task, leading to a large difference with control participants 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 and control participants 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 in one framing tasks are also more susceptible to a framing effect in another framing task. We can also investigate if such correlations between tasks are larger for depleted or control participants.

4. Results of Experiment One

We ran six sessions of experiment one 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

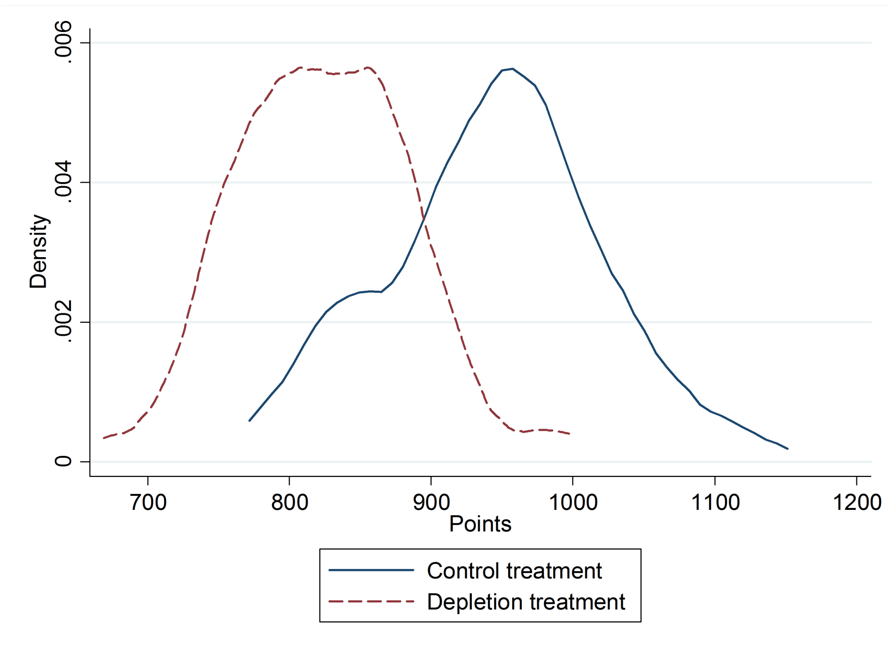

Figure 1: The Distribution of Earnings in the Stroop Task (in cents)

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 automatically determined and equal to 25.2 for the depletion treatment and 28.5 for the control treatment.

were male, with an average age of 21.17 and the majority (68%) were 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.

4.1. 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

Table 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 percentage of mistakes	1.33% (1.21%)	1.55% (0.87%)	0.22%	p=0.0572
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 percentage of mistakes and the average reaction time in the Stroop task separately for both treatments; the bracketed numbers are the between-subject standard deviations. P-values are calculated using Mann-Whitney tests.

Table 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 trials in which participants chose the decoy, dominating and non-dominating lotteries respectively; this uses the results from all five decision problems. 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 (and 104 overall).

perform worse on the Stroop task overall. Table 2 and figure 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 1 show that this difference is driven almost exclusively by depleted participants having a slower response time; the difference in the percentage of mistakes is significant at the 10% level only.[11] Moreover, according 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.[12] All in all, this is in line with previous studies and suggests that the Stroop manipulation was successful.

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 should be 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 3 shows that the dominating lottery was picked more often than the non-dominating lottery; this difference is significant both for the control treatment and the whole sample (but not for depleted participants).

Table 3 also shows that 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.

4.3. 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 would be more likely to choose option x or y if it was the mid-

---

[11] 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 software differences or all three.

[12] In terms of the questionnaire variables, we also found a negative correlation between the answer to the question "in general as a person I would describe myself as easily influenced" (p=.003, Pearson correlation) and performance in the Stroop task. Students of economics performed better in the Stroop task in the control treatment (p=.03, Mann-Whitney) but not in the depletion treatment (p=.3, Mann-Whitney). Men did better at the Stroop task in the control treatment (p=.046, Pearson correlation) but not in the depletion treatment (p=.68). Participants who performed well in 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 heavier smokers (p=.023).

Table 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 decoy, the compromise and the non-compromise lotteries respectively; this uses the results from all five decision problems. 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 (and 104 overall).

Table 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 in which participants chose the compromise, most risky and least risky lotteries respectively; this uses the results from all five decision problems.

dle 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 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 is 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.

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.[13] However, this apparent puzzle can be explained by risk attitudes. Table 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.[14] The least risky

[13]Herne (1999), however, used a within-subjects design that allowed her to look only at the small number of subjects (at most 12 and sometimes as few as two) 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 two observations (table 4, choice set e, p=.0784).

[14]We also observe a preference for the risk averse lottery in the attraction effect task, where the least risky lottery is chosen 62.7% of the time and the most risky option is chosen 35.6%

Table 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.

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.[15]

The finding that many participants tended to choose the least risky lottery does not change the fact that several participants still picked the compromise choice. Thus, in principle it would still be possible for the compromise choice 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. Note 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. Anchoring Effect Task

In 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

of the time.

[15]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.

Table 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 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.

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 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. 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).

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 within-subject variation 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 1.

Table 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.[16] Within either depleted or control participants, label

[16]For one participant in the banker/depleted group, the prisoner's dilemma task did not

Table 8: Prisoner's Dilemma 1a Probit Regression

	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.

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 1b in three of four cases. Similarly, the banker frame leads to slightly higher cooperation rates than the community frame in three of four cases. 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 for the combined data. Table 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 not significant for either depleted or control participants. Moreover, there was no significant depletion effect in the other six specifications we tested. Thus, there is little evidence that willpower depleted participants are differentially affected by framing effects.

set up correctly due to a software error, so her results were dropped from the analysis of the prisoner's dilemma task.

Table 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.

4.6. 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 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.

4.7. 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 changed over successive repetitions. Table 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 significant in all repetitions of the Stroop task ($p<0.0001$).

To trace framing effects over successive rounds, we need to use an individual 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 7 estimated at the individual level.[17] 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 a 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

[17]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 does not vary at the individual level and is therefore omitted.

Table 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.

Table 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 changed her choice between prisoner's dilemmas 1a and 1b. Both correlations are calculated overall and separately for depleted and control participants.

difference between depleted and control participants within these respective measures to check if depletion had a different effect in different rounds.

Figure 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.

4.8. Correlations between Tasks

Finally, our data also allow us to investigate whether participants' behavior is correlated between each of the six experimental tasks. To investigate this question, we use 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 11 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.[18] Most intuitive is the positive correla-

[18]In total, we computed 45 correlations. All results are identical if we use the Spearman

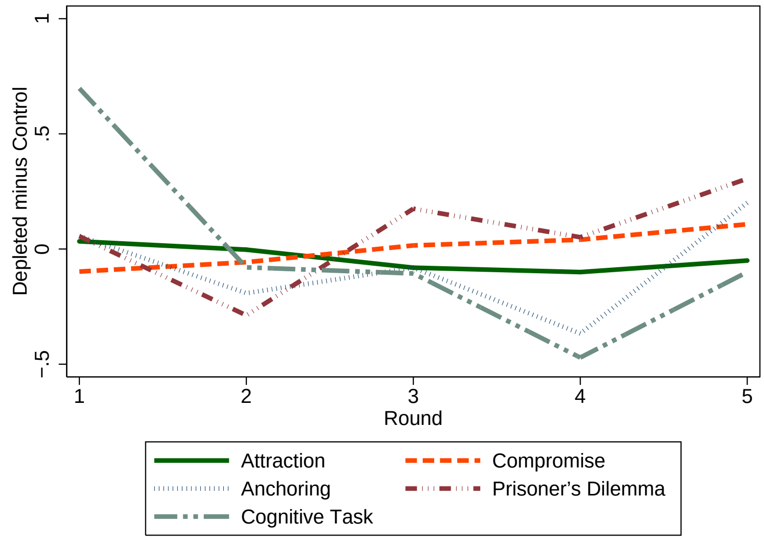

Figure 2: Depletion Effects over Successive Rounds

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).

tion 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 as well as 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 in the Stroop task were 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 each of the six experimental tasks.

5. Discussion of Experiment One

In experiment one we investigate the effect of willpower depletion on participants' susceptibility to framing effects. Framing effects appear in the attrac-

correlation instead or if we use OLS. 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.

tion, 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.

This raises the question why -in contrast with previous studies- there is no evidence of a willpower depletion effect in our study. One possibility is that our manipulation of participants' willpower by means of the Stroop task was not successful. In this light it is useful to compare our implementation of the Stroop task with previous studies. To our knowledge, experiment one is the first experiment that uses an incentivized version of the Stroop task to deplete participants' willpower. Though incentivizing participants for all tasks is common practice in economics, the incentives in the Stroop task may have inadvertedly increased the motivation of control participants to do well. Since the congruent (control) version of the Stroop task is less interesting than the incongruent version and doing well on the control version of the Stroop task still requires participants to maintain focus and concentration, it is possible that control participants were also depleted by the Stroop task. [19]

To investigate whether our lack of a willpower depletion effect was caused by our incentivized implementation of the Stroop task, we ran a second experiment in October 2012. For this experiment we kept the secondary tasks the same, but we used a non-incentivized version of the Stroop task that was almost identical to the version used by Pocheptsova et al (2009). Section 6 summarizes the changes in the design for experiment two as compared to experiment one and section 7 discusses the results.

6. Experimental Design for Experiment Two

The goal of experiment two was to give willpower depletion its best possible shot. For this purpose, we adapted the design of experiment one. The first change was that we adopted the design of the Stroop task administered by Pocheptsova et al. (2009). One reason was that we needed a version of the Stroop task with no monetary incentives. Adapting the Stroop task of experiment one to remove incentives would have required going from a time limit to

[19] Additionally, the average number of trials for a single Stroop task (180) is a lot 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. Taking into account that we repeated the Stroop task five times, the total number of trials in our experiment is larger; however, note that when we only looked at the results of the first secondary task for any of the four framing tasks we did not find any treatment effects either.

a trial limit and changing the text of the instructions. Rather than changing multiple things and using an untried version of the Stroop task, we decided to follow Pocheptsova et al.'s (2009) implementation. The second reason was that, after observing the results of experiment one, we wanted to give experiment two the best possible shot of replicating the results of Pocheptsova et al. (2009). By using the Stroop task of Pocheptsova et al. (2009), any differences between their results and ours could no longer be attributed to differences in the implementation of the Stroop task.

As a result, the differences between the Stroop task of experiment one and the Stroop task of experiment two are the following. Firstly and most crucially, there were no monetary incentives; this was also clearly communicated to the participants in the instructions. Secondly, folowing Pocheptsova et al. (2009) colors and words were now always incongruent in both the control and the depletion treatment. Participants in the control treatment now had to correctly report the word that was displayed; participants in the depletion treatment had to correctly report the print color of the word on the screen, as before. Thirdly, the Stroop task lasted for 40 trials instead of five minutes. Fourthly, subjects were given immediate feedback on whether their previous answer was correct or incorrect (as before), but were now also shown their reaction time for the last trial. Finally, we also changed the visual representation of the Stroop task by removing the colors from the bottom of the screen (see Appendix D). Again, all these changes were made to make our Stroop task implementation as close as possible to the implementation of Pocheptsova et al (2009).[20]

Secondly, the sequence of secondary tasks was no longer randomized, though participants did still alternate between the Stroop task and the secondary tasks. Participants started with the attraction task, followed by the compromise task, the anchoring task, the cognitive task, and the prisoner's dilemma task. We changed to a fixed order of tasks, since we wanted to give our experiment the best possible shot of replicating the results of the attraction effect task in Pocheptsova et al. (2009), which is administered immediately after a depletion task.

Other than these changes, the experimental design of experiment two was identical to the design of experiment one.

[20]There was a small software-related difference between the Stroop task of experiment two and Pocheptsova et al. (2009). Participants in this study used the keyboard to indicate a response instead of the mouse, which might explain the quicker response time we find in our study.

7. Results and Discussion for Experiment Two

Overall the results of experiment two are very similar to the results of experiment one, despite the different implementation of the Stroop task. We again find an attraction effect, an anchoring effect, and a 'reverse' compromise effect. The main difference with experiment one is that there are also framing effects in the prisoner's dilemma. Since the results are similar to experiment one, we refer the interested reader to appendix A for a more detailed overview of the results and the statistical analysis.

Importantly, as in experiment one, none of the framing effects are different in the depletion treatment relative to the control treatment. This is in spite of the fact that for experiment two we adopted an implementation of the Stroop task that has previously been demonstrated to lead to willpower depletion (Pocheptsova et al., 2009). Therefore, experiment two appears to rule out the explanation we previously postulated that the lack of a depletion effect on framing effects in experiment one is due to monetary incentives of the Stroop task of experiment one. More generally, the results of experiment two combined with Pocheptsova et al. (2009) suggest that the lack of differences in framing effects between depleted and control participants in this study is not due to our specific implementation of the Stroop task.

Relative to experiment one, a new finding in experiment two is that depleted participants in the attraction and compromise tasks are less risk averse than control participants. There are a number of possible post hoc explanations for this finding. One explanation is that depleted participants are less willing and able to spend cognitive resources on making a good decision and are more likely to make a random decision, increasing the share of otherwise 'low frequency' risky choices. An alternative explanation is that depleted participants have less self-control and are therefore drawn to the option with the highest possible payoff, (partially) neglecting the probabilities. However, the result is quite small and not present in experiment one. Moreover, the previous literature seems to be inconclusive on whether willpower depletion should lead to more or less risk aversion.[21] A related observation is that in the attraction effect task we find

[21] For example, Dohmen et al (2010) find in a survey study that higher cognitive ability is negatively correlated with less risk aversion. Burks et al. (2009) find a similar result in an experiment with truck driver trainees. However, these studies look at general cognitive abilities as tested by an IQ-test, rather than focusing on depletion. In a recent working paper Castillo et al. (2012) deplete people by first assessing whether subjects were 'morning' or 'evening' persons, and then randomly assigning them to make risky decisions either in the morning or in the evening. They found that incongruently assigned subjects (e.g., an evening person choosing in the morning) showed less risk aversion, in line with our results.

that depleted participants in experiment two more often chose the dominated 'decoy' option, which is more in line with depleted participants making more random decisions.

8. Conclusion

In this paper we investigated whether willpower depletion increases participants' susceptibility to framing effects in two separate experiments. Framing effects appear in the attraction, compromise, anchoring, and (for experiment two) prisoner's dilemma 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. Our results are in stark contrast with other studies that do document the willpower-depleting effects of the Stroop task. Most notably, our results are not in line with the results of Pocheptsova et al. (2009), who report an effect of depletion (also via the Stroop task) on the attraction effect and compromise effect. Moreover, when we specifically ran a follow up experiment (experiment two) to give willpower depletion in general and specifically the results of Pocheptsova et al. (2009) its best possible shot at replication, we still found no effects.

Where does this leave us? At the very least our results show that the effect of willpower depletion as administered by the Stroop task is not robust to the specific set-up used in our experiment. In fact, it is worthwhile emphasizing that we looked for a depletion effect in five different tasks, two different experiments, and many more subsets of the data (different combinations of prisoner's dilemmas, different periods in experiment one, etc.). In over 50 separate statistical tests, not once did we find a significant difference in framing effects between treatments. There was also no clear tendency among the non-significant effects: sometimes they were positive, sometimes they were negative.

Moreover, the fact that in experiment two we used the implementation of the Stroop task used by Pocheptsova et al. (2009) suggests that the lack of a depletion effect in our study is not due to the way willpower depletion was implemented. In fact, it suggests that the effect of willpower depletion might not be robust to the secondary tasks used in our experiment. Comparing our study with Pocheptsova et al. (2009) suggests a potential key role for incentives: in our study all five secondary tasks were incentivized, whereas in Pocheptsova et al. (2009), with one exception, the secondary tasks were not incentivized.[22]This

[22]This exception is an attraction effect task for which the depletion effect was significant at

leads to a potentially important implication: if depletion effects only appear when stakes are low and/or participants are indifferent, it is unlikely to have an impact in important economic decisions.

References

Ali, S. Nageeb. 2011. "Learning self-control." *The Quarterly Journal of Economics* 126 (2):857–893.

Ariely, Dan. 2008. *Predictably Irrational.* HarperCollins.

Ariely, Dan, George Loewenstein, and Drazen Prelec. 2003. ""Coherent arbitrariness": stable demand curves without stable preferences." *The Quarterly Journal of Economics* 118 (1):73–106.

Baumeister, Roy F., Erin A. Sparks, Tyler F. Stillman, and Kathleen D. Vohs. 2008. "Free will in consumer behavior: Self-control, ego depletion, and choice." *Journal of Consumer Psychology* 18 (1):4–13.

Bernatzi, Shlomo and Richard H. Thaler. 2002. "How Much Is Investor Autonomy Worth?" *Journal of Finance* 57:1593–1616.

Bettman, James R., Mary Frances Luce, and John W. Payne. 1998. "Constructive Consumer Choice Processes." *Journal of Consumer Research* 25 (3):187–217.

Bucciol, Alessandro, Daniel Houser, and Marco Piovesan. 2010. "Willpower in children and adults: a survey of results and economic implications." *International Review of Economics* 57 (3):259–267.

———. 2011a. "Temptation and productivity: A field experiment with children." *Journal of Economic Behavior & Organization* 78 (1-2):126–136.

———. 2011b. "Temptation at Work." *Harvard Business School Working Paper* (11-090).

Burger, Nicholas, Gary Charness, and John Lynham. 2011. "Field and online experiments on self-control." *Journal of Economic Behavior & Organization* 77 (3):393–404.

---

the 10% only.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini. 2009. "Cognitive skills affect economic preferences, strategic behavior, and job attachment." *Proceedings of the National Academy of Sciences of the United States of America* 106 (19):7745–50.

Busemeyer, Jerome R., Rachel Barkan, Shailendra Mehta, and Alok Chaturvedi. 2007. "Context effects and models of preferential choice: implications for consumer behavior." *Marketing Theory* 7 (1):39–58.

Castillo, Marco, David Dickinson, and Ragan Petrie. 2012. "Cognitive Resource Depletion, Choice Consistency, and Risk Preferences."

Croson, Rachel and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2):448–474.

De Martino, Benedetto, Dharshan Kumaran, Ben Seymour, and Raymond J Dolan. 2006. "Frames, biases, and rational decision-making in the human brain." *Science* 313 (5787):684–7.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review* 100 (3):1238–1260.

Druckman, James L. 2001. "Evaluating framing effects." *Journal of Economic Psychology* 22:91–101.

Epley, N. and T. Gilovich. 2001. "Putting Adjustment Back in the Anchoring and Adjustment Heuristic: Differential Processing of Self-Generated and Experimenter-Provided Anchors." *Psychological Science* 12 (5):391–396.

Fudenberg, Drew and David K. Levine. 2012. "Timing and self-control." *Econometrica* 80 (1):1–42.

Furnham, Adrian and Hua Chu Boo. 2011. "A literature review of the anchoring effect." *Journal of Socio-Economics* 40 (1):35–42.

Hagger, Martin S., Chantelle Wood, Chris Stiff, and Nikos L. D. Chatzisarantis. 2010. "Ego depletion and the strength model of self-control: a meta-analysis." *Psychological bulletin* 136 (4):495–525.

Herne, Kaisa. 1999. "The effects of decoy gambles on individual choice." *Experimental Economics* 2 (1):31–40.

Huber, John, John W. Payne, and Christopher Puto. 1982. "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." *Journal of Consumer Research* 9:90–98.

Kahneman, Daniel. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93 (5):1449–1475.

Levin, Irwin P., Sandra L. Schneider, and Gary J. Gaeth. 1998. "All frames are not created equal: a topology and critical analysis of framing effects." *Organizational Behavior and Human Decision Processes* 76:149–188.

MacLeod, Colin M. 1991. "Half a century of research on the Stroop effect: an integrative review." *Psychological Bulletin* 109 (2):163–203.

Mead, Nicole L, Roy F Baumeister, Francesca Gino, Maurice E Schweitzer, and Dan Ariely. 2009. "Too Tired to Tell the Truth: Self-Control Resource Depletion and Dishonesty." *Journal of experimental social psychology* 45 (3):594–597.

Müller, Holger, Eike B. Kroll, and Bodo Vogt. 2010. "Fact or artifact? Empirical evidence on the robustness of compromise effects in binding and non-binding choice contexts." *Journal of Retailing and Consumer Services* 17 (5):441–448.

———. 2011. "Do real payments really matter? A re-examination of the compromise effect in hypothetical and binding choice settings." *Marketing Letters* 23 (1):73–92.

Muraven, Mark and Roy F. Baumeister. 2000. "Self-regulation and depletion of limited resources: Does self-control resemble a muscle?" *Psychological Bulletin* 126 (2):247–259.

Mussweiler, Thomas and Fritz Strack. 2001. "The Semantics of Anchoring." *Organizational Behavior and Human Decision Processes* 86 (2):234–255.

Ok, Efe, Pietro Ortoleva, and Gil Riella. 2011. "Revealed (P)Reference Theory." *Working Paper* .

Ozdenoren, Emre, Stephen W. Salant, and Dan Silverman. 2011. "Willpower and the optimal control of visceral urges." *Journal of the European Economic Association* (Forthcoming):1–27.

Pocheptsova, Anastasiya, On Amir, Ravi Dhar, and Roy F. Baumeister. 2009. "Deciding without resources: resource depletion and choice in context." *Journal of Marketing Research* 46 (3):344–355.

Pruitt, Dean G. 1967. "Reward Structure of Cooperation: The Decomposed Prisoner's Dilemma Game." *Journal of Personality and Social Psychology* 7 (1, Pt.1):21–27.

Ross, Lee and Andrew Ward. 1995. "Psychological Barriers to Dispute Resolution." *Advances in Experimental Psychology* 27:255–304.

Simonson, Itamar. 1989. "Choice Based on Reasons: The Case of Attraction and Compromise Effects." *Journal of Consumer Research* 16 (2):158.

———. 1999. "The effect of product assortment on buyer preferences44The paper has benefited from the comments of Ran Kivetz." *Journal of Retailing* 75 (3):347–370.

Simonson, Itamar and Aimee Drolet. 2004. "Anchoring Effect on Consumers Willingness to Pay and Willingness to Accept." *Journal of Consumer Research* 31 (3):681–690.

Simonson, Itamar and Amos Tversky. 1992. "Choice in Context: Tradeoff Contrast and Extremeness Aversion." *Journal of Marketing Research* 29 (3):281.

Sloof, Randolph and C. Mirjam Van Praag. 2010. "The effect of noise in a performance measure on work motivation: A real effort laboratory experiment." *Labour Economics* 17 (5):751–765.

Stroop, J. R. 1935. "Studies of interference in serial verbal reactions." *Journal of Experimental Psychology* 18 (6):643–662.

Tversky, Amos and Daniel Kahneman. 1974. "Judgment under uncertainty: heuristics and biases." *Science* 185 (4157):1124–31.

———. 1981. "The framing of decisions and the psychology of choice." *Science* 211 (4481):453–458.

Wansink, Brian, James Painter, and Koert V. Ittersum. 2001. "Descriptive Menu Labels' Effect on Sales." *The Cornell Hotel and Restaurant Administration Quarterly* 42 (6):68–72.

Webb, Thomas L. and Paschal Sheeran. 2003. "Can implementation intentions help to overcome ego-depletion?" *Journal of Experimental Social Psychology* 39 (3):279–286.

Table 12: Stroop Task Statistics for Experiment Two

	Control	Depletion	Difference	P-value
Percentage of mistakes	1.64% (1.55%)	3.02% (2.68%)	1.38%	p=0.0008
Average reaction time in seconds	0.780 (0.097)	0.977 (0.125)	0.197	p<0.0001

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

A. Results for Experiment Two

We ran six sessions for experiment two in October 2012 at the CREED laboratory of the University of Amsterdam. A total of 122 people participated; half of the participants were assigned to the depletion treatment and half to the control treatment. Sixty-five percent of participants were male, with an average age of 22.11 with 48% studying economics and/or business administration.[23]

A.1. Stroop task

As a manipulation check, we once again compare the performance of participants in the depletion and control participants. As in experiment one, participants have a faster response time in the control treatment than in the depletion treatment. Participants in the depletion treatment also make more mistakes; this difference is significant at the 1% level. Comparing the results with experiment one, reaction times are very similar but the percentage of mistakes in experiment two is larger, particularly in the depletion treatment. This could be due to the removal of incentives, but also due to the smaller number of trials participants in experiment two had to go through. Overall, the Stroop task manipulation seems to have a strong effect on performance in the Stroop task, in line with previous studies and experiment one.

A.2. Attraction effect task

To identify an attraction effect, we again compare the fraction of times the dominating lottery was chosen to the number of times the non-dominating lottery was chosen. As in the first experiment we find a clear attraction effect for both treatments. More importantly, once again there is no significant difference in the size of the attraction effect between the two treatments (Mann-Whitney,

[23]Unfortunately, one participant in the control treatment misunderstood the instructions for the Stroop task. As a result, she identified the word color rather the word itself, as if she was in the depletion treatment. Despite the fact that she got every word wrong in the first round, she only realized her mistake in the second round of the Stroop task; she then corrected herself. Hence we excluded her from the analysis, so the data below are for 121 participants only. We have made a note of when excluding this participant led to different results.

Table 13: Attraction Effect Task Statistics for Experiment Two

	Decoy	Dominating	Non-Dominating	P-value (Attraction Effect)
Control subjects	0.33%	55.3%	44.3%	.009
Depleted subjects	3.28%	54.4%	42.3%	.011
All subjects	1.82%	54.8%	43.3%	.000

Notes. The table gives the percentage of trials in which participants chose the dominating, non-dominating, and decoy lotteries respectively; this uses the results from all five decision problems. 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 equal to 121.

p-value=0.7857). Two extra things to note are that in experiment 2 participants in the depletion treatment were more likely to choose the decoy lotteries (Mann-Whitney, p-value=0.0166) and were also less likely to choose the least risky lottery (75% versus 64%, Mann-Whitney, p-value=0.0297), effects that were both not present in experiment one. Overall, however, we still find no effect of willpower depletion on the size of the attraction effect.

A.3. The compromise task

For the compromise task, we again compared the number of 'compromise' lottery choices with the number of 'non compromise' choices to determine whether there was a compromise effect. As we see in table 14 for both the control and depletion treatment, we again find a reverse compromise effect, which is very similar in size to experiment one. This reverse compromise effect shows no difference in size between the two treatments (Mann-Whitney, p-value=0.5044).

As in the first experiment, the reverse compromise finding seems to be driven by risk aversion. Table 15 shows that both in the control and depletion treatment, most of the time participants choose the least risky option , and significantly more than both the most risky and compromise option. As with the attraction effect task, there is a small effect that depleted participants choose the least risky lottery less often, although this effect is significant at the 10% level only (Mann-Whitney, p=.0911).[24] Overall, we again find no effect of willpower depletion on the size of the compromise effect.

Table 14: Compromise Effect Task Statistics for Experiment Two

	Decoy	Compromise	Non-Compromise	P-value (Comp. Effect)
Control subjects	40.7%	20.0%	39.3%	.000
Depleted subjects	38.4%	23.3%	38.4%	.003
All subjects	39.5%	21.7%	38.8%	.000

Notes. The table gives the percentage of trials in which participants chose the compromise, non-compromise, and decoy lotteries respectively; this uses the results from all five decision problems. 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 equal to 121.

Table 15: Compromise Effect Task Statistics by Riskiness for Experiment Two

	Compromise	Most risky choice	Least risky choice
Control subjects	20.0%	18.7%	61.3%
Depleted subjects	23.3%	23.9%	52.8%
All subjects	21.7%	21.3%	57.0%

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

Table 16: Anchoring Effect Task Estimates for Experiment Two

	Dependent variable: estimate of the statistic		
	Coefficient	Std Error	P-value
Anchor	0.16	0.05	0.002
True answer	0.40	0.06	0.000
Depleted	-39.33	61.60	0.524
Depleted×Anchor	0.07	0.08	0.367
Depleted×True answer	0.03	0.09	0.740
Constant	164.25	48.98	0.001

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 121 independent observations.

A.4. The anchoring effect task

To investigate differences in the anchoring effect, we once again regress participants' estimates on the value of their anchor, the true answer, a treatment dummy and interactions between the treatment dummy and the other two variables. The results (table 16) are very similar to those of the first experiment. We find a significant and substantial anchoring effect, although the coefficient is substantially smaller than in experiment one. Importantly, we again find no interaction effect between the anchor and the treatment. Thus, there is no difference in the strength of the anchoring effect between depleted participants and controls.

A.5. Prisoner's dilemma task

Table 17 shows the summary statistics from the prisoner's dilemma task for the second experiment. Our first observation is that in contrast to experiment one there is now a difference between the cooperation rates between the banker and community label (Mann-Whitney p-value=0.0279). We also find a significant decomposition framing effect à la Pruitt: participants are significantly more cooperative in Prisoner's Dilemma 1b (Wilcoxon p-value=0.0053). Both framing effects are significant for control participants (p=0.0350 for the label and p=.0076 for the decomposition framing respectively) and not significant for depleted participants (p=0.3024 for the label and p=0.2482 respectively).[25] However, the effect goes in the same direction in both treatments and the difference-in-differences is hence not significant.[26]

Overall, we find significant framing effects in the prisoner's dilemma in experiment two, but these still do not differ between depleted and control participants.

A.6. Cognitive task

The results for the cognitive task in experiment two are displayed in table 18. Although differences in performance are somewhat more pronounced than

---

[24] If we pool the data from the attraction and compromise tasks, the effect is significant at the 1% level (Mann-Whitney, p=.0015). Thus, participants in the depletion treatment appear to be less risk averse than control participants.

[25] If we include the excluded participant in the analysis here, the labeling framing is only significant at the 10% level both for controls and overall; the decomposition framing is significant at the 5% level only.

[26] For the decomposition framing, the Mann-Whitney p-value is 0.1648. For the Community versus Banker framing, we run a probit regression of the cooperation rate on the frame, the treatment and the interaction of the two. The interaction term is not significant (p-value=0.464).

Table 17: Prisoner's Dilemma Task Statistics for Experiment Two

% Cooperation	PD 1a	PD 1b	PD 2
	Control Participants		
Community label (n=32)	31.3%	46.9%	28.1%
Banker label (n=28)	10.7%	32.1%	7.1%
	Depleted Participants		
Community label (n=34)	29.4%	41.2%	23.5%
Banker label (n=27)	25.9%	25.9%	14.8%

Notes. This table gives the percentage of cooperative choices for the three prisoner's dilemma tasks 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.

Table 18: Cognitive Task Statistics for Experiment Two

	Control	Depletion	Difference	P-value
Cognitive task earnings	244.33 (77.77)	224.26 (70.49)	20.57	p=0.1772
Number of exercises completed	26.93 (7.50)	25.77 (7.45)	1.16	p=0.3266
Number of mistakes	1.25 (1.39)	1.67 (1.79)	0.42	p=0.1163

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.

in experiment one, and suggest that depleted participants performed worse than controls, the treatment difference is once again not significant.

A.7. Depletion over Successive Tasks

Since the Stroop task in experiment two was not incentivized, in this section we will only examine reaction times and how these evolved over rounds. Reaction times seem to gradually decrease over the course of the experiment, the treatment difference remains largely constant and strongly significant in all periods.[27]

A.8. Correlations between tasks

As in experiment one we analyze if there are any significant correlations in individual behavior between different tasks. For the secondary tasks we use the same measures of the size of the different framing effects as in the first experiment. For the Stroop task we use a participant's average reaction time multiplied by -1 to ensure that the quickest performer had the highest score, similar to experiment one. When we look again at the 45 correlations, we

[27] Since the order of tasks was fixed, we cannot examine how depletion effects differed with task over different periods.

Table 19: Stroop Task Average Reaction Time per Round for Experiment Two

Order	Control	Depletion	Mann-Whitney P-value
1	0.826 (0.104)	1.093 (0.169)	p=0.0000
2	0.795 (0.108)	0.987 (0.146)	p=0.0000
3	0.772 (0.102)	0.946 (0.123)	p=0.0000
4	0.763 (0.098)	0.936 (0.120)	p=0.0000
5	0.745 (0.097)	0.925 (0.136)	p=0.0000

Notes. The table gives the average reaction time for the Stroop task per treatment separately for each successive repetition of the Stroop task (or round); the bracketed numbers are between-subject standard deviations.

Table 20: Correlations between Tasks

	Overall		Depleted		Control	
	corr.	p-value	corr.	p-value	corr.	p-value
Stroop/Cognitive	0.219	0.016	0.265	0.039	0.075	0.567
Stroop/Prisoner's Dilemma	0.177	0.052	0.218	0.092	0.072	0.587
Stroop/Attraction Effect	0.149	0.102	0.315	0.014	-0.040	0.759

Notes. The first row of the table gives the Pearson correlation coefficients between average reaction time in the Stroop task (multiplied by minus 1) and the cognitive task. The second row gives the Pearson correlation coefficient between average reaction time in the Stroop task (multiplied by minus 1) and a dummy that indicates whether a participant changed his or her choice between prisoner's dilemmas 1a and 1b. The third row gives the correlation between average reaction time in the Stroop task (multiplied by minus 1) and the number of times the participant chose the dominant option in the attraction effect task. All three correlations are calculated both overall and separately for depleted and control participants.

find five to be significant at the 10% level or better. As in experiment one, participants who do well in the Stroop task are likely to do well in the cognitive task and change their decision between prisoner's dilemmas 1a and 1b. As in experiment one, this effect is present only among participants in the depletion treatment. What is new is that depleted participants who do well in the Stroop task are more likely to choose dominating lotteries in the attraction effect task. However, we once again find no correlations between performance on secondary tasks.

B. Lotteries of the Compromise and Attraction Effect Tasks

Table 21: 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 of the experiment. Every participants went through five lotteries; half the participants went through the lottery 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 22: 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 of the experiment. Every participants went through five lotteries; half the participants went through the lottery 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.

C. Experimental Instructions[28]

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".

C.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 paper 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.

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).[29] The key-color combinations will also be visible at the bottom of the screen throughout

[28]These are the instructions for Experiment One. The original instructions (in Dutch) and the instructions for Experiment Two (which differ only for Round 1) are available on request.

[29]The Dutch word for yellow is 'geel', hence the abbreviation of yellow.

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.

C.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 length 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.

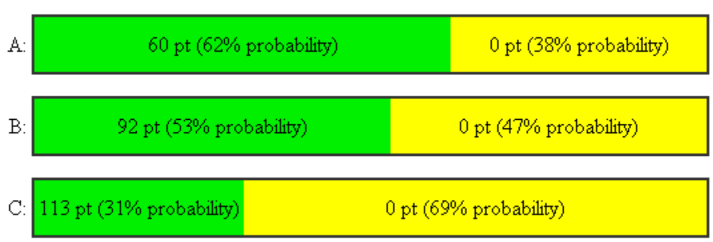

Figure 3: Example Used for the Compromise and Attraction Effect Tasks

C.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.

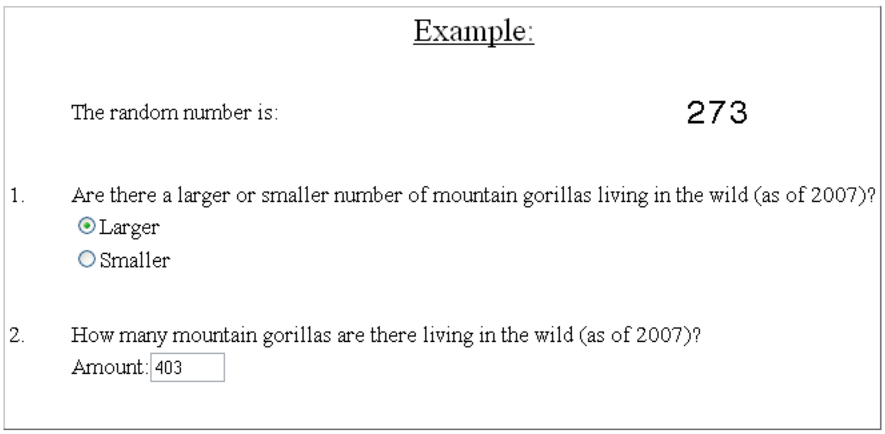

Figure 4: Example Used for the Anchoring Effect Task

C.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 5: Example Used for the Prisoner's Dilemma Task

C.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.

Example of a possible exercise:

What is the sum of the following numbers?

Number A: 16
Number B: 72
Number C: 23

A+B+C=
Your Answer: 111

Figure 6: Example Used for the Cognitive Task

D. Screen Shot of the Stroop Task

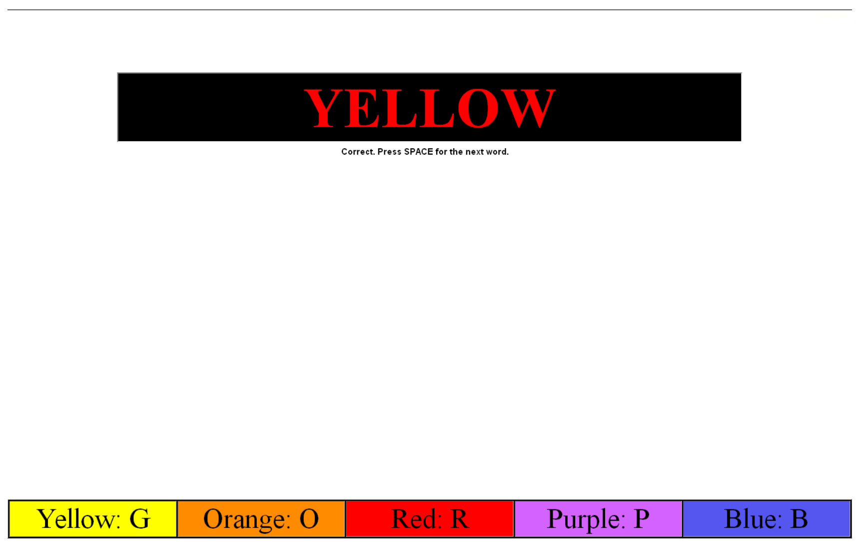

Figure 7: Screen Shot of the Stroop Task

Notes: This figure is a screen shot 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'. The screen shot corresponds to experiment one, a screen shot for experiment two is available on request.

Discussion Papers of the Research Area Markets and Choice 2013

Research Unit: **Market Behavior**

Nadja Dwenger, Dorothea Kübler, Georg Weizsäcker — SP II 2013-201
Preference for Randomization: Empirical and Experimental Evidence

Kai A. Konrad, Thomas R. Cusack — SP II 2013-202
Hanging Together or Being Hung Separately: The Strategic Power of Coalitions where Bargaining Occurs with Incomplete Information

David Danz, Frank Hüber, Dorothea Kübler, Lydia Mechtenberg, Julia Schmid — SP II 2013-203
'I'll do it by myself as I knew it all along': On the failure of hindsight-biased principals to delegate optimally

David Hugh-Jones, Morimitsu Kurino, Christoph Vanberg — SP II 2013-204
An Experimental Study on the Incentives of the Probabilistic Serial Mechanism

Yan Chen, Onur Kesten — SP II 2013-205
From Boston to Chinese Parallel to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms

Thomas de Haan, Roel van Veldhuizen — SP II 2013-206
Willpower Depletion and Framing Effects

Research Unit: **Economics of Change**

Luisa Herbst, Kai A. Konrad, Florian Morath — SP II 2013-301
Endogenous Group Formation in Experimental Contests

Kai A. Konrad, Florian Morath — SP II 2013-302
Evolutionary Determinants of War

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice