Social and monetary rewards in economic decision making – behavioral and neural findings

Inaugural-Dissertation

Zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von:

Tina Strombach
aus Wipperfürth

Düsseldorf, im Juni 2015

aus dem Institut für Experimentelle Psychologie

der Heinrich-Heine Universität Düsseldorf

Table of Content:

Abstract

One of the most important questions in human and animal behavior is, without a doubt, the question of why we behave the way we do. A major aspect of the discussion can be narrowed down to the concept of motivation and incentives. That means that our behavior is often guided by the expectation of a positive reward, whether monetary or social. Thus, whenever a reward is available, it is assumed that we behave a particular way in order to receive it. One important field of research within neuroeconomics and behavioral economics is, therefore, the study of how changes in incentives can shape behavior and how social reward is integrated and valuated in our decision-making processes.

Two main drivers—monetary and social reward—influence our behavior and decision making. On the one hand, we derive rewards from social interaction, and social considerations modulate decision making to a significant degree, e.g. sharing a good bottle of wine instead of drinking it alone. The pure value of a bottle of wine might be higher when I drink it alone because I get all the wine for myself. However, I might have a hangover on the next day. This fact might also affect my decision. It is likely that when I opt to share the bottle of wine with my dearest friends, I get less wine, but I also experience the wine as being more rewarding. Thus, although the absolute reward is lower, the subjective value is significantly higher, adding a social reward to the reward of the wine itself. Of course, sharing it with a friend is more rewarding than sharing it with a stranger. Thus, my willingness to forego a reward (that is, the wine) is dependent on the perceived social distance between me (the decision-maker) and the recipient. Overall, we make social decisions constantly and can consider them trade-offs between selfish goals and social considerations. At the same time, we often use monetary rewards to modulate behavior in favorable ways, for instance, bonus systems in companies and monetary incentives for good grades at school. The monetary incentive is likely to modulate the amount of effort that one puts into a good performance. Overall, these two reward systems can be considered the main drivers of human behavior.

The two drivers demonstrate how rewards, either social or monetary, influence our behavior and show that these incentives are constantly present. I am investigating the effect of changes in these incentives in a series of three studies. In the first study, we remove and introduce monetary incentives, leading to changes in the motivation to

perform well. We show that the changes in the incentives correlate with changes in BOLD activity in the ventral striatum. Thus, the BOLD activity increases when monetary incentives are introduced and decreases when monetary incentives are removed. Task-related activity does not alter in response to changes in incentives. We suggest that the reward system initiates the change in performance, leading to a different degree of engagement and motivation to perform.

In the second study, we focus on the social aspect of reward processing and investigate a biologically plausible model of prosocial choice by applying a social discounting experiment to the fMRI scanner environment. By varying social distance and selfish reward magnitude, we show that the temporoparietal junction (TPJ) tracks the temptation to be selfish, suggesting that the TPJ balances prosocial and selfish considerations. Also, connectivity between the TPJ and the ventromedial prefrontal cortex (VMPFC), a region associated with subjective value computation, is higher for social decisions than for egoistic ones. We suggest that the TPJ modulates basic value signals in the VMPFC, upregulating the value of generous choices.

Another open question in the decision sciences is that of how environmental distractions affect decision making, that is, prosocial choice behavior. Therefore, in the third study, we investigate the influence of cognitive load on social distance-dependent generosity. Previous research indicates ambiguous results for the effect of cognitive load. We suggest that the diverse findings could be explained using the social distance approach. While generosity decreases at closer social distances, generosity increases at larger social distances. The results indicate that this is indeed the case. However, men drive the effect. In fact, we find an interaction between the level of cognitive load and gender.

With the studies I present in this dissertation, I contribute to the development of a multidimensional view of the processing of rewards in general and performance-based monetary and social distance-dependent social rewards in particular.

Outlook

In this dissertation, I focus on two important drivers of human behavior. On the one hand monetary rewards are key drivers in the world today. On the other hand, social considerations and prosocial behavior are cornerstones of human society. In the first part of this dissertation, I introduce the idea of behavioral economics and neuroeconomics. I also introduce important features of functional magnetic resonance imaging (fMRI), which allows the investigation of changes in brain activity in a non-invasive manner. Moreover, I present the concept of reward and value in human decision making and examine monetary rewards from a psychological and economic perspective. I present previous work on the neural underpinnings of monetary incentives. Ultimately, I introduce social rewards and previous research on the neural response to social rewards.

In the second part of this dissertation, I present three studies that were conducted to investigate the processing of rewards and the underlying valuation processes of monetary and social rewards and discuss the conclusions and impact of the studies. I end with a general conclusion, discussing the contribution of the dissertation to the scientific community and future research opportunities.

PART I

1. Behavioral Economics and Neuroeconomics

"Almost all economic models assume that all people are exclusively pursuing their material self-interest and do not care about "social" goals per se. This may be true for some (may not many) people, but this is certainly not true for everybody" (Fehr & Schmidt, 1999, p. 817).

In canonical models of economic behavior, the individual is expected to solely have self-regarding preferences to maximize his or her own reward (Camerer & Fehr, 2006; Camerer & Loewenstein, 2004; Camerer, Loewenstein, & Prelec, 2005). Those theories are based on the assumption that humans are rational and unemotional in their decision making and decisions are exclusively directed towards optimizing the cost-benefit function of expected choice outcomes (Becker, 1974; Fehr & Schmidt, 1999). Economic theory offers tools to estimate how people will change their behavior in response to changes in incentives (Falk, Gächter, & Kovács, 1999; Fehr & Falk, 2002) .

In this context, all behavior is performed to increase one's own supply of food, money, or even well-being. Indeed, these incentives modulate behavior to a great extent and self-interest is undoubtedly a strong motivator that can explain a multitude of economic phenomena. However, researchers agree that more factors influence behavior. Human beings show non-opportunistic and reciprocal behavior and other "anomalies" or "paradoxes" that are not explicable by the traditional concept (Kenning & Plassmann, 2005; Tversky & Kahneman, 1981). As a consequence, researchers have formalized psychological ideas and ultimately translated them into testable predictions. The emerging field of research might be called behavioral economics. It aims to develop more realistic theories of economic choice. Therefore, it has added parameters to the standard economic models. As a result, the models are more complex but more realistic (Camerer & Loewenstein, 2004). The additional parameters might include preferences such as fairness, reciprocity, and equality, and they may be able to explain complex behavior, such as strategic interactions, that pure self-interest cannot explain. Adam Smith was the first to integrate psychological principles into economic theory (Smith, 1759). In one instance, he stated that "we suffer more […] when we fall from a better to a worse situation, than we ever enjoy when we rise from a worse to better" (Smith, 1759, p. 311). Thus, Adam Smith indicated that loss aversion was a determinant of behavior and decision making.

Tversky and Kahneman (1974) developed a theory that systematically integrated psychological principles into economic theory. Prospect theory states that not only is the expected utility of an option taken into account (Morgenstern & Von Neumann, 1953), but potential losses and gain are also evaluated differently using certain heuristics (Tversky & Kahneman, 1974, 1981). Moreover, according to the theory, losses are weighted more heavily than gains, changing the expected utility from a pure cost-benefit analysis to a subjective value context. Overall, psychological principles became more and more important for developing more realistic models of economic choice and human decision making.

Ernst Fehr extended the theoretical framework further and was one of the first to integrate social preferences such as fairness and reciprocity as parameters in a theory of economic decision making (Camerer & Fehr, 2006; Fehr & Falk, 2002; Fehr & Gächter, 1998, 1999, 2000, 2002; Fehr & Schmidt, 1999; Henrich et al., 2001). Social interactions, like social contact and closeness, as well as other-regarding preferences like fairness, modulate behavior and make us social beings (Fehr & Schmidt, 1999; Glimcher & Fehr, 2013; Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006). For example, sharing a bottle of wine might be more rewarding than drinking it alone. So, in absolute measures, we may have less of the reward, i.e., wine, but we experience it as more rewarding when sharing it with others. Thus, other-regarding behavior or social preferences are a "characteristic of an individual's behavior or motives, indicating that the individual cares positively or negatively about others' material payoff or well-being" (Glimcher, Camerer, Fehr, & Poldrack, 2009, p. 216).

In recent years, Fehr and his colleagues also played key roles in the development of the field of neuroeconomics (Eisenegger, Haushofer, & Fehr, 2011a, 2011b; Fehr & Camerer, 2007; Glimcher & Fehr, 2013; Haushofer & Fehr, 2008; Knoch et al., 2006; Morishima, Schunk, Bruhin, Ruff, & Fehr, 2012). Neuroeconomics is the interdisciplinary combination of three originally distinct disciplines (Camerer et al., 2005; Glimcher, 2004; Glimcher et al., 2009; Glimcher & Fehr, 2013; Glimcher & Rustichini, 2004; Kenning & Plassmann, 2005; Sanfey, Loewenstein, McClure, & Cohen, 2006): neuroscience, psychology, and economics. Economics is composed of theories that capture generalized behavior, which means that these approaches do not incorporate individuals' variability (Fehr & Falk, 2002). Nevertheless, it offers a powerful approach to the econometric reconstruction of individuals' decision making by investigating an accumulated level of decisions (Kenning & Plassmann, 2005). Economic theories lack the psychological insights that aim to explain individual

decisions. Thus, while economists are primarily interested in the accumulated level of behavior, psychologists are interested in the motivation behind individuals' decisions. As a result, behavioral economics aims at integrating psychological approaches into economic theory. It seems compelling that the combination of psychology and economics makes economic theory more realistic.

The integration of neuroscientific methods into behavioral economics offers the opportunity to describe decision-making processes on an additional, more objective level: that of brain activity. The origin of every decision we make is in the brain. Therefore, it seems natural to study the brain to understand how decisions are formed and how options in a decision problem are compared, for example, choosing money over food or a Snickers bar over a Mars bar. Over recent years, neuroeconomics has garnered attention as it combines methods from neuroscience, theories from the field of psychology, and models from economics to model behavior on an individual level. However, neurobiological methods offer a possibility for investigating cognitive processes at an objective level and for observing brain activities that underlie behavior. It has also been shown that the brain computes the values of rewards in a way that can be econometrically modulated and tested (Kable & Glimcher, 2007; McClure, Laibson, Loewenstein, & Cohen, 2004). Economics offers a possibility for formally describing brain mechanisms and cognitive functioning and demonstrates that the interdisciplinary approach can contribute to the creation of models of economic behavior that are based on more realistic assumptions about human behavior.

1.1 Methods used in Neuroeconomics

The earliest insights into the relationship between specific brain regions and cognitive functions implicated in decision making came from patients with frontal lobe damage (Damasio, Tranel, & Damasio, 1990; Rudebeck, Bannerman, & Rushworth, 2008). The studies showed that damage in the frontal regions led to a significant change in (economic) decision making. Nowadays, neuroscience offers a variety of methods for investigating the underlying cognitive processes and identifying the role of specific brain regions in the decision-making process in healthy subjects non-invasively. One method that is used on a frequent basis is functional magnetic resonance imaging (fMRI), which I briefly describe in the next chapter.

1.1.1 Neuroscientific Methods: Functional Magnetic Resonance Imaging

The aim of neuroeconomic research is to understand the underlying principles that drive our decision making, for example, how specific features are integrated into one value signal. Therefore, it is necessary to be able to look into the functioning brain while the subject makes decisions. fMRI offers such insights by depicting the changes in blood flow in the brain (Ashby, 2011; Huettel, Song, & McCarthy, 2004).

Magnetic resonance imaging (MRI) is an imaging technique that relies on the physical principles of magnetic fields and radiofrequencies. The imaging procedure is based on the magnetic properties of water molecules, which are aligned along the magnetic field and then subjected to resonance by radiofrequencies. When the radiofrequency is turned off, the molecules revert to their stable initial positions along the magnetic field and release some of the electromagnetic energy they absorbed before (Faro & Mohamed, 2010). The coil can depict the release of energy and, using mathematical calculations, can help determine the origin of the energy. The strength of the signal varies for the different tissues in the brain: Different shades of grey indicate distinct brain regions (fig. 1; Huettel et al., 2004).
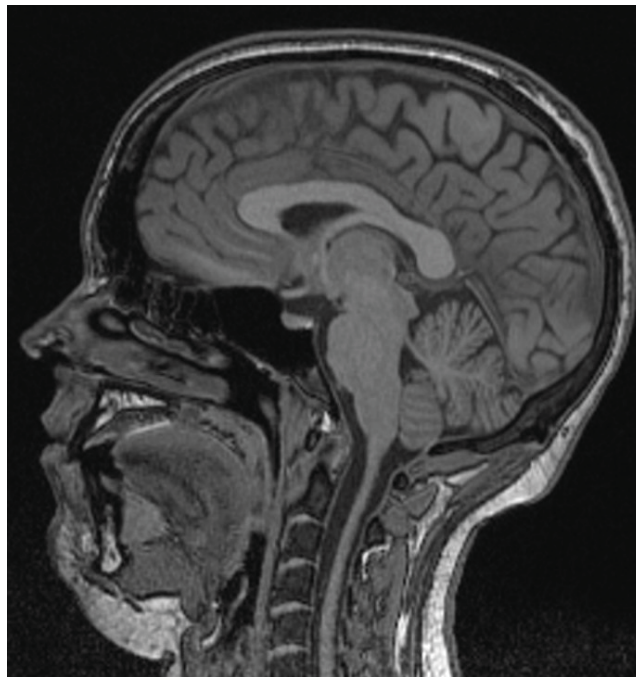

Figure 1: Sagittal plane of the T1 weighted MRI image of Tina Strombach's brain. The MRI image shows that the PhD student's brain is normally developed. However, the structure of the brain does not allow any conclusions about intelligence or cognitive abilities. The MRI was taken at the Life&Brain Center, University Hospital Bonn, 3T Siemens Trio.

fMRI is based on these technical principles. The imaging technique depicts the relative change in neural activity in the brain. The basic assumption is that the activation of a brain region leads to an increase in blood flow. The brain consumes more energy (glucose) and oxygen when neurons are active. Thus, if a region is implicated in a specific task, it becomes active and its energy consumption increases. The increased demand leads to an increase in the blood flow in the active regions to supply needed nutrients (Huettel et al., 2004).

The fMRI image depicts the blood oxygenation level dependent (BOLD) signal. One of the main components of the red blood cells (erythrocytes) is hemoglobin. Hemoglobin is an iron-containing protein that transports oxygen in the human body. The strength of the magnetization of the hemoglobin depends on whether the hemoglobin is carrying oxygen or not. The deoxygenated hemoglobin is paramagnetic; thus, it is more sensitive to the external magnetic field, while the oxygenated hemoglobin is diamagnetic. When a region is active, not only does the proportion of oxygenated hemoglobin increase, but the proportion of deoxygenated hemoglobin also does. Overall, the ratio of oxygenated to deoxygenated hemoglobin changes. The fMRI scanner is able to register this difference. The images that are collected during an fMRI scan are thus the relative change in deoxygenated and oxygenated hemoglobin in the brain (Huettel et al., 2004).

To make proper inferences about the changes, it is important to collect information about different states of activity in the region of interest. Since we are measuring the relative change in the BOLD signal, there must be at least two conditions that we can compare. Therefore, we do not measure absolute levels of activity. Rather, we investigate a relative change in activity that can be attributed to a specific cognitive task. However, it is important to keep in mind that using fMRI does not allow one to make causal inferences about the involvement of a brain region in a specific task. Using fMRI, we can only show that a specific region contributes to the cognitive task, not whether it is also necessary for task completion (Ashby, 2011; Huettel et al., 2004). Developments in the analysis tools allow now also inferences about neural networks that contribute to a specific task, developing biologically plausible models of decision making and other cognitive tasks (Bzdok et al., 2013; Eickhoff et al., 2009; Heim et al., 2009; Marreiros, Stephan, & Friston, 2010).

In the following chapter, I present economic theories of decision making. I introduce the concept of subjective value and value-based decision making and present

neurobiological findings from the field of neuroeconomics. I discuss the effects of changes in monetary incentives on behavior and the neural underpinnings. Secondly, as suggested by previous research, I do not restrict myself to the motivational aspects of monetary incentives. Rather, I focus on social preferences as one additional parameter that influences our behavior. I discuss subjective values in the context of social decisions.

2. Value-Based Decision Making

Although standard economic theories have developed and integrated other-regarding preferences into models of decision making, they still follow the simple idea that humans compare options that are available and choose the ones with the highest values (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Lim, O'Doherty, & Rangel, 2011; Padoa-Schioppa & Assad, 2006; Rangel, Camerer, & Montague, 2008; Samuelson, 1937; Tversky & Kahneman, 1974, 1981). However, the computation is not restricted to self-oriented reward maximization. In economic and neuroeconomic theory, value is conceptualized as a measure of the benefit that people can gain from choosing an option (Brosch & Sander, 2013). Decision neuroscience now offers the techniques and possibilities to characterize the exact computational properties of the process responsible for value computation. Consequently, they allow a focus on how these properties are able to generate reward-maximizing choices. The integration of additional parameters into the econometric models enhanced the precision of the subjective value computation of the options available. Therefore, a subject facing a choice determines the subjective value of each available alternative and transfers the value into a common currency, allowing a comparison of options with different qualities. Ultimately, the alternative with the greatest value is chosen (Bartra, McGuire, & Kable, 2013).

The idea that humans compare options and consider their individual value to drive a decision is dubbed value-based decision making. In general, there are in general two different kinds of values. On the one hand, absolute values are just the absolute values of options. That might translate into the amount of money if you have to choose between €5 and €10, for example. The decision might be easy, since the value of each option is clear and the measures are objective. However, if you have to choose between an apple and €2 cash it might get more difficult. If you are really hungry, the apple's value might be very high. However, if you have just had lunch, then you might appreciate the €2. Thus, not only does the subjective value integrate the value of the option itself, but it also incorporates it into a coherent value that includes your actual state of mind and situation. One might consequently argue that all values are subjective. Therefore, you must always consider values along with the associated circumstances: The value of €2 might be different for a millionaire than it is for a poor gamin.

To formalize the process of value-based decision making, I am introducing a formal model (Rangel et al., 2008). According to this model, decision making can be broken down into five basic processes (fig. 2; Rangel et al., 2008). In the first step, the subject has to identify the decision making problem. Thus, the subject forms a representation is formed, which builds the foundation for the computation of the subjective value. That includes the identification of one's internal states (e.g. the level of hunger), external states (such as the threat level), and potential courses of action (such as pursuing prey). It might also include identifying the options that are available.
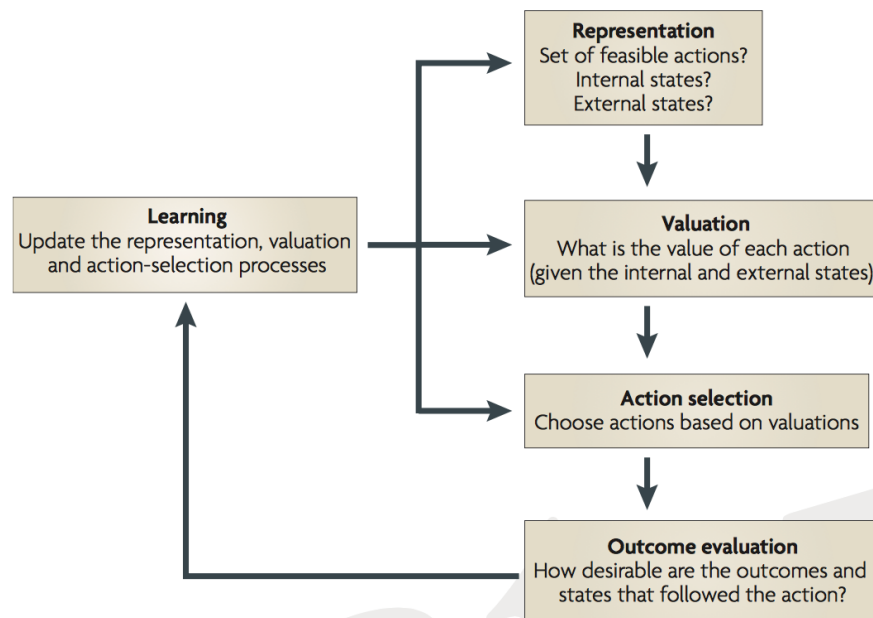
Figure 2: Five stages in the value-based decision making process. Previously experienced outcomes influence the decision making process by modulating the representation, valuation, and selection of the outcomes available. The actual outcome is then used again to update the representation, which is a constant learning process (reprinted from Rangel et al., 2008, p. 546).

Thus, first of all, one gathers and systematizes the available information concerning the decision problem itself and the information about the internal and external state of the decision maker. Following that first step, one assigns the different options and assigns a value to each one. The valuation process is especially important for making an appropriate decision that also meets one's preferences and satisfies one's expectations. The premise is of course, that these values are reliable predictors of the benefit associated with chossing that option. The third step involves the comparison of different values in the common currency to make a choice. By building an order from the most preferred option to the least preferred one, the decision maker is able to

choose the one that has the highest expected subjective value for himself/herself. Fourth, after implementation of the decision, the brain measures the desirability of the actual outcomes. Fifth, these feedback measures are relevant for updating the representation to improve the quality of the outcome of future decisions. The last step follows the revelation of the subject's choice. The information from the evaluation of the outcome is entered into a learning feedback loop to improve future decisions and facilitate learning from mistakes (Rangel et al., 2008).

The core concept in this framework of decision making is the valuation process. In general, it is possible to differentiate between three types of valuation. These entail different strategies to evaluate the option or potential course of action at hand, e.g. staying versus leaving. They are Pavlovian, habit, and goal-directed systems. Pavlovian processes are those that assign value based on a "small set of behaviors that are evolutionarily appropriate responses to particular environmental stimuli" (Rangel et al., 2008, p. 2). Thus, certain stimuli are associated with specific behavioral responses and include certain preparatory behaviors, for instance salivation when food is available (Balleine, Daw, & O'Doherty, 2008; P. Dayan, 2008; Peter Dayan, Niv, Seymour, & Daw, 2006; O'Reilly, Frank, Hazy, & Watz, 2007; Pavlov, 1941; Rangel et al., 2008). In this way, the Pavlovian system approaches rewards and avoids punishments. One can reduce the behavior to simple reactions without the evaluation of the options available to form the decision. Changing of learned and hard-wired association is difficult.

The second system that is expected to drive valuation processes is the habit system. The habit system is not limited to a set of specific behavioral responses but involves learning through repetition and training (Rangel et al., 2008). Trial-and-error helps one to learn which potential outcomes are desirable and which ones to avoid (Boswell, 1947; Robert & John, 1908). As long as the environment is stable, one can connect the association between certain actions or behaviors to a certain value. As a consequence, the habit system is relatively slow and dependent on the repetition effects. Thus, in situations with novel options and courses of action, the subject is dependent on a "generalization" of previously learned habits, which might be maladaptive in a given context. Habits are also often used as an explanation for abusive behavior like drug addiction (Gerdeman, Partridge, Lupica, & Lovinger, 2003; Volkow, Fowler, Wang, Baler, & Telang, 2009; Williamson, Cheng, Etchegaray, & Meck, 2008).

The last and most commonly referred to system in behavioral economics and neuroeconomics is the goal-directed system. This system is the most complex one as it assigns values to actions and options by computing action-outcome associations and evaluates the rewards in comparison to other options that are available and associated with other outcomes. However, it is also the most flexible one, allowing the explanation of variance in many kinds of behaviors. Thus, the ideal decision making situation occurs when all information is available and one can integrate it into a coherent value. In that context, the subject is able to make the best decision that is in line with his or her goals. In turn, these goals experience the influence of many factors. If you are hungry, for example, your goal might be to get food. However, if you are well rested and sated, your goal might be to receive an education. Thus, your goals are not static, but vary with your state. It is not possible to explain these changes with the habitual system. Once you learn that pressing a certain button delivers food, the goal-directed system might change your behavior when you are sated. The habit system would continue pressing the button, because it delivered food. This would fall under positive reinforcement. Thus, the goal-directed system updates the value of the action and assigns a current value. One important component of the goal-directed system is the computational aspect. The system needs to be capable of integrating different aspects into one common (currency) value signal.

These three approaches to valuation processes are all able to explain a multitude of behaviors. However, it is assumed that these processes do not work in isolation. Rather they interact (de Wit et al., 2012; Loewenstein & O'Donoghue, 2007; Schulz, Fischbacher, Thöni, & Utikal, 2014). The question of which system "wins" depends on a variety of factors. If you are well rested and your goal is to lose weight, you might eat healthy food and do sports, thus following the goal-directed system. However, if you are exhausted, tired, and hungry, you are more to follow the habit system and buy unhealthy food in the supermarket. Thus, it is likely that multiple systems simultaneously active and compete against each other. The following chapter will evaluate the valuation process that is part of the goal-directed learning approach in more detail and introduce the neural underpinnings of the common currency.

The conceptuality of the stages presented above is especially useful for the field of decision neuroscience. In general, the valuation process gains special interest, especially since it offers the opportunity to investigate specific valuation processes using econometric models and to separate the decision making process into the distinct valuation processes and stages. The process is often unconscious, and therefore,

neuroimaging methods offer the possibility of investigating the underlying cognitive processes during decision making. In recent years, many published studies have tackled precisely these problems: 1) how the subjective values of the various options under consideration are learned, stored, and represented and 2) how decision are made and implemented.

2.1 The Neurobiology of Value-Based Decision Making

As indicated above, the valuation process in the goal-directed system is basically a transformation to a common currency that one can use to compare different types of rewards. As a result, they are comparable when one is deciding between several options (Brosch & Sander, 2013; Lim et al., 2011; Rangel et al., 2008). Over the last decade, our knowledge of the underlying mechanisms of those valuation processes has increased substantially (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Ballard et al., 2011; Balleine et al., 2008; Balleine, Delgado, & Hikosaka, 2007; Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Brosch & Sander, 2013; Crockett et al., 2013; Daniel & Pollmann, 2014; Delgado, Gillis, & Phelps, 2008; Glimcher, 2004; Grabenhorst & Rolls, 2009; Haber & Knutson, 2010; Hare, Hakimi, & Rangel, 2014; Hare et al., 2008; Knutson & Cooper, 2005; Lim et al., 2011; O'Doherty, 2004; Padoa-Schioppa & Assad, 2006; Robbins & Everitt, 1996; Schultz, 1998, 2006). One of the major purposes of neuroeconomics is to study the neural signal associated with the computation and representation of subjective value. As the introduction asserts, one of the methods most commonly used to study the neural underpinnings of cognitive processes is fMRI.
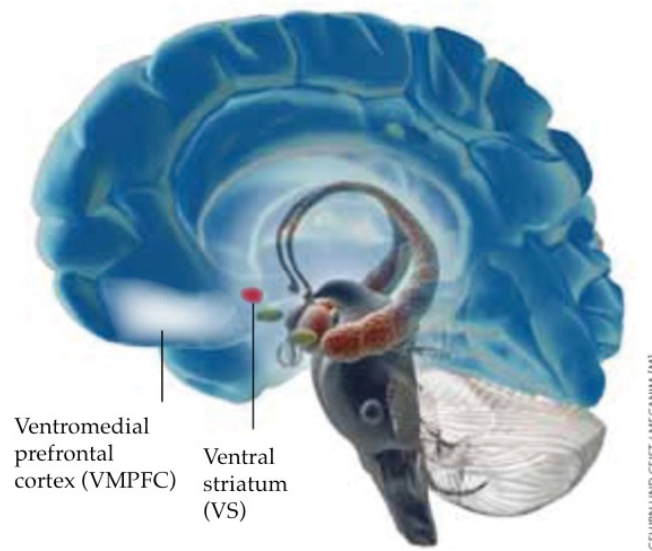
Figure 3: Schematic illustration of the brain. Sagittal plane depicting the VMPFC and the VS (adapted from Kalenscher & Strombach, 2013, p. 64).

Previous work on subjective value computation came to the conclusion that (1) BOLD signal in a small number of regions, including the striatum and VMPFC scales with subjective value of the options available during decision making (Bartra et al., 2013). In addition, there seem to be more regions associated with and implicated in the decision making process. For example, the dorsolateral prefrontal cortex contributes to the decision making process by exerting self control (Crockett et al., 2013; Hare, Camerer, & Rangel, 2009); (2) the same regions (striatum and VMPFC) process the act of receiving a rewards, implying that they are implicated in the computation of the values, decision formation, and the evaluation of the outcome. These steps all contribute to decision making in the formal model of Rangel et al. (2008); and (3) responses are modality independent. Thus, they respond to rewards from multiple domains, for instance, food, money, and social rewards.

In the following chapters, I will introduce two central regions associated with reward processing and value-based decision making–the striatum and the VMPFC. Both regions have been shown to process value and to contribute to the computation of subjective value. However, they also perform slightly different forms of mathematical calculations to obtain the aspects of subjective value.

2.1.1   The Striatum

The striatum is the primary input structure of the basal ganglia, a group of brain areas that also include the globus pallidus, the subthalamic nucleus, and the dopamine neuron-containing substantia nigra, located in the midbrain (see fig. 3 & 4; Alexander & Crutcher, 1990; Delgado, 2007; Graybiel, 2000). The striatum receives highly convergent projections from the prefrontal cortex, as well as afferents from the amygdala, the hippocampal formation, and the midbrain (Clithero & Rangel, 2013; Haber, 2003; Haber & Knutson, 2010; Robbins & Everitt, 1996). The striatum sends information back to the prefrontal cortex via the thalamus, forming a frontal – striatal loop (Robbins & Everitt, 1996). The tight connection between the striatum and frontal regions allow fast and flexible integration and involvement in motor, cognitive, and affective components of behavior. This also makes it highly adaptive for decision making processes.
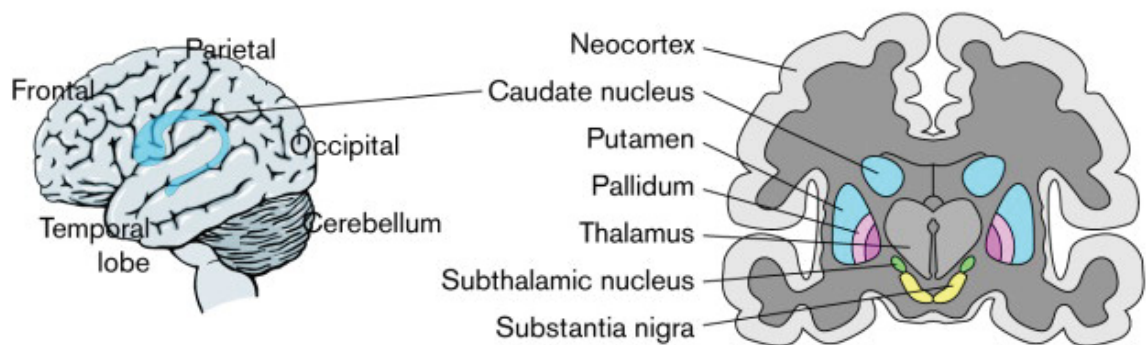
Figure 4: The basic anatomy of the brain showing the major regions within the basal ganglia: the striatum (blue), which is made up of the caudate nucleus and the putamen; the pallidum (pink), which is made up of outer and inner segments; the subthalamic nucleus (green); and the substantia nigra (yellow) (reprinted from Graybiel, 2000, p. 509).

The striatum can be divided into a dorsal (upper) and a ventral (lower) part (Alexander & Crutcher, 1990; Graybiel, 2000; Haber, 2003). This can be attributed to different aspects of the decision making process. While the ventral region is often associated with the expectation of a reward, the dorsal part is more involved in the experience of rewards (Balleine et al., 2007; Delgado, 2007).

Work on the neurobiology of behavior in non-human primates (macaques) has shown that the dopaminergic neurons in the striatum signal rewards (Schultz, 1998; Schultz, Apicella, Scarnati, & Ljungberg, 1992; Schultz, Dayan, & Montague, 1997). Schultz et al. showed that, in macaques, the dopaminergic neurons in the striatum respond more strongly to large rewards and less strongly to smaller rewards, thus signaling the size of the expected reward (Schultz et al., 1992). They were among the first to investigate the role of the dopaminergic neurons in striatum in reward signaling. In a later study, Schulz et al. refined these findings and determined that the responses also differed between unexpected rewards and expected rewards (fig. 5).
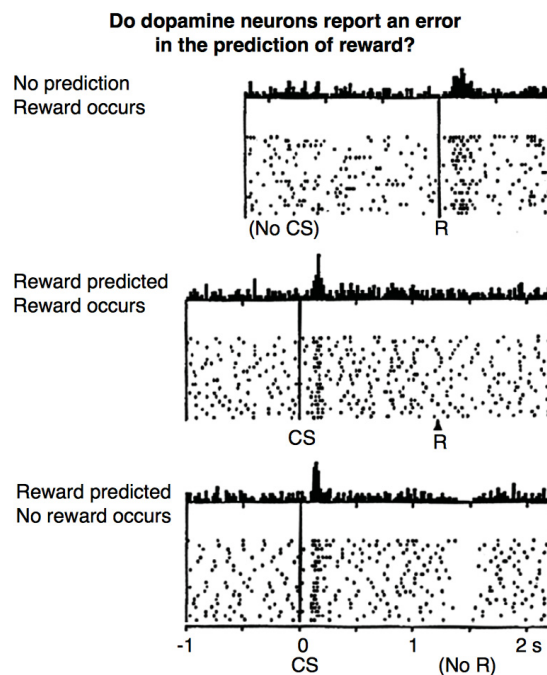
Figure 5: Dopamine neurons respond to errors in the prediction of an outcome (RPE). Top: Before learning, the monkey received an unpredicted drop of juice (R). This refers to a positive RPE. Middle: After learning, the monkey associated a signal with a reward, i.e. a drop of juice. The dopamine neurons react to the signal (CS), but not to the reward itself (R). Bottom: After learning, the conditioned stimulus indicates the upcoming administration of a reward. The reward fails to occur and the dopamine neurons react with a decreased firing rate during the time the reward was expected. The dopamine neurons thus code for the negative RPE. Each panel shows a time histogram and raster of impulses from the same neuron in different conditions. (reprinted from commSchultz et al., 1997, p. 1594)

They concluded that the striatum signaled rewards, but more importantly, that the signal reflected reward-prediction errors (RPE). Thus, when a reward was

administered, the monkeys' dopaminergic neurons fired as a response to the reward. However, when they expected the reward and it was administered, the neurons did not respond to it with increasing firing rates any more. The reason for this was that the RPE equaled zero. Thus, the expectation was equal to the actual outcome. However, when a reward was expected and none was delivered, the firing of the neurons reduced as a reaction to the negative prediction. That absence of a signal was likely to be implicated in learning processes, reflecting an update in the representation of the value of the chosen option.

In humans, almost every study investigating subjective values of rewards identifies the VS (Bartra et al., 2013; Daniel & Pollmann, 2014; Delgado, 2007; Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Diekhof, Kaps, Falkai, & Gruber, 2012; Fehr & Camerer, 2007; Hare et al., 2009; Hare et al., 2008; Kable & Glimcher, 2007; Kuss et al., 2013; Lim et al., 2011; O'Doherty et al., 2004; Pagnoni, Zink, Montague, & Berns, 2002). This also seems to be independent of the type of reward used in the experiment (Bartra et al., 2013). Kable and Glimcher (2007) showed that the VS and the VMPFC reflected the subjective value of a reward. By integrating a potential delay into their experimental design, they showed that subjects did not track the absolute monetary reward in the VS but the subjective value of that reward, discounting the delay from the monetary reward value.

Thus, the striatum, and especially the ventral part (ventral striatum, VS), are implicated in reward processing and constitute a major input unit to the basal ganglia and the dopaminergic reward system. The findings of Schultz et al. show that the dopaminergic neurons play an important role in value learning and shaping decisions. They do not play this role through hedonic signals but by coding positive and negative RPEs. Previous experiences and learned associations are integrated into the value signal, and the RPE reflects the deviation of the actual reward from expectations. Thus, the striatum is important in several stages of the decision making process. It values the options available, measured by the expected outcome, and then evaluates the actual outcome, updating value representations when necessary. The projection of inputs from regions like the VMPFC, the orbitofrontal cortex (OFC), and the anterior cingulate cortex (ACC) to the striatum highlight the central role of the striatum in the reward processing network and, consequently, in decision making.

2.1.2 The Ventromedial Prefrontal Cortex

The ventromedial prefrontal cortex (VMPFC) is part of the prefrontal cortex and is located at the bottom of the cerebral hemisphere (fig. 3). The VS has been identified in value-based decision making as often as the VMPFC has been named (Bartra et al., 2013; Clithero & Rangel, 2013). These findings corroborate to the theory that there is a fronto-striatal network strongly implicated in valuation processes and value-based decision making. Previous studies identified the VMPFC as being involved in the computation of subjective values both during the decision phase and outcome evaluation (Grabenhorst & Rolls, 2009; Hare, Camerer, Knoepfle, O'Doherty, & Rangel, 2010; Hare et al., 2009; Levy & Glimcher, 2012; Padoa-Schioppa & Assad, 2006; Rangel & Hare, 2010). In a review paper, Clithero and Rangel (2013) showed that the VMPFC was part of several distinct sub-networks, suggesting an involvement in the computation of various value-related signals. However, it has been suggested that, contrary to the VS, the VMPFC tracks the difference between values of the options in a decision making task. A subjective value thus integrates the other options available, not just the attended or favored option, but builds relative values (Lim et al., 2011).

It is likely that the VMPFC receives information from many cortical regions, depending on the context of the valuation process. In a study, Hare, Camerer and Rangel (2009) showed that activity in the DLPFC increased when self-control was exerted, and DLPFC modulated the value signal encoded in the VMPFC. On the one hand, that supports the idea that DLPFC is implicated in self-control but also suggests that different brain regions can modulate VMPFC activity (Chib, Rangel, Shimojo, & O'Doherty, 2009; Crockett et al., 2013; Hare et al., 2010; Hare et al., 2009; Hare et al., 2014; Kable & Glimcher, 2007; Rangel & Hare, 2010). The VMPFC thus receives input from different regions to form the value signal of a given option. To summarize, the VMPFC is thus strongly and reliably implicated in the decision process by reflecting value. The VMPFC is assumed to be a central orchestrating region, receiving input from several brain regions and integrating them into a coherent common value signal. Thus, while the VS is mainly associated with RPEs and, thus, the learning mechanisms in decision making, the VMPFC reflects the subjective value of a given option, integrating different information into one coherent "common currency".

3. Psychological and Economical Theories on Monetary Incentives

Initially, everyone would suggest that receiving monetary incentives was considered positive, motivating you to perform at your personal best and giving the maximum possible on the task you have been paid for. That is also common in the corporate world. Workers receive bonuses, stock options, and other monetary incentives to encourage them to perform better at their jobs (Camerer & Hogarth, 1999). However, that is a rather economic perspective. As I stated in the beginning, standard economic theory assumes that people only perform to get food, money, or other positive reinforcement for themselves. In this view, performance is tightly coupled with the incentives that people receive as compensation. Incentives thus allow modulation in a favorable way (Berridge, 2004; Camerer & Hogarth, 1999; Fehr & Falk, 2002).

Psychologists challenge this economical view of the positive effect of incentives. Psychological theories state that intrinsic motivation should be generally high enough to produce steady effort (Camerer & Hogarth, 1999; Deci, Koestner, & Ryan, 1999). Thus, even in the absence of financial or other non-monetary rewards, people will engage in a task and fulfill expectations out of the intrinsic drive to perform well. Psychological theories state that it is more likely that performance is even reduced when monetary incentives are present. Thus, contrary to the economic view, psychological approaches state that monetary incentives have a negative effect on performance (Deci, 1971; Deci et al., 1999; Mobbs et al., 2009). This detrimental effect of the incentives is especially visible when the intrinsic motivation to perform is initially high. Then the motivation to perform is crowded out by the external motivator, i.e., the performance-contingent reward (Arkes, Dawes, & Christensen, 1986; Ashton, 1990; Camerer & Hogarth, 1999). For example, you are painting pictures during your leisure time. A gallery shows interest in your pictures and even sells some for a reasonable price. After a while, you only paint the pictures because of the monetary perspective, not because you enjoy doing it. Thus, the initially intrinsic motivation to paint has been crowded out by the monetary incentive.

Psychologists agree that higher incentives might produce higher levels of effort. However, the increase in effort does not necessarily lead to an increase in performance (Baumeister & Showers, 1986; Camerer & Hogarth, 1999; Mobbs et al., 2009). A study by Baumeister (1984) showed that when the importance of an external incentive is

highlighted, performance drops because of the increased pressure to perform well. Baumeister refers to the fact that the incentives lead to a shift in attention to the internal performance process. This is often the case in important sports championships.

Setting aside the discrepancies about the directionality of the effect the incentives have on performance, psychologists and economists agree that removing the monetary incentive harms performance tremendously. Thus, although psychological and economic theories disagree about the effect of the introduction of monetary incentives, they agree that removing the incentives leads to disengagement in the tasks and therefore reduced performance on a given tasks (Camerer & Hogarth, 1999).

A meta-analysis by Camerer and Hogarth reviewed previous studies in the search for commonalities. It came to the conclusion that "for easy and hard jobs, and intrinsically motivated workers, marginal changes in incentives will not improve performance much. However, for boring jobs, unmotivated workers, or tasks in which variance is bad, incentives are likely to have a positive effect" (Camerer & Hogarth, 1999, p. 35). Thus, there is no coherent opinion on the effect of incentives on performance. Neuroscience might offer alternative approaches to investigate the effect of performance-based incentives. Neuroimaging techniques offer an opportunity to investigate the underlying mental processes and to study the unconscious processes that might drive changes in behavior. Thus, while performance is a direct measure of behavior, the neural correlates are a proxy of the underlying cognitive processes. Therefore, the following chapter will discuss neuroimaging studies on the effect of incentives on behavior.

3.1 Neural Correlates of Monetary Incentives

From a neuroeconomical point of view, it seems natural to assume that a change in performance due to a change in incentives arises from changes in activity in a dedicated brain region. Thus, if a monetary incentive is introduced, the neural representation transforms as a reaction to the change and modifies performance.

As indicated earlier, the VMPFC and the VS are central parts in the processing of value. Therefore, these regions are also assumed to be involved in the processing of performance-based incentives. The VS is sensitive to rewards and RPEs, and these

reinforcements are often confounded with motivation, which drives performance (Arana et al., 2003; Berridge, 2004; Miller, Shankar, Knutson, & McClure, 2014). Therefore, the aim is to differentiate between valuation and motivational processes. When a reward is applied, the motivation to perform increases. Therefore, a positive firing rate in the VS might indicate the initiation of higher levels of motivation on a higher level of cognitive processing, leading to goal-directed behavior (Soutschek, Stelzel, Paschke, Walter, & Schubert, 2014). Previous studies found that higher expected rewards translate into higher effort exertion (Berridge, 2004; Schmidt, Lebreton, Cléry-Melin, Daunizeau, & Pessiglione, 2012). Certain reward amounts anticipated after successful performance on given tasks elicit increases in activity in the VS proportional to reward amounts. Also, the degree of signal change correlates to the degree of performance enhancement (Miller et al., 2014; Pessiglione et al., 2007; Schmidt et al., 2012). An interaction between these common motivational signals in the VS and different task-specific systems underpinning behavioral performance is suggested to drive motivational processes and to modulate performance on specific tasks (Schmidt et al., 2012). Schultz et al. (Schultz, 1998; Schultz et al., 1992; Schultz et al., 1997) indicate that the brain and particularly dopaminergic neurons in the striatum respond to errors in the prediction of rewards and use the information to update future behavior. That might also be an important underlying premise when studying the effect of incentives. While the constant application of rewards does not necessarily affect motivational processes, the change in incentives is associated with a change in the underlying neural correlates and, ultimately, in the performance on a behavioral level. Thus, the main focus in research on incentives is the disentanglement of the simultaneous processes of an upcoming reward, the implementation of motivational signals, and learning, which requires constant updates on decision-outcome relationships (Gilbert & Fiez, 2004).

The value of the incentive and, thus, the degree of the incentive and the motivation that it initiates are also processed in regions associated with subjective value, i.e., the prefrontal regions in general and the medial prefrontal cortex in particular (Arana et al., 2003). The medial regions were shown to react more to subjective value computation and choice behavior, while the lateral fraction responded primarily to the suppression of a desired outcome in a decision. In an fMRI study, Kouneiher et al. (2009) showed that connectivity between several prefrontal regions associated with subjective value computations changed as a function of motivational factors, e.g., monetary incentives. However, these changes were rather slow, reacting not to single trials but to changes between blocks (Kouneiher et al., 2009). Therefore,

motivation was suggested not to change on-line but as a prudent process, reflecting incentive-induced fluctuations over time.

Paradoxically, the option to earn very large rewards has been shown to diminish performance. Baumeister (Baumeister, 1984; Baumeister & Showers, 1986) showed that the crowding out effect led to lower levels of performance when incentives were present. Neuroimaging studies suggested that the VS was a core component in these processes. An fMRI study indicated that monetary incentives led to increased midbrain activity, associated with reduced performance (Mobbs et al., 2009). The association between midbrain activity and the decrease in performance was suggested to be an "overmotivation" signal for the high reward and the perceived pressure to perform well. The suggested mechanism driving this process was that incentives were first encoded as potential gains associated with successful task performance. However during task completion, the VS encoded potential losses that might arise from failure in the task (Chib, De Martino, Shimojo, & O'Doherty, 2012; Mobbs et al., 2009). Thus, the VS contributed to the detrimental effect of monetary incentives by mediating the interaction between incentives and behavioral performance.

In addition, several other regions might play a role in the processing of incentive motivation. Research reliably indicated that the dorsolateral prefrontal cortex (DLPFC) was involved in executive and self-control (Crockett et al., 2013; Gilbert & Fiez, 2004; Hare et al., 2010; Hare et al., 2009; Hare et al., 2014; Hare, Malmaud, & Rangel, 2011). It is likely that these processes are also affected by changes in incentives. Neurophysiological studies in nonhuman primates showed that task incentives can affect performance via increased activity in DLPFC neurons, demonstrating greater accuracy on rewarded trials (Kobayashi, Lauwereyns, Koizumi, Sakagami, & Hikosaka, 2002; Leon & Shadlen, 1999; Roesch & Olson, 2003). The studies suggest that DLPFC signals the "effort" that is put into solving the task or the facilitation of the cognitive processes needed for task fulfillment but not the pure anticipation of incentives (Gilbert & Fiez, 2004).

To summarize, the VS seems to be the core component in incentive motivation processing. However, the VMPFC and other subjective value-related regions are likely to be involved in the process by integrating motivational signals or mediating the signals to task-related regions. The mechanism suggested involves the VS serving as

the initiative node in the network modulating activity in task- and subjective value-related regions. This ultimately leads to changes in the performance in the task.

4. Social Reward and Social Decision Making

The previous chapter focused on individual decisions and, thus, on actions that only affected the participant and not his/her social environment. However, we live in a highly complex social world in which many decisions are made in the context of social interactions (Rilling & Sanfey, 2011). Our society is built upon the premise that everyone contributes to the common good, e.g., to pay for pensions, unemployment benefits and infrastructure. So cooperation and social exchange are important cornerstones of human society. In our personal lives, many actions also take place within a social context, for example, collaborating with colleagues on a project at work or sharing lunch with a friend. The pursuit of others' approval and acceptance may reflect the desire to fulfill social needs. The prerequisite for that is the ability to form relationships and social networks (Baumeister & Leary, 1995; Fareri & Delgado, 2014b; Van Winden, Stallen, & Ridderinkhof, 2008).

Axelrod and Hamilton (1981) stated that, against the idea of survival of the fittest individual, cooperation is much more common within, but even between, species. They highlighted the significance of forming social groups. From a value-based decision-making perspective, it seems natural to assume that social interaction, cooperation, or prosocial behavior is processed as some kind of value. In other words, we experience pleasure when interacting with other people and behavior is not solely driven by material self interest (Fehr & Fischbacher, 2002), indicating that a reward that another person consumes appears in our own utility function to some extent (Camerer & Fehr, 2006). This principle refers to social preferences, which means that we care positively or negatively for the well-being of other people.

"A person exhibits social preferences if the person does not only care about the material resources allocated to her but also cares about the material resources allocated to relevant reference agents" (Fehr, Fischbacher, & Gächter, 2002, p. C2). The standard economic theories as presented in the introduction and highlighting pure self-interest have no valid explanation for the emergence of cooperation and competition in markets and organizations. Social preferences like fairness shape the functioning of competition and cooperation (Fehr & Fischbacher, 2002; Fehr et al., 2002).

4.1 Social Preferences

One of the core social preferences is *reciprocity*. Reciprocity implies responding to "actions that are perceived to be kind in a kind manner, and to actions that are perceived to be hostile in a hostile manner" (Fehr et al., 2002, p. 2). What exactly is regarded as kind or hostile is based on the (un-)fairness of the consequences and the intentions associated with the actions of the opponent. "Fair" can be considered an equitable distribution of payoffs, relative to the set of possible payoff distributions. Thus, a fair split of a reward might be 50/50. The difference between cooperation and reciprocity is the anticipation of future material benefit in cooperation. Cooperation, thus, implies an interest in the maximization of the material output from the interaction with another actor (Falk, Fehr, & Fischbacher, 2003; Fehr & Gächter, 1998). Thus, people expressing high reciprocity will behave prosocially if they have previously experienced prosocial behavior from the interaction partner.

Closely related to the social preference of reciprocity, is *inequity aversion*. As the term implies, a person with high aversion for inequity will be generous toward another person if the other person's material payoff is below the equitable benchmark. However, the former will feel envy if the latter's payoff is higher than the equitable level (Fehr & Schmidt, 1999). Thus, although the underlying reasons for prosocial behavior might differ, people with high reciprocity and inequity aversion preferences behave in a similar manner to maintain a fair share of the resources available. Social preferences are thus strong influences in our decision-making processes, highlighting the importance of incorporating them into theories of decision making and economic models.

4.2 Neural correlates of social decision making

The most prominent question in social neuroscience is whether the same regions are involved in the computation of social and non-social value. Ruff and Fehr suggested two distinct models for how social values might be integrated into the valuation process (Ruff & Fehr, 2014, compare fig. 6). As outlined before, the VS and the VMPFC are core elements of the reward-related brain network. Therefore, one idea is that the same regions process different kind of rewards, implying that there is a common currency that makes options with different qualities comparable ("extended common

currency schema", fig. 6a). An alternative approach is that there are two distinct brain systems that process either social or non-social rewards ("social-valuation-specific schema", fig. 6b). While the common currency idea implies that there are some regions involved in both types of reward, the social-valuation-specific schema states that there are two distinct networks, which are specialized for specific input (Ruff & Fehr, 2014).
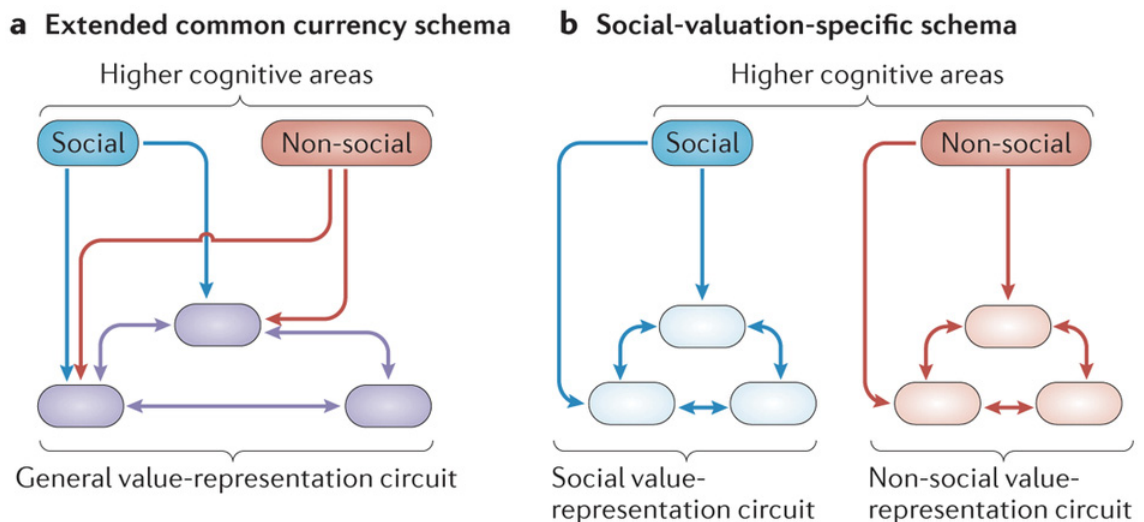
Figure 6: Two schemas for neural value computation in social versus non-social context. Two different ideas might be plausible about how values of different kinds are computed might be plausible. (a) The "extended currency schema" suggests that there is one common node in the brain that reflects integrated values and, thus, reflects social and non-social reward values. However, the information that is relevant for the value computation might differ between social and non-social choices and may therefore be provided by distinct brain regions that are either specific to social or to non-social information processing. Social and non-social rewards can thus result in identical activity in reward-related brain areas, but differ in their connectivity with other brain regions. (b) The "social-valuation-specific schema" is based on the social brain hypothesis. The hypothesis states that there are specific regions in the brain that selectively respond to social stimuli. Therefore, it is suggested that there are two rather identical processes, using similar neural computations, that either process social or non-social rewards. That includes the idea that there are specific and distinct brain regions that either process social or non-social valuation (reprinted from Ruff & Fehr, 2014, p. 3).

The social-specific-valuation-specific schema was built upon the idea that there is social-specific cognition and, thus, a dedicated network that is only responsive to social stimuli. The study of the visual system and findings indicating that some regions, such as the temporoparietal junction and the fusiform face area, only responded to social stimuli like faces or bodies (Gauthier et al., 2000; Kanwisher & Yovel, 2006; Morishima et al., 2012; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014)

led to the conclusion that specific regions involved in representing the intentions, emotions or actions of other people and the control of social and non-social behavior should differ fundamentally in terms of neural architecture. However, more recent neuroeconomical studies challenge this approach by suggesting a common region that reflects different kinds of values (Chib et al., 2009; Hare et al., 2010; Hare et al., 2009). The general common currency approach, thus, suggests that there is a value-specific region and that the processing of social and non-social values is similar. However, the input for these value computations might come from distinct regions. Thus, in this context, the TPJ might be an input region, providing information for value computation.

The idea of a common currency implies that the same regions are involved in social and non-social valuation processes. Lin et al. (2012) suggested that the VS and VMPFC computed the value of the options available to make a decision. The authors suggested that the same concept might be applied to social and non-social rewards, thus analogously weighting egoistic and prosocial considerations to make a decision. To do that, it is necessary to convert the options with different qualities into a common value signal. Lin et al. come to the conclusion that this is indeed the case. Thus, the subjective value and the experiences utility of a selected option are represented in the VMPFC, while the RPE correlates with VS activity. That entails strong support for a neural signal that converts options of different qualities into a coherent common currency, also allowing comparison between different kinds of rewards. The core regions in social reward processing are analogous to monetary reward processing, i.e. the VMPFC and the VS. These findings were confirmed in a multitude of experimental studies (Chib et al., 2009; Fareri, Chang, & Delgado, 2012; Fareri & Delgado, 2014a; Fareri, Niznikiewicz, Lee, & Delgado, 2012; Hare et al., 2010; Izuma, Saito, & Sadato, 2008, 2010; Levy & Glimcher, 2011; Lin et al., 2012; Morishima et al., 2012; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Smith, Clithero, Boltuck, & Huettel, 2014; Winecoff et al., 2013) and meta-analytic reviews (Bartra et al., 2013; Bhanji & Delgado, 2014; Carter & Huettel, 2013; Clithero & Rangel, 2013; Declerck, Boone, & Emonds, 2013; Levy & Glimcher, 2012; Rilling & Sanfey, 2011; Ruff & Fehr, 2014). To summarize, based on the neuroimaging studies, monetary and social reward values are processes in overlapping brain regions. However, it is likely that the various forms of input for the computation of a common currency come from different regions in the brain, depending on the kind of input (Crockett et al., 2013; Hare et al., 2010; Hare et al., 2009; Park, Kahnt, Rieskamp, & Heekeren, 2011; Smith et al., 2014).

One core candidate for the input region for social information in value computation is the TPJ. The TPJ has often been tied to tasks associated with the theory of mind, which refers to the ability to put oneself in someone else's shoes (Frith & Frith, 2005; Schurz et al., 2014; Stone, Baron-Cohen, & Knight, 1998), and social cognition, for instance, mentalizing (Van Overwalle, 2009), perspective taking (Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006; Strang, Utikal, Fischbacher, Weber, & Falk, 2014), or the detection of social agents in the environment (Schultz, Friston, O'Doherty, Wolpert, & Frith, 2005; Tankersley, Stowe, & Huettel, 2007). The TPJ is, thus, capable of processing diverse contextual cues about the social environment. This might influence our interpersonal interactions and guide social decisions (Smith et al., 2014). Previous research showed that social valuation relied on the interaction between regions like the VMPFC and specific regions, including the TPJ, that were modulated by social information (Carter, Bowling, Reeck, & Huettel, 2012) and that the strength of the connectivity between these two regions relied on the subjective value of the social reward (Hare et al., 2010; Smith et al., 2014; van den Bos, Talwar, & McClure, 2013). For these reasons, it is likely that there is 1) a common currency used to make decisions involving options with different qualities, and 2) the VMPFC receives input from interacting brain systems, contributing to the computation of subjective value.

PART II

1. Concluding Remarks on the studies

In the second part of my dissertation, I present three studies on the processing of changes in monetary incentives and varying social reward. In all three studies, monetary reward is used as an objective measure of a reward. However, the subjective value of the social or monetary reward is dependent on the environmental and situational circumstances. The first study uses fMRI to study the effect of the introduction and removal of performance-contingent rewards. The second study is also an fMRI study, and it presents novel findings on the neural correlates of social discounting and the role of the TPJ in prosocial choice behavior. Finally, the third study is purely behavioral and uses the social discounting paradigm to systematically study the effect of cognitive load on prosocial behavior.

On the following pages, I present the scope of the studies, discuss potential conclusions, and evaluate the contributions the studies make to the scientific community. There are detailed descriptions of the studies and discussions in the appendix.

2. Study 1: Neural Underpinnings of Performance-Based Incentives

Strombach, T., Hubert, M., and Kenning, P. (in press). Neural underpinnings of performance-based incentives. *Journal of Economic Psychology.*

The study addresses a central question in the contemporary corporate world, i.e., the modulatory effect of performance-based incentives. In today's economy, incentives are among the most commonly used motivators to drive employees' performance (Colin Camerer & Hogarth, 1999). However, it is not well understood how these performance-contingent incentives modulate the neural underpinnings of the cognitive processing of a given task (Albrecht, Abeler, Weber, & Falk, 2014; Murayama, Matsumoto, Izuma, & Matsumoto, 2010). We investigate the extent to which performance-based incentives change the neural representation and perception of a specific task. In this study we combine economic theory with a psychological framework and use the advantages of

neurobiological methods to get a more profound understanding of an economic phenomenon, i.e., the effect of monetary incentives. We hypothesize that the changes in monetary incentives affect cognitive processing on three different levels: 1) reward related areas should process the newly introduced reward, 2) task-related areas become more active as a response to the monetary reward, and 3) effort-related regions should reflect a higher effort by increased activity. To do that, we change the monetary incentives for making the correct decision in an arithmetic calculations task. In a first block, without incentives, then with incentives for correct performance and in a third block incentives were removed again.

We use the direct measure of performance and measure the underlying cognitive processes by using fMRI simultaneously. We find that incentives induce changes in reward-related brain regions but not in task-related neural representations. Interestingly, when monetary incentives are introduced, blood oxygenation level-dependent activity increases in the VS, which is sensitive to reward (Clithero & Rangel, 2013; Delgado, 2007), and decreases in the VMPFC (Chib et al., 2009; Hare et al., 2009), which has been reliably shown to be involved in the calculation of subjective values. This corroborates the idea that the VS and VMPFC are central regions in motivation and reward processing (Albrecht et al., 2014; Bartra et al., 2013; Clithero & Rangel, 2013). We suggest that pay-for-performance does not directly affect performance by modulating neural activity in task-relevant regions but affects reward representation during task completion. Therefore, we conclude that performance-contingent incentives need to be applied carefully. This finding is of special interest for employee compensation in unstable economic environments, as companies in unstable environments might not permanently maintain pay-for-performance.

In the research article, we try to present one way of using neuroscientific methods to get a deeper understanding of economic behavior and to integrate neural findings into economic theory. Importantly, our article is also intended as an appeal to the broader research community, conveying the message that biologically based models can provide a framework for studying economic processes in the corporate world today. Although we study fundamental processes involving incentives and motivational processes, the paradigm is suited to getting a grasp on real-world economic questions. The study, therefore, contributes to the further development of neuroeconomic theory and the question of how incentives affect our performance and perception tasks in general and to the understanding of incentive-induced changes in motivation and performance in particular.

In summary, we were thus able to replicate previous findings on reward processing and the neural underpinnings of changes in incentives (Albrecht et al., 2014; Murayama et al., 2010) and to extend knowledge of the underlying cognitive processes. Nevertheless, the paradigm includes some weaknesses that need to be discussed critically. It is clear from previous studies that motivation and reward processing are strongly correlated in most cases (Arana et al., 2003; Kouneiher et al., 2009; Liljeholm & O'Doherty, 2012; Miller et al., 2014). With the current design, we are not able to differentiate between these two processes. A study by Miller et al. (2014) attempted to dissociate the processing of reward and motivation. The authors suggested that striatal activation outside the nucleus accumbens was associated more with motivation while the caudate and putamen scale were associated more with the expected reward outcome. However, a coherent model of motivation and reward still needs to be developed and tested in different settings. Therefore, for future research, the distinction between the cognitive processes associated with motivation and reward processing should be investigated.

3. Study 2: Social Discounting Involves Modulation of Neural Value Signals by Temporoparietal Junction

Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I. I., Tobler, P. N., & Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences, 112(5),* **1619-1624.**

One of the most relevant topics in decision making and society is the issue of why people are altruistic towards some individuals and selfish towards others (Jones & Rachlin, 2006; Jones & Rachlin, 2009; Rachlin & Jones, 2008). In this study, we tackle that question from a neuroeconomic point of view, combining psychological theories of prosocial behavior, an economic model that reconstructs decision making, and neuroscientific imaging methods to integrate another level of detail.

People often consider the well-being of others in their decision making (Falk et al., 2003; Fehr & Fischbacher, 2002; Fehr et al., 2002; Fehr & Schmidt, 1999). Nevertheless, they differentiate between people they feel close to (e.g., their partners or

mothers) and people they do not have any relationship with (e.g., strangers or people they only see sporadically on the street). Thus, perceived social distance from one's interaction partner seems to influence willingness to forego a reward in order to be generous to that specific person. This phenomenon—the decline in generosity across social distance—is called social discounting (Jones & Rachlin, 2006). In the present study, we investigate how the integration of social distance into the valuation process of decision making is realized in the brain. Social discounting refers to the idea that social decision making balances egoistic, self-regarding motives and generous, other-regarding considerations; this balance tilts towards selfishness with increasing social distance between the donor and recipient (Strombach et al., 2014). In our social discounting experiment, we systematically and independently varied other-regarding preferences and the strength of the egoism bias as a function of social distance so that we were able to discriminate between the influence of social distance and purely selfish, reward-maximizing strategies or social preferences like fairness.

A region that is often associated with mentalizing and perspective taking is the temporoparietal junction (TPJ) (Mars et al., 2012; Morishima et al., 2012; Schurz et al., 2014). We hypothesized that the TPJ would orchestrate the balance between generous and selfish considerations depending on social distance. Furthermore, the design allowed discrimination between two competing ideas about the role of the TPJ in prosocial choice. On the one hand, the theory of mind suggests that TPJ reflects the closeness between the donor and recipient and, thus, the ability to put oneself in someone else's shoes (Frith & Frith, 2007; Saxe & Kanwisher, 2003; Young, Dodell-Feder, & Saxe, 2010). Conversely, a recently proposed model suggests that the TPJ reflects the temptation to be selfish to facilitate overcoming the selfish motives in the decision-making process (Morishima et al., 2012; Silani, Lamm, Ruff, & Singer, 2013).

Using functional magnetic resonance imaging (fMRI), we investigate the neural underpinnings of social discounting. In our study, we do not find support for a correlation of TPJ activity with the econometrically reconstructed other-regarding value or social distance. However, we find support for the second idea about the TPJ, suggesting that TPJ activity reflects the temptation to be selfish. In fact, the TPJ is more active during generous choices than during selfish choices and covaries with the temptation to choose selfishly. Additionally, we find greater connectivity between the TPJ and the VMPFC, a region associated with subjective value, during generous decisions than during selfish ones (Bartra et al., 2013; Hare et al., 2009). Thus, the

higher the temptation to be selfish while making a prosocial choice, the greater the upregulation of the activity in the brain's valuation system by the TPJ.

The results are interesting for a couple of reasons. First of all, we propose a neural model about how prosocial considerations are integrated into the valuation and decision making process. Therefore, the study contributed to the further development of a biologically plausible model of decision making. That, of course, has implications for how social decision making is conceptualized in psychological models. In the present study, we show that TPJ does not merely reflect mentalizing and perspective tasking processes (actions associated with the theory of mind). We contribute to a more concrete picture of how social decisions are made—balancing egoistic and prosocial, social distance-dependent considerations—and this input is likely to originate from the TPJ. This contributes to the extension of the common currency idea. Besides the input from DLPFC in self-control processes (Hare et al., 2009; Hare et al., 2014), it is likely that the VMPFC receives input also from other regions to form a subjective value. We suggest that, in social decision making, the input might come from the TPJ. Thus, the study contributes to the refinement of psychological theories on prosocial choice. However, it also contributes to economic theory by highlighting the influence of social distance on economic decision making. Until this point, social distance had primarily been ignored in economic approaches. Here, we show that social distance modulates prosocial behavior, which is also a prerequisite for cooperation between actors (Axelrod & Hamilton, 1981; Fehr & Gächter, 1999). The study is, thus, an example of the successful integration of multidisciplinary approaches to develop a more detailed and elaborate picture of economic and social behavior.

The findings can also be used to develop application-oriented approaches. While clinical scientists and psychologists can use the data to develop models of antisocial behavior and ultimately develop strategies and therapies for treating antisocial behavior, economists can get insight into the variability of social preferences and their underlying neurobiology.

Future studies should focus on the replication of the patterns identified in the study. Thus, repeating the study in different social settings with diverse incentives might help to test whether we identified a specific form of social decision making or if the proposed model can be generalized to social decision making.

4. Study 3: Gender-specific effects of cognitive load on social discounting

Strombach, T., Gorczyca, B. & Kalenscher, T. (submitted). Social Discounting & limited cognitive resources: gender interacts with cognitive load.

Multitasking has become increasingly normal to us. At the same time, our cognitive resources are limited (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Gailliot & Baumeister, 2007). Consequently, cognitive load is high in many situations—including those where we have to make decisions. This has given rise to the question of how these occupied limited resources affect decision making. There is an ongoing discussion about the effect that cognitive load has on social decision making (Camerer & Hogarth, 1999). While some researchers report that subjects become more generous towards their interaction partners when cognitive load is high (Roch, Lane, Samuelson, Allison, & Dent, 2000; Schulz, Fischbacher, Thöni, & Utikal, 2014), others report more egoistic behavior—less willingness to share with an opponent (Crelley, Lea, & Fischer, 2008; Moore & Loewenstein, 2004).

The inconclusive findings on the effect on prosocial choice suggest that there is at least one factor missing in the analysis that might explain the diverse results. In the present study, we hypothesize that the missing factor might be social distance. The analysis of previous studies indicates that those studies reporting more generosity with high cognitive load use anonymous interaction partners: thus, people with very high social distance. By contrast, studies reporting less generous behavior use non-anonymous recipients of prosocial behavior: thus, interaction partners with lower social distances. Consequently, the perceived social distance is hypothesized to modulate the effect of cognitive load on prosocial behavior. To test this, high or low cognitive loads are induced in subjects, who subsequently perform a social discounting task, comparable to the one used in study 2.

In the present study, we varied the level of cognitive load by using different levels of difficulty in demanding cognitive tasks. Our data indicates that social distance is indeed an important factor modulating the influence of the cognitive load level. However, it interacts with gender. While women are not significantly affected by cognitive load, men show less steep discounting when cognitive load is high. We

interpret our results within the framework of the dual process approach (Chaiken & Trope, 1999; Schulz et al., 2014), which states that decision making is always a balance of deliberate and automatic processes. When someone is well rested and cognitive load is low, the deliberate system prevails in the decision-making process. However, when the cognitive load is high and cognitive resources are exhausted, the automatic processes dominate the decision-making process. Similarly, going grocery shopping sated or hungry will influence the food items you choose. Referring to our study, men with high cognitive loads appear to use the deliberate processes less than women with high cognitive loads.

Referring to the systems introduced in part I of this dissertation (Pavlovian, habit and goal-directed; Rangel et al., 2008), the habit system is likely to use more automatic or model-free processes, while the goal-directed one makes use of deliberate and model-based processes. Therefore, the present study suggests that, although different terminologies might be used, it is likely that there are different systems which control behavior dependent on the circumstances of the decision maker. It seems also reasonable to suggest that the different systems lead to different behavioral outcomes. Cognitive load is likely to be one factor that modulates which system is applied to the decision making problem. When cognitive load is high more automatic strategies or habits are used when cognitive resources are scarce, especially in men.

The study contributes to the ongoing discussion about the effect of cognitive load on decision making by integrating social distance into the theoretical approach. A better understanding of this construct might help to develop more effective strategies for dealing with the risks of cognitive exhaustion and improving the quality of decisions made when cognitive capacity is scarce, e.g., during jobs that demand multitasking or when time pressure is high.

5. General Conclusion

All three studies described in Part II investigate reward processing using different methods and experimental designs. Together, they provide an improved understanding of the processing of incentives in individual situations (performance-based incentives) and social reward (neural correlates of social discounting and

behavioral effects of cognitive load on social discounting), contributing to a multidimensional view of subjective valuation processes.

Study 1 shows that neural correlates in regions associated with reward processing react differently to cognitive tasks, depending on whether or not a monetary incentive is available. Previous research already indicated that the VS tracks the availability of performance-contingent rewards. However, we show that the removal of performance-contingent incentives changes neural activity in regions associated with reward processing, even when the task stays the same over time. We hypothesize that the VS modulates performance without directly affecting activity in task-related regions. We conclude that reward is a strong modulator and that changes in rewards have an especially strong influence on behavior and should, therefore, be applied carefully.

One of the goals of this dissertation is to establish a valid and reliable paradigm for measuring social distance-dependent prosocial decision making. We are able to show that the findings on social discounting are stable over experiments and therefore our paradigm offers a valid alternative to the original paradigm by Jones and Rachlin (Jones & Rachlin, 2006; Rachlin & Jones, 2008). The different methods and environmental settings highlight the importance of integrating social distance into models of decision making and economic theory. With this dissertation, I have, therefore, contributed to refining economic theory. Based on my results, I suggest extending the concept of social preferences by including a social distance dimension. We also show that the TPJ contributed to the social decision making process by modulating basic value signals in the VMPFC. The design of the experiment allows the investigation of the systematic influence of social distance and the interaction of prosocial and egoistic considerations. We show that the social decision is heavily reliant on the balance between other-oriented and self-oriented considerations, which can be distorted when cognitive load is high. We are therefore able to suggest a biologically plausible model of prosocial choice, which seems to be reliant on the environmental circumstances.

Nevertheless, there are still some unanswered questions that should be addressed in future research. Manipulation studies should be used to investigate whether the TPJ is not only involved in the prosocial choice behavior but also causally involved. Using for example transcranial magnetic stimulation (TMS), for example, we can disrupt the functioning of that specific region. This might indicate a potential

causal influence of TPJ in prosocial choice behavior. The combination of several methods (high temporal resolution (e.g., EEG) with high spatial resolution (e.g., fMRI)) might be fruitful, increasing the level of description in both the first and second fMRI studies.

One of the main weaknesses of fMRI is the question of ecological validity. The subjects are in the scanner and answer very basic questions with a limited number of options available. Therefore, future research should try to develop methods for testing neural underpinnings in real life settings to validate the findings. Although there is no method to reliably test that, neuroscience needs to keep in mind that the experiments are looking at very reduced and limited behaviors and decisions.

To sum up, subjective valuation is a highly complex process in both social and non-social contexts. With the studies included in this dissertation, I contribute to the advancement of neuroeconomic theory on value-based decision making and reward processing. The dissertation may help to refine and advance existing approaches on the neural underpinnings of choice behavior and economic theory.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron, 50*(3), 507-517.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *Neuroimage, 30*(3), 1059-1068.

Albrecht, K., Abeler, J., Weber, B., & Falk, A. (2014). The brain correlates of the effects of monetary and verbal rewards on intrinsic motivation. *Frontiers in neuroscience, 8*.

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in neurosciences, 13*(7), 266-271.

Arana, F. S., Parkinson, J. A., Hinton, E., Holland, A. J., Owen, A. M., & Roberts, A. C. (2003). Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *The journal of neuroscience, 23*(29), 9632-9638.

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes, 37*(1), 93-110.

Ashby, F. B. (Ed.). (2011). *Statistical Analysis of fMRI Data*. Cambrige: MIT Press.

Ashton, R. H. (1990). Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research*, 148-180.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390-1396.

Ballard, I. C., Murty, V. P., Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2011). Dorsolateral prefrontal cortex drives mesolimbic dopaminergic regions to initiate motivated behavior. *The Journal of Neuroscience, 31*(28), 10340-10346.

Balleine, B. W., Daw, N. D., & O'Doherty, J. P. (2008). Multiple forms of value learning and the function of dopamine. *Neuroeconomics: decision making and the brain*, 367-385.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience, 27*(31), 8161-8165.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76*, 412-427.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology, 46*(3), 610.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *Journal of personality and social psychology, 74*(5), 1252.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin, 117*(3), 497.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*(4), 361-383.

Becker, G. S. (1974). A Theory of Social Interactions. *Journal of Political Economy, 82*(6).

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & behavior, 81*(2), 179-209.

Bhanji, J. P., & Delgado, M. R. (2014). The social brain and reward: social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(1), 61-73.

Boswell, F. P. (1947). Trial and error learning. *Psychological review, 54*(5), 282.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron, 30*(2), 619-639.

Brosch, T., & Sander, D. (2013). Neurocognitive mechanisms underlying value-based decision-making: from core values to economic value. *Frontiers in human neuroscience, 7*.

Bzdok, D., Langner, R., Schilbach, L., Jakobs, O., Roski, C., Caspers, S., . . . Eickhoff, S. B. (2013). Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. *Neuroimage, 81*, 381-392.

Camerer, C., & Fehr, E. (2006). When does" economic man" dominate social behavior? *science, 311*(5757), 47-52.

Camerer, C., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty, 19*(1-3), 7-42.

Camerer, C., & Loewenstein, G. (2004). *Behavioral economics: Past, present, future*: Princeton: Princeton University Press.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of economic Literature*, 9-64.

Carter, R. M., Bowling, D. L., Reeck, C., & Huettel, S. A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science, 337*(6090), 109-111.

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal–parietal junction. *Trends in cognitive sciences, 17*(7), 328-336.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chib, V. S., De Martino, B., Shimojo, S., & O'Doherty, J. P. (2012). Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron, 74*(3), 582-594.

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *The Journal of Neuroscience, 29*(39), 12315-12320.

Clithero, J. A., & Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. *Social cognitive and affective neuroscience*, *9(9),* 1289-1302..

Crelley, D., Lea, S., & Fischer, P. (2008). Ego Depletion and the Tragedy of the Commons: Self Regulation Fatigue in Public Goods Games. *Unpublished manuscript*.

Crockett, M. J., Braams, B. R., Clark, L., Tobler, P. N., Robbins, T. W., & Kalenscher, T. (2013). Restricting temptations: Neural mechanisms of precommitment. *Neuron, 79*(2), 391-401.

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural brain research, 41*(2), 81-94.

Daniel, R., & Pollmann, S. (2014). A universal role of the ventral striatum in reward-based learning: Evidence from human studies. *Neurobiology of learning and memory, 114*, 90-100.

Dayan, P. (2008). Better than conscious? Implications for Performance and Institutional Analysis. In C. Engel & W. Singer (Eds.). Cambridge, Massachusetts: MIT Press.

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural networks, 19*(8), 1153-1160.

de Wit, S., Standing, H. R., DeVito, E. E., Robinson, O. J., Ridderinkhof, K. R., Robbins, T. W., & Sahakian, B. J. (2012). Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. *Psychopharmacology, 219*(2), 621-631.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology, 18*(1), 105.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin, 125*(6), 627.

Declerck, C. H., Boone, C., & Emonds, G. (2013). When do people cooperate? The neuroeconomics of prosocial decision making. *Brain and cognition, 81*(1), 95-117.

Delgado, M. R. (2007). Reward-Related Responses in the Human Striatum. *Annals of the New York Academy of Sciences, 1104*(1), 70-88.

Delgado, M. R., Gillis, M. M., & Phelps, E. A. (2008). Regulating the expectation of reward via cognitive strategies. *Nature neuroscience, 11*(8), 880-881.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of neurophysiology, 84*(6), 3072-3077.

Diekhof, E. K., Kaps, L., Falkai, P., & Gruber, O. (2012). The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude–An activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia, 50*(7), 1252-1266.

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human brain mapping, 30*(9), 2907-2926.

Eisenegger, C., Haushofer, J., & Fehr, E. (2011a). No sound evidence for a gender-specific effect of testosterone administration on aggressive motivation exists: reply to Josephs et al. *Trends in Cognitive Sciences, 15*(11), 510-511.

Eisenegger, C., Haushofer, J., & Fehr, E. (2011b). The role of testosterone in social interaction. *Trends in Cognitive Sciences, 15*(6), 263-271.

Falk, A., Fehr, E., & Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry, 41*(1), 20-26.

Falk, A., Gächter, S., & Kovács, J. (1999). Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. *Journal of Economic Psychology, 20*(3), 251-284.

Fareri, D. S., Chang, L. J., & Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in neuroscience, 6*.

Fareri, D. S., & Delgado, M. R. (2014a). Differential reward responses during competition against in-and out-of-network others. *Social cognitive and affective neuroscience, 9*(4), 412-420.

Fareri, D. S., & Delgado, M. R. (2014b). The importance of social rewards and social networks in the human brain. *The Neuroscientist*, 1-16.

Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., & Delgado, M. R. (2012). Social network modulation of reward-related signals. *The Journal of neuroscience, 32*(26), 9045-9052.

Faro, S. H., & Mohamed, F. B. (Eds.). (2010). *BOLD fMRI - A Guide to Functional Imaging for Neuroscientists*. New York: Springer.

Fehr, E., & Camerer, C. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences, 11*(10), 419-427.

Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review, 46*(4–5), 687-724.

Fehr, E., & Fischbacher, U. (2002). Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives. *The economic journal, 112*(478), C1-C33.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature, 13*(1), 1-25.

Fehr, E., & Gächter, S. (1998). Reciprocity and economics: The economic implications of Homo Reciprocans1. *European Economic Review, 42*(3–5), 845-859.

Fehr, E., & Gächter, S. (1999). Cooperation and punishment in public goods experiments. *Institute for Empirical Research in Economics working paper*(10).

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives*, 159-181.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*(3), 817-868.

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology, 15*(17), R644-R645.

Frith, C., & Frith, U. (2007). Social cognition in humans. *Current Biology, 17*(16), R724-R732.

Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review, 11*(4), 303-327.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of cognitive neuroscience, 12*(3), 495-504.

Gerdeman, G. L., Partridge, J. G., Lupica, C. R., & Lovinger, D. M. (2003). It could be habit forming: drugs of abuse and striatal synaptic plasticity. *Trends in neurosciences, 26*(4), 184-192.

Gilbert, A. M., & Fiez, J. A. (2004). Integrating rewards and cognition in the frontal cortex. *Cognitive, Affective, & Behavioral Neuroscience, 4*(4), 540-552.

Glimcher, P. W. (2004). *Decisions, uncertainty, and the brain: The science of neuroeconomics*. New York: MIT press.

Glimcher, P. W., Camerer, C. F., Fehr, E., & Poldrack, R. A. (2009). *Introduction: a brief history of neuroeconomics.* In: Glimcher, P.W., Camerer, C.F., Fehr, E., and Poldrack, R.A. (eds.) *Neuroeconomics: Decision Making and the Brain.* New York: Academic Press, 1-12.

Glimcher, P. W., & Fehr, E. (2013). *Neuroeconomics: Decision making and the brain*. London: Academic Press.

Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science, 306*(5695), 447-452.

Grabenhorst, F., & Rolls, E. T. (2009). Different representations of relative and absolute subjective value in the human brain. *NeuroImage, 48*(1), 258-268.

Graybiel, A. M. (2000). The basal ganglia. *Current Biology, 10*(14), R509-R511.

Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of chemical neuroanatomy, 26*(4), 317-330.

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology, 35*(1), 4-26.

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., & Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *The Journal of Neuroscience, 30*(2), 583-590.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science, 324*(5927), 646-648.

Hare, T. A., Hakimi, S., & Rangel, A. (2014). Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. *Frontiers in neuroscience, 8*.

Hare, T. A., Malmaud, J., & Rangel, A. (2011). Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. *The Journal of Neuroscience, 31*(30), 11077-11087.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience, 28*(22), 5623-5630.

Haushofer, J., & Fehr, E. (2008). You Shouldn't Have: Your Brain on Others' Crimes. *Neuron, 60*(5), 738-740.

Heim, S., Eickhoff, S. B., Ischebeck, A. K., Friederici, A. D., Stephan, K. E., & Amunts, K. (2009). Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM. *Human brain mapping, 30*(2), 392-402.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 73-78.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1). Sunderland: Sinauer Associates.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron, 58*(2), 284-294.

Izuma, K., Saito, D. N., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience, 22*(4), 621-631.

Jones, B., & Rachlin, H. (2006). Social discounting. *Psychological Science, 17*(4), 283-286.

Jones, B. A., & Rachlin, H. (2009). Delay, probability, and social discounting in a public goods game. *Journal of the Experimental Analysis of Behavior, 91*(1), 61-73.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature neuroscience, 10*(12), 1625-1633.

Kalenscher, T., & Strombach, T. (2013). Die Macht der Versuchung [The power of temptation]. *Gehirn&Geist, 11*, 62-67.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*(1476), 2109-2128.

Kenning, P., & Plassmann, H. (2005). NeuroEconomics: An overview from an economic perspective. *Brain Research Bulletin, 67*(5), 343-354.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *science, 314*(5800), 829-832.

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current opinion in neurology, 18*(4), 411-417.

Kobayashi, S., Lauwereyns, J., Koizumi, M., Sakagami, M., & Hikosaka, O. (2002). Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *Journal of Neurophysiology, 87*(3), 1488-1498.

Kouneiher, F., Charron, S., & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature neuroscience, 12*(7), 939-945.

Kuss, K., Falk, A., Trautner, P., Elger, C. E., Weber, B., & Fliessbach, K. (2013). A reward prediction error for charitable donations reveals outcome orientation of donators. *Social cognitive and affective neuroscience, 8*(2), 216-223.

Leon, M. I., & Shadlen, M. N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron, 24*(2), 415-425.

Levy, D. J., & Glimcher, P. W. (2011). Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *The Journal of neuroscience, 31*(41), 14693-14707.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology, 22*(6), 1027-1038.

Liljeholm, M., & O'Doherty, J. P. (2012). Contributions of the striatum to learning, motivation, and performance: an associative account. *Trends in cognitive sciences, 16*(9), 467-475.

Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *The Journal of Neuroscience, 31*(37), 13214-13223.

Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social cognitive and affective neuroscience, 7*(3), 274-281.

Loewenstein, G., & O'Donoghue, T. (2007). The heat of the moment: Modeling interactions between affect and deliberation. *Unpublished manuscript*.

Marreiros, A. C., Stephan, K. E., & Friston, K. J. (2010). Dynamic causal modeling. *Scholarpedia, 5*(7), 9568.

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral cortex, 22*(8), 1894-1903.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503-507.

Miller, E. M., Shankar, M. U., Knutson, B., & McClure, S. M. (2014). Dissociating motivation from reward in human striatal activity. *Journal of cognitive neuroscience, 26*(5), 1075-1084.

Mobbs, D., Hassabis, D., Seymour, B., Marchant, J. L., Weiskopf, N., Dolan, R. J., & Frith, C. D. (2009). Choking on the money reward-based performance decrements are associated with midbrain activity. *Psychological Science, 20*(8), 955-962.

Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research, 17*(2), 189-202.

Morgenstern, O., & Von Neumann, J. (1953). Theory of games and economic behavior.

Morishima, Y., Schunk, D., Bruhin, A., Ruff, Christian C., & Fehr, E. (2012). Linking Brain Structure and Activation in Temporoparietal Junction to Explain the Neurobiology of Human Altruism. *Neuron, 75*(1), 73-79.

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences, 107*(49), 20911-20916.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. J., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452-454.

O'Reilly, R. C., Frank, M. J., Hazy, T. E., & Watz, B. (2007). PVLV: the primary value and learned value Pavlovian learning algorithm. *Behavioral neuroscience, 121*(1), 31.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology, 14*(6), 769-776.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*(7090), 223-226.

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature neuroscience, 5*(2), 97-98.

Park, S. Q., Kahnt, T., Rieskamp, J., & Heekeren, H. R. (2011). Neurobiology of value integration: when value impacts valuation. *The Journal of neuroscience, 31*(25), 9307-9314.

Pavlov, I. P. (1941). Lectures on conditioned reflexes. Vol. II. Conditioned reflexes and psychiatry.

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science, 316*(5826), 904-906.

Rachlin, H., & Jones, B. A. (2008). Social discounting and delay discounting. *Journal of Behavioral Decision Making, 21*(1), 29-43.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545-556.

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current opinion in neurobiology, 20*(2), 262-270.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual review of psychology, 62*, 23-48.

Robbins, T. W., & Everitt, B. J. (1996). Neurobehavioural mechanisms of reward and motivation. *Current opinion in neurobiology, 6*(2), 228-236.

Robert, M. Y., & John, D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology, 18*, 459-482.

Roch, S. G., Lane, J. A. S., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive Load and the Equality Heuristic: A Two-Stage Model of Resource Overconsumption in Small Groups. *Organizational Behavior and Human Decision Processes, 83*(2), 185-212.

Roesch, M. R., & Olson, C. R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *Journal of Neurophysiology, 90*(3), 1766-1789.

Rudebeck, P., Bannerman, D., & Rushworth, M. (2008). The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cognitive, Affective, & Behavioral Neuroscience, 8*(4), 485-497.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience, 15*(8), 549-562.

Rushworth, M. F., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Current opinion in neurobiology, 22*(6), 946-955.

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies, 4*(2), 155-161.

Sanfey, A. G., Loewenstein, G., McClure, S. M., & Cohen, J. D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in cognitive sciences, 10*(3), 108-116.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755-1758.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind". *Neuroimage, 19*(4), 1835-1842.

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS biology, 10*(2), e1001266.

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron, 45*(4), 625-635.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology, 80*(1), 1-27.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol., 57*, 87-115.

Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *The Journal of Neuroscience, 12*(12), 4595-4610.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Schulz, J. F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology, 41*(0), 77-87.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9-34.

Silani, G., Lamm, C., Ruff, C. C., & Singer, T. (2013). Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. *The Journal of Neuroscience, 33*(39), 15466-15476.

Smith, A. (1759). The theory of moral sentiments, ed. *A. Millar, A. Kincaid & J. Bell. AM Kelley.(Originally published in 1759)*.

Smith, D. V., Clithero, J. A., Boltuck, S. E., & Huettel, S. A. (2014). Functional connectivity with ventromedial prefrontal cortex reflects subjective value for social rewards. *Social cognitive and affective neuroscience, 9*(12), 2017-2025.

Soutschek, A., Stelzel, C., Paschke, L., Walter, H., & Schubert, T. (2014). Dissociable Effects of Motivation and Expectancy on Conflict Processing: An fMRI Study. *Journal of Neuroscience, 27(2),* 409-423.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of cognitive neuroscience, 10*(5), 640-656.

Strang, S., Utikal, V., Fischbacher, U., Weber, B., & Falk, A. (2014). Neural Correlates of Receiving an Apology and Active Forgiveness: An fMRI Study. *PloS one, 9*(2), e87654.

Strombach, T., Jin, J., Weber, B., Kenning, P., Shen, Q., Ma, Q., & Kalenscher, T. (2014). Charity begins at home: cultural differences in social discounting and generosity. *Journal of Behavioral Decision Making, 27*(3), 235-245.

Tankersley, D., Stowe, C. J., & Huettel, S. A. (2007). Altruism is associated with an increased neural response to agency. *Nature neuroscience, 10*(2), 150-152.

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

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

van den Bos, W., Talwar, A., & McClure, S. M. (2013). Neural correlates of reinforcement learning and social preferences in competitive bidding. *The Journal of Neuroscience, 33*(5), 2137-2146.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human brain mapping, 30*(3), 829-858.

Van Winden, F., Stallen, M., & Ridderinkhof, K. R. (2008). *On the nature, modeling, and neural bases of social ties* (Vol. 20). In: Houser, D., & McCabe, K.(ed.) *Neuroeconomics: Advances in Health Economics and Health Services Research.* Emerald Group Publishing, 125-159.

Volkow, N., Fowler, J., Wang, G., Baler, R., & Telang, F. (2009). Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology, 56*, 3-8.

Williamson, L. L., Cheng, R.-K., Etchegaray, M., & Meck, W. H. (2008). "Speed" warps time: methamphetamine's interactive roles in drug abuse, habit formation, and the biological clocks of circadian and interval timing. *Current drug abuse reviews, 1*(2), 203-212.

Winecoff, A., Clithero, J. A., Carter, R. M., Bergman, S. R., Wang, L., & Huettel, S. A. (2013). Ventromedial prefrontal cortex encodes emotional value. *The Journal of Neuroscience, 33*(27), 11032-11039.

Young, L., Dodell-Feder, D., & Saxe, R. (2010). What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia, 48*(9), 2658-2664.

APPENDIX

1. Study 1: Neural Underpinnings of Performance-Based Rewards

Strombach, T.[a,b*], Hubert, M.[c], and Kenning, P.[b,c] (in press). Neural underpinnings of performance-based incentives. *Journal of Economic Psychology.*

[a] Comparative Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[b] Chair of Marketing, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[c] Department for Corporate Management and Economics, Zeppelin University, Friedrichshafen, Germany

[*] Corresponding author (tina.strombach@hhu.de)

Letter of acceptance

Betreff: **Your Submission JOEP-D-14-00205R2**
An: Tina.Strombach@hhu.de

Datum: 06.06.15 21:49
Von: Journal of Economic Psychology <joep@elsevier.com>
Absender: ees.joep.0.31c0a7.654d70dc@eesmail.elsevier.com

Ms. Ref. No.: JOEP-D-14-00205R2
Title: The neural underpinnings of performance-based incentives
Research Paper, Journal of Economic Psychology

Dear Ms. Strombach,

I am pleased to confirm that your paper "The neural underpinnings of performance-based incentives" has been accepted for publication in Journal of Economic Psychology.

You will shortly receive an acknowledgement letter from the Editorial-Production Journals Department with further administrative instructions.

When your paper is published on ScienceDirect, you want to make sure it gets the attention it deserves. To help you get your message across, Elsevier has developed a new, free service called AudioSlides: brief, webcast-style presentations that are shown (publicly available) next to your published article. This format gives you the opportunity to explain your research in your own words and attract interest. You will receive an invitation email to create an AudioSlides presentation shortly. For more information and examples, please visit http://www.elsevier.com/audioslides

On our authors' website you can find more information on how to improve visibility of your published article: http://www.elsevier.com/journal-authors/getting-your-paper-noticed.

Thank you for submitting your work to this journal.

With kind regards,

Carlos Alos-Ferrer, Dr.
Associate Editor
Journal of Economic Psychology

Abstract

Pay-for-performance is commonly applied in order to favorably modulate behavior and increase performance. However, the removal of an incentive leads to a significant decrease in performance. Although there is behavioral evidence that incentives have an effect on performance, the neural underpinnings of the underlying cognitive processes are not certain. We hypothesize that performance is affected by monetary incentives, and that these changes are reflected by the dopaminergic reward system. In this study, we use the direct measure of performance and measure the underlying cognitive processes by using functional magnetic resonance imaging (fMRI) simultaneously. We find that incentives induce changes in reward-related brain regions, but not in task-related neural representations. Interestingly, when monetary incentives are introduced, blood oxygenation level dependent activity increases in the ventral striatum, being sensitive to reward, and decreases in the ventromedial prefrontal cortex, representing the subjective value. We suggest that pay-for-performance does not directly affect performance by modulating neural activity in task-relevant regions, but affects the reward representation during task completion. Therefore, we conclude that performance-contingent incentives need to be applied carefully. This finding is of special interest for employee compensation in unstable economic environments, as companies in unstable environments might not permanently ensure pay-for-performance.

Highlights

- The introduction of pay-for-performance increases activation in the ventral striatum.
- The introduction of pay-for-performance decreases activation in the ventromedial prefrontal cortex.
- No support that monetary incentives affect task-relevant brain regions.

1. Introduction

That behavior is often guided by incentives in the form of rewards is a widely recognized condition (Locke, 1968; Schultz, 2006). Specifically, rewards are understood to initiate and modulate behavior (Schultz, 2006), and substantial research evidence demonstrates that such incentives affect performance (Bloom, 1999; Goldman, 2005; Jenkins Jr, Mitra, Gupta, & Shaw, 1998; Locke, Feren, McCaleb, Shaw, & Denny, 1980; Rynes, Gerhart, & Parks, 2005). However, opinions about the impact of rewards on performance remain far from unanimous, as the direction of the effect continues to be a topic of controversy both in psychology and economics (Albrecht, Abeler, Weber, & Falk, 2014; Camerer & Hogarth, 1999; Kamenica, 2012). When monetary incentives are introduced, intrinsic motivation to perform is overshadowed by the money, resulting in behavior that is predominantly elicited by financial incentives. Many theories indicate that this crowding out of motivation by financial incentives leads to a drop in performance (Deci, 1971; Deci, Koestner, & Ryan, 1999; Frey & Oberholzer-Gee, 1997; Murayama, Matsumoto, Izuma, & Matsumoto, 2010; Ryan & Deci, 2000). Nevertheless, an equally substantial amount of research demonstrates that monetary incentives can have a positive impact—possibly by increasing the effort that is invested in solving a task (Bloom, 1999; Jenkins Jr et al., 1998; Locke et al., 1980). Although the effect of incentives has not yet been definitely established, performance-based incentive systems (e.g., performance-based compensation plans, cash gifts, Christmas bonuses, annual bonuses, stock options, etc.) are widespread in today's corporate world (Camerer & Hogarth, 1999; Le Blanc & Mulvey, 1998; Merchant & Van der Stede, 2007; Rosenthal, Frank, Li, & Epstein, 2005).

Because the behavioral findings regarding performance-based incentives have not been definitively established, the ways in which performance-based monetary incentives change the neural representation of a specific task need to be investigated. A better understanding of the effects may reveal how such measures impact performance at a neural level. Neuroscience offers a method for investigating the latent but significant effects of monetary incentives on performance. Such insights may be very beneficial in a variety of corporate and institutional settings where performance measures are implemented on a regular basis. Because we are interested in the sequential effect of incentives, we observed reactions to the introduction of a monetary incentive, as well as subsequent responses to the removal of the incentive, noting, particularly, how these changes are reflected at the neural level. Subjects were presented with a cognitively engaging, but solvable, arithmetic task while measuring

blood oxygenation level dependent (BOLD) signals, which are an indirect measure of relative neural activation, and which allow inferences about the regions activated during a specific cognitive task (Huettel, Song, & McCarthy, 2004). Comparing BOLD signals that are recorded before performance-based rewards are introduced (baseline), during the administration of a pay-for-performance and subsequent to the removal of the incentives could be expected to reveal brain regions associated with processing such changes in incentives. As the task will remain the same in all conditions (i.e., before the incentive has been offered, when incentives are introduced and after the incentive has been removed), this experimental design is well suited to examining both the extent to which monetary incentives affect performance on a behavioral level, and how this is reflected in neural activation.

2. Theory and hypothesis development

2.1 Incentives and performance

Since the early 1960s, researchers have been interested in investigating the relationship between payment and performance (Lawler, 1971; Locke, 1968), so far with inconclusive results. While many studies have shown that money has a positive effect on motivation and performance (Bloom, 1999; Jenkins Jr et al., 1998; Lazear, 2000; Locke et al., 1980), others have demonstrated that performance-based payments are disadvantageous as motivational factors (Frey & Oberholzer-Gee, 1997; Locke et al., 1980). Though most researchers generally agree that there is a connection between monetary incentives and performance (Camerer & Hogarth, 1999; Kamenica, 2012; Locke et al., 1980), research in psychology and economics presents divergent opinions about such a relationship. While the greater number of researchers in psychology have concluded that monetary incentives undermine performance (Deci, 1971; Deci et al., 1999), economic researchers generally endorse the opposite view. The research on the psychology of motivation (Deci & Ryan, 1985, 2000, 2012; Ryan & Deci, 2000) highlights the importance of the autonomous self and the freedom to make personal decisions as a prerequisite for a high motivation to perform. They suggest that rewards that are contingent upon engagement, competition, or performance undermine intrinsic motivation and, consequently, lead to reduced performance in a given task. The introduction of incentives might also be detrimental in the same way that added incentives make people self-conscious about an automatic activity (Baumeister, 1984; Baumeister & Showers, 1986; Camerer & Hogarth, 1999). A series of behavioral

observations and experiments indicated that the introduction of larger-than-average monetary incentives reduces performance in, for example, automatic actions or sporting events (Baumeister, 1984; Mobbs et al., 2009). These experiments suggest that the increase in incentives may lead to arousal beyond an optimal level, and may result in explicit monitoring that can interfere with the activity itself. In other words, preoccupation with a reward may lead to pressure-induced performance decrements (Baumeister, 1984; Baumeister & Showers, 1986; Schlenker, Phillips, Boniecki, & Schlenker, 1995). Thus, in psychological terms, intrinsic motivation is expected to be high enough to produce steady effort even in the absence of financial rewards, and incentives are likely to reduce performance (Camerer & Hogarth, 1999).

In contrast, the conventional view proposed by economists states that behavior is initiated only when a reward is available, thereby implying that monetary incentives can be used to initiate a desired behavior (Camerer & Hogarth, 1999). When focusing on economic experiments to investigate the role of incentives, if subjects earn money for good performance, they work harder, more persistently, and more effectively (Camerer & Hogarth, 1999). A field study by Libby and Lipe (1992) also shows that the introduction of monetary incentive systems induces greater effort and results in better performance. Thus, economic and psychological theories make different predictions about the effect that monetary incentives have on performance. The general view of relationships between incentives, motivation, and performance is inconclusive. However, the initial level of intrinsic motivation to complete a task seems to be a core component to determine the effect of incentives on performance. Thus, the inconclusive results may be driven by the nature of the task or by the amount of intrinsic motivation and personal interest that the task elicits. According to a meta-analysis developed by Jenkins and colleagues (1998), the effect of money on motivation to perform is modulated by the level of personal interest a person has in the task. While performance in low-interest activities increases with payment (Locke et al., 1980; Rynes et al., 2005), the motivation to perform highly interesting activities seems to diminish after receiving monetary incentives (Frey & Oberholzer-Gee, 1997).

In the experiment conducted for our study, we examine the effects of introducing, and of removing, extrinsic monetary incentives for performance on a mental arithmetic task. We hypothesize that the introduction of monetary incentives will increase performance, as our task cannot be considered an interesting activity, so that incentives can be expected to increase the effort invested in solving the tasks rather than diminishing intrinsic motivation and performance (Camerer & Hogarth, 1999). Additionally, taking an incentive away has been widely shown to decrease

levels of motivation—diminished motivation that may be permanent, and that is independent of the personal level of interest in the task (Gneezy, Meier, & Rey-Biel, 2011; Rynes et al., 2005). These results cause us to expect the removal of a monetary incentive to decrease performance in our experiment as well. For the purposes of this paper, we argue that performance is a direct measure, while motivation to perform is the underlying construct that drives outcome-oriented behavior. We do not evaluate whether participants are intrinsically motivated; rather, we focus on the effect of incentives on performance. The motivation to perform is defined as being engaged in solving a given problem. Thus, in this context, motivation is tied to the concept of motivation to perform, independent of the intention behind it.

2.2 Neural correlates of changes in performance-based rewards

When task-related and incentive-related behaviors are combined, three processes are expected to modulate the underlying neural correlates: 1) reward-related activation as a response to the monetary incentives, 2) task-related activity to complete the task, and 3) effort-related changes in BOLD signals that reflect an increase in cognitive control intended to result in task success.

First, we expect activation in reward-related areas such as the corticobasal ganglia valuation system, including the ventral striatum (vStr), to reflect the introduction and removal of monetary rewards (Bartra, McGuire, & Kable, 2013; Chib, De Martino, Shimojo, & O'Doherty, 2012; Haber & Knutson, 2010; Murayama et al., 2010; O'Doherty et al., 2004; Robbins & Everitt, 1996). We come to this conclusion, in part, on the basis of recent studies in this area that have used functional magnetic resonance imaging (fMRI) to investigate the effect of monetary incentives on performance and motivation. Murayama and colleagues (2010) showed that performance-based rewards modulate activation in the vStr. They state that this activation reflects the integration of extrinsic reward value and intrinsic task value. However, their design was limited to a comparison between incentivized participants and those who were not incentivized which, because of variability in performance-based incentives, did not permit testing of within-subject changes. The authors suggest that the introduction of monetary incentives also modulates the effort that is invested in solving a given task—a response that can be observed by BOLD signal changes in the vStr. Although other studies focused more on money's detrimental effect on intrinsic motivation and performance (Albrecht et al., 2014; Chib et al., 2012; Mobbs et

al., 2009), a common feature in all studies is the important role of the striatum, notably as it is associated with the introduction and removal of monetary rewards. In the present experiment we anticipate increased BOLD responses in the vStr during performance-based monetary rewards. Should this finding be observed, it would support the idea that the vStr is indeed involved in processing extrinsic motivation and reward, independent of the level of intrinsic motivation or personal interest.

Second, we hypothesize that incentives drive changes in BOLD signals in task-related regions. Throughout the experiment, as subjects are asked to solve arithmetic calculations, we expect to find activation in regions known to be associated with arithmetic calculations, such as the intraparietal sulci (IPL), inferior prefrontal, and prefrontal cortices that have been found to be involved independent of the type of calculation (i.e., addition, subtraction, multiplication, division) (Arsalidou & Taylor, 2011; Dehaene, Molko, Cohen, & Wilson, 2004; Rivera, Reiss, Eckert, & Menon, 2005). Monetary incentives might also modulate BOLD activity in these regions.

Third, in addition to changes in the reward- and task-related regions, monetary incentives may induce changes in effort and in attention to the task that are reflected in variable intensity of BOLD signals. These neural changes are expected to be paralleled by changes in performance in the experiment. Previous research shows that increasing the reward received from a desired outcome can facilitate cognitive processes that are required for the achievement of that outcome (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Delgado, Gillis, & Phelps, 2008; Engelmann, Damaraju, Padmala, & Pessoa, 2009; Jimura, Locke, & Braver, 2010; Kouneiher, Charron, & Koechlin, 2009; Krawczyk, Gazzaley, & D'Esposito, 2007; Taylor et al., 2004). Thus, we assume that when performance-based rewards are introduced, activation in task-related areas such as the IPL will increase, reflecting the increased effort and cognitive engagement. When the reward is removed, activation in those regions is expected to decrease. The increase in effort may be accompanied by increased cognitive control and by behavior that is more consciously controlled. The dorsolateral prefrontal cortex (dlPFC) has reliably been found to be involved in self-control and executive processes (Crockett et al., 2013; Hare, Camerer, & Rangel, 2009; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006). An increase in dlPFC activation is expected after introducing monetary rewards, reflecting increased self-control and cognitive/executive control.

3. Material and Methods

3.1 Participants

Thirteen female and seven male, healthy, right-handed subjects participated in the fMRI study (age: $M = 42.10$, $SD = 4.83$, range = 34 to 50 years; no gender differences in age, $t = -0.31$, p = .758). All participants were native German speakers and gave written consent prior to the scanning sessions. Participants were informed that the examination could potentially reveal medically significant findings, and they were asked whether they would like to be notified in such a case. None of the participants was excluded from further analysis. Participants received a 30€ show-up fee at the end of the experiment. Additionally, participants received different amounts of money, based on the number of correct answers in the performance-based rewards condition. The amounts varied between 16€ and 48€ (benefit: $M = 37$€, $SD = 7.83$€). An institutional review board[1] approved the study.

3.2 Experimental design

The fMRI experiment consisted of four blocks in which an LCD beamer projected arithmetic calculations onto a transparent screen that was viewed by participants via a 45° mirror mounted on an element phased array coil. In order to prevent external confounding visual stimulation, the questions were equal in length, size, position, background, and luminance. Each question was visible for at least 10 seconds. Questions used in the experiment included simple addition, subtraction, multiplication and division, with numbers not higher than 221. In each question, subjects were shown a solution and were asked to indicate whether or not it was correct. Answers were given by pressing a corresponding button on a magnetic resonance compatible response box. In total, subjects had to solve 65 calculations, 20 within each main block (Blocks 1 – 3) and five in a control condition (Block 4). Short breaks between each block were applied. To ensure a comparable level of difficulty across blocks, the types of calculation (addition, subtraction, multiplication and division) and number of digits

---

[1] The study was approved by an external institution—the Freiburg Ethics Commission

were balanced across the three 20-trial blocks[2]. In the first block, after participants pressed a button to answer the question, the calculation disappeared and they saw a fixation cross for the rest of the trial within 10 seconds. If no answer was given, the trial ended automatically after 10 seconds. Questions were presented in succession until the end of the block. In the second block, subjects were again asked to evaluate the correctness of a number of calculations and to press a button to indicate their decisions. The difference between the first and second block was that, while participants received neither feedback on their performance nor a performance-based reward in block one, after the calculations in block two they received a performance-based reward of 2€ for each correct answer, comparable to the study by Murayama et al. (2010). Thus, an external, quantifiable incentive was introduced. Block three was identical to block one, allowing within-subject comparisons to evaluate the effect of the removal of monetary incentives. In a fourth block, participants were presented with five more calculations and were again paid 2€ for each correct answer. Though Block 4 was not included in the fMRI analysis, it was designed to control for potential fatigue effects and to allow participants to leave the setting with a positive feeling about the experience. (See Figure 1 for the experimental setup.)

---

[2] We have to mention that the necessary mix of different operations for the present research – due to possible differences of mental strategies – might have lowered the signal-to-noise ratio (see for example Fehr, 2013)
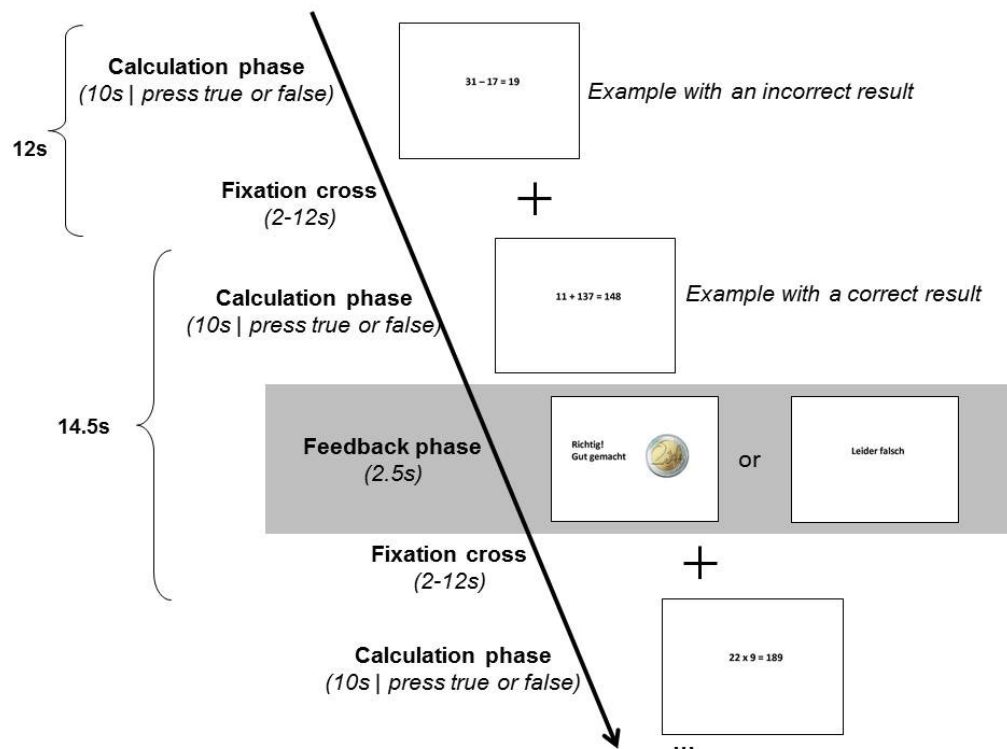
Figure 1: Experimental setup for Block 1 – Block 3 with feedback phase only for block 2.

We recorded the responses using Cogent (www.vislab.ucl.ac.uk/cogent.php). Average performance (individual fraction of correct answers ($indSC_{block}$ = *number of correct answers divided by the total number of calculations*) was calculated for each of the three blocks. Consequently, values ranged between 0 [all answers were incorrect] and 1 [all answers were correct]. After the scanning session, participants were asked to complete a questionnaire that included demographic data (e.g., age, gender, net income, work status). Additionally, participants completed a self-report questionnaire that assessed their motivation with respect to monetary incentives ($Q_{motivation}$), their level of pleasure, over time, for performing the calculations (Q_{fun}), and any demotivation that resulted from the removal of monetary incentives ($Q_{demotivation}$). A 7-point scale ranging from 1 (totally disagree) to 7 (totally agree) was used.

3.3 Imaging protocol and fMRI analysis

The study was executed on a 3T fMRI scanner (MAGNETOM Trio, Siemens, Erlangen, Germany). The dataset used consisted of 36 transversal slices of 3.6 mm thickness

without a gap, a field of view of 230 mm x 230 mm, and an acquired 64 x 64 matrix (i.e., isotropic voxels with 3.6 mm edge length). Contrast parameters were a signal response time of 3000 ms, echo time of 50 ms, and flip angle of 90°. Data analysis was conducted with SPM8-freeware (Friston, 1996; Friston et al., 1994) using MATLAB as a working basevand following procedures described in Poldrack et al. (2007). The data preprocessing consisted of three initial steps. First, to correct for artifacts due to head movement in the scanner, all images (Block 1 – Block 3) were realigned and resliced by a "rigid body" transformation to the mean image of all sessions (realignment). Second, to compare all participants within the group analysis, all images were normalized and re-sampled to the standard Montreal Neurological Institute (MNI) template (normalization). Third, to prepare the data for statistical analysis, all images were smoothed with an 8-mm Gaussian kernel (smoothing) (Ashburner, Neelin, Collins, Evans, & Friston, 1997) followed by a file set split, in order to re-assign the images to their related sessions (blocks).

For the main data analysis, we estimated a general linear model (GLM with Block 1, Block 2, and Block 3 separately) using the robust Weighted Least Squares Estimation (rWLS). The GLM consisted of the following independent variables: (R1) indicator variable for Block 1; (R2 – R7) movement regressors for Block 1; (R8) indicator variable for Block 2; (R9) indicator variable for the feedback phase; (R10) parametric modulator for feedback phase with regard to correct and incorrect answers; (R11 – R16) movement regressors for Block 2; (R17) indicator variable for Block 3; (R18 – R26) movement regressors for Block 3 and session constants for Blocks 1 – 3. The regressors capturing each block (R1, R8, R17) were modeled using a boxcar function with the individual response time as duration. The regressor capturing the feedback phase (R9) was modeled using a boxcar function with a fixed duration of 2.5 seconds. Each of the regressors was convolved with a canonical hemodynamic response function (HRF). We calculated the following first-level single-subject contrasts of interest: Block 1 versus Block 2; Block 1 versus Block 3; and Block 2 versus Block 3. On the second level, to extract differences of single-subject contrasts between subjects, we computed a one-sample t-test. We generated statistical parametric maps for the given contrast that displayed the *t*-value of each voxel meeting a $p < .0001/p < .001$ (uncorrected) significance level with an extent threshold voxel of $k = 10$, if not indicated elsewhere. Results are reported using the MNI coordinate system.

4. Results

4.1 Behavioral results

Behavioral data were analyzed to study the effects of incentives on performance. As indicated previously, we used the fraction of correct answers ($indSC_{block}$) as a direct measure of performance. To check for behavioral differences in the number of correct answers between the three blocks, we analyzed individual fractions of correct answers ($indSC_{block}$: M_{block1} = .74, SD =.16, range = .35 to .95; M_{block2} = .72, SD =.15, range = .45 to 1; M_{block3} = .42, SD =.07, range = .35 to .55) using a one-way repeated measures ANOVA (with block numbers 1, 2, and 3 being the levels of the within-subject factor) corrected for multiple comparisons by the Greenhouse-Geisser (GG) correction criterion. Analysis revealed a significant main effect of block number ($F(1.378, 26.174) = 50.21$, p < .001, η_p^2 = .703). Post-hoc tests revealed a significant drop in performance from Block 2 to Block 3 when the monetary incentive was removed, but no increase in performance from Block 1 to Block 2. Therefore, thus no increase in performance is observed after the introduction of the incentive (Block 2) (see Figure 2). A bivariate analysis of performance levels between Block 1 and Block 3 (r = -.27, p = .25) and between Block 2 and Block 3 (r = -.29, p = .212) revealed no indication of a bias regarding participants' individual performance level.
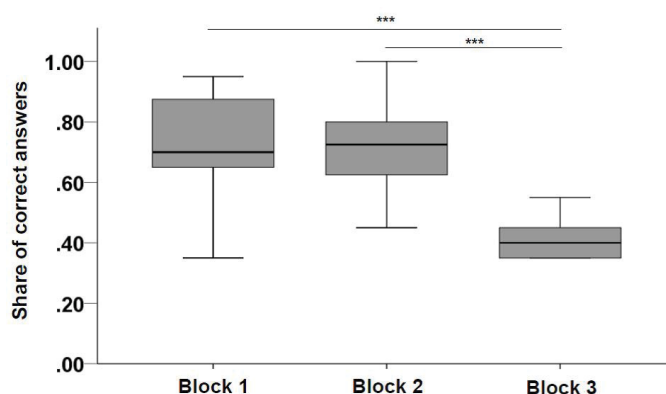
Figure 2: Boxplot with mean and standard deviation for block1-3 based on $indSC_{block}$ (***=p <.001).

After analyzing the post-scanning questionnaire, we found significant correlations (one-tailed) between an increase in motivation because of monetary incentives ($Q_{motivation}$) and a decrease in the pleasure of calculation over time (Q_{fun}; r = .43, p = .030), as well as between an increase in motivation because of monetary incentives ($Q_{motivation}$)

and a demotivation after the removal of monetary incentives ($Q_{demotivation}$; $r = .48$, $p = .015$).

4.2 fMRI results

To investigate the effects of the introduction and removal of monetary incentives on neural activation, we first analyzed the contrasts using a threshold of $p < .001$ and $k \geq 10$ voxel. In a second step, we used masks of chosen regions as a region of interest based on a coordinate-based meta-analysis of BOLD fMRI experiments that examine neural correlates of subjective value (Bartra et al., 2013).

When comparing BOLD activity before (Block 1) and during (Block 2) performance-based incentives, we found higher BOLD activation during the incentive-free phase (Block 1) within the right orbitofrontal cortex (OFC; BA11, x=2, y=36, z=-18), the left dorsal anterior cingulate cortex (dACC; BA32, x=-6, y=36, z=-8), and the right ventromedial prefrontal cortex (vmPFC; BA10, x=2, y=36, z=-18). These regions are located in the medial prefrontal cortex and have previously been associated with the brain's valuation network (Bartra et al., 2013). Using a mask of the vmPFC based on Bartra and colleagues (2013), we found that this region, which is known to be involved in modality-independent subjective value representation, shows a significant decrease in BOLD activity as a response to the introduction of the monetary reward (x=2, y=38, z=-18; p=0.0353, t=4.1815, small-volume (SV) familywise error (FWE) corrected) (see Figure 3).

Using the same contrast to identify brain regions associated with processing of the monetary performance-based reward, we compared neural activity, as measured by the BOLD activity, during incentivized (Block 2) and non-incentivized (Block 1) trials. Of the notable observations for Block 2, we found significantly higher BOLD signals, as compared to Block 1, within the left parahippocampal gyrus (BA36, x=-40, y=-30, z=-16), the caudate head (especially the left vStr) (vStr; x=-4, y=12, z=-6), the left insula (BA13, x=-28, y=18, z=-2), and the bilateral dorsolateral prefrontal cortex (dlPFC; BA46, x=42, y=16, z=18; BA9, x=-42, y=4, z=42). Using a mask of the vStr from the meta-analysis (Bartra et al., 2013), we found higher activity in the vStr, a region that is reliably found to be involved in reward processing (x=-4, y=12, z=-6; t=6.400, p=.0113, SV FWE corrected) during the introduction of the monetary incentive, as compared to non-incentivized trials in Block 1 (see Figure 4).
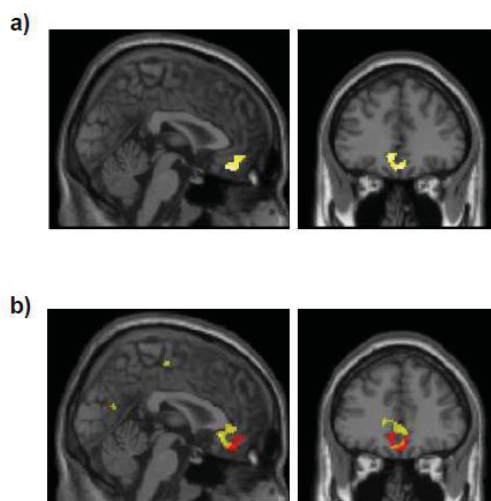
Figure 3: Visualization of activity within a) the vmPFC for block 1 (x = 2, y = 36, z = ·18; contrast of block 1 versus block 2) and b) a conjunction between the contrasts of block 1 versus block 2 (red clusters) and block 1 versus block 3 (yellow clusters), with an overlapping activation pattern of a cluster (orange) within OFC, vmPFC and ACC.
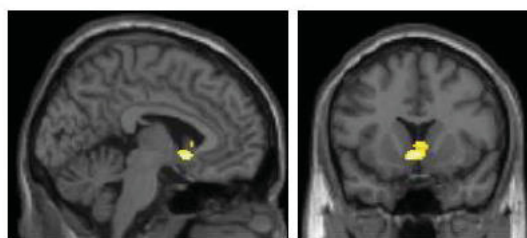
Figure 4: Visualization[3] of activity within the ventral striatum for block 2 (x = · 4, y = 12, z = ·6; contrast of block 2 versus block 1).

Contrasting BOLD activity during (Block 2) and after (Block 3) performance-based incentives, we found significantly higher BOLD signals in Block 2 compared to Block 3 within the right parahippocampal gyrus (x=38, y=-2, z=-24), the bilateral superior temporal gyrus (BA38, x=42, y=20, z=-6; x=-46, y=16, z=-20), the caudate head (especially the left vStr) (vStr; x=-4, y=10, z=-4), the left insula (BA13, x=-32, y=6, z=18), and the bilateral dorsolateral prefrontal cortex (dlPFC; BA9, x=38, y=20, z=36; x=-34, y=12, z=42) (see Figure 5). For Block 3, we found no significant increases in activity compared to Block 2 in any region.

---

[3] p-level of .005 [unc.] only for visualization
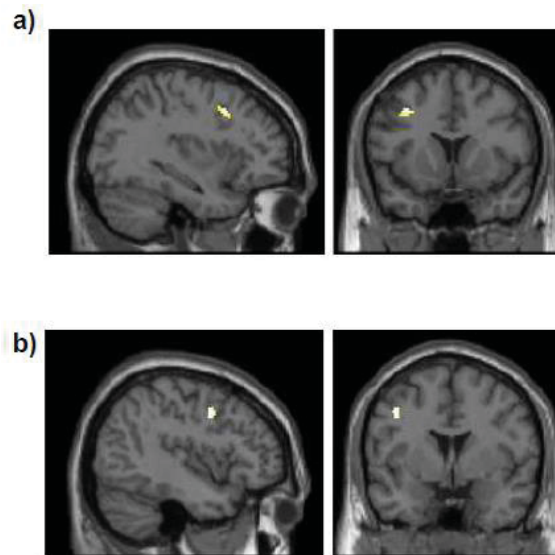
Figure 5: Visualization of activity within a) the dlPFC for block 2 (x = -42, y = 2, z = 42; contrast of block 2 versus block 1) and b) the dlPFC for block 2 (x = -34, y = 12, z = 42; contrast of block 2 versus block 3).

For the contrast of BOLD activity before (Block 1) and after (Block 3) performance-based incentives, we found significantly higher BOLD activity in Block 1 compared to Block 3 in BOLD signals within the left amygdala (x=-28, y=-2, z=-28), the left dorsal anterior cingulate cortex (dACC; BA32, x=-2, y=18, z=-8), and the left insula (BA13, x=-42, y=2, z=-8). Additionally, for an ROI analysis using the mask for the vmPFC (Bartra et al., 2013), we found significantly lower BOLD signals after removing the monetary incentives (x=-2, y=28, z=-8, t=5.0383, p=.0161) than before they were introduced in Block 1 (see Figure 3). However, we did not find any differences in the vStr when using that mask for ROI analyses. For Block 3, we found no significant increase in BOLD signals compared to Block 1. Thus, the removal of the monetary incentives decreased neural activity when contrasted with the baseline condition in Block 1, representing the changes that can be attributed to the changes in the monetary incentives and not the task itself (see Table 1 for a detailed report of the fMRI results).

Table 1: *Overview of activation patterns during the experiment*

Block	Region	Side	No. of voxels	BA	MNI Coordinates			t-value of peak voxel
Main clusters for the contrast of Block 1 versus Block 2								
Block 1	Orbitofrontal cortex	R	25	11	2	36	-18	4.39
	Anterior cingulate	L	9	32	-6	36	-8	4.14
Block 2	Parahippocampa gyrus	R	20		38	-24	-14	5.36
	Parahippocampa gyrus	L	34	36	-40	-30	-16	6.07
	Middle occipital gyrus	R	6	19	42	-74	-14	4.81
	Hippocampus	L	9	36	-28	-36	-8	6.55
	Hippocampus	R	29		30	-32	-6	6.36
	Thalamus	L	261		-4	-24	12	6.34
	Caudate head, ventral striatum	L	55		-4	12	-6	6.40
	Thalamus	L	26		-16	-32	0	5.39
	Caudate head	R	12		6	16	4	-5.25
	Inferior frontal gyrus	R	12	45	56	20	20	5.89
	Caudate body	R	18		20	-16	18	4.98
	Insula	L	99	13	-28	18	-2	4.41
	Dorsolateral prefrontal gyrus	R	19	46	42	16	18	4.84
	Dorsolateral prefrontal gyrus	L	25	9	-42	4	42	4.11
	Medial frontal gyrus	L	15	6	-14	-14	56	4.95
Main clusters for the contrast of Block 1 versus Block 3								
Block 1	Amygdala	L	12		-28	-2	-28	4.07
	Parahippocampa gyrus	L	45		-30	-4	-12	4.80
	Anterior cingulate	L	40	32	-2	28	-8	5.04
	Middle temporal gyrus	L	26		-58	-32	-6	4.28
	Insula	L	7	13	-42	2	8	4.10
	Cuneus	R	7	18	16	-86	18	3.96
Block 3	No significant activity changes							
Main clusters for the contrast of Block 2 versus Block 3								
Block 2	Parahippocampa gyrus	R	71		38	-2	-24	5.27
	Superior temporal gyrus	R	9	38	42	20	-26	5.16
	Superior temporal gyrus	L	21	38	-46	16	-20	5.94
	Parahippocampa gyrus	R	31		18	-26	-18	4.07
	Cuneus	L	137	18/17	-8	-94	8	4.51
	Caudate head, ventral striatum	L	11		-4	10	-4	4.13
	Thalamus	L	48		-18	-32	-2	4.43
	Posterior cingulate	R	11		10	-56	4	3.71
	Insula	L	6	13	-42	0	6	3.93
	Inferior frontal gyrus	R	26	45	56	20	20	4.33
	Insula	L	81	13	-32	6	18	5.36
	Cingulate gyrus	R	64	31/24	12	-24	40	4.57
	Cingulate gyrus	L	26	23	-2	-20	32	4.51
	Dorsolateral prefrontal gyrus	R	37	9	38	20	36	4.59
	Cingulate gyrus	R	122	24	2	-6	42	4.41
	Dorsolateral prefrontal gyrus	L	40	9	-34	12	42	5.68
Block 3	No significant activity changes							

Height threshold T = 4.5899, p < .0001 [uncorrected], k =5; T = 3.505, p < 0.001 [uncorrected], k =5

5. Discussion

The objective of this study was to explore how monetary incentives affect actual performance, and to determine how changes in incentives are reflected in the underlying neural activation. Our experimental design allowed a within-subject examination of this effect by using a cognitively engaging, but solvable, arithmetic task. As a reaction to the introduction and removal of monetary performance-based rewards, results revealed distinct changes in several brain regions associated with reward processing. Subjects completed three consecutive conditions to examine the systematic influence of changes in performance-based payments on neural activity and behavioral performance.

On a behavioral level, we found a significant decrease in performance in response to the removal of monetary incentives. Because the task itself did not change throughout the experiment, and given that no general interaction with block-independent individual performance level was found, this decrease is most likely attributable to the removal of the reward, while fatigue effects can be excluded as an explanation (fraction of right answers (block 4): M_{block4} = .76, SD =.23; interaction between blocks with performance-based payments: $r_{block2/4}(20)$ =.62, p = .004). Overall, we were able to confirm our hypothesis that taking a monetary reward away leads to decreased performance. Interestingly, no changes in performance due to the introduction of monetary rewards were observed. Because subjects did not anticipate the performance-based rewards, they may have already been performing at the maximum level when the experiment began. A performance level pre-set at maximum implies that, though the additional incentive could not further improve performance, it could at least maintain performance at the initial level.

While performance was a direct measure, neural correlates served as a proxy for the motivation to perform and for the underlying cognitive processes. We concluded that as fatigue effects were excluded as an explanation for the observed behavioral changes, the accompanying changes in BOLD activation can be attributed to the monetary reward manipulation. Moreover, we suggest that the underlying changes in BOLD signals reflect three different processes. First, reward-related brain areas were expected to respond to the introduction and removal of the monetary incentive. Second, we expected to find task-related activation as a response to the cognitive tasks. Third, we hypothesized that effort and self-control changes would be detectable on a neural level. While we found evidence suggesting involvement of reward- and effort-

related changes in BOLD signals in the performance-based incentives processing, we did not find consistent support for a change in task-related areas. It may be the case that these regions were constantly active, so that no relative changes in BOLD signals were detectable. We therefore speculate that incentives do not directly affect task-related activity, but rather modulate the value of a task or the effort that is put into solving a task. This observation implies that several brain regions modulate the decision-making process without directly affecting task-related regions. As well, it may suggest that regions other than task-relevant areas contribute to the execution of a better-elaborated decision process.

We hypothesized that changes in monetary incentives are represented in the vStr. The vStr has repeatedly been implicated in general reward processing (Bartra et al., 2013; Haber & Knutson, 2010; Knutson & Cooper, 2005), and our findings support the idea that because the vStr is sensitive to rewards, observation of that region would track only the incentive. Our finding is also in line with studies suggesting that the vStr represents reward-prediction errors (RPE) in particular (Hare et al., 2008; Schultz, 1998; Schultz, Dayan, & Montague, 1997). RPEs are neural signals generated in reaction to a difference between an expected and an actual reward. In our experiment, RPE reflects reaction to unexpected change in reward when the monetary incentives are introduced or removed. These findings suggest that the incentives are, indeed, processed by participants and are perceived as rewarding (Balleine, Delgado, & Hikosaka, 2007; Bartra et al., 2013; Bloom, 1999; Jenkins Jr et al., 1998). The removal of the monetary incentive was also linked to a decrease in BOLD signals for reward-related brain regions, as well as to a decrease in performance. This is in line with Murayama and colleagues (2010), who found that the removal of monetary incentive leads to a decrease in activity in the striatum. They argue that the striatum may update the beliefs about an outcome of an action, which is equivalent to the idea of the RPE. Nevertheless, in the vStr we did not find a difference between the two non-incentivized conditions. We speculate that the vStr responds only to the reward and RPEs, but not to the demotivation and drop in performance that occurs after removal of the monetary incentives. For further research, it might be advantageous to use a method that is more sensitive to temporal resolution—for example, electroencephalography (EEG) or magnetic encephalography (MEG). Additionally, it might be interesting to investigate the constitution of the network contributing to the processing of the introduction and removal of the performance contingent rewards (Friston et al., 1997; Friston, Harrison, & Penny, 2003; Fuster, 2009; Marreiros, Stephan, & Friston, 2010).

Although the vmPFC is also part of the reward network, that region responded differently than the vStr to the introduction and removal of the monetary incentive. This was to be expected, as previous research has shown that these two regions are associated with different components of the task process. While the vStr responds to RPEs and is, therefore, often associated with updating and learning processes (O'Doherty et al., 2004), it is proposed that activity in the vmPFC reflects subjective value (Hare et al., 2009; Strombach et al., 2015). Subjective values refer to an individual value that is assigned to a reward or action (Hare et al., 2008). Thus, the subjective value reflects the value of an option that is most in accordance with the subject's interests, integrating several motivational factors in addition to monetary rewards. In the given case, as is suggested by the deactivation of the vmPFC in response to the introduction of the monetary incentive, the subjective value of the task is smaller after subjects are rewarded with money. The positive RPE in the vStr, coupled with the deactivation of the vmPFC, is in line with the findings of Hare et al. (2009), who state that the vStr reflects goals or absolute value, and the vmPFC reflects decision or relative value. The observed neural activation patterns also support the behavioral model of Deci et al. (1999), who propose that the introduction of monetary incentives crowd out the intrinsic motivation to perform, thus decreasing the inherent value of a specific action. Although it would be speculative to attribute the changes in BOLD activity to variability in intrinsic motivation, we assume that the perception of the task changes. It might be that the participant focuses on the monetary incentive when rewards are available. Removing the monetary incentives again leads to a decreased BOLD signal in the vStr. Thus, although the vStr and the vmPFC are both involved in reward processing, they show divergent activation patterns over task progression. This divergent pattern suggests that incentives affect the dopaminergic reward system in different ways.

In addition to changes in the vStr and the vmPFC, we also found changes in the activation patterns of the dlPFC. The dlPFC is an area known to be involved in self-control and attention (Crockett et al., 2013; Hare, Hakimi, & Rangel, 2014). Its activation thus reflects cognitive engagement in the task. We hypothesized that the dlPFC may reflect the cognitive engagement to solve the task, and the cognitive engagement seems to be dependent on the monetary incentives. The higher activation in the dlPFC suggests increased cognitive engagement and enhanced self-control when monetary incentives are present (Crockett et al., 2013; Hare et al., 2009). Ballard and colleagues (2011) point out that the dlPFC drives the mesolimbic dopaminergic regions (i.e., the vStr, or nucleus accumbens) to initiate motivated behavior. In line with this, our data shows that the introduction of a monetary, performance-based reward

initiates stronger activation in networks that deal with motivational changes and modulation of cognitive control, supplementing the pure reward-processing signals. Neural activation patterns found in the present experiment correspond to the idea that the removal of the monetary incentive decreased the effort that was put into solving the tasks, along with reduced cognitive control and motivationally driven behavior. Removing the monetary incentives appears to de-motivate participants to the extent that performance-level drop below the initial performance baseline, although the difficulty of the arithmetic calculations is the same throughout the experiment.

Overall, our findings are in line with the results of previous studies investigating the detrimental effect of monetary incentives on motivation (Albrecht et al., 2014; Mobbs et al., 2009; Murayama et al., 2010). Our results suggest that monetary incentives or pay-for-performance do not necessarily modulate task-specific activity, but may change neural activity in regions that are associated with reward-processing and cognitive control. This might lead to a more attentive completion of a given task, without necessarily affecting performance independent of the initial interest in the task itself— and thus also independent of the degree of change in motivation to complete that task.

It is important to consider that the task used in the present experiment may, to some extent, lack ecological validity. However, due to the complex nature of using fMRI in a real-life work setting, application of such methods is currently not possible. Therefore, research methods in this area are currently reliant on behavioral tasks suitable for an fMRI environment, which compromises ecological validity. In future research, methods should be developed to transfer these findings to real life and to organizational situations—functional near-infrared spectroscopy, for example (see Kopton & Kenning, 2014). Furthermore, the use of a strict follow-up order of our experimental design was necessary for our given research question but might have produced different results than a different order or a more randomly chosen order. Future research could investigate possible inherent sequence effects. Nevertheless, our findings strongly support the finding that the removal of monetary rewards leads to a decrease in performance and neural effort. Therefore, incentives should be applied carefully, and only if the continuance of those incentives can be assured.

6. Conclusion

The purpose of the study was to examine whether monetary incentives alter the neural underpinnings of a cognitively engaging task. The dopaminergic reward system plays a major role in the processing of performance-based incentives. Two regions in that system, the vStr and vmPFC, appear to be in opposition to each other. While incentives increase neural activity in the vStr, which is supposed to be sensitive to reward in general, activity decreases in the vmPFC, a brain region associated with the processing of subjective value. However, our findings do not support the idea that the incentives change neural activity in task-relevant regions. We suggest that pay-for-performance does not directly affect performance by modulating neural activity in task-relevant regions, but affects the reward representation during task completion. Therefore, we conclude that performance-contingent incentives need to be applied carefully. This finding is of special interest for employee compensation in unstable economic environments, as companies in unstable environments might not permanently ensure pay-for-performance.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron, 50*(3), 507-517.

Albrecht, K., Abeler, J., Weber, B., & Falk, A. (2014). The brain correlates of the effects of monetary and verbal rewards on intrinsic motivation. *Frontiers in neuroscience, 8*.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+ 2= 4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage, 54*(3), 2382-2393.

Ashburner, J., Neelin, P., Collins, D., Evans, A., & Friston, K. (1997). Incorporating prior knowledge into image registration. *Neuroimage, 6*(4), 344-352.

Ballard, I. C., Murty, V. P., Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2011). Dorsolateral Prefrontal Cortex Drives Mesolimbic Dopaminergic Regions to Initiate Motivated Behavior. *The Journal of Neuroscience, 31*(28), 10340-10346. doi: 10.1523/jneurosci.0895-11.2011

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The journal of neuroscience, 27*(31), 8161-8165.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76*, 412-427.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology, 46*(3), 610.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*(4), 361-383.

Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal, 42*(1), 25-40.

Camerer, C., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty, 19*(1-3), 7-42.

Chib, V. S., De Martino, B., Shimojo, S., & O'Doherty, J. P. (2012). Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron, 74*(3), 582-594.

Crockett, M. J., Braams, B. R., Clark, L., Tobler, P. N., Robbins, T. W., & Kalenscher, T. (2013). Restricting temptations: Neural mechanisms of precommitment. *Neuron, 79*(2), 391-401.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and social psychology, 18*(1), 105.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin, 125*(6), 627.

Deci, E. L., & Ryan, R. M. (1985). *Self-Determination*: Wiley Online Library.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry, 11*(4), 227-268.

Deci, E. L., & Ryan, R. M. (2012). Overview of self-determination theory. *The Oxford Handbook of Human Motivation*, 85-107.

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current opinion in neurobiology, 14*(2), 218-224.

Delgado, M. R., Gillis, M. M., & Phelps, E. A. (2008). Regulating the expectation of reward via cognitive strategies. *Nature neuroscience, 11*(8), 880-881.

Engelmann, J. B., Damaraju, E., Padmala, S., & Pessoa, L. (2009). Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Frontiers in human neuroscience, 3*.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American economic review*, 746-755.

Friston, K. J. (1996). Statistical parametric mapping and other analyses of functional imaging data. *Arthur W. Toga and John C. Mazziotta, editors, Brain Mapping: The Methods*, 363-386.

Friston, K. J., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6*(3), 218-229.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage, 19*(4), 1273-1302.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping, 2*(4), 189-210.

Fuster, J. M. (2009). Cortex and memory: emergence of a new paradigm. *Journal of Cognitive Neuroscience, 21*(11), 2047-2072.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives, 25*(4), 191-209.

Goldman, R. S. (2005). Payment for performance: in sickness and in health. For better or for worse? *Blood purification, 24*(1), 28-32.

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology, 35*(1), 4-26.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science, 324*(5927), 646-648.

Hare, T. A., Hakimi, S., & Rangel, A. (2014). Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. *Frontiers in neuroscience, 8*.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The journal of neuroscience, 28*(22), 5623-5630.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1): Sinauer Associates Sunderland.

Jenkins Jr, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology, 83*(5), 777.

Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences, 107*(19), 8871-8876. doi: 10.1073/pnas.1002007107

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annu. Rev. Econ., 4*(1), 427-452.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science, 314*(5800), 829-832.

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current opinion in neurology, 18*(4), 411-417.

Kopton, I. M., & Kenning, P. (2014). Near-infrared spectroscopy (NIRS) as a new tool for neuroeconomic research. *Frontiers in human neuroscience, 8*.

Kouneiher, F., Charron, S., & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature neuroscience, 12*(7), 939-945.

Krawczyk, D. C., Gazzaley, A., & D'Esposito, M. (2007). Reward modulation of prefrontal and visual association cortex during an incentive working memory task. *Brain research, 1141*, 168-177.

Lawler, E. E. (1971). *Pay and organizational effectiveness: A psychological view*: McGraw-Hill New York.

Lazear, E. P. (2000). Performance Pay and Productivity. *The American Economic Review, 90*(5), 1346-1361. doi: 10.2307/2677854

Le Blanc, P. V., & Mulvey, P. W. (1998). Research Study How American Workers See the Rewards of Work. *Compensation & Benefits Review, 30*(1), 24-28.

Libby, R., & Lipe, M. G. (1992). Incentives, effort, and the cognitive processes involved in accounting-related judgments. *Journal of Accounting Research*, 249-273.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational behavior and human performance, 3*(2), 157-189.

Locke, E. A., Feren, D. B., McCaleb, V. M., Shaw, K. N., & Denny, A. T. (1980). The relative effectiveness of four methods of motivating employee performance. *Changes in working life, 363*, 388.

Marreiros, A. C., Stephan, K. E., & Friston, K. J. (2010). Dynamic causal modeling. *Scholarpedia, 5*(7), 9568.

Merchant, K. A., & Van der Stede, W. A. (2007). *Management control systems: performance measurement, evaluation and incentives*: Pearson Education.

Mobbs, D., Hassabis, D., Seymour, B., Marchant, J. L., Weiskopf, N., Dolan, R. J., & Frith, C. D. (2009). Choking on the money reward-based performance decrements are associated with midbrain activity. *Psychological Science, 20*(8), 955-962.

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences, 107*(49), 20911-20916.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. J., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452-454.

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and affective neuroscience, 2*(1), 67-70.

Rivera, S., Reiss, A., Eckert, M., & Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex, 15*(11), 1779-1790.

Robbins, T. W., & Everitt, B. J. (1996). Neurobehavioural mechanisms of reward and motivation. *Current opinion in neurobiology, 6*(2), 228-236.

Rosenthal, M. B., Frank, R. G., Li, Z., & Epstein, A. M. (2005). Early experience with pay-for-performance. *JAMA: the journal of the American Medical Association, 294*(14), 1788-1793.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68.

Rynes, S. L., Gerhart, B., & Parks, L. (2005). Personnel psychology: Performance evaluation and pay for performance. *Annu. Rev. Psychol., 56*, 571-600.

Schlenker, B. R., Phillips, S. T., Boniecki, K. A., & Schlenker, D. R. (1995). Championoship Pressures: Choking or Triumphing in One's Own Territory? *Journal of Personality and Social Psychology, 68*(4), 632-643.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology, 80*(1), 1-27.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol., 57*, 87-115.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I. I., Tobler, P. N., & Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences, 112*(5), 1619-1624.

Taylor, S. F., Welsh, R. C., Wager, T. D., Luan Phan, K., Fitzgerald, K. D., & Gehring, W. J. (2004). A functional neuroimaging study of motivation and executive function. *Neuroimage, 21*(3), 1045-1054.

Supplementary Material
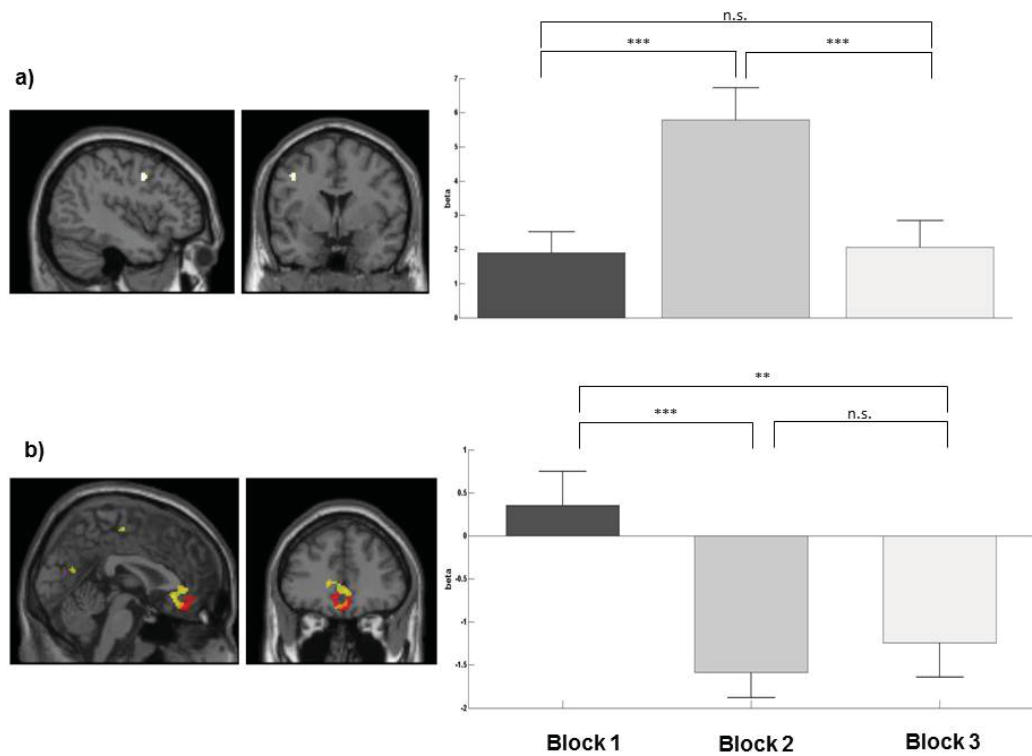
Figure S1: a) visualization and post-hoc ROI-analysis of DLPFC (block 2 minus block 1); b) visualization of conjunction OFC/vmPFC/ACC for contrasts block 1 minus block 2 and block 1 minus block 3)
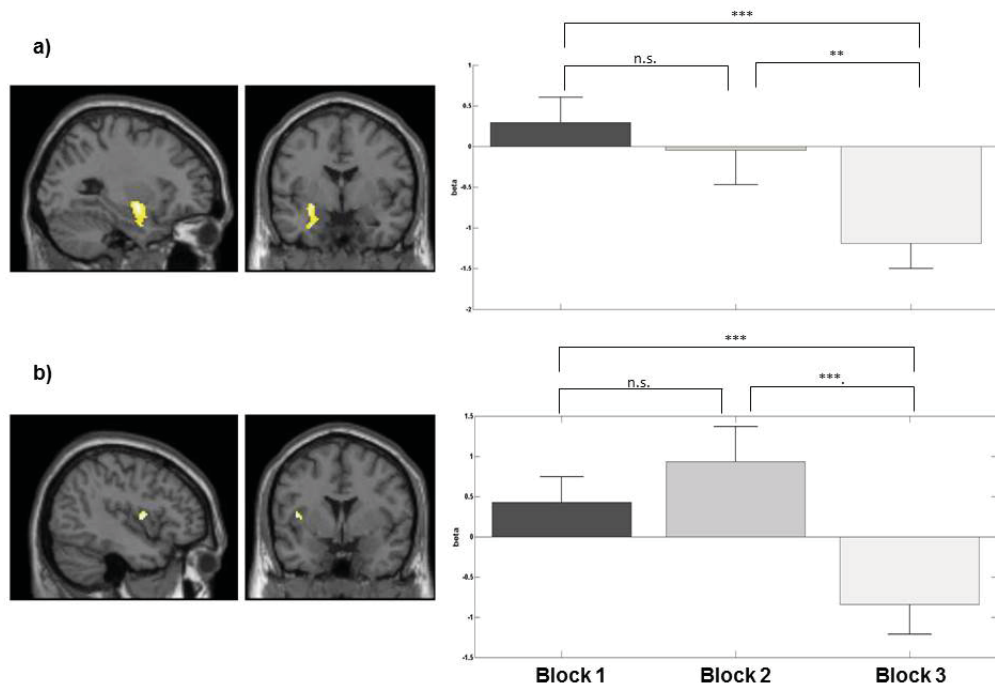
Figure S2: a) visualization and post-hoc ROI-analysis of amygdala (block 1 minus block 3); b) visualization and post-hoc ROI-analysis of insula (block 3 minus block 1)

2. Study 2: Social Discounting Involves Modulation of Neural Value Signal by Temporoparietal Junction

Strombach, T.[a, b, c*]**, Weber, B.**[b,c]**, Hangebrauk, Z.**[a]**, Kenning, P.**[d,e]**, Karipidis, I. I.**[f]**, Tobler, P. N.**[f]**, & Kalenscher, T.**[a] **(2015). Social discounting involves modulation of neural value signals by temporoparietal junction.** *Proceedings of the National Academy of Sciences, 112(5)*, **1619-1624.**

[a] Comparative Psychology, Heinrich Heine University, 40225 Düsseldorf, Germany;

[b] Department of Epileptology, University Hospital Bonn, 53105 Bonn, Germany;

[c] Center for Economics and Neuroscience, University of Bonn, 53127 Bonn, Germany;

[d] Department of Corporate Management and Economics, Zeppelin University, 88045 Friedrichshafen, Germany;

[e] Chair of Marketing, Heinrich Heine University, 40225 Düsseldorf, Germany; and

[f] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zürich, CH-8006 Zürich, Switzerland

* Corresponding author (tina.strombach@hhu.de)

Abstract

Most people are generous, but not towards everyone alike: generosity usually declines with social distance between individuals, a phenomenon called social discounting. Despite the pervasiveness of social discounting, social distance between actors has been surprisingly neglected in economic theory and neuroscientific research. We used fMRI to study the neural basis of this process to understand the neural underpinnings of social decision making. Participants chose between selfish and generous alternatives, yielding either a large reward for the participant alone, or smaller rewards for the participant and another individual at a particular social distance. We found that generous choices engaged the temporo-parietal junction (TPJ). In particular, the TPJ activity was scaled to the social-distance-dependent conflict between selfish and generous motives during prosocial choice, consistent with ideas that the TPJ promotes generosity by facilitating overcoming egoism bias. Based on functional coupling data we propose and provide evidence for a biologically plausible neural model according to which the TPJ supports social discounting by modulating basic neural value signals in the ventromedial prefrontal cortex to incorporate social-distance-dependent other-regarding preferences into an otherwise exclusively own-reward value representation.

Significance Statement

People often consider the well-being of others. However, they are more likely to be generous towards individuals they feel close to than to those they only meet sporadically. Using neuroimaging tools, we show that the decline in generosity across social distance is realized by the interplay of two brain structures – the ventromedial prefrontal cortex coding the relative appeal of a selfish or a generous option, and the temporo-parietal junction modulating appeal signals of the generous outcome, depending on social distance between participant and beneficiary. Based on these findings, we developed a biologically plausible model explaining social discounting in particular, and prosocial behavior in general. Our study opens up new avenues to understand and tackle frictions arising in social networks.

1. Introduction

Prosociality is one of the most fundamental qualities of all human societies. Without the ability to take other people's interests into account human relationships would disintegrate and societies would malfunction. It has been widely demonstrated in laboratory and field experiments that individuals consider the welfare of others in their decisions and the consequences a decision has on them (Fehr & Schmidt, 1999; Jones & Rachlin, 2006; Strombach et al., 2014). Although almost all of us behave prosocially at times, it is clear that people are not equally generous to everyone alike. Rather, generosity decreases as a function of the closeness of the relationship between two individuals (Goeree, McConnell, Mitchell, Tromp, & Yariv, 2010; Jones & Rachlin, 2006). However, it is currently unknown how social distance contributes to the decision process on a neural level. In the present study we set out to address this question.

Our first aim was to investigate the systematic influence of social-distance dependent levels of generosity on neural activation. This was investigated using a social discounting experiment adapted to the functional magnetic resonance imaging (fMRI) environment (Strombach et al., 2014). We measured blood oxygen-level dependent (BOLD) responses while subjects made choices between selfish and generous rewards for themselves and for other people that varied in social distance. Choosing selfishly yielded a payoff only for the subject, while making a generous choice resulted in a lower payoff for the subject coupled with a reward for another person at a specific social distance (fig. 1). Next, based on the individual choices, we reconstructed the social-distance-dependent other-regarding utility (ORU), that is, the value participants attached to increasing the wealth of another person at a given social distance.
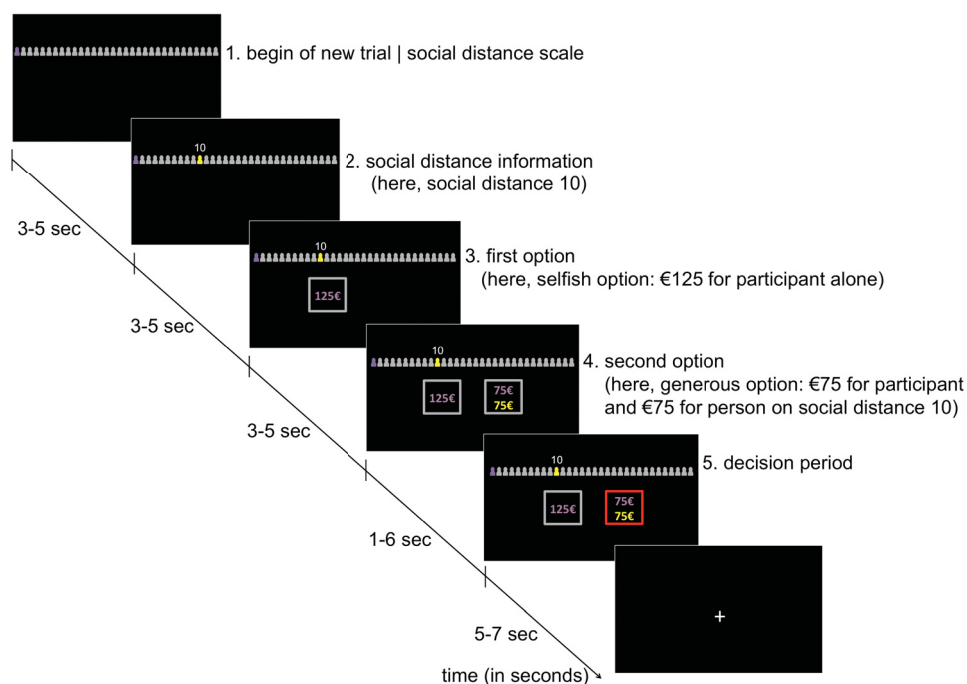
Figure 1: Participants received task-relevant information sequentially. First, social distance information was given on a scale consisting of 101 icons (100 icons representing 100 social distance levels plus one icon, shown in purple on the left end, representing the participant himself). The social distance information for a specific trial was indicated by a yellow icon and, additionally, presented numerically as a number on top of the yellow icon (here: social distance 10). Participants chose between a selfish (here: V125 only for themselves) and a generous option (here: V75 for the participant and V75 for a recipient on the specific social distance). The generous and selfish options were then presented sequentially and in random order. All ISIs had a mean duration of 4 s (jittered by ±1 s). Participants indicated their pref- erence during the decision period within a maximum time frame of 6 s. The trials were separated using a fixation cross with a mean ITI of 6 s (jittered by ±1 s). Note that this figure has been adjusted for illustration purposes; stimulus size and screen format are not to scale with the presentation dimensions used during fMRI scanning. In addition, the figure displays only 21 icons, instead of 101 icons shown during scanning, to facilitate perceptibility.

We then asked which brain regions showed activity that correlated with the difference between other- and self-regarding utilities. Our paradigm was designed so that the degree of generosity varied systematically as a function of social distance while objective economic outcome parameters – own- and other-person payoffs – were kept constant. This allowed us to identify the neural correlates of social-distance-dependent other-regarding preferences independent of objective payoffs.

We hypothesized, based on existing literature (Bartra, McGuire, & Kable, 2013; Fareri, Niznikiewicz, Lee, & Delgado, 2012; Haber & Knutson, 2010; Hare, Camerer, &

Rangel, 2009; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; McClure, Laibson,
Loewenstein, & Cohen, 2004), that own reward values are represented in the brain's
valuation system, specifically in the ventromedial prefrontal cortex (VMPFC).
Furthermore, changes in other-regarding value would recruit areas associated with
theory of mind (ToM) and altruistic choice, such as the temporoparietal junction (TPJ)
(Saxe & Kanwisher, 2003; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014). Should
this be the case, this would show that social distance is indeed systematically
integrated into the neural underpinnings of the decision process.

Our second aim was to investigate the role of the TPJ in prosocial behavior in
more detail. To this end, we tested the predictions of two competing ideas on the role
of the TPJ during prosocial choice in general, and social discounting in particular.
Previous research showed that this region is involved in tasks requiring the ability to
represent and understand others' perspectives (Saxe & Kanwisher, 2003; Saxe, Moran,
Scholz, & Gabrieli, 2006) and in social and selfish decisions (Carter, Bowling, Reeck, &
Huettel, 2012; De Quervain, Fischbacher, Treyer, & Schellhammer, 2004; Krajbich,
Adolphs, Tranel, Denburg, & Camerer, 2009; Morishima, Schunk, Bruhin, Ruff, & Fehr,
2012). Thus, the TPJ's implication in prosocial choice, perspective-taking, empathizing
and ToM suggests that it plays a role in putting oneself in someone else's shoes. In
other words, the TPJ may encode the other-regarding value participants attach to
increasing the well-being of another person. Individuals empathize more with people
they feel close to than with more distant others. Therefore, if this hypothesis is true,
TPJ activation should correlate positively with the social-distance-dependent ORU.
This view of the role of the TPJ is challenged by more recent studies postulating that
TPJ activation solves the conflict between generous and selfish motives (Morishima et
al., 2012). According to this hypothesis, to make a generous decision, the putatively
natural bias to maximize own-payoff needs to be overcome (Knoch, Pascual-Leone,
Meyer, Treyer, & Fehr, 2006). If the TPJ enables overcoming egoism bias, activation
should be high when the temptation to be selfish is high (i.e. large social distance
and/or large selfish reward) and low when there is little conflict between selfish and
generous motives (i.e. small social distance and/or relatively small selfish reward).

Our results confirmed the latter hypothesis according to which the TPJ plays a
role in overcoming the default response of maximizing one's own profit and thus
behaving selfishly, rather than in representing other-regarding value. We also asked
how the brain implements generous decisions. Specifically, we propose that the TPJ
facilitates generous decisions by modulating basic reward signals in the VMPFC,
incorporating other-regarding preference signals into an otherwise exclusive own-

reward value representation, thus computing the subjective value for a social reward. Thus, the TPJ supports prosocial choice by shaping neural value signals in the VMPFC whenever the temptation to be selfish needs to be overcome; the stronger the temptation to be selfish, the more the TPJ upregulates VMPFC activity in favor of generous choices.

A mechanistic model that integrates these data makes two predictions for which we provide empirical support: first, VMPFC activity should be higher during generous than selfish decisions. Second, connectivity between the TPJ and the VMPFC should be stronger during generous than during selfish decisions. Together, our findings suggest that prosocial decisions arise from a refined interplay between the VMPFC and the TPJ. In particular, value signals in the VMPFC are orchestrated by the TPJ according to the social distance between the decision maker and the recipient of generous decisions.

2. Results

2.1 Behavioral results

We inferred social discounting parameters based on the participants' individual choices, and used the obtained social discount functions to econometrically reconstruct the social-distance-dependent ORU for each individual (Strombach et al., 2014). To this end, we first determined, for each social distance level, the point at which a participant was indifferent between the selfish (yielding a larger reward for the participant) and the generous alternative (yielding a smaller reward for the participant plus a reward for the other person) using logistic regression. The difference in reward magnitudes for the participant between the two alternatives at the indifference points represented the amount of money a subject was willing to forego to increase the wealth of another person at a given social distance, and could be construed as a social premium equivalent to the utility of increasing the other person's well-being. For example, if a participant was indifferent between €125 own-reward and €75 own-reward and €75 for a person at social distance 1, this participant was willing to forego €50 (the social premium) to increase the wealth of the other person by €75. Subsequently, we fit a standard hyperbolic model (Jones & Rachlin, 2006; see methods) to the individual social-distance-dependent social premiums with the parameters k (Mdn=0.078) and V (Mdn=74.15). As expected, the magnitude of the social premium subjects were willing to pay for someone else's benefit declined with increasing social distance. The standard

hyperbolic model captured the individual discounting behavior well (mean r^2=0.72, SD=0.242; fig. 2 shows the median social premiums together with the best-fitting hyperbolic function).
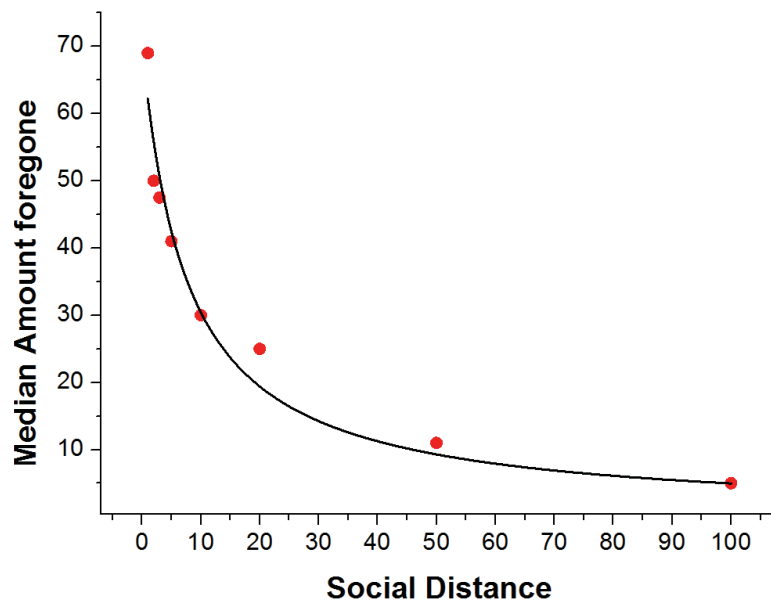
Figure 2: We determined, for social distance levels 1, 2, 3, 5, 10, 20, 50, and 100, the individual payoff magnitudes at which a participant was indifferent between the selfish (yielding a larger reward for the participant) and the generous alternative (yielding a smaller reward for the participant plus a reward for the other person). The amount foregone, i.e., the difference in own-reward magnitude between the selfish and generous option at in-difference point, indicates the willingness to sacrifice a reward to give to another person at a specific social distance. The amount foregone can be interpreted as a social premium that reflects the utility a participant attaches to increasing a recipient's payoff. A standard hyperbolic model was fit to the individual social-distance–dependent amounts foregone to reconstruct the participant's ORU function. The figure shows the best-fitting hyperbolic function to the median amounts foregone across all participants.

These findings replicated those of previous studies on social discounting (Jones & Rachlin, 2006; Strombach et al., 2014) suggesting that the scanner environment did not substantially affect discounting behavior compared to studies carried out in more natural surroundings. The obtained individual hyperbolic fits served as estimates of the decline in other-regarding value across social distance, and were used to estimate the individual ORUs, which corresponded to the hyperbolic fit plus the sure €75 (constant, within-subject) for the participant herself.

2.2 Neural mechanisms of social discounting

Social discounting is the consequence of the social-distance-dependent balance between generous and selfish motives. The fundamental premise in our study is that, in essence, a prosocial decision results from the balance between generous and selfish motives. Due to social discounting, the balance between generous and selfish motives is increasingly tilted towards selfishness as social distance increases. With this premise in mind, we tested the hypothesis that the TPJ is involved in orchestrating the balance between generous and selfish motives across social distance. We also hypothesized that the TPJ would perform this function together with classical value coding regions such as the VMPFC (Hare et al., 2009; Kenning & Plassmann, 2005).

Neural value signals in the valuation network. First, we investigated the neural correlates of selfish rewards. To identify brain regions associated with own-reward value coding, we examined neural activity during the decision period of selfish decisions only. For this analysis we concentrated on the VMPFC, and asked whether BOLD activity in the VMPFC covaried with selfish reward magnitude, thus with the value of the selfish decision. We used a region of interest (ROI) based on a meta-analysis [-2, 40, -4] (Clithero & Rangel, 2013), that suggests this part of the VMPFC plays a role in value processing. Using a 6 mm sphere around the ROI, we found significant correlations within the VMPFC (-6, 41, -5; $t_{(22)}=3.10$, p=0.017, small volume (SV) family-wise error (FWE) corrected, SI Appendix, fig. S3, table S1).

Next, we also included generous decisions in our analysis. Interestingly, we found that activity in the VMPFC was significantly higher during generous than during selfish choices (0, 47, -20; $t_{(22)}=4.21$, p=0.028, whole brain FWE corrected; fig. 3A, SI Appendix, table S2). Thus, in line with other findings (Harbaugh, Mayr, & Burghart, 2007), the VMPFC coded not only the own-reward value of a selfish choice, but also generosity in addition to own-reward value, possibly reflecting the satisfaction derived from increasing someone else's wealth (Harbaugh et al., 2007). Generous decisions also elicited stronger responses than selfish decisions in the right (60, -58, 31; $t_{(22)}=5.15$, p<0.001, whole brain FWE corrected; SI Appendix, table S2) and left TPJ (-24, -79, 52, $t_{(22)}=4.51$, p=0.002, whole brain FWE corrected). This section of TPJ has previously been shown to be associated with ToM and altruistic choice (Schurz et al., 2014). We used these brain areas as ROIs in all subsequent analyses to further characterize their contribution to social discounting.
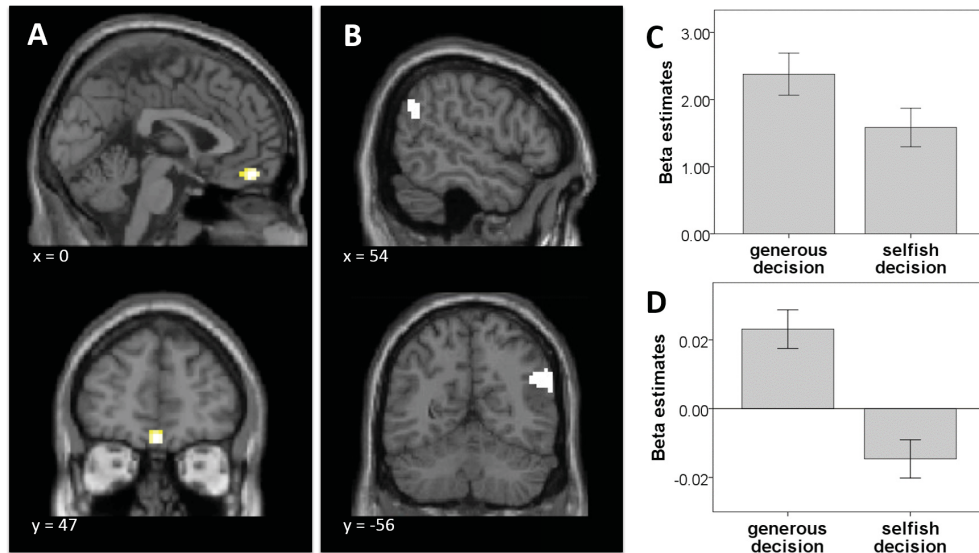
Figure 3: Brain activations during social discounting. (A) BOLD responses in the VMPFC were stronger during generous than during selfish decisions [0, 47, −20; t(22) = 5.47, P = 0.028, whole-brain FWE corrected; displayed at P < 0.005, uncorrected, k ≥ 10 voxel]. (B) Generous decisions elicited activation in the posterior part of the rTPJ. (C) Beta estimates within the rTPJ . The rTPJ was more activated during generous than selfish decisions. (D) Activity in the rTPJ was more strongly modulated by the temptation to be selfish during generous than selfish decisions. Error bars indicate ±1 SE.

Generous decisions recruit the TPJ to resist the temptation to be selfish. Our design allowed us to shed light on two competing hypotheses about the role of the TPJ in social decision-making. If the TPJ was important for encoding the social-distance-dependent value participants derive from increasing someone else's well-being, we would expect a positive correlation between TPJ activity and the ORU. To test the first hypothesis, we searched for brain regions whose activity correlated with the social-distance-dependent other-regarding value, using the individual ORUs as parametric regressors at decision onset. Inconsistent with our predictions, the parametric analysis revealed no activation in the TPJ, even at very liberal thresholds (p<0.1, uncorrected; see supplemental material for more analyses; SI Appendix, table S3).

Next, we tested the second hypothesis, that the TPJ is associated with overcoming egoism-bias. We reasoned that the temptation to make a selfish choice should be stronger, the larger the utility of the own reward relative to the social-distance-dependent other-regarding utility. By extension, the stronger the temptation to be selfish, the more effort should have been exerted in order to overcome this temptation when a generous decision had been taken. Thus, we hypothesized that

activity in brain regions important for overcoming egoism bias would scale to the difference between own-reward and other-regarding values when a generous choice is revealed. We therefore searched for BOLD signals that correlated with the difference between own-reward and other-regarding values during generous decisions. This analysis revealed, among others, activation in the right parietal cortex, expanding into the parietal part of the TPJ (rTPJ; 42, -79, 46; $t_{(22)}$=5.55, p=0.019, whole brain FWE corrected; SI Appendix, table S4). A conjunction analysis, revealing the strict intersection between this contrast and the contrast yielding generosity-related activations confirmed that this was indeed the same region as the one engaged during generous decision making. Additionally, a ROI-analysis using a mask for the parietal subsection of rTPJ (Crockett et al., 2013), an area known to be involved in social cognition, confirmed that it was activity in this 'social' part of the TPJ that correlated with the parametric modulation of the temptation to be selfish during generous decisions (p=0.032, SV FWE corrected, fig 3B).

Importantly, according to the second hypothesis, the rTPJ should be active when the conflict between selfishness and generosity is resolved in favor of the latter (i.e. in generous choices), while activity should be less when selfish choices are made. In support of this idea, rTPJ activation survived when contrasting the difference in own-reward and ORU during generous against selfish decisions, suggesting that the rTPJ was more active during generous (i.e., when the temptation to be selfish had been overcome) than during selfish choices (42, -79, 46; $t_{(22)}$=7.10, p<0.001, whole brain FWE corrected; SI Appendix, table S5). Thus, rTPJ activation correlated with the difference between own reward values and ORU when generous choices were made, but not with ORU in general, irrespective of the actual decisions taken. Although the lack of evidence in favor of the first hypothesis of TPJ function is not evidence against ToM, it is worth noting that rTPJ activity was in fact *positively* correlated with the difference between own-reward and other-regarding value. In other words, rTPJ activity was *negatively* correlated with other-regarding utility alone, which is difficult to reconcile with the ToM-based hypothesis. Conversely, our results are in line with the hypothesis of overcoming egoism bias, and are therefore more consistent with the idea that the rTPJ facilitates generous choice whenever a conflict between egoistic and selfish motives needs to be resolved.

The TPJ was functionally connected with the VMPFC when egoism bias was overcome. So far, our data have shown that the VMPFC encoded the value of own-reward during

selfish decisions, and both the VMPFC and the TPJ were more engaged during
generous than during selfish choices. Furthermore, the rTPJ activation pattern was
consistent with the idea that the rTPJ facilitates prosocial choice by overcoming the
temptation to maximize own payoff. We next asked how a social-distance-dependent
prosocial choice is implemented in the brain. We propose a model of generous choice
according to which the VMPFC encodes goal values. We hypothesize that the TPJ
suppresses egoism bias by modulating basic value signals in the VMPFC to incorporate
other-regarding preferences into an otherwise exclusively own reward value
representation.

If our hypothesis was true, functional connectivity between the rTPJ and the
VMPFC should be higher during generous than during selfish decisions. Our idea was
inspired by recent models of self-control (Crockett et al., 2013; Hare et al., 2009)
according to which basic value representations in the VMPFC are modulated by
superordinate brain regions encoding higher-order considerations, such as long-term
goals or other-regarding preferences. Thereby, these higher-order factors are
incorporated into basic valuation signals of the VMPFC. Adapting this approach to
social decision-making, we propose that the rTPJ modulates activation in the valuation
network and orchestrates social decision-making in favor of other-regarding instead of
individual preferences.

To test the hypothesis, we conducted a psychophysiological interaction analysis
(PPI) to identify which brain regions were functionally more strongly connected with
the rTPJ during generous than during selfish decisions. To this end, we placed a seed
(fig 4A) in the individual peak activations in the right TPJ cluster associated with
overcoming egoism bias. The PPI analysis identified functional connectivity between
the rTPJ and the VMPFC ($t_{(14)}$= 6.61, p = 0.031, whole-brain FWE corrected; fig. 4B, SI
Appendix, table S6). A conjunction analysis confirmed that the activated VMPFC
cluster was indeed the same region as the VMPFC ROI that coded own-reward value
during selfish decisions. Thus, compared to selfish decisions, the TPJ increased
functional connectivity with regions associated with own-reward processing during
generous decisions.
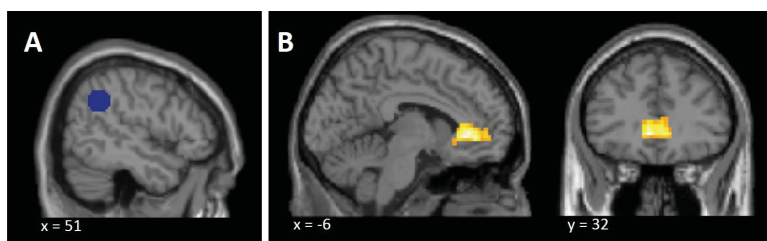
Figure4: (A) ROI in the rTPJ (51, −49, 34; 10-mm sphere) as the seed region for the PPI. (B) Positive functional connectivity of the rTPJ with VMPFC during generous decisions. The PPI analysis revealed that connectivity between the VMPFC and the rTPJ was stronger during prosocial than selfish choices [t(14) = 6.61, P = 0.031, whole-brain FWE corrected; displayed at P < 0.005, un-corrected, k ≥ 10 voxel].

3. Discussion

In order to function well in our society it is important to share resources with others. The closer our interaction partners are to us, the more likely we are to be generous towards them. Thus, the social context in which a decision is made strongly affects how the information is processed, making it essential for our brain to be able to encode such social context factors (Rushworth, Kolling, Sallet, & Mars, 2012). However, neuroeconomic theories have so far neglected social distance in models of decision-making. The current experiment investigated the neural correlates of social discounting and aimed to provide support for a neural model of social distance-dependent generous decision-making. Generosity requires overcoming egoistic motives (Declerck, Boone, & Emonds, 2013; Jones & Rachlin, 2006; Morishima et al., 2012; Strombach et al., 2014), and the temptation to be selfish grows with increasing own-reward magnitude, but also with increasing social distance. We were able to demonstrate, that a region associated with ToM, social cognition and decision making, the TPJ, is involved in this process (Declerck et al., 2013). However, contrary to the predictions of the first, ToM-based, hypothesis we find no evidence that the TPJ computes other-regarding value. Instead, we propose that the TPJ facilitates overcoming egoistic motives to maximize own-payoff during generous decisions by modulating basic value signals in the VMPFC through integrating other-regarding preferences into an otherwise exclusively own-reward value representation (Declerck et al., 2013; Smith, Clithero, Boltuck, & Huettel, 2014). Thus, the stronger the social-distance-dependent temptation to be selfish, the more the TPJ is engaged and VMPFC-value signals become upregulated to facilitate a generous decision.

On a behavioral level, we replicated existing findings on social discounting, confirming that generosity declines hyperbolically across social distance, with individuals being more willing to forego a reward for recipients at close social distances. Hyperbolic social discounting has been observed in diverse locations, populations, cultures as well as under different experimental conditions and with various methods of implementing social distance and eliciting social preferences (Jones & Rachlin, 2006; Jones & Rachlin, 2009; Rachlin & Jones, 2008a, 2008b; Strombach et al., 2014). This suggests that a hyperbolic discount function is an accurate, valid and useful description of social discounting behavior, even though it is likely that there is a large range of individual motives underlying the actual decisions during social discounting.

To reveal the neural mechanisms underlying social discounting, we first identified regions that code value signals. When participants made selfish choices, activity in the VMPFC reflected own-reward value, replicating findings on valuation processes from a multitude of studies (Bartra et al., 2013; Hare et al., 2009; Hare et al., 2008; Kurniawan, Guitart-Masip, Dayan, & Dolan, 2013; McClure et al., 2004). Additionally, we found that the VMPFC was more active during generous than during selfish choices, even though selfish decisions yielded higher payoffs for the participants. This is in line with studies which postulated that VMPFC activity represents the extra value obtained from charitable giving (Harbaugh et al., 2007). Thus, it is likely that the increased activation during generous choices reflected the additional satisfaction derived from sharing a reward during generous decision-making.

Our results suggest that the TPJ is important for overriding selfish impulses during prosocial decisions. Note that recent research on TPJ function suggests that it is not a monolithic structure that supports one single cognitive function, but is more likely to be composed of anatomically and functionally distinct subdivisions that may subserve different computational roles such as value, salience and ToM (Declerck et al., 2013; Kahnt, Park, Haynes, & Tobler, 2014; Kahnt & Tobler, 2013; Schurz et al., 2014; Smith et al., 2014). Although we cannot rule out that our subjects used ToM or other mechanisms to make their decisions, it is possible that we found no evidence in favor of the first, ToM-based hypothesis simply because we did not explicitly elicit ToM-cognition. Thus, we are not rejecting the wealth of evidence from previous work relating the TPJ to ToM and mentalizing. Instead, we complement existing literature by lending support to the idea that subparts of the TPJ have the additional role of overcoming egoism bias and thus facilitating prosocial choice. Future research should

study the TPJ in more detail and map the putative different cognitive functions to its different subdivisions.

In summary, the present findings provide new insights into social decision making. We were able to characterize the role of the TPJ and proposed a neural model of prosocial choice. Our data identify the TPJ as the core component in overcoming egoism bias. This finding has a significant impact on neuroeconomic theory that has so far neglected the effect of social distance on prosocial decision making. Having shown that social distance is an important component of an individual's decision-making process it should be integrated into future models of decision making. Furthermore, using social discounting to understand the influence of social factors and individual differences in generosity and other-regarding behavior opens up new opportunities to evaluate psychopathologic decision-making and antisocial behavior in more detail.

4. Materials and Methods

Participants. 27 subjects (mean$_{age}$=25.03, 14 men) were tested. Participants received a €10 show-up fee and an additional amount depending on their decisions in the experiment (own reward: €7,50 - €16,50, other reward: €2,50 or €7,50), determined by a random draw of one of the trials. All subjects were native German speakers. Participants had no history of psychiatric or neurological disorders. Written consent was obtained according to the Declaration of Helsinki (BMJ 1991; 302: 1194). The study was approved by the ethics committee of the University of Bonn. Subjects were acquired using the subject database of the Life&Brain Centre, University Hospital Bonn.

Stimuli and task. The experimental paradigm was adapted from a cross-cultural study on social discounting (Strombach et al., 2014). During the preparation phase participants received verbal and written instructions for the tasks they carried out during the experiment. Participants started with a self-representation task in which they specified closeness to people in their social environment (Strombach et al., 2014). Using a 20 point scale (1=very close; 20= not close), participants were asked to rate their closeness to the following people: mother, father, siblings, grandparents, family, kin, best friend, circle of friends, colleagues, neighbors, acquaintances, partner, children, and stranger. In case some of these people did not exist in a participant's social environment, the corresponding trial was skipped. This task was aimed at

getting subjects used to the idea of social distance and to think about their social
network.

In the fMRI scanner, social distance was transformed into a scale consisting of
100 icons (fig. 1). Participants were informed that the purple icon at the left end of the
scale represents themselves and the yellow icon stands for a specific person in their
social environment. For example, if the yellow icon is directly next to the purple one
(social distance 1) this shows the person they feel closest to, e.g. mother or partner. If
the yellow icon is at social distance 50 (the middle of the scale) this symbolizes an
acquaintance while at social distance 100 it would represent the most socially distant,
but emotionally neutral person, such as a stranger. Before entering the scanner,
participants were asked to choose and write down names of representatives from their
social environment, one for each of the following social distances: 1, 2, 3, 5, 10, and 20.
We also included social distance levels 50 and 100 in the experiment, however, as these
distance levels represent remote acquaintances or strangers, subjects were not required
to indicate a name. Thus, eight social distances were included. We used a network-
based approach, according to which one might mentally assign more than one person
to a particular social distance (Harrison, Sciberras, & James, 2011). However, in the
experiment participants were asked to choose just one person for each distance. The
network based approach is important since we assume that an individuals' social
network is adaptive and might change within days. For example, bad experience with
a friend might result in a readjustment of the perceived social distance to that friend.
Furthermore, as negative emotions can interfere with prosocial behavior subjects were
explicitly asked to only include individuals they did not have a negative attitude
towards (Bechara, 2004; Lerner & Tiedens, 2006; Sanfey, Rilling, Aronson, Nystrom, &
Cohen, 2003). Subjects did not indicate any problems understanding the scale and the
idea of social distance.

In the scanner, subjects were asked to make 160 decisions involving the eight
social distances (fig. 1). The task in each trial was to think about the person previously
chosen for the specific social distance relevant to the actual trial. Each of the 160 trials
started with the presentation of the scale indicating the relevant social distance
followed by the generous and selfish options with a mean ISI of 4 sec (jittered: +/- 1
sec) and a mean ITI of 6 sec (jittered: +/- 1 sec). The temporal and spatial ordering of
the selfish and generous option presentation was pseudo-randomized. For the specific
social distance, subjects had to choose between these two alternatives. The selfish
alternative always yielded a large reward for the participant alone, while the generous
option yielded a smaller reward for the participant and an additional reward for the

person at the indicated social distance. The selfish reward varied between €75 and €165, changing in increments of €10, resulting in ten selfish alternatives. The generous option was a fixed reward of €75 for the participant and €75 for the other person (high other-reward trials), or, €25 for the other person, respectively (low other-reward trials). The presentation order of social distances, selfish and generous alternatives was fully randomized. Subsequent to the experimental part in the scanner, subjects were asked to name the people they assigned to the specific social distances again, serving as a manipulation check and indicate demographic information.

After completing the last questionnaire, subjects received their payment. In addition to the €10 show-up fee, a randomly chosen trial was paid out. Depending on whether the participant chose the selfish or the generous alternative in that specific trial, she received 10% of the selfish or the generous reward respectively. The money for the selfish option was paid directly to the participant, and for the generous option subjects were asked to indicate the address of the other person. If the randomly chosen trial was about a person at social distance 50 or 100, a random person on the campus of the University of Düsseldorf, Germany received the reward. Thus, the experiment was fully incentive-compatible, did not include deception and met the experimental standards of behavioral economics.

Social discount function. Since we aimed to quantify the degree of generosity as a function of social distance, we estimated the amount of money a participant was willing to forego in order to benefit a specific other at a given social distance (Jones & Rachlin, 2006). We first determined, for each social distance level, the point at which a participant was indifferent between the selfish and the generous alternatives, using logistic regression as described above. The decision maker switched from being generous to being selfish as selfish rewards increased. If the decision maker switched from generous to selfish decisions between a selfish reward of €135 and €145, the indifference point would be determined to be €140, thus a 50% probability of choosing generous and 50% of choosing selfish. We interpreted the amount foregone (the indifference point minus the €75 the subject would certainly get if he chose the generous option) as a social premium the participant was willing to pay in order to benefit the other. We fit the following standard hyperbolic model to the individual, social-distance-dependent social premiums (Jones & Rachlin, 2006):

$$v = \frac{V}{(1+kD)}$$
                                          Equation 1: Hyperbolic discount function

where *v* symbolizes the magnitude of a reward received by another person at social distance *D*. The parameter *V* refers to the social premium a subject is willing to pay in exchange for endowing another person with reward *v*. Thus, *V* can be interpreted as the socially discounted other-regarding utility of improving the wealth of another individual at social distance *D*. *V* is equal to the self-regarding utility at social distance *D*=0, thus the intercept with the y-axis, and determines the height of the social discount function. The degree of discounting is described by the parameter *k*, which indicates the steepness and shape of the curve (B. Jones & Rachlin, 2006; Strombach et al., 2014). The individual hyperbolic fits were used to estimate individual ORUs.

fMRI data acquisition and preprocessing. Scanning was performed on a 3 Tesla Trio Scanner (Siemens, Erlangen, Germany) using an 8-channel head coil. Functional data were acquired using EPI-sequences with a repetition time (TR) of 2.5 s, an echo time (TE) of 30 ms, and a Flip angle of 90 degrees. Each volume comprised 37 slices acquired in an axial orientation covering all of the brain, including midbrain, but sparing parts of the cerebellum. The presentation of the task and recording of behavioral responses were performed using Presentation® software version 14.9 (Neurobehavioral Systems, Albana, Canada). Subjects saw the experiment via video goggles (Nordic NeuroLab, Bergen, Norway) and gave their responses by response grips (Nordic NeuroLab, Bergen, Norway) using their index fingers of both hands.

Neural data of 23 participants were analyzed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK) software. The results are visualized using the xjview toolbox. Three subjects had to be excluded due to extreme head-movements during the experiment (> 4mm translation, > 4° rotation). One subject had to be excluded who made exclusively selfish decisions, even when being generous did not involve any reduction in own-reward, as the aim of this study was to investigate pro-social behavior. Indeed, no social discount function can be fit to the data when a participant shows no variation in other-regarding utility.

The following pre-processing steps were carried out: slice timing correction, motion correction, segmentation using the T1 weighted image, linear trend removal, high pass temporal filtering with a filter size of 128 seconds, spatial smoothing using a Gaussian kernel with full-width at half-maximum (FWHM) of 8mm, spatial segmentation and spatial normalization by co-registering the functional with the individual structural data and then transforming it into the Montreal Neurological Institute (MNI) space.

General linear model. We regressed fMRI time series onto two separate general linear models (GLMs). With the first GLM, we aimed to identify brain regions whose activity correlated with selfish reward magnitude and the econometrically reconstructed, social-distance-dependent ORU. In the second GLM, we searched for neural activity correlating with the relative value, i.e. the social-distance-dependent difference between the selfish value and ORU.

For both GLMs we defined the following five onset regressors:1) onset of the social distance information at the beginning of each trial, 2) onset of the generous option, 3) onset of the selfish option, 4) onset of the button press when deciding generous, and 5) onset of the button press when deciding selfish. We modeled BOLD responses at these onsets as stick functions. For the first GLM, we used the following three parametric modulators to assess brain activation (36): 1) the social distance level during the onset of the social distance information (onset regressor 1), 2) the econometrically reconstructed ORU, given a generous choice (onset regressor 4), and 3) the selfish reward magnitude given a selfish choice (onset regressor 5). To obtain commensurability, ORU and selfish reward magnitudes were transformed and normalized to a common scale for all analyses.

In the second GLM we used the following parametric modulators: 1) the social distance level during the onset of the social distance information, 2) the difference between own-reward magnitude and ORU, given a generous decision (i.e., the strength of the temptation to choose selfish, given a generous decision; see main text), and 3) the difference between own-reward magnitude and ORU, given a selfish decision. Both GLMs additionally included six movement regressors of no interest, three for translational movements (x, y, z) and three for rotation movements (pitch, roll, yaw). All regressors were convolved with the canonical hemodynamic response function (HRF). For each event, onset regressor parameter estimates were obtained and contrast images of each of the parameters against zero were generated. Furthermore, we obtained contrast images of "deciding generous versus deciding selfish". The obtained images were transferred to a second level random effects analysis using one-sample t tests on the single-subject contrasts. We performed whole brain corrections for multiple comparisons at the cluster level. For all of the main contrasts reported in the results section and figures, the individual voxel threshold was set to $p > 0.005$ with a minimal cluster extent of $k \geq 10$ voxel (Lieberman, & Cunningham, 2009). Results are reported using the MNI coordinate system.

Psychophysiological Interactions (PPI). We performed a whole-brain PPI analysis
with the TPJ as seed region (Crockett et al., 2013; Friston et al., 1997; Hare et al., 2009).
The location of the TPJ seed ROI was based on a 10mm sphere (Eickhoff et al., 2009)
around the peak activation within the conjunction between the contrasts of generous
versus selfish decisions (first GLM), and the parametric modulation of the temptation
to be selfish (second GLM; 51, -49, 34; rTPJ; fig. 4A). We computed individual average
time series within a 4 mm sphere surrounding (Eickhoff et al., 2009) the individual
subject peak activations within the TPJ seed ROI. Seven participants had to be
excluded from the PPI analysis because they did not show any individual activation
above threshold in the TPJ ROI at $p<0.05$, uncorrected. This exclusion criterion is the
standard for identifying the location of corresponding activations in individual
subjects as needed to extract time courses for connectivity analyses (Booth, Wood, Lu,
Houk, & Bitan, 2007; Bzdok et al., 2013; Eickhoff et al., 2009; Heim et al., 2009). We
created two PPI regressors by computing an interaction regressor between the
normalized time-series and the respective condition, i.e. one regressor for generous
and one for selfish decisions.

Second, we estimated a GLM with the following regressors: 1) a physiological
regressor (i.e., the entire time-series of the seed region over the whole experiment, 2) a
psychological regressor for the onset of the generous choices, 3) the PPI regressor for
the generous choices, 4) a psychological regressor for the onset of the selfish choices,
and 5) a PPI regressor for the selfish choices. The onset and PPI regressors were
convolved with a canonical form of the hemodynamic response. The model also
included the six motion parameters as regressors of no interest.

In a third step, to identify regions whose connectivity was higher during
generous than during selfish choices, single subject contrasts were calculated for the
contrast between the PPI regressors, i.e. the contrast between the PPI regressor of the
generous compared to the PPI regressor of the selfish choices. Then a second level
analysis was performed by calculating a one-sample *t*-test on the single-subject
contrast coefficients. We then identified voxels with significantly higher connectivity
difference during generous compared to selfish choices.

References

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76*, 412-427.

Bechara, A. (2004). The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain and cognition, 55*(1), 30-40.

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain research, 1133*, 136-144.

Bzdok, D., Langner, R., Schilbach, L., Jakobs, O., Roski, C., Caspers, S., . . . Eickhoff, S. B. (2013). Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. *Neuroimage, 81*, 381-392.

Carter, R. M., Bowling, D. L., Reeck, C., & Huettel, S. A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science, 337*(6090), 109-111.

Clithero, J. A., & Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. *Social cognitive and affective neuroscience*, 106.

Crockett, M. J., Braams, B. R., Clark, L., Tobler, P. N., Robbins, T. W., & Kalenscher, T. (2013). Restricting temptations: Neural mechanisms of precommitment. *Neuron, 79*(2), 391-401.

De Quervain, D. J., Fischbacher, U., Treyer, V., & Schellhammer, M. (2004). The neural basis of altruistic punishment. *Science, 305*(5688), 1254.

Declerck, C. H., Boone, C., & Emonds, G. (2013). When do people cooperate? The neuroeconomics of prosocial decision making. *Brain and cognition, 81*(1), 95-117.

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random effects approach based on empirical estimates of spatial uncertainty. *Human brain mapping, 30*(9), 2907-2926.

Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., & Delgado, M. R. (2012). Social network
modulation of reward-related signals. *The Journal of neuroscience, 32*(26), 9045-
9052.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation.
Quarterly Journal of Economics, 114(3), 817-868.

Friston, K. J., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997).
Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage,
6*(3), 218-229.

Goeree, J., McConnell, M., Mitchell, T., Tromp, T., & Yariv, L. (2010). The 1/d Law of
Giving. *American Economic Journal: Microeconomics, 2*(1), 183-203.

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and
human imaging. *Neuropsychopharmacology, 35*(1), 4-26.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and
voluntary giving reveal motives for charitable donations. *Science, 316*(5831),
1622-1625.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making
involves modulation of the vmPFC valuation system. *Science, 324*(5927), 646-
648.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating
the role of the orbitofrontal cortex and the striatum in the computation of goal
values and prediction errors. *The Journal of Neuroscience, 28*(22), 5623-5630.

Harrison, F., Sciberras, J., & James, R. (2011). Strength of social tie predicts cooperative
investment in a human social network. *PLoS One, 6*(3), e18338-e18338.

Heim, S., Eickhoff, S. B., Ischebeck, A. K., Friederici, A. D., Stephan, K. E., & Amunts,
K. (2009). Effective connectivity of the left BA 44, BA 45, and inferior temporal
gyrus during lexical and phonological decisions identified with DCM. *Human
brain mapping, 30*(2), 392-402.

Jones, B.A., & Rachlin, H. (2006). Social discounting. *Psychological Science, 17*(4), 283-
286.

Jones, B. A., & Rachlin, H. (2009). Delay, probability, and social discounting in a public goods game. *Journal of the Experimental Analysis of Behavior, 91*(1), 61-73.

Kahnt, T., Park, S. Q., Haynes, J.-D., & Tobler, P. N. (2014). Disentangling neural representations of value and salience in the human brain. *Proceedings of the National Academy of Sciences, 111*(13), 5000-5005.

Kahnt, T., & Tobler, P. N. (2013). Salience signals in the right temporoparietal junction facilitate value-based decisions. *The Journal of Neuroscience, 33*(3), 863-869.

Kenning, P., & Plassmann, H. (2005). NeuroEconomics: An overview from an economic perspective. *Brain Research Bulletin, 67*(5), 343-354.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *science, 314*(5800), 829-832.

Krajbich, I., Adolphs, R., Tranel, D., Denburg, N. L., & Camerer, C. F. (2009). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *The Journal of Neuroscience, 29*(7), 2188-2192.

Kurniawan, I. T., Guitart-Masip, M., Dayan, P., & Dolan, R. J. (2013). Effort and valuation in the brain: the effects of anticipation and execution. *The Journal of Neuroscience, 33*(14), 6160-6169.

Lerner, J. S., & Tiedens, L. Z. (2006). Portrait of the angry decision maker: How appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making, 19*(2), 115-137.

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social cognitive and affective neuroscience, 4(4),* 423-428.

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral cortex, 22*(8), 1894-1903.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503-507.

Morishima, Y., Schunk, D., Bruhin, A., Ruff, Christian C., & Fehr, E. (2012). Linking
Brain Structure and Activation in Temporoparietal Junction to Explain the
Neurobiology of Human Altruism. *Neuron, 75*(1), 73-79. doi:
http://dx.doi.org/10.1016/j.neuron.2012.05.021

Rachlin, H., & Jones, B. A. (2008a). Altruism among relatives and non-relatives.
Behavioural processes, 79(2), 120-123.

Rachlin, H., & Jones, B. A. (2008b). Social discounting and delay discounting. *Journal of
Behavioral Decision Making, 21*(1), 29-43.

Rushworth, M. F., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-
making in frontal cortex: one or many serial or parallel systems? *Current opinion
in neurobiology, 22*(6), 946-955.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The
neural basis of economic decision-making in the ultimatum game. *Science,
300*(5626), 1755-1758.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the
temporo-parietal junction in "theory of mind". *Neuroimage, 19*(4), 1835-1842.

Saxe, R., Moran, J., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping
brain regions for theory of mind and self-reflection in individual subjects. *Social
Cognitive and Affective Neuroscience, 1*(3), 229-234.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating
theory of mind: A meta-analysis of functional brain imaging studies.
Neuroscience & Biobehavioral Reviews, 42, 9-34.

Smith, D. V., Clithero, J. A., Boltuck, S. E., & Huettel, S. A. (2014). Functional
connectivity with ventromedial prefrontal cortex reflects subjective value for
social rewards. *Social cognitive and affective neuroscience, 9*(12), 2017-2025.

Strombach, T., Jin, J., Weber, B., Kenning, P., Shen, Q., Ma, Q., & Kalenscher, T. (2014).
Charity begins at home: cultural differences in social discounting and
generosity. *Journal of Behavioral Decision Making, 27*(3), 235-245.

Supporting Information

Materials and Methods.

Additional analyses. In addition to those presented in the main text, we performed further analyses to explore alternative explanations of TPJ functionality. In particular, we aimed to elaborate on the possibility that the TPJ represents the value attached to increasing another person's well-being.

These analyses were variants of (1) the GLMs described in the main text (decision onset split according to generous and selfish decisions), and (2) an additional GLM in which all decision onsets were collapsed into one vector, thus not conditioned by choice.

For the GLM presented in the main text, we used the following parametric modulations for the onset regressor of generous decisions to model other-regarding utilities:

1. the amounts foregone at the indifference point
2. the ORU minus the selfish reward magnitude, thus the relative value of the chosen, generous option
3. the inverse of the social distance as absolute numerical value
4. z-scores of the social distance

For GLM (2) in which all decisions were collapsed into one vector, we used the following parametric modulations:

5. the ORU in all trials, independent of whether the decision was generous or selfish
6. the ORU independent of the actual decision as a first parametric modulator, selfish reward magnitude as a second parametric modulator
7. the selfish reward magnitude as a first parametric modulator, ORU (independent of the actual decision) as a second parametric modulator
8. the inverse of the social distance as absolute numerical value
9. the value of the chosen minus the value of the unchosen option, independent of whether the decision was generous or selfish

None of these contrasts revealed significant activation in the TPJ, even at very liberal thresholds ($p < 0.01$, uncorrected).

Onset of the second option. All information necessary to make a decision was already available after presentation of the second option (generous or selfish alternative; see figure 1 in main text for details). Thus, it is possible that our participants already integrated all decision-relevant information at this time point, and formed their decision before it was revealed at decision onset. To test this possibility, we modulated the BOLD response at the onset of the second option to estimate the neural response to social discounting in addition to the GLM reported in the main text in which we used the time point of the revealed decision,. Since the order of the options was randomized, either the generous or the selfish option was presented second. We estimated the following GLM, consisting of four onset regressors: 1) onset of the 2nd option, generous decision, 2) onset of the 2nd option, selfish decision, 3) onset of a generous decision, and 4) onset of a selfish decision. For all four regressors, we included the temptation to be selfish (own-reward value minus ORU) as parametric modulators. We were especially interested in comparing the parametric modulators for onset regressors 1) and 3) as well as 2) and 4) in order to examine whether we can identify similar activation patterns at the two distinct time points.

A conjunction analysis revealed a significant overlap between the parametric modulation of the temptation to be selfish, given a generous decision at the time of presentation of the second option (onset regressor 1) and at the time of the revealed decision (onset regressor 3; fig. S1). We find similar patterns also for the temptation to be selfish, given a selfish decision during both onsets.

Thus, the data suggest that similar neural networks were active at onset of the second option and decision onset. However, as it is difficult to isolate decision making processes from potentially choice-irrelevant sensory or computational processes related to the presentation of the second option, and since the consideration of all relevant choice-information is only ultimately evident when a decision is revealed we decided to report neural activations at decision onset in the main manuscript.

Reaction times. To control for a possible effect of reaction times on BOLD responses, we performed additional analyses on behavioral and neural data. First, a paired sample t-test was carried out to compare the reaction times (time between decision prompt and decision onset) of generous (M=0.726, SD=0.226) and selfish decisions (M=0.731, SD=0.250). The test did not indicate a significant difference in RT between the two types of decisions (t(22)=-0.176, p=0.862; fig. S2).

Second, to explore the possibility that reaction times were correlated with BOLD responses, we included individual reaction times at decision onset as additional parametric modulators in the first GLM reported in the main text. Our aim was to investigate whether activity in the VMPFC might be influenced by variability in reaction times, as VMPFC has been shown to be sensitive to response speed (Büchel, Holmes, Rees, & Friston, 1998; Frackowiak et al., 2004; Grinband, Wager, Lindquist, Ferrera, & Hirsch, 2008). Thus, we added two parametric modulators to the GLM in the main text for the regressor for the onset of the selfish decisions: 1) the magnitude of the selfish reward and 2) the *reaction time*.

Using the same VMPFC ROI as in the manuscript (Clithero & Rangel, 2013), in the correlation between VMPFC-BOLD signal and the selfish reward magnitude at decision onset remained significant (-6, 41, 5; t(22)=3.09, p=0.034) even after controlling for reaction times. Thus, it is unlikely that differences in RTs accounted for VMPFC activity, and conclude that VMPFC activity genuinely reflected the selfish utility.

High vs. low other-regarding trials. In our experiment, we included generous trial types with high (€75 own-reward / €75 other-reward) and low (€75 own-reward and €25 other-reward) other-reward alternatives. In our main analyses, we pooled across high- and low other-reward trial types for the following reasons:
First, we found no significant differences in neural activation between the low and high other-reward trials anywhere in the brain, including VMPFC and TPJ. More specifically, we used an additional GLM to compare high and low other-reward trials. This GLM contained the following onset regressors: 1) onset of the decision, given a high other-reward trial and 2) onset of the decision, given a low other-reward trial. With an initial threshold of p<0.005, k ≥10 voxel, the contrast between the high and low other-reward regressors revealed no significantly activated voxel anywhere. Also when using a mask of both the parietal part of rTPJ and the whole rTPJ (Mars et al., 2012), we found no significant small volume corrected differences in activation (all p>0.05, FWE SV corrected).
Second, we found no differences in activation between the high and low other-reward trials in valuation networks, including the VMPFC: using a 6mm sphere around the VMPFC-ROI identified in a meta-analysis (-2, 40, -4; Clithero & Rangel, 2013), that was

also used in the main analysis, revealed no significant activation after small volume correction (p>0.05, FWE SV corrected).

Third, in a further analysis, we investigated whether TPJ, in particular, was differentially activated in high and low other-reward trials given a generous decision. To this end, we calculated an additional GLM with one onset regressor for low other-reward trials, given a generous decision and one onset regressor for high other-reward trials, given a generous decision. Neither a whole-brain analysis (p>0.005, uncorrected) nor a small volume correction using the mask from Mars et al. (2012) revealed any significant difference between high and low other-regarding trials in the TPJ.

Because these analyses revealed no meaningful difference in neural activity between high and low other-reward trials, and in order to preserve statistical power, we pooled across high- and low other-reward trial types in our main analyses.

Social distance or general metric of magnitude? It is possible that the participants may not have been considering social distance during the task but rather a more general metric of magnitude or distance. However, we think it is unlikely that our participants disregarded the social nature of the distance scale. This is evident from the participants' behavior. Our results showed that participants were always less generous towards recipients at remote compared to close social distances. Our participants would not have shown this decrease in generosity across social distance had they disregarded social distance. It is important to highlight that the experiment was carried out in an incentive compatible manner and that subjects were asked to indicate representatives for several social distances (including name and their relationship to that person). Therefore, we assume that subjects were aware of the fact that their decisions could affect the well-being of another person. Moreover, if the TPJ was only processing a general numeric metric of social distance, independent of its social significance, TPJ activation should have been unmodulated by the type of decision (selfish/generous), but this is not what we found. Our results showed that TPJ activity was higher for generous vs. selfish decisions (see main text), although the numerical metric of social distance was identical in both generous and selfish decisions.

References.

Büchel, C., Holmes, A., Rees, G., & Friston, K. (1998). Characterizing stimulus–response functions using nonlinear regressors in parametric fMRI experiments. *NeuroImage, 8*(2), 140-148.

Clithero, J. A., & Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. *Social cognitive and affective neuroscience*, nst106.

Frackowiak, R. S., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., . . . Penny, W. D. (2004). *Human brain function*: Academic Press.

Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P., & Hirsch, J. (2008). Detection of time-varying signals in event-related fMRI designs. *NeuroImage, 43*(3), 509-520. doi: http://dx.doi.org/10.1016/j.neuroimage.2008.07.065

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral Cortex*, 22(8), 1894-1903.

Figures.
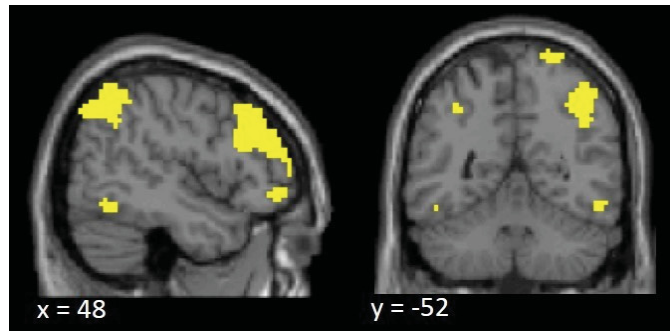
Figure S1: Conjunction analysis of the parametric modulation of the temptation to be selfish after a generous decision and after presentation of the second option in that trial, given a generous decision. Overlap is seen especially within the boundaries of the TPJ (48, -52, 37).
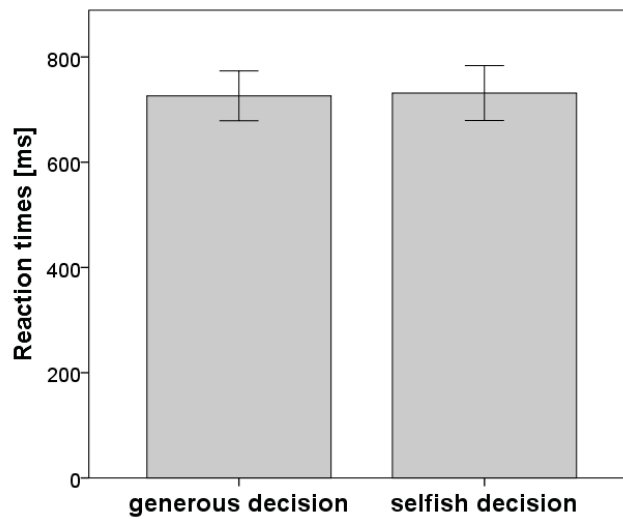
Figure S2: Mean reaction times, split by decision. Mean reaction times were calculated first on an individual level and then summarized on a group level. Error bars indicate +/- 1 standard error.

Figure S3: Parametric modulation of activity by selfish reward magnitude after selfish decisions in the VMPFC (and the ventral striatum).

Tables.

Table S1. GLM1. Regions showing parametric modulation by the selfish reward magnitude during selfish decisions.

Region	Side	Cluster size	MNI Coordinates	t-Value
Inferior frontal gyrus	R	87	27, 8, -17	4.91
Brainstem, Cerebellum	R	22	-9, -28, -14	3.71
Middle temporal gyrus	R	23	57, -7, -14	3.37
Hippocampus	R	21	30, -25, -11	3.67
Anterior cingulate	R	28	9, 29, -5	3.55
Anterior cingulate	L	47	-15, 44, 4	4.03
Insula	R	17	42, -16, 7	3.69
Superior temporal gyrus	R	13	57, -61, 19	3.26
Medial frontal gyrus	R	66	15, 41, 22	4.43
Supramarginal gyrus	R	43	69, -40, 37	3.88
Supramarginal gyrus	L	12	-63, -43 40	3.33

Threshold: $k \geq 10$ voxel, $t_{(22)} > 2.8188$, $p < 0.005$

Table S2. GLM1. Regions showing more activity for generous versus selfish decisions at the onset of the decision (button press).

Region	Side	Cluster size	MNI Coordinates	t-Value
Positive				
Cerebellum	R	97	36, -58, -32	4.22
Occipital Lobe	L	219	-12, -103, -5	5.48
Inferior frontal cortex	R	241	45, 44, -17	4.64
Medial frontal cortex	R/L	28	0, 47, -20	4.21
Occipital Lobe, Cuneus	R/L	390	24, -100, 13	4.67
Fusiform Gyrus	R	14	51, -61, -14	3.24
Middle frontal gyrus	L	161	-39, 53, 1	4.25
Inferior frontal gyrus	L	49	-33, 23, -2	3.45
Insula	R	17	42, -7, 7	3.42
Caudate	R	21	12, 8, 13	3.58
Middle frontal gyrus	R	113	36, 35, 16	4.22
Anterior cingulate	L	25	-6, 44, 13	3.28
Superior frontal gyrus	L	282	-15, 59, 22	5.44
Middle frontal gyrus	R	116	30, 5, 43	4.41
Limbic lobe, cingulate cortex	R/L	431	3, -10, 28	5.87
Anterior cingulate	R	40	6, 47, 22	4.04
Inferior parietal lobule, angular gyrus	R	690	60, -58, 31	5.11
Supramarginal gyrus	R	16	48, -19, 28	3.05
Inferior parietal lobe	L	486	-24, -79, 52	4.49
Middle frontal gyrus	L	328	-42, 20, 43	4.71
	R	15	39, 20, 37	3.29
Superior frontal gyrus	R	21	18, 32, 52	3.31
Precentral Gyrus	R	26	18, -25, 61	4.04
Negative				
Temporal lobe	R	24	24, -46, 10	3.92

Threshold: $k \geq 10$ voxel, $t_{(22)} > 2.8188$, $p < 0.005$

Table S3. GLM1. Regions showing parametric modulation by the econometrically reconstructed ORU during generous decisions.

Region	Side	Cluster size	MNI Coordinates	t-Value
Rolandic operculum	L	69	-42, -19, 19	4.65
Precentral gyrus	R	21	15, -19, 64	3.80
Precentral gyrus	L	10	-18, -19, 67	3.38
Postcentral lobe	R	10	24, -31, 82	3.56

Threshold: k ≥ 10 voxel, $t_{(22)} > 2.8188$, p < 0.005

Table S4. GLM2. Regions showing parametric modulation by the difference between own-reward value and ORU (i.e. the temptation to choose the selfish option) at decision onset, given a generous choice.

Region	Side	Cluster size	MNI Coordinates	t-Value
Cerebellum	R	10	27, -67, -32	4.16
Superior temporal gyrus	L	10	-33, 20, -29	3.53
Brainstem, Cerebellum	R	18	6, -31, -23	4.33
Occipital lobe	L	186	-48, -64, -14	4.78
Inferior frontal gyrus	R	185	30, 20, -14	5.19
Inferior temporal lobe	R	34	51, -58, -14	3.73
Occipital lobe	R	37	30, -88, -5	3.52
Middle frontal gyrus	R	475	51, 41, 25	5.65
Middle orbito-frontal lobe	L	15	-48, 47 -5	3.55
Medial frontal gyrus	R/L	924	9, 23, 34	5.38
Occipital lobe	L	58	-33, -94, 16	4.50
	R	13	33, -85 10	3.16
Posterior cingulate	R	24	9, -61, 13	3.43
Middle frontal gyrus	L	37	-21, 62, 25	3.44
Superior medial frontal	R	10	12, 68, 22	3.12
Middle frontal gyrus	L	117	-51, 17, 34	4.37
Inferior parietal lobule, angular gyrus	R	309	42, -79, 46	5.98
			42, .49, 46	4.23
			54, -61, 43	4.20
Precentral	L	14	-42, 2, 43	4.01
Inferior parietal lobule	L	32	-36, -49, 43	3.57

Threshold: k ≥ 10 voxel, $t_{(22)} > 2.8188$, p < 0.005

Table S5. GLM2. Regions showing stronger parametric modulation by the difference in own-reward value and ORU during generous than selfish decisions.

Region	Side	Cluster size	MNI Coordinates	t-Value
Positive				
Cerebellum	R	19	36, -58, -41	3.37
	L	19	-6, -85, -32	3.25
Occipital lobe	L	677	-33, -97, 16	5.13
Inferior temporal lobe	R	103	54, -43, -14	4.77
Inferior frontal gyrus	R	244	33, 26, -5	6.26
	L	117	-33, 20, 4	5.28
Middle occipital gyrus	R	319	36, -91, -8	4.76
Superior frontal gyrus	R/L	2190	51, 44, 25	6.98
Occipital lobe	R	20	36, -70, -8	4.15
Cuneus	L	14	-12, -76, 10	3.26
Superior medial frontal	L	16	-6, 50, 19	3.21
Middle frontal gyrus	L	345	-51, 23, 34	5.26
Precuneus	R	114	6, -67, 40	3.53
Parietal lobe, angular gyrus	R	547	42, -79, 46	7.09
Parietal lobe	L	167	-33, -52, 43	5.01
Negative				
Temporal lobe	R	83	36, -37, -8	3.96
Frontal lobe	L	372	-21, -4, 28	5.03
Frontal lobe	R	72	21, 35, -5	4.37
Temporal lobe	L	23	-30, -49, 1	3.92
Inferior parietal lobule	L	420	-57, -34, 22	4.91
Rolandic operculum	R	15	66, 8, 4	3.85
Supra marginal gyrus	R	318	63, -28, 31	4.50
Cerebrum	R	172	18, -1, 28	4.06
Medial frontal gyrus	R	110	9, -13, 61	4.13
Parietal lobe	R	56	18, -34, 49	3.63
Postcentral	R	313	24, -52, 76	5.75
Postcentral	L	111	-21, -49, 76	3.77

Threshold: $k \geq 10$ voxel, $t_{(22)} > 2.8188$, $p < 0.005$

Table S6. PPI. Regions showing stronger task relevant functional connectivity with the right
TPJ during generous than selfish decisions.

Region	Side	Cluster size	MNI Coordinates	t-Value
Cerebellum	L	15	-18, -40, -44	3.89
	L	16	-27, -34, -32	4.09
Superior temporal gyrus	L	22	-51, 17, -23	3.87
Middle temporal gyrus	R	191	57, 11, -26	4.71
Inferior frontal gyrus	L	34	-33, 20, -20	3.58
Anterior cingulate	R/L	293	-6, 32, -2	6.60
Insula	L	1522	-36, -16, 19	5.90
	R	226	39, 11, 7	4.91
Superior temporal gyrus	R	37	72, -34, 19	4.24
Supramarginal gyrus	R	21	51, -52, 25	3.85
Cingulate gyrus	L	10	-15, -22, 25	3.89
Precuneus	L	18	-9, -55, 34	4.03
Inferior parietal lobule	R	315	69, -37, 37	7.42
Middle frontal gyrus	L	162	-27, 29, 43	4.80
Medial frontal gyrus	R	139	9, -22, 55	4.52
Precentral gyrus	L	108	-42, -22, 49	4.05

Threshold: $k \geq 10$ voxel, $t_{(14)} > 2.8188$, $p < 0.005$

3. Study 3: Gender-specific effects of cognitive load on social discounting

Strombach, T.[a,b*], GorczycaB. [a], & Kalenscher T. [a]

[a] Comparative Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[b] Chair of Marketing, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[*] Corresponding author (tina.strombach@hhu.de)

Abstract

We live busy, social lives, and meeting the challenges of our complex environments puts strain on our cognitive systems. However, cognitive resources are limited. It is unclear how cognitive load affects social decision making. Previous findings on the effects of cognitive load on other-regarding preferences have been ambiguous, allowing no coherent opinion whether cognitive load increases, decreases or does not affect prosocial considerations. Here, we suggest that social distance between individuals modulates whether generosity towards a recipient increases or decreases under cognitive load conditions. Participants played a financial social discounting task with several recipients at variable social distance levels. In this task, they could choose between a generous alternative, yielding a medium financial reward for the participant and a recipient on a given social distance, or a larger reward for the participant alone. We show that the social discount function was significantly flattened under high cognitive load conditions, suggesting that participants became less generous towards recipients at close social distance, but more generous towards socially distant recipients, e.g., strangers. Unexpectedly, the cognitive-load effect on social discounting was gender-specific: while social discounting was strongly dependent on cognitive load in men, women were nearly unaffected by cognitive load manipulations. We interpret these results within a dual-process framework and suggest that cognitive load leads men, but not women, to switch from deliberate processing to more automatic decision making, neglecting the social distance information. Our study illustrates the importance of considering social distances as well as gender in research on prosocial choice.

1. Introduction

In today's societies, we have to accomplish a multitude of different tasks in parallel. We are on the phone with a business partner while simultaneously scanning through the headlines of the newspaper, drinking our coffee and keeping upcoming duties in mind, like trying to get our kids ready for school – all at the same time. In this state of mind, we have to make decisions that might affect our own future, but very often, our choices impact also other people in our social environment. The question arising is how we deal with the strain of cognitive load, given that our cognitive resources have limited capacity (Sweller, 1994), to decide and act efficiently. And how does our behavior change when the cognitive capacity is exceeded? Here, we aim at investigating the role of depleting cognitive resources on social decision making (Strombach et al., 2014, 2015).

Most studies on the effects of limited cognitive resources on behavior investigate decision making in an isolated social environment. For instance, it has been shown that cognitive load manipulations alter learning (Sweller, 1994), aviation (Wilson, 2002) and user interface design (Saadé & Otrakji, 2007). However, in real life, many, if not most of our decisions are influenced by the consideration of the well-being of other people. But, unfortunately, the role of cognitive load and depleting cognitive resources on social decisions is much less understood, and previous research found ambiguous results on the role of cognitive load on social preferences (Benjamin, Brown, & Shapiro, 2006; Cappelletti, Güth, & Ploner, 2011; Hauge, Brekke, Johansson, Johansson-Stenman, & Svedsäter, 2009; Roch, Lane, Samuelson, Allison, & Dent, 2000; Schulz, Fischbacher, Thöni, & Utikal, 2014; Shiv & Fedorikhin, 1999): while some studies on prosocial choice behavior, in which subjects make decisions affecting the payoff of other participants, report that subjects became more generous towards their interaction partner under cognitive load conditions (Roch et al., 2000; Schulz et al., 2014), others demonstrated increased selfishness (Crelley, Lea, & Fischer, 2008; Moore & Loewenstein, 2004), yet most publications do not report any effect of cognitive load on social decision making (Benjamin et al., 2006; Cappelletti et al., 2011; Cornelissen, Dewitte, & Warlop, 2011; Hauge et al., 2009). Next to procedural differences, the social choice designs used in these studies differed in the degree of familiarity, or social closeness, between the participants and their interaction partners; e.g., in some studies, the interaction partner was anonymous, whereas in others, he/she was introduced to the participants, and yet in other studies, the interaction partners were actual acquaintances. Interestingly, whether the interaction partner was anonymous or not appeared to determine the sign of the cognitive load effects on generosity: when the

interaction partner was an anonymous stranger, subjects often became more generous under cognitive load conditions (Roch et al., 2000; Schulz et al., 2014), but when the interaction partner was familiar to the subject, e.g. the interaction partner was a co-student or introduced before, he/she became more selfish (Crelley et al., 2008; Moore & Loewenstein, 2004). When familiarity was not well controlled, any putative effects of cognitive load on social preferences might have been obscured by the uncontrolled variability in anonymity. We therefore hypothesize that the ambiguity in previous results might be evoked by differences in the degree of social distance between participants and their interaction partners. This hypothesis blends in with recent theories on prosociality and cognitive control. These theories suggest that prosocial behavior requires self-control to resolve the conflict between selfish and other-regarding motives (Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; Strombach et al., 2015). Because evidence suggests that self-control capacities become exhausted with increasing cognitive load (Baumeister, 1984; Baumeister, Bratslavsky, Muraven, & Tice, 1998; Baumeister & Showers, 1986; but see Carter, Kofler, Foster, & McCullough, in press), putting strain on the cognitive control system is therefore likely to change social-distance-dependent generosity profiles.

In the present study, we investigate the effect of cognitive load on prosocial decision making with socially close, socially distant and strange interaction partners. We systematically vary social distance using a social discounting paradigm (Jones & Rachlin, 2006; Strombach et al., 2014; Strombach et al., 2015). Social discounting refers to the idea that generosity towards others diminishes systematically over social distance between donor and recipient, with social distance indicating how much, or how little, the donor cares about the recipient.

We hypothesize that cognitive load affects prosocial, other-regarding decision making, and that this effect is modulated by the social distance between donor and recipient. More specifically, we expect that, under high cognitive load conditions, individuals become less generous towards people at closer social distance, but more generous towards people at large social distance. We use a psychometric approach to address this question. In a financially incentivized social discounting task, participants indicated their level of generosity towards recipients at variable social distances. We fitted a hyperbolic social discount function (Jones & Rachlin, 2006; Strombach et al., 2014; Strombach et al., 2015) to our participants' choice data to mathematically capture their social discounting behavior. We expected that cognitive load flattened the social discount function, reflecting the hypothesized social-distance-dependent cognitive-load effects on generosity.

The present study has important implications for theories of social decision making, but will also inform research on business settings. Insights into the effect of cognitive load on social preferences might lead to a better understanding why high workload, stress and other states characterized by cognitive preoccupation often also result in interpersonal distress. Moreover, a multitasking environment is common in the business world. Understanding the impact of cognitive load on decision making and other-regarding behavior might help to create working environments that are more productive and less prone to exhaustion, occupational stress and work-related depression.

2. Material and Methods

2.1 Participants

89 Participants (34 male, M_{age}=23.09, SD_{age}=2.69, Range$_{age}$: 18-30) were recruited at the University of Düsseldorf. The participants were randomly assigned to either the control condition (low cognitive load ; N=44, 17 men) , or the experimental condition (high cognitive load; N=45, 17 men). Participants who had previously participated in psychological experiments as well as students enrolled in Psychology or Economics were excluded from participation. One subject stopped the experiment during the procedure and was excluded from the analysis. Written consent was obtained before the experiment started. The study was approved by the local ethics committee and conformed to the guidelines of the Declaration of Helsinki. Participants were financially compensated for their participation, as outlined below. The study was fully incentive-compatible, did not involve deception and thus met the standards in psychology and behavioral economic research.

2.2 General experimental procedure

Participants were randomly assigned to a high (experimental) and low (control) cognitive load manipulation. Before the start of the main experiment, a brief self-control scale was administered (see below for details). After completing the scale, participants received all instructions and information about the procedure.

Subsequently, they were subjected to the cognitive load manipulation (high versus low cognitive load) and then performed the social discounting task (see below for details on the cognitive load manipulation and social discounting task). They were debriefed after they finished the experiment.

2.2.1 Cognitive load manipulation

Cognitive load was elicited using two different treatments. Both treatments were similar to previous tasks used in studies on self-regulation and ego-depletion (Baumeister et al., 1998; Fennis, 2011; Wheeler, Briñol, & Hermann, 2007). For the stimulus-detection task, all participants received typewritten sheets of paper with a text extracted from an advanced machine learning book (Eifler, 2009). In the control condition, participants were asked to cross off every *e* they found in the text. In the experimental condition, instructions were similar, but participants received an extra set of rules as follows: they were instructed to cross off every *e*, except if the *e* was followed by a vowel and except if the *e* was the beginning letter of a word. However, when the word that began with an *e* was at the beginning of the sentence, the *e* had to be crossed off. In an unstructured interview after the procedure, participants in the experimental group indicated more often than participants in the control group that the procedure was exhausting.

In addition to the instructions, the control and the experimental conditions also differed with regard to font size and font transparency of the text (Control: font size: 14, brightness: 0%, experimental condition: font size: 9, brightness: 75%). This made it more difficult and effortful for the participants in the experimental condition to read the text. Both groups had five minutes to work on the task.

The stimulus-detection task was followed by a computer-based stroop task (Goto & Kusumi, 2013; Gwizdka, 2010; Inzlicht & Gutsell, 2007; Soutschek, Strobach, & Schubert, 2013; Stroop, 1935), programmed in Presentation (Neurobehavioral Systems). The stroop task has often been validated to induce cognitive load (Gwizdka, 2010; Soutschek et al., 2013; Stroop, 1935). Subjects in the experimental group saw a color word displayed with differently colored fonts on a computer screen (six different colors). They were asked to indicate, by pressing a corresponding button on a keyboard, either the semantic meaning of the displayed word, or the name of the color of the font used to display the word, respectively. The meaning of the presented word

was sometimes incongruent with the color of the font used. For example, the word 'blue' may have been presented in green fonts. Congruent (font color and semantic meaning are identical) and incongruent (font color and semantic meaning are dissimilar) trials were presented in 12 blocks (subjects had to indicate the words' semantic meanings in six blocks, and their font color name in six other blocks) with 24 trials in random order. In order to perform this task, subjects in the experimental condition had to suppress the automatism to read the semantic meaning of the word. In the high cognitive load condition, incongruent trials were presented in 50% of the trials. The incongruence between font-color and word-meaning was not present in the control group, where subjects were always asked to indicate the semantic meaning of the color-word, independent of the font-color used. To simplify the task even more, the words were always presented in grey fonts to avoid incongruences. Performance in the control condition therefore required less suppression of the automatism to read out the word instead of indicating the color. In both groups, words were presented on a white screen. Inter-stimulus intervals had duration of 500ms in which a fixation cross was presented. Maximum response time was limited to 5000ms.

After the stimulus detection task and the Stroop task were completed, the experiment continued with the social discounting task (Strombach et al., 2014; Strombach et al., 2015).

2.2.2 Social discounting task

Social discounting was measured with the same paradigm used in our previous studies on social discounting (Strombach et al., 2014, 2015; see Strombach, 2014 for a discussion of the elicitation procedure of social distance). To introduce the concept of social distance, each participant was shown a scale consisting of 101 icons, with the leftmost icon representing the participant and the others representing his social environment. Participants were told that social distance 1 (the most leftward icon closest to the participant) represents the socially closest person, while distance 100 (the most rightward icon) would be a stranger who they may have randomly met on the street. Social distance 50 stands for a distant acquaintance, whose name they may not know. Once participants were familiar with the concept of social distance, they were asked to write down the names of representatives for the following social distances: 1, 2, 3, 5, 10, 20. Although distances 50 and 100 were also included in the experiment, participants could, but were not required to provide a name, as these distance levels often represent

remote individuals. Participants were specifically asked not to include anyone in their list against whom they have negative feelings.

In each trial, a yellow icon on the social distance scale, indicating the social distance of the recipient, was shown to the participants. To avoid perceptual issues with the visual representation of social distance, the social distance information was additionally indicated by a number on top of the yellow icon (cf. figure 1). Participants had to choose between a selfish option, yielding a large reward for themselves, and a generous option, yielding a smaller reward for the them and the same amount for the recipient on the indicated social distance. The selfish reward varied between €75 and €165, with increments of €10. The generous option was identical in all trials, yielding €75 for the participant and €75 for the recipient on the specific social distance. For example, in a given trial, a subject may choose between a €125 reward only for herself (selfish option), or a €75 reward for herself and a €75 reward for a recipient on social distance 20 (generous option). In total, the participants made 160 decisions – 8 social distances, 10 selfish rewards and all combinations were presented twice. The order of trials, as well as the side of the presentation of the selfish and generous choice alternatives, was fully randomized (cf. fig. 1).

Participants were informed during the instructions before the experiment that, at the end of the task, one of their decisions would be randomly chosen and 10% of its payoff would be paid out, therefore they and potentially another person would be able to earn money based on their decisions. The money the participant allocated to herself was paid out directly after the experiment, and for the money shared, subjects were asked to indicate the address of the other person in the randomly chosen trial. If the randomly chosen trial was about an anonymous person or stranger, e.g. at higher social distances, a random person on the campus of the University of Düsseldorf, Germany received the reward.
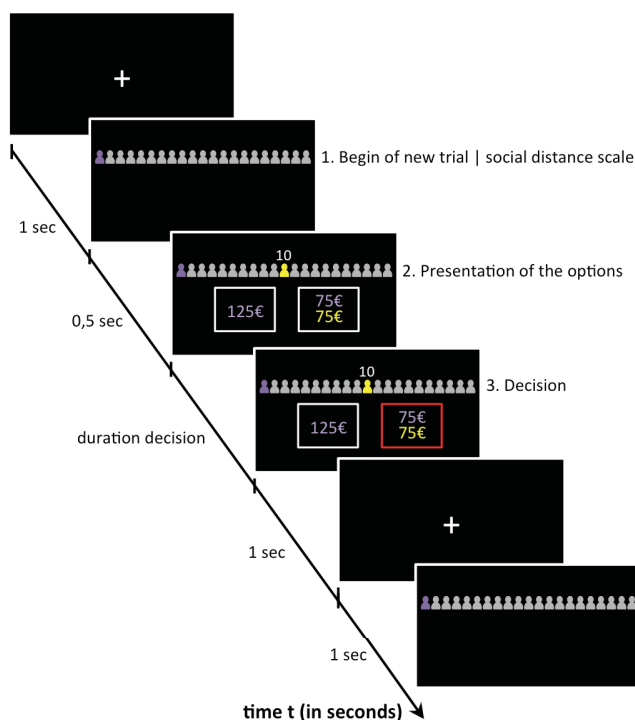
Figure 1: An exemplar trial in the social discounting task. Social distance information was given on top of the screen, the options below. The subject could choose between a selfish reward just for herself or a generous option, yielding a reward for herself and another person on the indicated social distance. The side of presentation of the options was randomized. As soon as both options were presented, the participant could make her decision. The final choice was fed back by a red box around the chosen option. Note that this figure has been adjusted for illustration purposes; stimulus size and screen format are not to scale with the presentation dimensions used testing. In addition, the figure displays only 21 icons, instead of 101 icons shown during scanning, to facilitate perceptibility.

2.2.3 Brief self control scale

Before the experimental procedure started, self-control was measured with a German translation of the Brief Self-Control Scale (BSCS; Tangney, Baumeister, & Boone, 2004; available at www.uni-konstanz.de/diagnostik/research_measures.htm). The BSCS consists of 13 items indicating agreement with given statements on a 5-point-Likert scale to quantify a subjective measure of self-control. The scale was intended as a control for potential inhomogeneous group differences.

2.3 Data analysis

The analysis procedure was identical to procedures used previously (Jones & Rachlin, 2006; Strombach et al., 2014; Strombach et al., 2015). First, we determined, for each subject and each social distance, the point at which a subject was indifferent between the selfish and the generous option. To this end, logistic regression was used to identify the point at which the probability of answering generously or selfishly was 50%. For each indifference point, we calculated the individual amount foregone, i.e., the difference between own-reward of the selfish alternative and own-reward of the generous alternative. For example, if a participant was indifferent between receiving €125 just for herself and €75 for herself and €75 for a recipient at a specific social distance, she was willing to forego €50 to increase the wealth of the recipient by €75. The individual amount foregone at a given social distance level measures how much it was worth to the participant to endow the recipient with €75; it can therefore be construed as a social premium a subject was willing to pay to improve the recipient's wealth. This social premium served as a social-distance-dependent estimate of the other-regarding value a subject attaches to increasing a recipient's wealth by €75. If the participant made exclusively selfish or generous choices at a given social distance, indifference points were determined to be €170 or €70, respectively (for more details on the estimation of the discount curve see Strombach, 2014). After determining the individual amounts foregone for each social distance played, the following standard hyperbolic model was fitted to the social-distance-dependent social premiums (Jones & Rachlin, 2006; Strombach et al., 2014; Strombach et al., 2015):

$$v = \frac{V}{(1+kD)} \qquad\qquad \text{(equation 1)}$$

where *V* symbolizes the magnitude of a reward received by a recipient at social distance *D*. The parameter *v* refers to the amount foregone, i.e., the social premium a subject is willing to pay in exchange for endowing the recipient on social distance *D* with reward *v*. Thus, *v* can be interpreted as a proxy of the socially discounted other-regarding value of improving the wealth of another individual at social distance *D*. *V* is the intercept with the y-axis, and determines the height of the social discount function. Thus, *V* can be interpreted as the level of generosity towards socially close recipients. The degree of discounting is described by the parameter *k*, which indicates the steepness and shape of the curve.

4. Results

The goal of the current experiment was to identify the effect of cognitive load on social discounting, i.e. on social distance-dependent generosity. We hypothesized that a higher cognitive load alters the social discount function. More specifically, we expected flatter social discounting after high cognitive load manipulations.

4.1 Cognitive Load Manipulation

There was no significant difference in the self-control scores of the BSCS between participants of the control and the experimental conditions (BSCS: $M_{control}$= 3.13, $SD_{control}$ = 0.60; $M_{CognitveLoad}$ = 3.17, $SD_{Cognitive\ Load}$ = 0.59; t-test: $t(85)$ = 0.282, p = 0.779, η_p^2 = 0.001).

As manipulation check for the cognitive load manipulation, we compared the reaction times in the Stroop task between experimental and control group. We assumed that higher cognitive load would go along with longer reaction times (Gwizdka, 2010; Schulz et al., 2014). We found a significant difference between experimental and control subjects in reaction times ($M_{control}$= 792.50 ms, $SD_{control}$ = 132.78; $M_{CognitveLoad}$= 862.71 ms, $SD_{Cognitive\ Load}$ = 155.00; $t(85)$ = -2.27; p = 0.026, η_p^2 = 0.057) and errors made in the Stroop task ($M_{control}$ = 4.34, $SD_{control}$ = 5.26; $M_{CognitveLoad}$ = 10.79, $SD_{Cognitive\ Load}$ = 8.94; $t(85)$=67.68; p < 0.001, η_p^2 = 0.166), supporting our assumption that cognitive load was higher in the experimental Stroop condition compared to the control condition. We additionally tester whether performance in the Stroop task differed between men and women. However, a mixed 2x2 ANOVA with the factors group and gender revealed no significant main effect of gender, or interaction effects between gender and condition, for reaction times (all p>0.40). For the e-crossing task we checked whether subjects in the experimental condition made less progress in identifying "*e*"s compared to subjects in the control condition because of differences in task difficulty and perceptibility of the text. To this end, we counted the letters that were processed by the participant until time-out. We found a significant difference in the total number of words processed between experimental and control subjects ($M_{control}$ = 1478.36, $SD_{control}$ = 466.37; $M_{CognitveLoad}$ = 2071.18, $SD_{Cognitive\ Load}$ = 436.40; $t(85)$ = 6.09; p < 0.001, η_{p^2} = 0.306). Again, there was no indication of gender main and interaction effects on performance in the e-crossing task (all p > 0.20). Thus, in line with others (Baumeister et al., 1998; Salmon, Adriaanse, De Vet, Fennis, & De Ridder, 2014; Vohs,

Baumeister, & Schmeichel, 2012) we assumed that the differences in complexity of the e-crossing task between experimental and control groups translated into differences in cognitive load.

4.2 Social Discounting

In both conditions, generosity levels, measured as the amount foregone at indifference points (the social premiums, see methods), decreased across social distance, replicating previous studies on social discounting (Hangebrauk et al., in press; Jones & Rachlin, 2006; Strombach et al., 2014; Strombach et al., 2015).

For each subject, we fitted the hyperbolic social discount model (eq. 1) to the individual amounts foregone. We used the best-fitting discount parameters V and $\log(k)$ to quantify and compare social discounting between experimental and control groups. k was log transformed to approximate a normal distribution. As stated earlier, V can be interpreted as the level of generosity towards socially close recipients, and k, or $\log(k)$ respectively indicates the steepness of the curve, thus how steeply generosity decays across social distance. Six subjects were excluded because they selected the identical options across all social distances, and therefore had a k-value of zero, which cannot be log-transformed. We hypothesized that cognitive load flattens the social discount function, reflecting the predicted decrease in generosity towards socially close recipients, and increase in generosity towards socially more distant people.

To test this hypothesis, we first compared the log-transformed k-values between the two experimental conditions. Contrary to our prediction, a t-test did not indicate a significant difference in $\log(k)$-values between the groups ($M_{control}$=1.617, $SD_{control}$=6.846; $M_{cognitive\ load}$=-3.305, $SD_{cognitive\ load}$=2.405; $t(79)$=1.458, p=0.149, η_p^2 = 0.026). It has been suggested that men and women are differently affected by cognitive load (Lighthall et al., 2012; Ptacek, Smith, & Dodge, 1994), and also differ in their social preferences (Olson, Rosso, Demers, Divatia, & Killgore, 2015). Thus, to further inspect our data, we ran additional analyses including gender as additional factor in a 2x2 analysis of variance (ANOVA) with cognitive load and gender as fixed factors. The ANOVA revealed a significant interaction effect between cognitive load and gender on the $\log(k)$-values ($F(77,1)$ = 6.390, p = 0.014, η_p^2 = 0.077; fig. 2, 3), while there was no significant main effect for gender ($F(77,1)$ = 0.383, p = 0.538, η_p^2 = 0.005). The lack of a main effect for gender may explain why also previous work did not report any gender

effects (Jones & Rachlin, 2006; Jones & Rachlin, 2009; Strombach et al., 2014; Strombach et al., 2015). However, the significant interaction between gender and cognitive load on the steepness of the social discount function seems to corroborate our hypothesis that cognitive load affects social preferences differently in men and women.

To further characterize the gender x cognitive load interaction, post-hoc analyses revealed that cognitive load effects on log(k) were most pronounced in men: while women did not show significant differences in log(k)-values between the cognitive load treatments ($M_{women\ \&\ control}$ = -2.95, $SD_{women\ \&\ control}$ = 2.08, $M_{women\ \&\ cognitive\ load}$ = -2.55, $SD_{women\ \&\ cognitive\ load}$ = 1.60, p > 0.1), men had significantly lower log(k)-values under high- compared to low-cognitive-load conditions. Furthermore, compared to log(k)-values of women, men's log(k)-values were significantly lower in the high cognitive load condition ($M_{men\ \&\ cognitive\ load}$ = -4.81, $SD_{men\ \&\ cognitive\ load}$ = 3.05, $F(37, 1)$ = 9.288, p = 0.004, η_p^2 = 0.201). Thus, while social discounting behavior in women seemed relatively unaffected by cognitive load, men showed flatter social discounting under high cognitive load (see fig. 2, 3).

Next, we repeated the fixed factor 2x2 ANOVA to investigate the effects of gender and cognitive load on the second parameter in the hyperbolic function V. The ANOVA did neither indicate significant main effects of cognitive load or gender on V, nor an interaction effect (all p > 0.1).
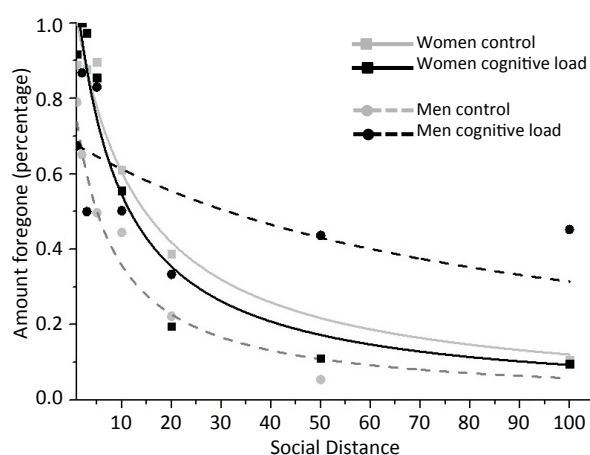
Figure 2: A standard hyperbolic model was fitted to the amounts foregone (social premiums). Social discount curves are presented separately for gender (men; women) and high and low cognitive load. While there was no significant difference in social discounting in women between high and low cognitive load conditions, men under high cognitive load were less sensitive towards variations in social distance, reflected by a considerably flatter discount curve.
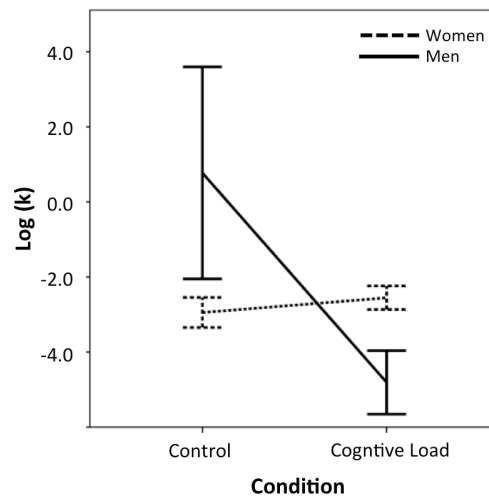
Figure 3: Interaction effect of gender and cognitive load on the averaged log transformed discount parameter k, reflecting the steepness of the social discount curve. Errorbars indicate standard errors of the mean.

In sum, our data suggest that cognitive load affected the steepness, but not necessarily the height, of the social discount function in a gender-specific way. However, there might be several alternative accounts for the interaction effect of cognitive load and gender on $\log(k)$ (Franco-Watkins, Pashler, & Rickard, 2006; Hinson, Jameson, & Whitney, 2003). For instance, it is possible that men, unlike women, simply made more noisy decisions under high- than under low-cognitive load conditions. To rule out this possibility, we performed additional analyses to establish putative effects of cognitive load on decision noise as follows: as outlined earlier, we used binary logistic regression to determine the individual indifference points between generous and selfish options for each social distance level. Decision noise as well as high choice variability and/or inconsistent decision patterns should be reflected in poorer goodness-of-fit measures of the individual logistic regressions. We therefore used the individual goodness-of-fit estimates R^2 as a measure of decision noise and choice inconsistency. In cases subjects always and invariantly selected the selfish over the generous option, or vice versa respectively, R^2 was set to one. To test whether men differed from women in the number of noisy or inconsistent decisions under high compared to low cognitive load, we calculated the mean adjusted R^2 across all indifference points for each subject, and used it as the dependent variable in a 2x2 ANOVA with gender (men/women) and level of cognitive load (high/low) as fixed factors. The ANOVA revealed no significant main or interaction effects of gender and cognitive load on adjusted R^2-values (all $p > 0.1$). This suggests that we have no

evidence to assume that the above reported gender-specific cognitive-load-effects on social discounting can be explained by cognitive-load-dependent decision noise and/or choice inconsistency.

Another potential explanation for the low $\log(k)$-values in men in the high cognitive load condition might be differential attention-effects: after high cognitive challenge, men may simply pay less attention to the task and just "click" through the experiment, randomly choosing any option, or perseverating on one response option. Subjects had to indicate their selection of the generous or selfish option by clicking on the left or right option presentation (cf. fig. 1). Since the side of the presentation of the options was randomized, perseveration on one side, or random clicking, would result, for each social distance level, in a medium indifference point estimate that would be similar for all social distances. Because of the fact that indifference point estimates would be similar across social distance, fitting a hyperbolic function to the choice data obtained from a perseverating or random "clicker" would yield flatter discounting, and consequently lower $\log(k)$-values. We assumed that such mindless, random or perseverating "clicking"-behavior should be reflected by reduced reaction times (Lighthall et al., 2012). To address the possibility that men became mindless "clickers" under high cognitive load conditions, we calculated mean reaction times per individual and ran another 2x2 ANOVA with gender and cognitive load as independent factors and mean reaction time as dependent variable. Again, we did not find any significant main or interaction effects (all $p > 0.1$). Thus, mindless random or perseverating clicking behavior is not likely to be the reason for above reported gender-specific cognitive-load-effects on social discounting.

Finally, we checked for correlations between the discount parameters $\log(k)$, V, the number of errors and reaction times in the Stroop task. None of the correlations reached significance (all $p > 0.1$).

5. Discussion

We investigated the effect of cognitive load on social discounting. We exposed participants to either a high or a low cognitive load manipulation in which they had to perform two tasks requiring cognitive control (the e-crossing and the Stroop task). They subsequently played a social discounting task in which they repeatedly decided between a large reward for themselves (the selfish option), or a smaller reward for

them plus an additional reward for recipients on variable social distances. We replicated previous findings on social discounting showing that a generosity metric – the willingness to forego a reward in exchange for increasing the wealth of the recipient (the social premium) - decreased over social distance. In line with our prediction, we found that high cognitive load flattened the social discount function, but, unexpectedly, this cognitive-load effect on social discounting was only found in men: under control conditions, men showed steeper social discounting than women, suggesting that their generosity towards others decreased very steeply with increasing social distance to the recipient. But following the cognitive load manipulation, men showed considerably flatter social discounting than women and under control conditions, indicating that, after increased cognitive load, the typical decrease in generosity across social distance was much less pronounced. By contrast, the discount function of female subjects was similar in the high- and low-cognitive-load conditions.

Gender differences in decision making and cognitive style have been reported before (Lighthall et al., 2012; Preston & de Waal, 2002; Van den Bos, Harteveld, & Stoop, 2009). For instance, a neuroimaging study by Lighthall and colleagues (2012) on reward-processing under stress identified opposing effects of stress on reward-related brain activation in men and women. They suggested that men and women use different strategies when cognitive resources became depleted due to the stress manipulation. While men switched to more automatic processing under stress (Lighthall et al., 2012; Porcelli & Delgado, 2009), women relied more on deliberate processing in stress- and no-stress situations, making use of greater explicit knowledge about the task contingencies (Lighthall et al., 2012; Preston & de Waal, 2002). This finding is corroborated by recent evidence for gender-specific effects of stress on performance in the IOWA gambling task (Van den Bos et al., 2009): while performance in men deteriorated after stress, presumably due to stress-related downregulation of prefrontal brain regions responsible for cognitive control and deliberative reasoning, stress effects on IOWA gambling performance in women was less unidirectional, and more complex overall. Taken together, these findings suggest that manipulations known to affect cognitive control and reasoning, such as stress as and cognitive load, change decision making patterns in men and women in opposite ways.

The idea that cognitive load effects on social discounting are the consequence of gender-specific exhaustion of cognitive control capacities blends in with recent dual process models of decision making (Chaiken & Trope, 1999; Evans, 2008; Johnson & Weber, 2009; Loewenstein & O'Donoghue, 2007; Schulz et al., 2014). These dual process

models postulate that decisions are based on the interplay between two complementary mental processes: while affective processes deal with emotional and automatic behavioral responses, cognitive processes are responsible for controlled, deliberated behavior. One important difference between the two processes is that controlled processes have a limited capacity (but see Carter et al., in press; Johnson & Weber, 2009; Schulz et al., 2014). Thus, when self-control or cognitive effort is exerted, the cognitive processes might eventually get exhausted and the affective processes might gain stronger influence on decision making. Previous research indicates that there are a multitude of factors influencing whether automated or cognitive processes are used (Chaiken & Trope, 1999; Cornelissen et al., 2011; Evans, 2008; Loewenstein & O'Donoghue, 2007; McClure, Laibson, Loewenstein, & Cohen, 2004; Schulz et al., 2014; Shiv & Fedorikhin, 1999). Amongst others, mood, low blood glucose levels, exhaustion of willpower and cognitive load determine whether the cognitive or affective processes have a greater impact on decision making (Gailliot & Baumeister, 2007; Gailliot et al., 2007; McClure et al., 2004; Schulz et al., 2014). The dual process approach thus states that, when cognitive capacity is high, decisions can be made in a more controlled way and presumably in line with long-term interests. However, when the cognitive resources are exhausted, the automatic processes dominate. Thus, the amount of available cognitive control seems to shape which preferences are revealed in a decision.

Recent theories on altruism and other-regarding preferences suggest that prosocial behavior requires cognitive control to overcome selfish motives (Knoch et al., 2006; Strombach et al., 2015). Combining this insight with the evidence from stress-research cited above, suggesting that men are more prone to cognitive depletion then women, we speculate that the results in our present study reflect a gender-dependent switch from controlled to automatic processing: while men use more automatic processes to make social decisions when cognitive load is high, women consistently rely on deliberate control processing independent of cognitive load. The disuse of the deliberative system in men is likely to lead to a reduction in the information complexity that is processed to form a decision. In order to come up with a sound decision in the social discounting task, participants have to consider own- and other-reward magnitudes as well as social distance to the recipient. One way to simplify the decision problem is to reduce the dimensionality of the social choice alternatives. It is therefore possible that men under high cognitive load conditions neglected the social distance dimension of the choice problem. Because women were presumably less sensitive to the cognitive load manipulation, the switch from deliberative to automatic processes was less pronounced, and women consequently showed no difference in social discounting behavior between cognitive load conditions.

Finally, it is important to mention that a recent meta-analysis challenges the idea that self-control relies on a limited resource that is subject to exhaustion by cognitive load (Carter et al., in press). Although the conclusion of this meta-analysis does not contest the validity of our results per se, it has potentially some implications for the interpretation of our data. If cognitive control depletion was indeed a myth, the difference in social discounting between high and low cognitive load conditions in men must have been caused by a different mechanism than cognitive exhaustion. However, although we cannot entirely rule out alternative explanations, we are hesitant to entirely dismiss the cognitive depletion account of our data: first, the meta-analysis (Carter et al., in press) mainly covered studies on will-power, which is a special case of self-control, which in turn is also a special case of cognitive control. Here, we refrain from specifying whether the change in social discounting was instigated by a drop in will-power, or by changes in other forms of cognitive control. The fact remains that men and women showed different social choice behavior, which is very likely the consequence of our cognitive load manipulation. Second, we show that cognitive load effects on (social) preferences are highly complex and depend on several factors, including social distance as well as gender. Studies that do not consider or control these modulators of behavior may yield inconsistent findings, which would factor in as a null result in any meta-analysis. Finally, the neural mechanisms underlying social discounting have little or nothing in common with the neural processes associated with the self-control capacities dissected in Carter et al., (in press). More specifically, we have recently provided a neurobiological model of social discounting according to which the social-distance-dependent conflict between selfish and generous motives is resolved by upregulating prosocial neural value signals in ventromedial prefrontal cortex through functional coupling with one of the core social brain regions – temporoparietal junction (Strombach et al., 2015). By contrast, will-power, self-regulation or other forms of self-control inspected in Carter et al. (in press) are usually associated with different, mainly frontal neural networks, including dorsolateral prefrontal cortex (Hare, Camerer, & Rangel, 2009), inferior frontal gyrus (Casey et al., 2011), or frontopolar cortex (Crockett et al., 2013). To our knowledge, no neuroimaging study on self-control has revealed an involvement of temporoparietal junction. Hence, the neural mechanisms, and thus possibly also the mental processes underlying social discounting (Strombach et al., 2015) are not reminiscent of the willpower mechanisms scrutinized in Carter et al. (in press), and the conclusions of their meta-analysis may therefore not apply to the current study. We suggest, therefore, that changes in cognitive control might have distinct effects on social preferences that are reliant on the neural network used to process social discounting.

In summary, we suggest that our findings that cognitive load affects social discounting in a gender-dependent way might be interpreted within the dual process framework (Evans, 2008; Johnson & Weber, 2009; Schulz et al., 2014). We propose that cognitive load diminishes deliberation capacities in men and, thus, the likelihood of activating the controlled, cognitive system. As such, decisions taken under additional cognitive load are governed to a greater extent by affective processes and automatic choice heuristics. This might lead to a reduction in the choice alternatives' dimensions that are considered to form a decision (Rachlin & Jones, 2008; Simon, 1995). In the present case, we propose that the social distance dimension is considered less under increased cognitive load. In line with previous research, we suggest that the cognitive-load-induced diminution of deliberation capacities is more pronounced in men than women.

Our data are interesting for a couple of reasons. First, we show that cognitive-load effects on prosocial sentiments are complex and dependent on several interacting factors, including social distance and gender. Since social distance and gender were often not, or only partially controlled in previous studies on cognitive load and social choice, our results may help reconcile the inconsistencies in earlier findings. Second, previous studies are incongruous regarding gender difference in cognitive load effects on cognition. We show that gender effects only become visible in prosocial choice tasks when social distance is taken into consideration. Thus, only the combination of social distance, gender and a cognitive load manipulation is able to unshadow the effect of cognitive load on social behavior. In sum, overall, we suggest that inclusion of social distance in social experiments might be advantageous.

Our results highlight the importance of research on the effects of cognitive load on social behavior. Insights in this field might lead to a better understanding of behaviors in situations where cognitive capacity is scarce, i.e. in jobs that demand multitasking faculties, or executive decisions that are made under time pressure and stress. A better understanding could help to develop strategies to deal with the risks of cognitive exhaustion to improve the quality of decisions. That might also lead to optimize the work environment to improve the quality of the decisions made in a work-related setting. Finally, we suggest that social preferences as well as social distance should be included in economic models and psychological theories to further their descriptive and predictive value.

6. Conclusion

The effects of cognitive load on generosity and prosocial behavior are complex, and depend on social distance between donor and recipient, as well as the gender of the donor. The gender-difference in the impact of cognitive load on social behavior suggests that men and women process social information differently when cognitive control becomes exhausted. We propose that our results can be interpreted within the dual system framework, which suggests two distinct processes that drive decision making – a deliberate and an automatic system. We suggest that cognitive load affects the interplay of these two systems differently between men and women. Our findings have implications for our understanding how a person's environment might influence her or his ability to make decisions (Porcelli & Delgado, 2009).

References

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology, 46*(3), 610.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *Journal of personality and social psychology, 74*(5), 1252.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*(4), 361-383.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2006). Who is' behavioral'? Cognitive ability and anomalous preferences. *Cognitive Ability and Anomalous Preferences (May 5, 2006)*.

Cappelletti, D., Güth, W., & Ploner, M. (2011). Being of two minds: Ultimatum offers under cognitive constraints. *Journal of Economic Psychology, 32*(6), 940-950.

Carter, E. C., Kofler, L. M., Foster, D. E., & McCullough, M. E. (in press). A Series of Meta-Analytic Tests of the Depletion Effect: Self-Control does not Seem to Rely on a Limited Resource. *Journal of experimental psychology*.

Casey, B. J., Somerville, L. H., Gotlib, I. H., Ayduk, O., Franklin, N. T., Askren, M. K., . . . Shoda, Y. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences, 108*(36), 14998-15003. doi: 10.1073/pnas.1108561108

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*: Guilford Press.

Cornelissen, G., Dewitte, S., & Warlop, L. (2011). Are social value orientations expressed automatically? decision making in the dictator game. *Personality and Social Psychology Bulletin, 37*(8), 1080-1090. doi: 10.1177/0146167211405996

Crelley, D., Lea, S., & Fischer, P. (2008). Ego Depletion and the Tragedy of the Commons: Self Regulation Fatigue in Public Goods Games. *Unpublished manuscript*.

Crockett, M. J., Braams, B. R., Clark, L., Tobler, P. N., Robbins, T. W., & Kalenscher, T. (2013). Restricting temptations: Neural mechanisms of precommitment. *Neuron, 79*(2), 391-401.

Eifler, W. (2009). *Küttner Kolbenmaschinen: Kolbenpumpen, Kolbenverdichter, Brennkraftmaschinen; mit 40 Tabellen sowie zahlreichen Übungen und Beispielen mit Lösungen*: Springer-Verlag.

Evans, J. S. B. T. (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Vol. 59. Annual review of psychology* (pp. 255-278).

Fennis, B. M. (2011). Can't get over me: Ego depletion attenuates prosocial effects of perspective taking. *European Journal of Social Psychology, 41*(5), 580-585.

Franco-Watkins, A. M., Pashler, H., & Rickard, T. C. (2006). Does working memory load lead to greater impulsivity? Commentary on Hinson, Jameson, and Whitney (2003).

Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review, 11*(4), 303-327.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., . . . Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *Journal of personality and social psychology, 92*(2), 325.

Goto, T., & Kusumi, T. (2013). How can reward contribute to efficient self-control? Reinforcement of task-defined responses diminishes ego-depletion. *Motivation and Emotion, 37*(4), 726-732.

Gwizdka, J. (2010). *Using stroop task to assess cognitive load.* Paper presented at the Proceedings of the 28th Annual European Conference on Cognitive Ergonomics.

Hangebrauk, Z., Strombach, T., van Wingerden, M., Joels, M., Schwabe, L., & Kalenscher, T. (in press). A Friend in Need: Time-Dependent Effects of Stress on Social Discounting in Men. *Hormones & Behavior*.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science, 324*(5927), 646-648.

Hauge, K. E., Brekke, K. A., Johansson, L.-O., Johansson-Stenman, O., & Svedsäter, H. (2009). Are social preferences skin deep? Dictators under cognitive load. *rapport nr.: Working Papers in Economics 371*.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive Decision Making and Working Memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 29*(2), 298-306. doi: 10.1037/0278-7393.29.2.298

Inzlicht, M., & Gutsell, J. N. (2007). Running on empty neural signals for self-control failure. *Psychological Science, 18*(11), 933-937.

Johnson, E. J., & Weber, E. U. (2009). Mindful judgment and decision making. *Annual review of psychology, 60*, 53.

Jones, B., & Rachlin, H. (2006). Social discounting. *Psychological Science, 17*(4), 283-286.

Jones, B. A., & Rachlin, H. (2009). Delay, probability, and social discounting in a public goods game. *Journal of the Experimental Analysis of Behavior, 91*(1), 61-73.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science, 314*(5800), 829-832.

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E. Y., . . . Mather, M. (2012). Gender differences in reward-related decision processing under stress. *Social cognitive and affective neuroscience, 7*(4), 476-484.

Loewenstein, G., & O'Donoghue, T. (2007). The heat of the moment: Modeling interactions between affect and deliberation. *Unpublished manuscript*.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503-507.

Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research, 17*(2), 189-202. doi: 10.1023/B:SORE.0000027409.88372.b4

Olson, E. A., Rosso, I. M., Demers, L. A., Divatia, S., & Killgore, W. D. S. (2015). Sex Differences in Psychological Factors Associated with Social Discounting. *Journal of Behavioral Decision Making*, n/a-n/a. doi: 10.1002/bdm.1876

Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science, 20*(3), 278-283.

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(1), 1-20. doi: 10.1017/S0140525X02000018

Ptacek, J. T., Smith, R. E., & Dodge, K. L. (1994). Gender differences in coping with stress: When stressor and appraisals do not differ. *Personality and Social Psychology Bulletin, 20*(4), 421-430.

Rachlin, H., & Jones, B. A. (2008). Social discounting and delay discounting. *Journal of Behavioral Decision Making, 21*(1), 29-43.

Roch, S. G., Lane, J. A. S., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive Load and the Equality Heuristic: A Two-Stage Model of Resource Overconsumption in Small Groups. *Organizational Behavior and Human Decision Processes, 83*(2), 185-212. doi: 10.1006/obhd.2000.2915

Saadé, R. G., & Otrakji, C. A. (2007). First impressions last a lifetime: effect of interface type on disorientation and cognitive load. *Computers in human behavior, 23*(1), 525-535.

Salmon, S. J., Adriaanse, M. A., De Vet, E., Fennis, B. M., & De Ridder, D. T. D. (2014). "When the going gets tough, who keeps going?" Depletion sensitivity moderates the ego-depletion effect. *Frontiers in Psychology, 5*, 647. doi: 10.3389/fpsyg.2014.00647

Schulz, J. F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology, 41*(0), 77-87. doi: http://dx.doi.org/10.1016/j.joep.2012.08.007

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research, 26*(3), 278-292.

Simon, J. L. (1995). Interpersonal Allocation Continuous with Intertemporal Allocation Binding Commitments, Pledges, and Bequests. *Rationality and Society, 7*(4), 367-392.

Soutschek, A., Strobach, T., & Schubert, T. (2013). Working memory demands modulate cognitive control in the Stroop paradigm. *Psychological research, 77*(3), 333-347.

Strombach, T., Jin, J., Weber, B., Kenning, P., Shen, Q., Ma, Q., & Kalenscher, T. (2014). Charity begins at home: cultural differences in social discounting and generosity. *Journal of Behavioral Decision Making, 27*(3), 235-245.

Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I. I., Tobler, P. N., & Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences, 112*(5), 1619-1624.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18*(6), 643.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction, 4*(4), 295-312.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality, 72*(2), 271-324.

Van den Bos, R., Harteveld, M., & Stoop, H. (2009). Stress and decision-making in humans: performance is related to cortisol reactivity, albeit differently in men and women. *Psychoneuroendocrinology, 34*(10), 1449-1458.

Vohs, K. D., Baumeister, R. F., & Schmeichel, B. J. (2012). Motivation, personal beliefs, and limited resources all contribute to self-control. *Journal of Experimental Social Psychology, 48*(4), 943-947. doi: http://dx.doi.org/10.1016/j.jesp.2012.03.002

Wheeler, S. C., Briñol, P., & Hermann, A. D. (2007). Resistance to persuasion as self-regulation: Ego-depletion and its effects on attitude change processes. *Journal of Experimental Social Psychology, 43*(1), 150-156.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *The International Journal of Aviation Psychology, 12*(1), 3-18.

4. List of Abbreviations

ACC	Anterior cingulate cortex
BOLD	Blood oxygenation level dependent
DLPFC	dorsolateral prefrontal cortex
EEG	Electroencephalography
fMRI	functional magnetic resonance imaging
FWE	familywise error
GG	Greenhouse Geisser (correction)
GLM	general linear model
HRF	hemodynamic response function
LCD	Liquid-crystal display
M	mean
Mdn	median
MNI	Montreal Neuroimaging Institute
MRI	magnetic resonance imaging
OFC	orbitofrontal cortex
ORU	other-regarding utility
PPI	psychophysiological interaction
ROI	region of interest
RPE	respose prediction error
s	seconds
SD	standard deviation
STG	superior temporal gyrus
SV	subjective value
TMS	transcranial magnetic stimulation
ToM	theory of mind
TPJ (rTPJ)	temporoparietal junction (right temporoparietal junction)
VMPFC	ventromedial prefrontal cortex
VS	ventral striatum
vStr	ventral striatum (abbreviation used in study 1)

5. Acknowledgements

I met Tobias in Amsterdam in 2010, when he was organizing a summer school on neuroeconomics. I was still a master's student at the time and was fascinated by neuroeconomics. Since then, Tobias has accompanied me in my scientific career: first, by collaborating on my master's thesis, then by supervising my PhD. I think it is no overstatement to say that he contributed tremendously to my becoming the researcher I am today, and I would like to thank him for that. I would also like to thank my second supervisor, Bernd Weber, for giving me the opportunity to do neuroscientific research, the freedom to work independently (also with the MRI scanner in Bonn), and help whenever I needed it. My colleagues at the CompPsycho Department supported my work constantly. Although most of you are working with much smaller subjects (rats to be precise), it was great interacting with you! Thank you for the last few years. I'd like to mention one person in particular – Julien. We started our PhD together and you were the best Büro-Mate ever (sorry again for killing Kaktüs!). Of course, the research would not have been possible without my amazing co-authors: Marco Hubert, Peter Kenning, Philippe Tober, Iliana Karapidis, Zsofia Hangebrauk, and Barbara Gorczyca. Thank you for putting so much effort into my PhD. ;)

In addition to the scientific support, I would like to thank my friends for making the last few years so special! Thank you so much for your constant support Anna, Anna-Lena, Julia, Moni, Anti, Steffi, Fabi, Simon, Marjolein, and Habs (and all the other people who supported me and are not mentioned here)! I would like to thank you for enduring all my ups and downs, celebrating every success, and giving me moral support when necessary. Special thanks go to my friend and colleague, Sabrina Strang, who discussed my research with me, travelled with me to conferences and "around" conferences, and made me watch soccer (even that was great!).

Last but not least, I'd like to thank my parents, my brother Jan and his wife Annemarie and my sister Maike and Gerrit for their remarkable support. You fevered with me every second during the last couple of years! My parents never questioned any of my decisions and assisted me without exception. Thank you for giving me the opportunity to make my own decisions and for the financial and mental support during all these years of studies! You are the most important contributors to my career!

And of course: Thank you for reading my thesis!

And always remember: "Sharing is caring!"

6. Curriculum Vitae

Tina Strombach

Eigelstein 34
50668 Cologne, Germany

Tel.: +49 172 3954794
Email: tina.strombach@icloud.com
Date of birth: November 29, 1985

Education

10/2011 - present	**PhD Candidate** at the **Heinrich-Heine University,** Düsseldorf, Germany PhD at the Institute of Comparative Psychology, Faculty of Mathematics and Natural Sciences on *Decision Neuroscience & Neuroeconomics* (Prof. Tobias Kalenscher)
10/2011 - present	**Graduate Student** at Interdisciplinary Graduate and Research Academy Düsseldorf (iGRAD)

09/2009 - 09/2011	**Zeppelin University**, Friedrichshafen, Germany (final grade: 1,7)
	Graduation: Master of Arts in Corporate Management and Economics
	       Topics of the studies:
	            - Marketing
	            - Behavioral economics and neuroscience
	            - Systems theory and networks
	       Subject of Master-Thesis: *"Cultural Differences in Social Discounting"* (grade: 1,3)

09/2005 - 08/2009	**Universiteit Maastricht**, Maastricht, the Netherlands (final grade: 2,0)
	Graduation: Bachelor of Science in Psychology
	      Major in biological psychology
	      Subject of the Bachelor-Thesis: *"The gut feeling – An Examination into the Role of Unconscious Processes in Decision Making"* (grade: 1,7)

01/2008 - 06/2008	Exchange Program at the **Université Pierre Mendes France** Grenoble, France
2002 - 2005	**Sankt Angela Gymnasium,** Wipperfürth, Germany

Work & Research Experience

04/2015 - present	**Research assistant** at the **Heinrich-Heine University**, Düsseldorf, Germany, Marketing (Chairholder: Prof. Dr. Peter Kenning) Research on consumer neuroscience, social neuroscience, acceptance of food innovation

10/2011 – 03/2015	**Research assistant** at the **Heinrich-Heine University**, Düsseldorf, Germany, Comparative Psychology (Chairholder: Prof. Dr. Tobias Kalenscher)
	Research on neuroeconomics, decision neuroscience & social neuroscience
02/2011 – 06/2011	**Internship** at the **Center for Economics and Neuroscience, University of Bonn, Germany** (Group leader: Prof. Bernd Weber)
02/2010 – 06/2011	**Student Assistant,** Zeppelin University, Germany, Department of Marketing
	(Chairholder: Prof. Peter Kenning),
	Research assistance within the field of Neuromarketing, Neuroeconomics and Consumer Neuroscience
21/06 - 05/08/2010	**Internship** at the **Freien Universität**, Berlin, Germany at the Department of Educational Sciences and Psychology (Group leader: Prof. Hauke Heekeren)
	Research on Neuroeconomics and Portfolio Theories
09/2008 - 03/2009	**Internship** at **Dr. Otto Training & Consulting**, Berlin, Germany
	Organizational development and change management, knowledge management, "evolutionary management"

Awards and Travel Grants

2011	**FUSY Research Fund**, Zeppelin University
	funding for Research Project
2011	**FUSY Research Fund**, Zeppelin University
	for attending an international conference (Society for Neuroeconomics, 2011, Evanston, USA)
2013	**DAAD Travel Grand** (German Academic Exchange Service), for attending an international conference (Society for Social and Affective Neuroscience, 2013, San Francisco, USA)

Conferences & Publications (selection)

Strombach, T., Hubert, M. & Kenning, P. (in press). Neural underpinnings of performance-based incentives. *Journal of Economic Psychology.*

Hangebrauk Z., Strombach T., van Wingerden, M., Joels M., Schwabe L., Kalenscher T. (in press). A Friend in Need: Time-Dependent Effects of Stress on Social Discounting in Men. *Hormones & Behavior.*

Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I.I., Tobler, P.M., Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction, *PNAS, 112(5).*

Strombach, T., Jin, J., Weber, B., Kenning, P., Shen, Q., Na, Q. & Kalenscher, T. (2014). Charity Begins at Home - Cultural Differences in Social Discounting and Generosity. *Journal of Behavioral Decision Making.*

Kalenscher, T. & **Strombach, T.** (2013). Die Macht der Versuchung [The power of temptation]. *Gehirn&Geist, 11,* 62-67.

Strombach, T., Weber, B., Kenning, P., Tobler, P., & Kalenscher, T. (2013). Neural Correlates of Social Discounting. Annual Meeting of the Society for Social and Affective Neuroscience, San Francisco, CA, USA.

Strombach, T., Weber, B., Kenning, P., Tobler, P., & Kalenscher, T. (2012). Neural Correlates of Social Discounting. Society for Neuroeconomics Annual Conference, Miami, FL, USA.

Strombach, T., Kalenscher, T., Kenning, P., Ma, Q., Jin, J., Shen Q. & Weber,B. (2011). Individualism vs. Collectivism: An Investigation into Cultural Differences in Social Discounting. Society for Neuroeconomics Annual Conference, Evanston, IL, USA.

Reviewer activities

Ad-hoc reviewer for the Journal of Behavioral Decision Making, European Research Council

Teaching

WS 2014	Psychological methods in market research
WS 2012 - SS 2013	Experimental Practicum (realization of a small research project)
SS 2012 - WS 2014	Statistics and SPSS
WS 2011	Neuroscientific Methods: functional magnetic resonance imaging
Since 2011	Supervision of several Bachelor and Master theses

Language skills

German	Native speaker
English	Business fluent (written and verbal)
Dutch	Fluent (written and verbal)
French	Good

Computer skills

SPSS	Statistical Package for the Social Sciences
Matlab	high-level programming language and interactive environment
SPM	Statistical parametric mapping – analysis of brain imaging data sequences

7. Erklärung (declaration)

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine- Universität Düsseldorf" erstellt worden ist.

Ich versichere, dass ich die Dissertation an keiner anderen Fakultät vorgelegt habe und auch sonst bisher keine Dissertation eingereicht habe.
Ich habe diese Dissertation weder in dieser, noch in irgendeiner anderen Fassung bereits einer anderen Fakultät vorgelegt. Ich habe darüberhinaus bislang auch keine andere Dissertation vorgelegt.

Düsseldorf, 15. Juni 2015