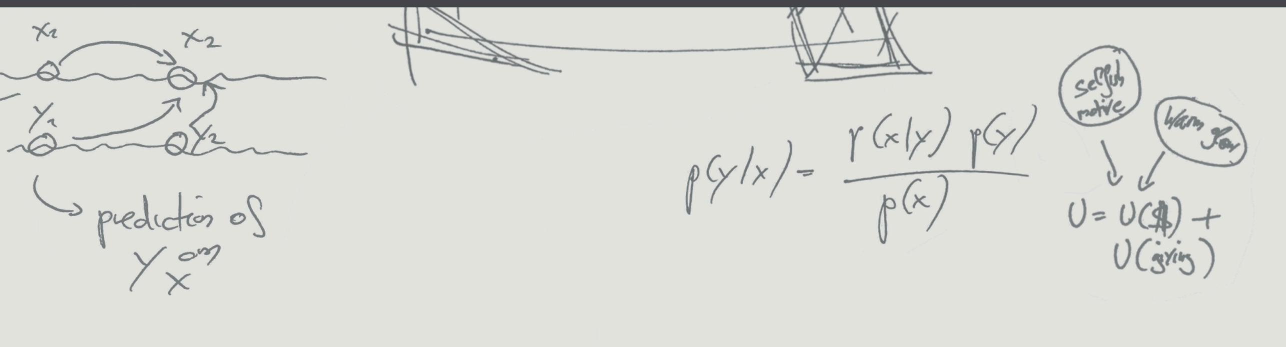
From Simple Choice to Social Decisions

On the Neurobiological and Evolutionary Roots of Decision Making

Jörg Gross



From Simple Choice to Social Decisions

On the Neurobiological and Evolutionary Roots of Decision Making

DISSERTATION

to obtain the degree of Doctor at Maastricht University
on the authority of the Rector Magnificus Prof. Dr. L.L.G. Soete,
in accordance with the decision of the Board of Deans,
to be defended in public on Wednesday, 25th of March 2015, at 14:00 hours

by Jörg Gross

Supervisors	Prof. Dr. Arno Riedl
	Prof. Dr. Frank Moers
	Prof. Dr. Rainer Goebel
Assessment Committee	Prof. Dr. Elia Formisano
	Dr. Tony Williams
	Prof. Dr. Tobias Kalenscher, University of Düsseldorf
	Prof. Dr. Simon Gächter, University of Nottingham

Contents

List of Figures

List of Tables

Introduction

The New Consilience in Decision Making Research

Understanding how humans and animals make decisions is the goal of multiple disciplines, including psychology, economics, sociology and biology. However, theories and methodologies have only recently started to cross discipline boundaries. Psychologists, on the one hand, begin to recognize the benefits and beauty of game theory as a formalized language to characterize decisions in strategic contexts. In economics, on the other hand, the acceptance of laboratory experiments is increasing and psychological concepts such as emotions or impulsivity are incorporated into microeconomic theories.

The convergence of different schools of thought can prove fruitful and lead to novel insights. A prime example for this is the research of the psychologists Daniel Kahneman and Amos Tversky. After working on visual perception and measurement theory, they became interested in decision making and turned to economic theory. For economists, decisions should be made by evaluating the gain in personal value or utility an outcome would yield and weighting it by the (estimated) probability that this outcome is realized. A series of experiments by Kahneman, Tversky and others, however, documented that human decision making often violates this principle. Humans fail to follow the rules of probability theory and systematically over- or underestimate the probability of certain events (Kahneman and Tversky, 1972). Further, in the dominant economic theory of decision making, expected utility theory (EUT), choice options should be evaluated based on how they are expected to change one's overall wealth level. However, according to Kahneman and Tversky, humans evaluate outcomes relative to a mentally constructed reference point, such as past experience or expectations. Depending on the reference point a nominal gain might be perceived as a loss. The price change of a stock can exemplify this. If the past value of the stock is used as a reference point, the psychological perception of a recent drop in price is that of losing money, even if the overall return of the stock is positive. On the contrary, an increase in price for a stock that is far in the negative, can be perceived as a gain. In an attempt to replace EUT, Kahneman and

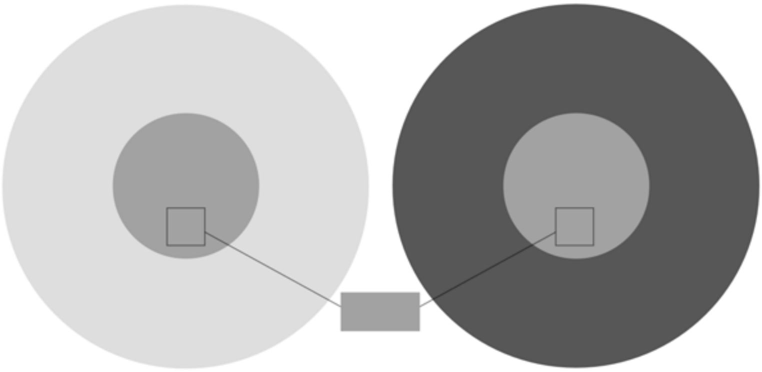


Figure 1.1: Contrast Illusion. The same grey tone appears darker when surrounded by a lighter grey tone (left), but lighter when surrounded by a darker grey tone (right).

Tversky formulated their own theory of decision making: prospect theory (Kahneman and Tversky, 1979).

Compared to EUT, prospect theory tries to capture how humans actually perceive and reformulate decision situations in their mind. Thinking about how the environment is mentally reconstructed is a typical approach in modern psychology, but is missing in economics. The idea of reference dependent utility was ultimately inspired by research on visual perception.[1] From optical illusions, it is long known that perception is reference dependent. Hue, saturation and contrast are not represented objectively in our visual system, but processed relatively to the surrounding. The same grey tone appears darker next to a light grey area, but lighter next to a dark grey area (see Figure 1.1). Kahneman and Tversky translated this phenomenon to choice behavior; What is perceived as a loss or a gain does not depend on changes of overall wealth, but on past experience, aspiration goals or other reference points, framing a decision and its prospects just like the outer circles in Figure 1.1.

Prospect theory is now widely accepted as an alternative account of human decision making. However, EUT is still the dominant framework for conceptualizing decisions in economics. The reason for that, I believe, can be found in the different approaches to the scientific exploration of the human mind, approaches that are deeply rooted in the history of the two disciplines and shape the interdisciplinary discourse today.

[1] "Knowledge of perception and ignorance about decision theory both contributed to a large step forward in our research" (Kahneman, 2011, p. 278).

A short walk through the history of thought

Economics – the measurement of value

The fundamental concept of subjective value, central to both EUT and prospect theory, gained popularity with the work of the English philosophers Jeremy Bentham and John Stuart Mill in the late 18th and early 19th centuries. Bentham believed that human behavior is driven by two fundamental forces: pleasure and pain. As a trained lawyer, he was convinced that human life should be governed by institutions that help to maximize pleasure and reduce pain (Bentham, 1789). This 'principle of utility' was revolutionary for its time. It implied that all men (including women) are equal and deserve to be happy. In his lifetime, he tried to influence the British law system, advocating the unification of laws under the principle of utility. But to translate the sensation of pleasure or pain into formal laws, Bentham needed a way to quantify and compare the experience of utility across individuals. He formulated a measurement theory of pleasure, 'hedonic calculus', based on psychological dimensions such as intensity, duration or purity and was confident that utility is quantifiable on a cardinal scale making it able to compare the experience of pleasure even across individuals. He envisioned that eventually, 'hedonic calculus' would guide policy-making in the social sphere.

While the 'principle of utility' had a major impact on philosophy and economics, his student John Stuart Mill already criticized Bentham's quantification attempt as too narrow-minded. For Mill, the sensation of pleasure can make qualitative jumps, rendering it impossible to represent it on one scale. In his view, reading a philosophy book (e.g., Bentham's "Introduction to Principles of Morals and Legislation"), although painful and challenging, can eventually yield a higher level of pleasure than an exciting but shallow novel can ever create.[2]

From cardinal utility to revealed preferences

Economists adapted the general idea of Bentham and Mill that human decisions can be described as attempts to maximize utility but struggled in their efforts to directly measure it. Ultimately, they discarded Bentham's vision of a calculus of hedonic sensation altogether.

In the late 19th century, after the death of John Stuart Mill, a new era in the history of economic thought began. During the 'Marginal Revolution', the utility concept underwent its first major overhaul. In its beginning, most economists, like the Irish economist Francis Edgeworth, still believed that utility might be directly measurable. Edgeworth,

[2] As Mill famously formulated it: "It is better to be a human being dissatisfied than a pig satisfied" (Mill, 1863, p. 260).

most notably known for the theory of indifference curves, proposed the idea of a hedonometer, a psychophysical machine designed to measure the sensation of utility. At the same time, Edgeworth also realized the problems for a theory that rests on a highly subjective experience like pleasure. He wrote poetically: "We cannot count the golden sands of life; we cannot number the innumerable smile of seas of love; but we seem to be capable of observing that there is here a greater, there a less [...] and that is enough" (Edgeworth, 1881, p. 8).

Marginal utility economists were the ones to realize that there might be a way to circumvent an exact quantification of utility. When making a decision between two options, it is not necessary to know the exact amount of pleasure of each option, but the difference in pleasure one option provides over the other. With the focus on marginal changes, trying to measure the total amount of pleasure is not necessary. Also the idea that utility can be regarded as cardinal, as Bentham envisioned it, was being more and more questioned. Even if measurable, are assertions like 'option A yields double the amount of pleasure than option B' meaningful? Can utils (the units of utility) be added and subtracted, and integral calculus performed on them? And can the experienced pleasure be compared across individuals? Most marginal economists of the second generation, like Alfred Marshall or Vilfredo Pareto, had serious doubts. Pareto rejected attempts to compare utility across individuals: "The utility [...] for one individual, and the utility [...] for another individual, are heterogeneous quantities. We can neither add them together nor compare them" (Pareto, 1909, p. 192). Further, he replaced cardinal utility with an ordinal concept of utility. When deciding between bundles of goods that vary in their quantity, the indifference curves between these bundles can be ordered from high to low. These numbers only represent the rank of preferences of an individual, not the absolute amount of utility. For Pareto, this is sufficient for the prediction of choice and a theory of demand.

Economists increasingly believed that economic theory should not rely on mental processes that are not directly observable. For Pareto, attempts to measure utility by self reports were not only unscientific, but also unnecessary for the science of economics (Pareto, 1909). Departing from the psychological dimension of value, such as sensation or experience, leading economists, like William Jevons or Léon Walras, were also convinced that for economics to become a 'real' science, often looking at Physics as a role model, it must be built on a formal mathematical foundation; "It is clear that if economics, if it is to be a science at all, must be a mathematical science. [...] it must reason by real equations" (Jevons, 1871, p. 78). In the next decades, economic theory became increasingly mathematical, and contemplations about inner experiences were replaced with the analysis of directly observable decisions. However, in its core there was still a psychological concept: utility. The assumption that humans strive for the maximization of utility was treated like a psychological axiom, an unverifiable truth that is not accessible to rigorous science.

Since Bentham, the chain of thought started from the decision maker that tries to anticipate the utility of different options and then decides according to the maximization principle. The American economist Paul Samuelson, in an attempt to completely reformulate the basis of economic theory, turned this upside down: "I propose [...] that we start anew in direct attack upon the problem, dropping off the last vestiges of the utility analysis" (Samuelson, 1938, p. 62). In order to remove the last remnants of the psychological utility concept from economic theory, Samuelson restricted his attention to what is accessible through direct observation: decisions. Starting from observable decisions, he derived a theory about the conditions that must be met in order to assume that the decision maker decides *as if* he would maximize some kind of utility (Samuelson, 1947). When observing someone choosing choice set A over choice set B, what does this tell us about the decision maker? For Samuelson, the chooser reveals her preference. A is chosen because it is preferred over B. But what pattern of decisions, Samuelson continued in his analysis, need to be observed, so that one can pretend that the decision maker is maximizing utility? First of all, if A is preferred over B, the decision maker should never choose B over A, keeping everything else constant (the weak axiom of revealed preference). Samuelson laid out the criteria under which choices are consistent with the idea of underlying stable preferences or an utility function, without the need to say anything about mental processes. Following Samuelson, the revealed preference approach was extended to rule out choice cycles. If A is preferred over B and B is preferred over C, then in a decision between A and C, A should be chosen (derived from the strong axiom of revealed preference).

With the work of Samuelson and his successors, economists found a way to remove any reliance on psychological concepts in economic theory. The decision maker transformed from the human who is governed by "two sovereign masters, pain and pleasure" (Bentham, 1789, p. 1) to the homo oeconomicus, an agent who behaves 'as if' utility was its driving force. Milton Friedman famously formulated this positivist approach to the science of decision making as follows:

> "Consider the density of leaves around a tree. I suggest the hypothesis that the leaves are positioned as if each leaf deliberately sought to maximize the amount of sunlight it receives [...]. Is the hypothesis rendered unacceptable or invalid because, so far as we know, leaves do not 'deliberate' or consciously 'seek,' have not been to school and learned the relevant laws of science or the mathematics required to calculate the 'optimum' position, and cannot move from position to position? Clearly, none of these contra-

dictions of the hypothesis is vitally relevant; the phenomena involved are not within the 'class of phenomena the hypothesis is designed to explain'; the hypothesis does not assert that leaves do these things but only that their density is the same *as if* they did." (Friedman, 1953, p. 14)

Psychology – the struggle with latent variables

Similarly to economists, psychologists also struggled with how to study mental processes and subjective sensations scientifically. From avoiding these latent unobservable constructs altogether to making them the central object of investigation, the history of psychology can be characterized as an alternating approach-avoidance conflict.

The beginning of psychology as an academic discipline is often dated back to 1879, the year in which Wilhelm Wundt founded the first experimental psychology laboratory at the University of Leipzig, Germany.[3] Wundt established psychology as an independent discipline and defined the agenda for the study of the mind. For Wundt, psychology has to be rooted in the natural sciences. Mental processes have to be explained on the basis of physiological changes. In the same vein, this link between environment and inner sensation has to be explored empirically. In his own experiments, Wundt's students often became the subjects, confronted with controlled visual, tactile or acoustic stimulation, while objective measurements like reaction time were taken (see Figure 1.2).

Wundt's numerous students and colleagues were also trained in introspection, the systematic examination of thoughts and feelings. Although there is a debate on how important introspection was for Wundt's scientific program (Danziger, 1980), it definitely became an important method for some of his students. Notably, Oswald Külpe, who later founded the 'Würzburg School of Psychology', believed that systematic introspection is the Via Regia to the scientific exploration of the conscious mind: "Like personal experience is the fundamental source, self-observation is the fundamental method of descriptive psychology".[4] Külpe and his colleagues diverged from the idea that the content of the mind can be traced back to mere physiological stimulation. Heavily relying on introspection, they tried to explore the complex inner workings of the mind.

[3] Psychology was a topic of great interest also outside Wundt's laboratory. Most notably, Ernst Heinrich Weber, a German physician, was interested in the relation of the physical properties of objects and their psychological representation, and explored it empirically. Later, Gustav Theodor Fechner, building on Weber's work, established it as a scientific subfield: psychophysics. Psychophysics was well equipped for the systematic exploration of the psychological dimensions of utility. Although some economists were aware of psychophysics, it never came to a broad interdisciplinary cooperation between these fields.

[4] Own translation. Original quote: "Wie die eigene Erfahrung der Grundquelle, so ist die Selbstbeobachtung die Grundmethode der beschreibenden Psychologie" (Külpe, 1920, p. 45).

Figure 1.2: Wilhelm Wundt (sitting) surrounded by colleagues in front of a psychophysical apparatus that created sounds while recording reaction times.

One of his most famous students, Karl Bühler, for example confronted his subjects with complicated metaphors and instructed them to think aloud. He observed that some participants, while first struggling with the meaning of the sentences, experienced an eureka effect, a sudden understanding of meaning (Bühler, 1907). Bühler concluded that thought processes often do not follow a logical order. Instead they can make unexpected jumps that are hardly reducible to changes in the outer physiological stimulation. Wundt, who, in addition to directly observable measures, accepted systematic introspection as one valid method of psychology, disagreed with the notion that self observation should play such a key role in the methodological repertoire of scientific psychology. He strongly criticized the studies of Bühler (now referred to as the Bühler-Wundt controversy), believing that such high order cognitive processes are not examinable with scientific rigor.

Nevertheless, 'Denkpsychologie', the school that set out to examine the complex realm of thoughts and consciousness, gained momentum with psychologists, like Edward B. Titchener, who exported this approach to the U.S. or Max Wertheimer, who used optical illusions like the Phi phenomenon, to demonstrate the constructive nature of our perception.

The retreat and revival of the mind

Introspection and the exploration of inner processes were, however, soon damned as unacceptable pseudoscience. Because of their Jewish background, many German psychologists had to flee Germany with the rise of German Nazism and Germany lost its position as the central hub of psychological research in the world. The U.S. became its new home, accompanied by a deep reformulation of the scientific program of human psychology.

This reformulation is mainly attributed to the American psychologist John Watson. Watson, who recognized the importance of Ivan Pavlov's conditional reflex theory for human psychology, was deeply dissatisfied with the state of his discipline. At the age of 34, in a speech given at Columbia University, he shared his concerns and presented his new direction for psychology: Behaviorism, a program that should dominate teaching and research at universities for the next thirty years. In his speech, he clearly defined what should not be part of psychology: "The time seems to have come when psychology must discard all reference to consciousness; when it need no longer delude itself into thinking that it is making mental states the object of observation" (Watson, 1913, p. 163). For Watson, there was no consensus on what concepts like 'perception', 'emotion' or 'volition', frequently used by Wundt, Bühler or Titchener, really mean. Nor did he expect such a consensus to be ever reached scientifically. Instead, psychologists, if they want to become 'real' scientists, should focus on the control and prediction of observable behavior: "Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the *prediction* and *control* of behavior. Introspection forms no essential part of its methods [...]" (Watson, 1913, p. 158).

Like Samuelson's new foundation for a theory of decision making in economics, behaviorists treated the human mind as a black box that is not part of rigorous science. Researchers shifted their attention to stimulus-response patterns and psychological theories were exclusively built on observable and exactly measurable behavior. Doing research on mental states or constructs like memory, emotion or motivation became increasingly difficult at psychology faculties in the next decades. In the view of behaviorism, human behavior is the product of positive and negative reinforcement and punishment.[5] Even the most complex behavior, behaviorists believed, was fundamentally driven by these forces, forces that existed outside of the individuals mind. Human personality, emotions and habits were merely the product of the learned contingencies of their environment. Whatever is reinforced will be shown more frequently, while the rate of behavior that is followed by punishment will decrease. For the prediction and manipulation of behavior, this 'law of effect', as Edward Thorndike (1927) called it, was

[5]Economists today would call it incentives.

enough. In radical behaviorism inner psychological processes played no determinant role for behavior.

However, the predictive power of behaviorist theories was questioned and eventually rejected by a new generation of researchers. The first weaknesses of a theory that tries to explain behavior solely by reinforcement contingencies were already pointed out by Edward C. Tolman in the 1930s. In one of Tolman's experiments, hungry rats were placed in a maze and could freely explore it (Tolman and Honzik, 1930). One group always found a food pallet (the reinforcer) in a corner of the maze and learned to directly walk to this spot. In the second group no food was placed in the maze. After some trials, the experimenter also placed a food pallet in the maze of the second group. What Tolman and his colleague Honzik observed was that the second group learned to walk to the reinforcer much faster compared to the initial learning rate of the first group. Tolman concluded that rats learned the structure of the maze even when they were not enforced to do so.[6] Based on this observation, he questioned the conviction that behavior is reducible to reinforcement learning. More importantly, he proposed two concepts that pointed to the revival of inner processes in psychology: latent learning and the cognitive map (Tolman, 1948). For Tolman, learning can take place without any external reinforcement. He further concluded that the rats in his experiments learned more then just stimulus-reaction patterns. They learned the structure of the maze by constructing a mental representation of its architecture that they can use to plan and flexibly react to changes in the maze.

However, Tolman's much broader cognitive learning concepts were way ahead of its time. It took another decade and an interdisciplinary effort of psychology, linguistics and the emergent computer sciences to open the black box again and replace 'mindless' psychology with cognitive psychology, the dominant paradigm in psychology until today. Starting in the late 1950s the so-called 'cognitive revolution' begun and researchers rediscovered the importance of concepts like memory, motivation, volition or planning not only for predicting, but foremost for accurately describing human behavior. Latent variables, although only indirectly observable and measurable, were again introduced as legitimate parts of psychological theories. In cognitive psychology, influenced by the rise of computers and work on artificial intelligence, humans are conceptualized as sophisticated information processing machines, that not only react to but act in a complex environment. Compared to behaviorism, behavior is not only shaped by reinforcers or incentives. Instead we also actively manipulate and influence our environment. Therefore, in the cognitive research agenda, understanding the interplay of the environment and the mental representation that is guiding behavior plays a major role.

[6] Economists would say that there was no incentive to learn.

The tradeoff between application and validity

Throughout their history, economics and psychology both struggled with how to approach the content of the human mind. How can unobservable quantities like utility or the consciousness be studied rigorously, if they are not directly accessible through observation? Even worse, how can testable theories be formulated, if the objects of interest depend on how they are defined and what assumptions are made about them? After a period of reducing their focus on behavioral responses, psychologists today acknowledge these challenges but also recognize that a psychological theory without any reference to inner processes lacks descriptive validity. On the other hand, in economics inner processes slowly retreated and mainstream economic theory bears flavors of behaviorism. Utility serves merely as a theoretical construct and a theory is judged by its ability to predict not by its descriptive accuracy.

One might assume that a theory that accurately describes mental processes is also good in prediction. However, especially in the domain of mental constructs, accurate description and applying a theory to forecast actual behavior do not go hand in hand. This is one of the problems with Tversky's and Kahneman's prospect theory. Prospect theory contains three mental constructs: subjective value, a nonlinear subjective weighting of probabilities (decision weights) and the reference point. In EUT on the other hand, it is assumed that decisions adhere the probability axioms and are based on overall wealth levels, not on a mentally constructed reference point. As discussed above, a large body of empirical evidence shows that humans often violate the principles of probability theory when confronted with uncertain events and frame decisions based on situation-dependent reference points. Thus, prospect theory is more accurately describing the way how humans choose. The problem is, that one has to know quite a lot about a particular individual in order to predict her choices: How exactly does the individual distort probabilities? And more importantly, what mental reference point is employed in a particular situation? This is one of the reasons why economists are wary of theories that are based on mental processes.[7] The applicability of a theory has to be traded off against its descriptive validity. And as of today, psychology and economics settled on two different exchange rates. In psychology, theories often try to reconstruct all causes, effects, and the internal representation of the environment as accurately as possible, while in economics, theories are preferred that are mathematically tractable and can be easily used for prediction and application.

[7] Another reason is that theories that rely on mental processes are hard to test and falsify. This is the problem that gave rise to behaviorism as well as the decline of the cardinal concept of utility. In a nutshell, this is also the problem psychologists have with psychoanalytical theory.

However, the history of economic and psychological thought is not only a history of changing the ratio of accurate description and prediction power. If one believes in scientific progress, applicability and validity can increase at the same time with more knowledge and better methods. The American biologist Edward O. Wilson is also convinced that both can increase with disciplines joining forces and researchers engaging in interdisciplinary endeavors. Wilson, referring to the philosopher of science William Whewell, calls the convergence of different disciplines consilience (Wilson, 1999). For example, in the beginning of the 20th century, chemistry and physics started to converge and together built the foundations of modern biology, molecular biology. Consilience, according to Wilson, could also foster progress in the social sciences. For Wilson, the particularization we observed across scientific disciplines since the Enlightenment could reverse, thereby uniting the natural sciences (physics, chemistry, biology) with the social sciences (psychology, economics, sociology). Theories will evolve that increase descriptive validity and predictive power at the same time. Eventually Wilson envisions theories that encompass smooth transitions from atoms to cells, to organisms and ultimately to complex social behavior.

Many famous psychologists and economists in history, like Wundt or Jevons, who admired the success and progress in physics, would probably cherish this idea. But so far, the methods to open the black box and root psychological constructs in the biological realm were missing.

Filling the gap – neuroscience and sociobiology

A lot of hope to establish this missing link between the natural and the social sciences rests on modern neuroscience. With the emergence of imaging techniques like functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) enabling the noninvasive recording of brain activity, as well as transcranial magnetic stimulation (TMS), enabling the systematic manipulation of brain areas and observation of its effect on behavior, many psychologists and economists have started to explore the biological substrates of mental processes and decision making.

Another attempt to embed complex human behavior and cognition in the natural sciences is sociobiology or evolutionary psychology. Although humans are unique in many respects in the animal kingdom, the human mind is nevertheless a product of evolution. The premise that human cognition and behavior is shaped by evolutionary principles should therefore help to understand the rules and constraints of psychological systems.

In my thesis, I use neuroscientific methods as well as concepts of evolutionary biology in combination with economic theory to investigate human decision making. Similarly to its use in economics, I use the term decision making in a broad sense, from simple choice (like deciding between an apple and an orange) to complex social decisions (like helping a stranger or cooperating with a group of people). My **second chapter** is concerned with utility, the most central concept of economics that transformed over the centuries from a psychological driving force to a mathematical thought experiment. By going back to Bentham, the study investigates if neurobiological signals exist in our brain that resemble a hedonic calculator, representing subjective value across decision domains. The **third chapter** documents a decision making anomaly. We demonstrate experimentally that participants' purchasing decisions can be manipulated by constraining their arm movement. We propose that this seemingly irrational behavior might be explainable by the biological heritage and the natural environment in which humans evolved. The fourth and fifth chapters report findings in the domain of social decisions. In the **fourth chapter**, we show that a brain area in the frontal cortex is causally involved in the ability to strategically adapt behavior by dividing a pile of money more fairly when a punishment threat is present. The **fifth chapter** investigates cooperative behavior in groups. Inspired by anthropological and primate research, we show how the establishment of hierarchies among group members can foster cooperation and prevent free riding.

References

Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation*. Clarendon Press, Oxford.

Bühler, K. (1907). Tatsachen und Probleme zu einer Psychologie der Denkvorgänge. W. Engelmann, Leipzig.

Danziger, K. (1980). Wundt's psychological experiment in the light of his philosophy of science. *Psychological research*, 42(2):109–122.

Edgeworth, F. Y. (1881). *Mathematical Psychics, an Essay on the Application of Mathematics to the Moral Sciences*. Kegan Paul, London.

Friedman, M. (1953). *Essays in Positive Economics*. University of Chicago Press, Chicago.

Jevons, W. S. (1871). *The Theory of Political Economy*. Macmillian, London.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan, London.

Kahneman, D. and Tversky, A. (1972). Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology*, 3:430–454.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Külpe, O. (1920). *Vorlesungen über Psychologie*. S. Hirzel, Leipzig.

Mill, J. S. (1863). *Utilitarianism*. Parker, Son, and Bourn, London.

Pareto, V. (1909). *Manuel d'économie politique*. Giard & Brière, Paris.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17):61–71.

Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Harvard University Press, Cambridge.

Thorndike, E. L. (1927). The Law of Effect. *The American Journal of Psychology*, 39(1/4):212–222.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208.

Tolman, E. C. and Honzik, C. H. (1930). Introduction and Removal of Reward, and Maze Performance in Rats. *University of California Publications in Psychology*, 4:257–275.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20(2):158–177.

Wilson, E. O. (1999). *Consilience: The Unity of Knowledge*. Vintage, New York.

Part I

Simple Choice

Chapter 2

Value Signals in the Prefrontal Cortex Predict Individual Preferences Across Reward Categories

---

Abstract. Humans can choose between fundamentally different options, such as watching a movie or going out for dinner. According to the utility concept, put forward by utilitarian philosophers and widely used in economics, this may be accomplished by mapping the value of different options onto a common scale, independent of specific option characteristics. If this is the case, value-related activity patterns in the brain should allow predictions of individual preferences across fundamentally different reward categories. We analyze fMRI data of the prefrontal cortex while subjects imagine the pleasure they would derive from items belonging to two distinct reward categories: engaging activities (like going out for drinks, daydreaming, or doing sports) and snack foods. Support vector machines trained on brain patterns related to one category reliably predict individual preferences of the other category and vice versa. Further, we predict preferences across participants. These findings demonstrate that prefrontal cortex value signals follow a common scale representation of value that is even comparable across individuals and could, in principle, be used to predict choice.

Based on Jörg Gross*, Eva Woelbert*, Jan Zimmermann*, Sanae Okamoto-Barth, Arno Riedl, Rainer Goebel (2014). Journal of Neuroscience, 34(22), 7580-7586. *shared first authorship

2.1 Introduction

Life can be seen as a series of decisions. Often, comparisons have to be made between qualitatively very different options, such as going on a vacation or buying a car. Following the utility concept put forward by utilitarian philosophers and used by economists, it has been suggested that the brain performs decisions between such dissimilar options by assigning a value to each of them, which is mapped on a common scale of desirability regardless of the specific type of option (Montague and Berns, 2002; Fehr and Rangel, 2011; Levy and Glimcher, 2012). In line with the hypothesis of a common scale of subjective value, human functional magnetic resonance imaging (fMRI) studies have observed an overlap of value-related signals in the medial prefrontal cortex (mPFC) for different types of reward, such as consumer goods, monetary rewards, and also social rewards (Chib et al., 2009; FitzGerald et al., 2009; Lebreton et al., 2009; Smith et al., 2010; Kim et al., 2011; Levy and Glimcher, 2011; Lin et al., 2012). Further, it has been demonstrated that value signals for money and food options do not only spatially overlap, but that equally preferred money and food options elicit comparable BOLD responses in the mPFC both univariately (Levy and Glimcher, 2011) and multivariately (McNamee et al., 2013).

In this study we want to rigorously test the common scale hypothesis on three different grounds. First, if there is a common neural code for value, this should also be the case for goods that are rarely traded, rarely used as substitutes, and whose value cannot be easily expressed in monetary terms. Second, abstract value signals should be detectable in the absence of a monetary evaluation task. Third, if humans employ the same, possibly innate, coding mechanism for value, value signals should be comparable across individual participants.

We therefore use multivariate analysis of fMRI data and test whether it is possible to predict individual preferences across fundamentally different categories in the absence of monetary evaluation. In contrast to univariate analysis, multivariate analysis of fMRI data allows exploring whether value signals are not only spatially overlapping, but also encoded in a similar way, which is a prerequisite for a common scale representation. As reward categories, we chose snack foods and engaging activities because they differ fundamentally with respect to the sensory and motivational systems involved. Snack foods (like donut, cheesecake, chocolate etc.) constitute primary rewards that serve energy intake and are directly linked to gustatory perception, whereas the engaging activities serve the pursuit of secondary goals as diverse as socializing, relaxing, fitness, and culture (like going out for drinks, daydreaming, playing tennis, visiting a museum). In addition, these activities are rarely traded and it is therefore unlikely that subjects associate a monetary value with them.

2.2 Materials and methods

2.2.1 Subjects and task

Eight healthy subjects (5 female, 3 male, age 25–30 years) without prior history of psychiatric or neurological disorders participated in the study. All subjects gave their written informed consent before participation and the study protocol was approved by the Ethics Committee of the Faculty of Psychology and Neuroscience, Maastricht University. Subjects were asked to refrain from eating for 4 hours before the experiment.

During functional imaging, subjects were asked to imagine the pleasure they would derive from (1) eating different snack food items (e.g., potato chips, a blueberry muffin, or chocolate ice cream) or (2) engaging in different activities (e.g., listening to music, having a nap, or window cleaning). There were 60 items per category (see below). Items were presented in written form and random order to each subject for three seconds separated by a variable intertrial interval of 10–14 seconds (Figure 2.1a). Each item was presented two times over the whole fMRI session.

Immediately after scanning, outside the scanner, subjects made a series of hypothetical binary choices. In total, 500 item pairs were presented, both within and across categories with the instruction to choose the item that would yield more pleasure to the participant right now. This choice data were used to infer a preference ranking over all 120 items. We refer to these ranks as observed subjective values. After the binary choice task, participants rated their imagined pleasure for each item using a visual analog scale (Figure 2.1b). This allowed us to validate the preference ranking obtained from the binary choice task. Consistency across these two preference measures was high (Pearson's correlation ranged between $r = 0.75$ and $r = 0.92$).

2.2.2 Stimuli

Sixty items per category were presented to the participants. Both categories were constructed such that they included items that were potentially liked a lot, but also items that may be disliked. As activities we used for example "playing tennis," "jogging," "listening to music," "observing animals," "daydreaming," "sitting in a park," "fixing a bike," "cleaning windows," and "taking an exam." Examples for snack foods are "croissant," "waffle," "blueberry muffin," "chocolate cookie," "wasabi nuts," "paprika potato chips," and "salmiak." The full list of items is available from the authors upon request.

Measurements were performed on a 3T TIM Trio scanner (Siemens). Functional responses were measured in four independent runs using a whole brain standard gradient echo EPI sequence (GRAPPA = 2, TE = 30 ms, slices = 32, TR = 2000 ms, FOV =

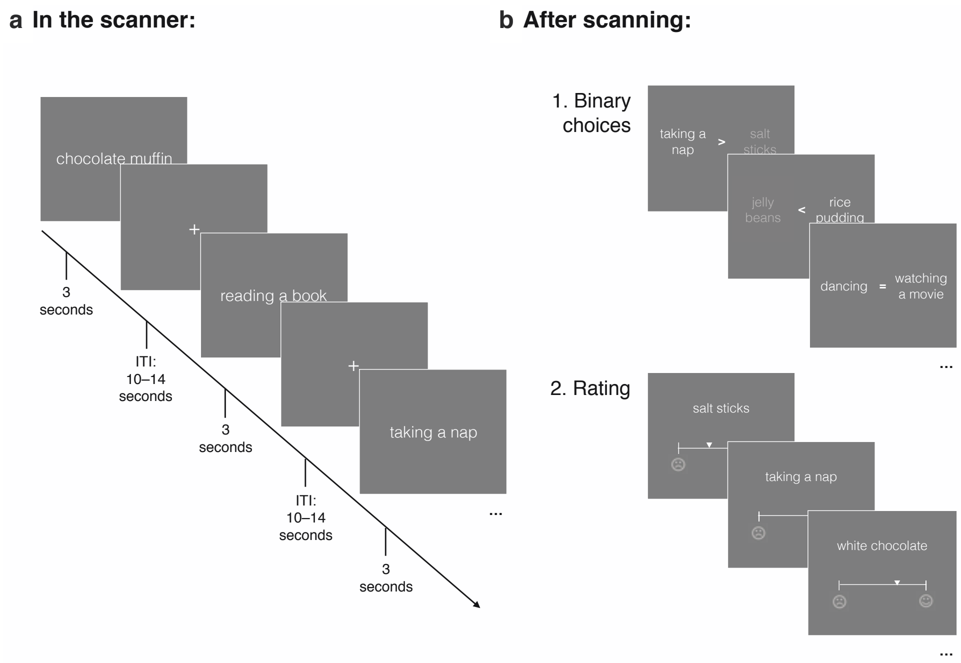


Figure 2.1: Experimental Procedure. (a) In the scanner participants were presented with 60 different activities and 60 different snack foods in written form, each presented for three seconds with the instruction to imagine how much pleasure they would derive from it. Each item was presented twice, resulting in 240 trials. (b) After scanning, participants made 500 binary choices between the presented items followed by a rating task in which they rated each of the 120 items on a visual analogue scale.

192×192 mm^2, voxel size = $3 \times 3 \times 3$ mm^3). A T1-weighted magnetization prepared rapid acquisition gradient echo (3D-MPRAGE, GRAPPA = 2, TR: 2050 ms, TE = 2.6 ms, FOV = 256×256 mm^2, flip angle = 9°, 192 sagittal slices, voxel size = $1 \times 1 \times 1$ mm^3) anatomical dataset was acquired for coregistration, segmentation and visualization of the functional results.

2.2.3 MRI data analysis

Functional and anatomical images were analyzed using BrainVoyager QX (Brain Innovation) as well as custom code in R (Team, 2008). Preprocessing of the functional data included interscan slice-time correction, rigid body motion correction, as well as temporal filtering using a Fourier basis set of five cycles per run. The functional data were then coregistered with the individual anatomical scan and transferred into Talairach space (Talairach and Tournoux, 1988). Anatomical data were segmented to identify gray matter for mask generation. All multivariate fMRI data analysis was performed on a single subject level and was restricted to an anatomically defined gray matter mask (Figure 2.2) entailing each subject's entire frontal cortex (corresponding to Brodman areas 9–13, 25, 32, 33, 46, 47). For comparison, univariate analysis was performed using the post-scanning observed subjective value as a linear parametric modulation of hemodynamic response.

2.2.4 Multivoxel pattern analysis

Feature creation. Each item was presented twice during scanning to improve data reliability. Functional data for each individual item was averaged across the two trials in which it was presented. Thus, for each item we obtained a single time course from each voxel by averaging across the two trials. This time course for each voxel was then collapsed across functional volumes to obtain the set of predictors (feature sets) for the model. Because participants did not perform any actions during fMRI but were instructed to just imagine subjective pleasure, it can be expected that the individual onset and duration of the BOLD response vary with participants. Therefore, for each subject multiple feature sets were created by computing the mean of the raw fMRI signal for slightly jittered time intervals. These time intervals differed with respect to the start-point after stimulus onset (1–4 volumes in steps of 1) and the length of the time interval (1–3 volumes in steps of 1). Thus 12 different feature sets were created for each participant (Figure 2.3a).

Feature selection. To select the best time interval for the final analysis, we evaluated model performance within one category for each feature set individually for each participant. For each feature set, recursive feature elimination (RFE) was used to reduce the number of voxels by 75%. Two linear support vector regression models (ϵ-SVR) were recursively fitted to the data. One model was fitted using all snack food items, the other one using all activity items (Figure 2.3b).

RFE was performed following Duan et al. (2005). In each RFE step, the dataset was randomly split into 10 mutually exclusive subsets (10-fold). In each subset, an ϵ-SVR was fitted to the data. Weight scores of the features were obtained and aggregated across the 10 subsets to stabilize the feature ranking (multiple SVM-RFE; Duan et al., 2005). Based on the aggregated feature ranking the 250 voxels with the lowest ranking scores were eliminated. For each model, this procedure was repeated until the 25% most informative voxels remained. Thus, multiple SVM-RFE was used to optimize the snack

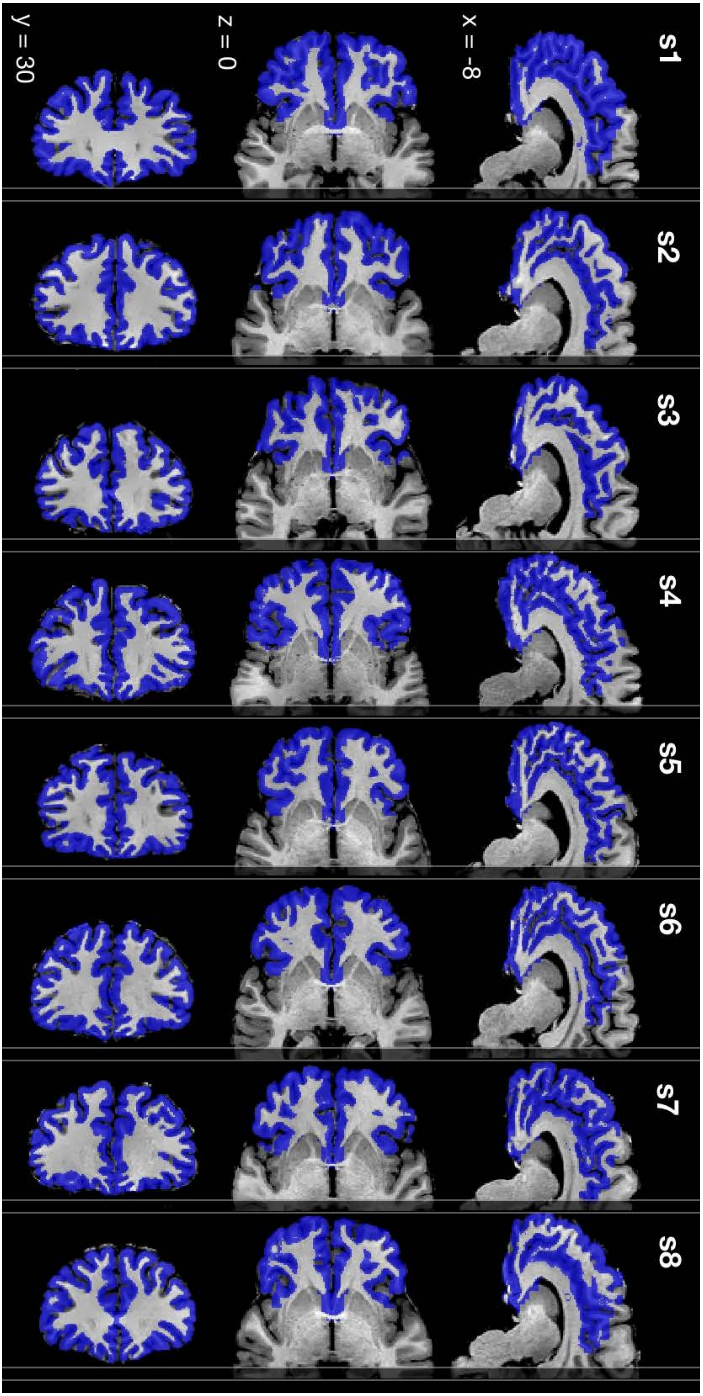


Figure 2.2: Frontal cortex masks. Anatomically defined gray matter mask including Brodman areas 9–13, 25, 32, 33, 46, 47 used in this study for each subject.

food model in predicting the value for other snack foods and the activity model in predicting the value for other activities. For each participant, the time interval underlying the feature set that yielded the highest mean performance of these two models was ultimately used for across-category predictions.

As within-category prediction performance is not the main focus of this paper, we do not report within-category performance measures, but only across-category performance. This approach circumvents any biases resulting from two-stage selection procedures that first pick the best-performing feature set and then fit a final model using the same data (Kriegeskorte et al., 2009; Vul et al., 2009).

Final model. Across-category performance was calculated at the very end of the analysis and was thus not part of any two-stage selection procedure. The SVR model trained on snack food items was used to predict the preference ranking over activities and vice versa (Figure 2.3c). Training and testing set were therefore entirely independent. Correlations between the observed preference ranks and predicted ranks were calculated to assess prediction accuracies between categories for each subject.

Permutation testing. Permutation testing was used to assess the statistical significance of the correlations between predicted and observed preference ranks. For each permutation, labels (in our case the rank-value) were randomly reassigned to the features. The data with permuted labels was then treated in exactly the same way as the original data. That is, the same multiple SVM-RFE procedure was used to reduce the number of voxels to the 25% most informative (within-category). On these voxels, the final model for across-category predictions was trained. This procedure was repeated 2000 times, yielding 2000 correlations between predicted and observed preference ranks based on randomly assigned rank-value. The p value for the actual correlation was determined based on this distribution.

Across subject models. For the across subject predictions we fitted one SVR model for each participant using all items ignoring item category (120 examples). Each SVR was then used to predict preferences over all items for the other seven participants, respectively. To not amplify the already existing differences in brain anatomy across participants, data were not further reduced by RFE for across subject predictions. Instead, all voxels in the individual frontal cortex mask were used.

Support vector regression machine. We used LIBSVM's (Chang and Lin, 2011) ϵ support vector regression machine as implemented in the 'e1071' library in R. Whenever a

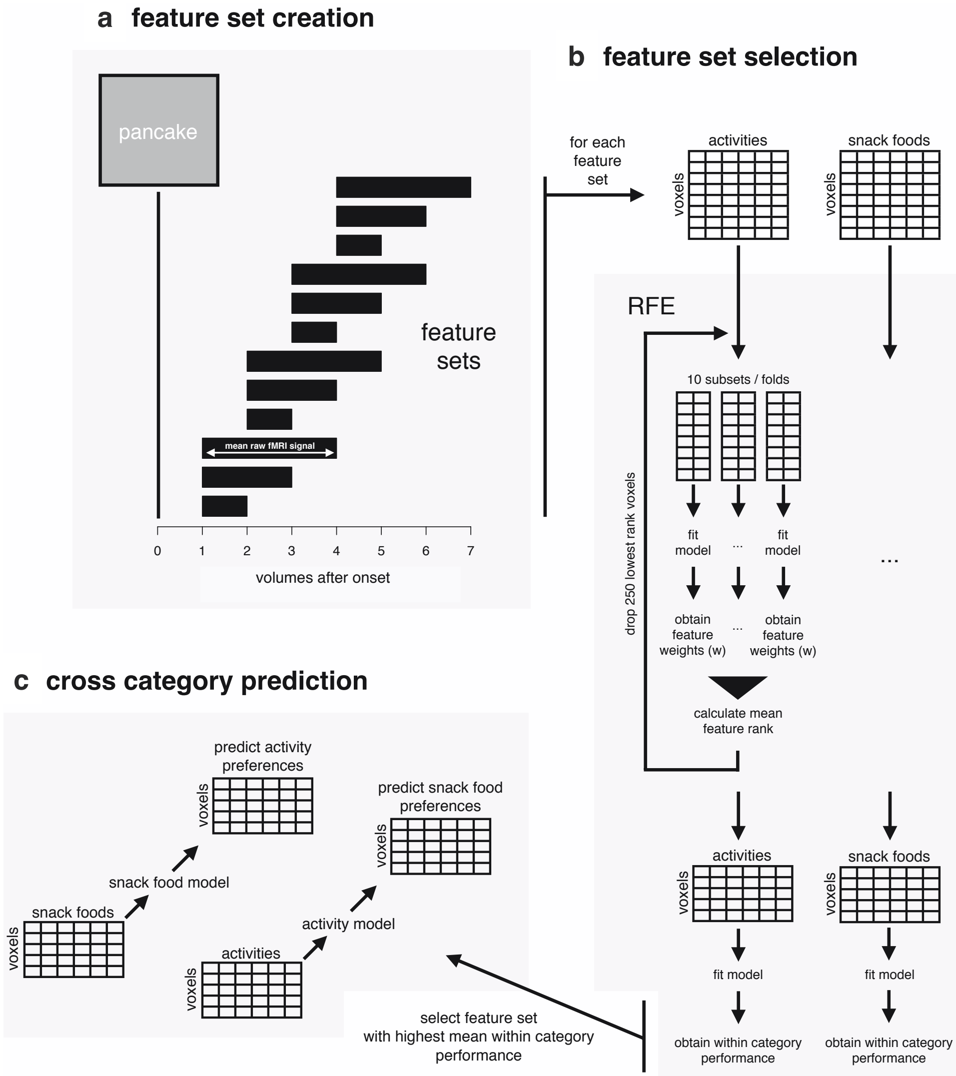


Figure 2.3: Multivoxel pattern analysis. Figure illustrating the steps of the main multivoxel pattern analysis employed in this study. (a) Twelve feature sets of different time intervals were created for each participant. (b) Recursive feature elimination (RFE) was used within each category for each feature set to obtain the 25% most informative voxels. (c) The feature set with the highest within-category performance was used to perform across category predictions by training a model on one item category (e.g., snack foods) and using it to predict preferences over items of the other category (e.g., activities) and vice versa.

SVR was fitted to the data it was first tuned with respect to the cost and ϵ parameter using tenfold cross-validation. For better interpretation of the feature weights a linear kernel was used.

2.3 Results

To test whether subjective value of snack foods and activities is represented on a common scale, we assessed whether similarly liked snack foods and activities evoke similar brain patterns. Therefore, after an SVR model was trained on one category of items (e.g., snack food), it was used to estimate the subjective value for items from the other category (e.g., activities). For each subject, we assessed each model's estimation performance in these across-category estimations by correlating the estimated subjective values with the observed subjective values obtained from the binary choices after scanning. Figure 2.4a illustrates these correlations for an exemplary subject.

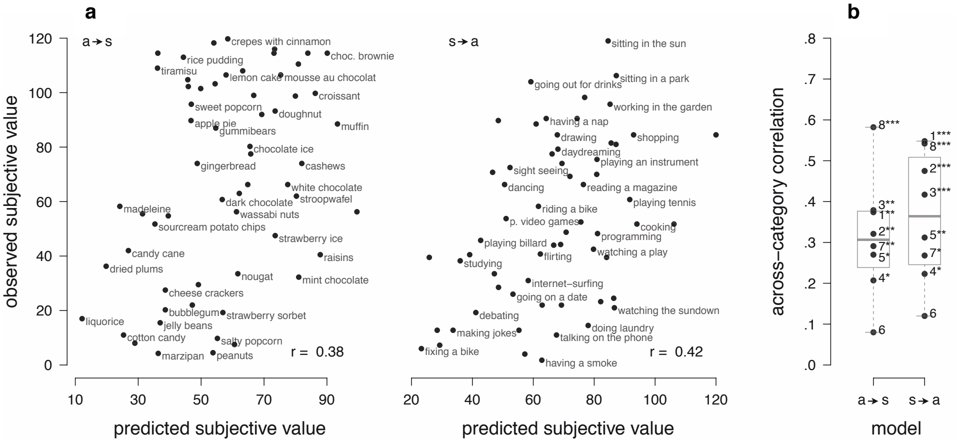


Figure 2.4: Estimation performance across reward categories. (a) Scatterplots of observed against estimated subjective values for one exemplary subject (subject 3) for snack foods predicted by a model trained on preferences over activities (a $\rightarrow$ s) and vice versa (s $\rightarrow$ a) and (b) correlation coefficients between observed and estimated subjective value for all eight subjects. Numbers identify the individual subjects. '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$ based on permutation testing.

Statistical significance was determined individually for each of these correlations based on permutation testing. Fourteen of 16 individual correlations were significantly higher than expected by chance on the 5% significance level. Figure 2.4b shows the estimation performance of each model for each participant. On average, estimated subjective values of activities, obtained by an SVR model trained on snack food items, correlated with observed subjective values by $r = 0.31$ (Pearson's correlation, $t(7) = 6.1$, $p < 0.001$, one sample t-test). Likewise, using the SVR model trained on activities, correlations between estimated and observed subjective values for snack food items reached an average of $r = 0.36$ (Pearson's correlation, $t(7) = 6.6$, $p < 0.001$, one sample t-test). These correlations demonstrate that value-related brain patterns are similar across categories.

As mentioned above, each item was presented twice during scanning and BOLD signals were averaged over presentations. We additionally analyzed each presentation set independently to investigate any systematic differences between the first and second presentation. For that we trained a model on BOLD signals recorded during the first presentation of each item from one category (e.g., snack foods) and used it to predict the value based on BOLD signals that were recorded during the first presentations of items from the other category (e.g., activities). We then compared this to the prediction accuracy of models that were trained on the second presentation of each item. Prediction performance was slightly lower when using the second trial (second presentation: mean $r = 0.27$, first presentation: mean $r = 0.33$, Pearson's correlation), presumably due to adaptation of the BOLD response or subjects' fatigue. However, in a regression analysis with prediction accuracy (correlations) as the dependent and a dummy variable coding for first versus second presentation and a dummy coding for whether the model was trained on snack foods (predicting activities) or activities (predicting snack foods), the difference between first and second presentation was not significant (coefficient = -0.06, $t(29) = -1.01$, $p = 0.32$).

2.3.1 Predicting choices

The estimated subjective values should not only carry information about which of two items will be preferred, but also about how strong the preference is. This can be tested by computing how well subjects' actual choices in the binary decision task can be predicted by using the estimated subjective values obtained from the brain data. For this, we simply assume that in a binary choice the item with the higher estimated subjective value will be chosen over the item with the lower value. Because there is noise in the estimations, choices between items with a similar estimated subjective value should be harder to predict correctly than choices between items where the estimated subjective values are very different. As shown in Figure 2.5, the proportion of correctly predicted binary choices indeed monotonically increased with the distance between the estimated subjective values of these items. The prediction accuracies were as high as 81% when subjects chose between items that were classified as highly disliked and highly liked by the SVR model.

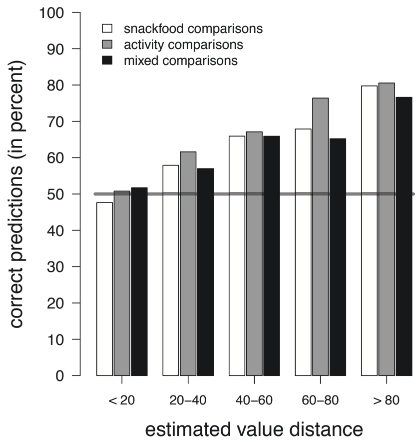


Figure 2.5: Correctly predicted binary choices, using the predicted subjective values. Bars show the percentage of correctly predicted binary choices depending on the distance between the estimated subjective values of two items. The horizontal line shows the accuracy expected by chance (50%).

Importantly, to show that the brain indeed encodes value on one common scale, estimated subjective values need not only be correlated with observed subjective values, as this merely indicates that items are ordered correctly. In addition, it is necessary to demonstrate that subjective values are estimated correctly in an absolute sense. To test for this, we assessed the accuracy of predicting binary choices that involved a snack food and an activity item. If estimated subjective values were only ordered correctly, then the prediction performance in these between-category choices would be worse than in choices involving only items from one category. We found that accuracies for binary choices between the two categories were not significantly different from the accuracies within each category ($\chi^2(2) = 0.41$, $p = 0.82$, Pearson χ^2 test). Thus, each SVR model did not only order items of a category correctly, as indicated by the significant correlations. Additionally, neither model over- or underestimated the subjective value ranks of the other category systematically, which further supports the hypothesis of a common currency of value.

2.3.2 Pattern and localization of value encoding

The multivariate analysis used in this study identified voxels which carry information about subjective value. In contrast to univariate analysis, it is not restricted to detecting aggregate levels of activity, but can also detect encoding which relies on more complex patterns. To better understand the pattern that encodes abstract subjective value, we analyzed the BOLD signal in the most informative voxels based on our RFE procedure. One possibility is that informative voxels exhibit increased activity with increasing subjective value. We therefore correlated the BOLD signal with observed subjective value. Aggregating across all voxels that survived RFE for each subject, average BOLD response was correlated moderately with subjective value ($r = 0.09$, Pearson's correlation). On the level of individual voxels, a majority of voxels (60%) showed a positive correlation but activity levels correlated with subjective value only with $r = 0.04$ (Pearson's correlation) on average. Another coding characteristic could be that an increasing number of voxels is activated as subjective value increases. To test for this possibility, we looked at the number of voxels contributing to coding of higher valued items. We first dichotomized each voxel's activity relative to its median BOLD value, so that each voxel that survived RFE had either value one or zero for each item: 0 if the BOLD signal during presentation of this item was below median signal magnitude of this voxel, and 1 if the BOLD signal was above the median BOLD value. This way, we could count how many voxels showed signals above the median in different observed value ranges. For low valued items (value rank 1–20), an average 48% of voxels showed a BOLD signal above the median. This steadily increased up to 53% for items that were most preferred (value rank 100–120). On average, with each increase of 10 value ranks, nine additional voxels showed BOLD signal above its median ($t(48) = 8.7$; $p < 0.01$, random-intercept regression). Thus, with higher subjective value an increasing number of voxels showed above-median BOLD response.

To localize the voxels in the frontal cortex that carry information about subjective value regardless of the reward category, we first identified the most informative voxels in each SVR model. For each participant, voxels were classified as carrying category-independent information if they were informative in both models. In Figure 2.6 we plot the overlap of these category-independent voxels across participants, revealing clusters in the ACC and the medial prefrontal cortex. The map shows only voxels with minimally 25% overlap (minimally 2 participants) and clusters with at least 162 anatomical (i.e., 6 functional) voxels. The highest overlap was observed in a cluster in the anterior portion of the mPFC with an overlap of 62.5% (5 participants). A list of peak voxel coordinates of all identified clusters can be found in Table 2.1.

Table 2.1
Brain regions carrying value-signals across reward categories.

	X	**Y**	**Z**	**# participants**	**# voxels**
right anterior mPFC	1	61	0	5	231
right dorsal mPFC	2	52	36	4	204
right mPFC	5	55	24	3	219
ACC 1	5	29	21	4	162
ACC 2	-7	37	31	4	169
left dorsal PFC	-7	52	36	3	202
left mPFC 1	-4	55	0	4	214
left mPFC 2	-6	58	18	3	279
left anterior PFC	-20	61	13	3	195

Note. Clusters are listed if they show an overlap for at least 2 participants and a cluster size of at least 162 anatomical (= 6 functional) voxels. No smoothing was applied to the data. Coordinates correspond to Talairach coordinates of the voxel with the highest overlap across subjects.

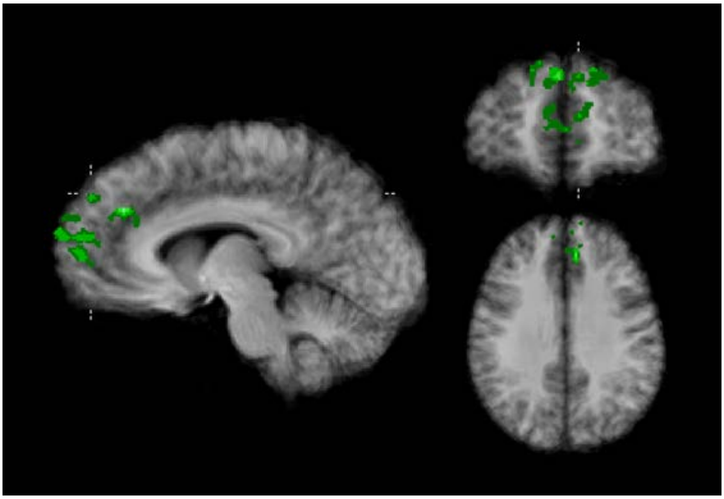


Figure 2.6: Voxel-clusters carrying across-category information based on the weights of the support vector regression models (Talairach x, y, z = -7, 51, 39) overlaid onto an anatomical average.

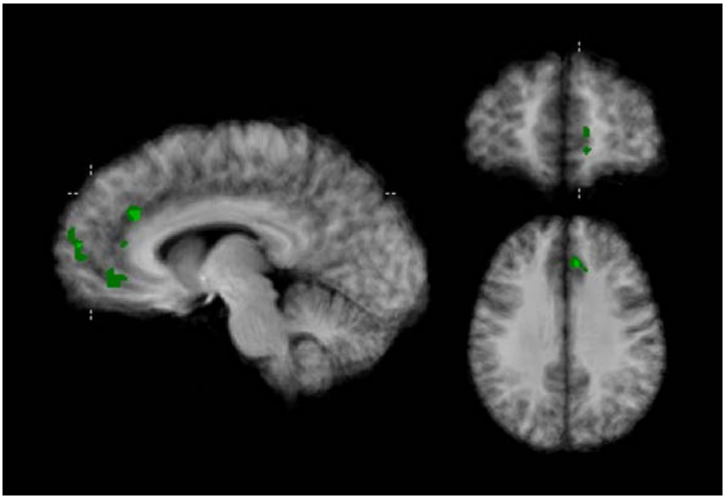


Figure 2.7: Voxel-clusters resulting from univariate analysis (Talairach x, y, z = -7, 51, 30) overlaid onto an anatomical average.

Surprisingly, among the 25% most informative voxels, there was no cluster in the ventral part of the mPFC. This region has been frequently implicated in common value computation in previous fMRI studies that used univariate data analysis (Levy and Glimcher, 2012). We therefore performed an additional univariate analysis by fitting observed subjective value to the hemodynamic response function's estimates in each participant. To maximize comparability with the multivariate results, we used the same probabilistic procedure as described above to aggregate the single subject maps. For each subject, we ranked the absolute model fits (betas) to reduce the number of voxels considered to 25% of subjects' frontal cortex mask, thus obtaining the same number of voxels per subject as resulted from the multivariate recursive feature elimination procedure. Figure 2.7 shows the overlap of the remaining voxels across subject (minimally 25% overlap; i.e., 2 participants). The obtained pattern is comparable to the multivariate patterns presented in Figure 2.6, but additionally includes a cluster in the more ventral part of the MPFC. We would like to stress that this result should be interpreted with caution because no significance standards were met by this analysis.

Given the spatial overlap of voxels carrying abstract subjective value across subjects, we hypothesized that value-representation across participants could be sufficiently similar to allow predicting subjective value of one person with a model trained on another person. To test this, we trained one SVR model for each participant, using all 120 items. We then used each participant's model to predict the preferences

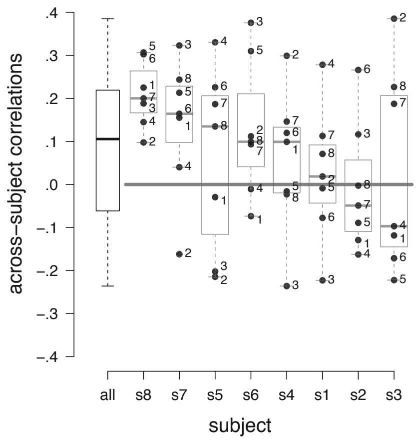


Figure 2.8: Prediction performance across participants. Each boxplot shows the prediction accuracy for an SVR model trained on one subject, predicting the preferences for all 120 items for the other seven subjects. Boxplot on the far left shows distribution of all 56 correlations obtained in this way. Bars in boxes show the median.

of the other seven participants, respectively, and correlated these with the observed preferences. Figure 2.8 summarizes all of the resulting 56 correlation coefficients in a boxplot, and also shows the coefficients by individual subject. Correlations were significantly above zero ($t(48) = 3.11$, $p < 0.01$, random-intercept regression), demonstrating that value patterns are to some extent comparable, even across subjects. As can be expected, given the anatomical variability across participants, performance was considerably lower (mean Pearson's correlation $r = 0.10$) than within individual across-category performance.

2.4 Discussion

We tested whether value signals in the frontal cortex are independent of specific reward characteristics and can therefore be used to infer the subjective value of inherently distinct reward types. A machine learning algorithm, which was trained only in decoding subjective value of snack food items, was capable of predicting the subjective value of engaging activities and vice versa. Hence, knowing the neuronal pattern that is associated with imagining the pleasure of, e.g., eating a donut or potato chips made it possible to reliably infer preferences over, e.g., playing tennis or shopping, and vice versa. This suggests that value signals in mPFC do not only spatially overlap, but also that the distributed pattern is similar across our two categories of reward. This is remarkable, since these categories differ fundamentally with respect to the associated sensory and motivational systems. If subjective value representation was entirely linked to specific properties of a stimulus this should not be possible.

Our activity category comprises items that are normally not traded and hence not priced. Therefore, our finding cannot be explained by cognitive processes that merely reflect learned market prices. Further, in contrast to previous studies on common-scale representation of value (Levy and Glimcher, 2011; McNamee et al., 2013), we refrain from using any type of monetary valuation task. It is thus unlikely that the similarity of value signals we observe across different categories is artificially caused by a task that encourages subjects to evaluate items in a common reference frame, such as monetary value or a Likert scale. Instead our data suggest that a common scale representation of subjective value is an inherent feature of our valuation system, much like the psychological concept of utility, first envisioned by Jeremy Bentham as anticipated pleasure or general satisfaction (Kahneman et al., 1997; Bentham, 2007).

Interestingly, we find that to a certain extent this representation is common even across individual brains, as is evident in the significant across-participant predictions. A machine-learning algorithm trained on brain patterns of one individual, when applied to the brain patterns measured in another participant, predicted subjective preferences above chance level. Although prediction power was considerably lower, knowing the neuronal pattern that is associated with imagining the pleasure of, e.g., eating a donut or playing tennis for Participant A makes it in principle possible to infer subjective preferences over these items for Participant B.

To explore the localization of the abstract value signals that enabled the across-category predictions, we mapped the voxels which were informative in predicting subjective value for both reward categories. Note that we did not restrict this analysis to a small region, but considered voxels from the entire frontal cortex of each individual participant. Given this large mask, it is very encouraging that we observe clusters of informative voxels in the mPFC in line with previous human fMRI studies (Tusche et al., 2010; Levy and Glimcher, 2012; McNamee et al., 2013). Our multivariate analysis, however, did not reveal a cluster of voxels in the most ventral part of the mPFC, where many previous studies reported peak voxels. Instead, our most prominent cluster corresponds well to a region reported to predict consumer choices using a similar multivariate approach as ours (Tusche et al., 2010). This discrepancy could be due to a number of fundamental differences between classical univariate analyses and the analysis presented in this study. Whereas most previous studies used a smoothing kernel, our results were obtained on a single-subject level without any image filtering. The multivariate results depicted in Figure 2.6 represent the probabilistic overlap instead of a multi-subject aggregate. Neighboring voxels that are usually smoothed in common approaches might therefore not survive our set inclusion threshold. Discrepancies between univariate and multivariate analyses could also arise because of distributed patterns in the data that influence the machine learning algorithm when picking the most informative voxels, but leave the univariate procedure unaffected (Haufe et al., 2014).

In addition to the medial prefrontal cortex, we find voxels carrying abstract value signals in the ACC. This is interesting, because value signals in the ACC have been reported in monkey studies using single-cell recordings (Wallis and Kennerley, 2010; Kennerley et al., 2011; Cai and Padoa-Schioppa, 2012), but not typically in human fMRI studies. The fact that we do observe value-signals in the ACC is likely due to the differential sensitivity of multivariate analysis compared with univariate analysis (Wallis, 2011), and suggests that abstract value signals are not limited to the mPFC.

In contrast to previous studies (Levy and Glimcher, 2011; McNamee et al., 2013), our tasks were of entirely hypothetical nature and without any reference to a monetary or another numerical frame. This design allowed us to employ a wide range of fairly abstract stimuli, and minimized the possibility that subjects evaluate items of different categories in terms of an externally imposed reference frame. It might still be argued that the revealed patterns are not generalizable to value computations during real everyday decisions. Previous research has shown, however, that imagined and experienced rewards elicit overlapping patterns (Bray et al., 2010). This suggests that hypothetical evaluation and evaluation during actual choices rely on similar mechanisms.

The possibility remains that subjects used their own reference frame regardless of the instruction to focus on the pleasure they would derive from an item. Such a reference frame could be, for example, the minimum/maximum pleasure they can imagine, and although unlikely for our activities category, it is even conceivable that subjects might spontaneously adopt a monetary evaluation scheme. Our across subjects predictions however provide evidence that even if this is the case, the value signal contributing to our models is abstract enough to generalize over such subject specific strategies. Future research could further clarify this question by refraining from giving any task instructions like imagining pleasure, or even using a distractor task while displaying different items, an approach that has been used for example by Tusche et al. (2010).

To conclude, our results provide strong evidence for the existence of abstract value signals in the mPFC and also the ACC. These value signals are comparable even for fundamentally different and immaterial categories of reward, and in principle, can be used to predict choice.

References

Bentham, J. (2007). *An Introduction to the Principles of Morals and Legislation*. Dover, Mineola, New York.

Bray, S., Shimojo, S., and O'Doherty, J. P. (2010). Human Medial Orbitofrontal Cortex Is Recruited During Experience of Imagined and Real Rewards. *Journal of Neurophysiology*, 103(5):2506–2512.

Cai, X. and Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *The Journal of Neuroscience*, 32(11):3791–3808.

Chang, C. and Lin, C. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27.

Chib, V. S., Rangel, A., Shimojo, S., and 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.

Duan, K.-B., Rajapakse, J. C., Wang, H., and Azuaje, F. (2005). Multiple SVM-RFE for gene selection in cancer classification with expression data. *IEEE Transactions on NanoBioscience*, 4(3):228–234.

Fehr, E. and Rangel, A. (2011). Neuroeconomic foundations of economic choice—recent advances. *The Journal of Economic Perspectives*, 25:3–20.

FitzGerald, T. H. B., Seymour, B., and Dolan, R. J. (2009). The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects. *The Journal of Neuroscience*, 29(26):8388–8395.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., and Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, 87:96–110.

Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of Experienced Utility. *The Quarterly Journal of Economics*, 112(2):375–406.

Kennerley, S. W., Behrens, T. E. J., and Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, 14(12):1581–1589.

Kim, H., Shimojo, S., and O Doherty, J. (2011). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cerebral cortex*, 21(4):769–776.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, 12(5):535–540.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. *Neuron*, 64(3):431–439.

Levy, D. J. and Glimcher, P. (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. and Glimcher, P. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, 22:1027–1038.

Lin, A., Adolphs, R., and Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, 7(3):274–281.

McNamee, D., Rangel, A., and O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, 16(4):479–485.

Montague, P. R. and Berns, G. S. (2002). Neural Economics and the Biological Substrates of Valuation. *Neuron*, 36(2):265–284.

Smith, D. V., Hayden, B. Y., Truong, T.-K., Song, A. W., Platt, M. L., and Huettel, S. A. (2010). Distinct Value Signals in Anterior and Posterior Ventromedial Prefrontal Cortex. *The Journal of Neuroscience*, 30(7):2490–2495.

Talairach, J. and Tournoux, P. (1988). Co-planar Stereotaxic Atlas of the Human Brain. *Thieme Medical: New York*.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing.

Tusche, A., Bode, S., and Haynes, J. D. (2010). Neural Responses to Unattended Products Predict Later Consumer Choices. *The Journal of Neuroscience*, 30(23):8024–8031.

Vul, E., Harris, C., Winkielman, P., and Pashler, H. (2009). Puzzlingly high correlations in fmri studies of emotion, personality, and social cognition. *Perspectives on Psychological Science*, 4(3):274–290.

Wallis, J. D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience*, 15(1):13–19.

Wallis, J. D. and Kennerley, S. W. (2010). Heterogeneous reward signals in prefrontal cortex. *Current Opinion in Neurobiology*, 20(2):191–198.

Supporting Information

Stimuli

Activities

“playing tennis,” “jogging,” “horse-riding,” “hiking,” “riding a bike,” “playing soccer,” “fitness training,” “having a nap,” “sitting in the sun,” “listening to music,” “observing animals,” “daydreaming,” “sitting in a park,” “watching the sundown,” “having a smoke,” “shopping,” “going for a walk,” “reading a book,” “reading a newspaper,” “reading a magazine,” “drawing,” “painting,” “playing computer games,” “playing an instrument,” “watching a movie,” “watching a play,” “working in the garden,” “cooking,” “baking,” “cleaning the house,” “programming,” “gossiping,” “playing board games,” “playing cards,” “dancing,” “making jokes,” “going to a party,” “going out for drinks,” “talking on the phone,” “playing billiard,” “playing with children,” “going on a date,” “flirting,” “sight-seeing,” “going to a lecture,” “solving puzzles,” “visiting a museum,” “debating,” “internet-surfing,” “writing a letter,” “driving a car,” “washing dishes,” “doing laundry,” “studying,” “giving a presentation,” “fixing a bike,” “cleaning windows,” “vacuuming,” “solving Sudoku’s,” “taking an exam”.

Snack foods

“strawberry sorbet,” “crepes with cinnamon,” “croissant,” “waffle,” “doughnut,” “blueberry muffin,” “chocolate muffin,” “cupcake,” “pancake,” “stroopwafel,” “chocolate brownie,” “chocolate cookie,” “speculaas,” “gingerbread,” “butter cookie,” “milk chocolate,” “dark chocolate,” “white chocolate,” “mousse au chocolat,” “truffles,” “nougat,” “mint chocolate,” “peanuts,” “pistachios,” “cashews,” “chocolate ice,” “vanilla ice,” “strawberry ice,” “wasabi nuts,” “raisins,” “dried plums,” “salty popcorn,” “sweet popcorn,” “apple pie,” “cheesecake,” “lemon cake,” “tiramisu,” “crème brulee,” “vanilla pudding,” “rice pudding,” “tortilla chips,” “salt sticks,” “cheese crackers,” “sourcream potato chips,” “paprika potato chips,” “jelly beans,” “gummibears,” “wine gums,” “candy cane,” “bubblegum,” “lollipop,” “sour candy,” “cotton candy,” “marshmallow’s,” “marzipan,” “fried banana,” “madeleine,” “cinnamon roll,” “liquorice,” “salmiak”.

Experimental instructions

In every trial of the following experiment, one item will be presented to you on the screen for approximately 3 seconds. An item can be a food item, such as “muffin”, or an activity you can engage in, such as “playing tennis”.

In every trial, please *imagine how much pleasure you would derive from this item right now*. That is, for food items please imagine the pleasure you would have while eating this food item, for activities please imagine the pleasure you would have while engaging in it.

Supplemental figures

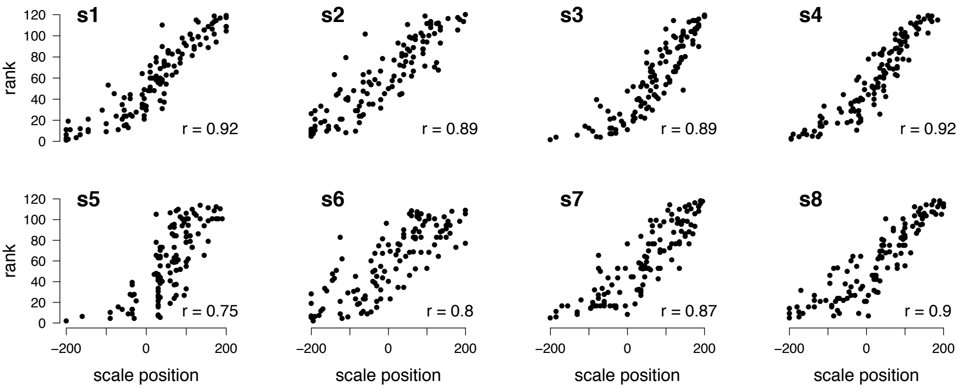


Figure 2.S1: Consistency across preference measures. Scatterplots showing the association between preference ranks, inferred by binary decisions, and the pleasure ratings made on a visual analogue scale for each participant.

Chapter 3

The Fox and the Grapes – How Physical Constraints Affect Value Based Decision Making

Abstract. One fundamental question in decision making research is how humans compute the values that guide their decisions. Recent studies showed that people assign higher value to goods that are closer to them, even when physical proximity should be irrelevant for the decision from a normative perspective. This phenomenon, however, seems reasonable from an evolutionary perspective. Most foraging decisions of animals involve the trade-off between the value that can be obtained and the associated effort of obtaining. Anticipated effort for physically obtaining a good could therefore affect the subjective value of this good. In this experiment, we test this by letting participants state their subjective value for snack food while the effort that would be incurred when reaching for it was manipulated. Even though reaching was not required in this experiment, we find that willingness to pay was significantly lower when subjects wore heavy wristbands on their arms. Thus, when reaching was more difficult, items were perceived as less valuable. Importantly, this was only the case when items were physically in front of the participants but not when the items were presented as text on a computer screen. Our results suggest automatic interactions of motor and valuation processes which are unexplored to this date and may account for irrational decisions that occur when reward is particularly easy to reach.

Based on Jörg Gross*, Eva Woelbert*, Martin Strobel (under review). *shared first authorship

3.1 Introduction

Driven by hunger, a fox tried to reach some grapes hanging high on the vine but was unable to, although he leaped with all his strength. As he went away, the fox remarked, "Oh, you aren't even ripe yet! I don't need any sour grapes."

– Aesop's fable

Every day we make choices between different options based on our preferences. Research on decision making suggests that this is accomplished by assigning a subjective value to each decision option (Montague and Berns, 2002; Levy and Glimcher, 2012). However, there is ample evidence that human decisions are influenced by seemingly irrelevant aspects of the choice situation (Tversky and Kahneman, 1981). One fundamental question in decision making research is therefore how humans compute the values that guide their decisions. Recent studies suggest that the subjective value of a good may depend on how this good is presented to the person at the time of the decision.

For example, Reb and Connolly (2007) asked subjects how much money they would exchange for chocolate bars and coffee mugs. Subjects were either told that they owned the good, or not, and the good was either physically in front of the subject or not. The physical presence of the good, but not ownership status, significantly increased the monetary value the subjects assigned to the goods. A similar result has been observed by Knetsch and Wong (2009). The authors randomly assigned one of two goods to subjects, and asked them whether they would like to swap it for the other good. Some subjects were physically possessing the good at the time of decision, others not. Subjects who were in physical possession of the good were biased towards keeping it, whereas subjects that were not in physical possession of the good were equally likely to keep it than to swap it for the alternative. In a study by Bushong et al. (2010), participants were asked how much they would be willing to pay for different snack food items under several display conditions. Subjects were willing to pay less when items were presented as pictures or words on the computer screen than when the items were physically present. This lower willingness to pay was also observed in a fourth condition when snack food items were physically present (at the same distance as before) but put behind a 9 by 9 feet Plexiglas barrier. Thus, it seems that the physical presence of a good increases valuation, but only if the good is within immediate reach.

In any of these experiments, obtaining an item was solely based on the stated preferences. According to normative theories of decision making (e.g. Von Neumann and Morgenstern, 1944; Debreu, 1972), the physical presence of the object is an irrelevant detail of the choice situation and should not influence the valuations or decisions. Therefore, the results of these experiments are quite puzzling. In this study we propose and test an interpretation of these findings, drawing from theories and empirical observations in the field of grounded cognition. Theories of grounded cognition posit that processes like memory, language, perception and decision making are not executed in encapsulated modular systems but are deeply interrelated (Gibson, 1986; Wilson, 2002; Barsalou, 2008). For example, a substantial body of research suggests that the visual perception of our surrounding is not just the product of the visual information entering the visual stream but also non-visual factors like the perceiver's physiological state, capabilities, and intentions (Proffitt, 2006). It has been demonstrated that hills appear steeper when subjects are fatigued (Proffitt et al., 1995), or are wearing a heavy backpack (Bhalla and Proffitt, 1999; but see also Woods et al., 2009). Likewise, older people perceive walkable distances as longer than younger people (Sugovic and Witt, 2013), and distances appear larger when overcoming them is associated with higher effort for the perceiver (Witt et al., 2004; Ramenzoni et al., 2008; Lourenco and Longo, 2009; Morgado et al., 2013). If, on the other hand, a tool is available that facilitates reaching an object, a decrease in distance to this object is perceived (Witt et al., 2004; Longo and Lourenco, 2006).

Based on these empirical findings it has been argued that judgments of physical properties of the surrounding environment, such as length, height, and slope, are influenced by the motor actions that would be afforded to overcome them (Cisek, 2007; Halász and Cunnington, 2012). Climbing a hill is energetically costlier when burdened with heavy load. According to Proffitt (2006), this burden is automatically integrated in how we perceive the hill. When energetic resources are scarce or effort costs are high, hills appear steeper and distances wider. Thus, our representation of the world is to some degree altered by transient changes in our capabilities to interact with it and the anticipated metabolic costs of doing so (Proffitt, 2006; Halász and Cunnington, 2012; Brockmole et al., 2013).

It seems plausible that not only perceptual judgments, but also preferences and decisions may be influenced by bodily states. In line with the broader literature on embodiment of attitudes and emotions (Strack et al., 1988; Tom et al., 1991; Niedenthal, 2007), it has been shown that items are evaluated more favorably when the bodily state signals agreement (e.g. nodding the head), or approach (e.g. flexing the arm), both of which frequently co-occur with positive stimuli (Wells and Petty, 1980; Cacioppo et al., 1993; Priester et al., 1996; Förster, 2004; Labroo and Nielsen, 2010).

From the grounded cognition perspective outlined above, anticipated effort costs could affect the valuation process in a similar fashion as perceptual judgements. The decrease in willingness to pay Bushong et al. (2010) observed when placing a barrier between the participant and the object could then be explained as follows: An object that is behind a physical barrier is more costly to reach. It takes more effort to make a grasping action circumventing the obstacle and therefore items placed behind a barrier would be perceived as less valuable, even when obtaining the item does not require any movement.

Here we test this hypothesis by directly manipulating the effort it would take to physically obtain an item. More precisely, we test whether increasing the effort it would take to reach for an item leads to lower valuations even when reaching is not required.

3.2 Materials and methods

3.2.1 Participants and materials

In individual sessions, 54 undergraduate students were presented with 44 different snack food items consisting of candy bars (e.g. "Mars" or "Snickers"), potato chips (e.g. "Lays"), gummibears, crackers, nut mixes, or licorice with the chance to buy one of the items at the end of the experiment. All of these items were available at local convenience stores (see Figure 3.S1 and Table 3.S1 in the Supporting Information for all items). The study protocol was approved by the Ethics Committee Psychology of the Faculty of Psychology and Neuroscience at Maastricht University, and all participants gave written informed consent. Participants were recruited from the subject pool of the behavioral and experimental economics lab (BEElab) and were invited via e-mail using the software ORSEE (Greiner, 2004). Before starting the experimental task, participants were endowed with 14.50 euro to compensate them for participation in the experiment and to make sure they had money to spend on snack food.

Since we were interested in measuring the value participants assigned to different snack food items, it was important that participants liked snack food and were motivated to obtain some during the experiment. The invitation e-mail therefore clearly stated that the experiment involved snack food and that one part of the compensation for participation was monetary, and another in the form of snack food. To further ensure that participants were motivated to obtain snack foods, they were asked to refrain from eating for 3 hours prior to the beginning of the experiment. Participants were also instructed that they would be asked to stay in an adjacent room for 30 minutes after completion of the task. During these 30 minutes they were only allowed to eat whatever snack food they bought in the experiment. This was done to limit the influence of the market price at which participants could acquire the snack food item immediately after the experiment, and is standard practice in incentivized decision experiments (for example see Plassmann et al., 2007; Bushong et al., 2010). Before the actual task, participants received detailed instructions, comprehension questions, and practice trials to ensure good understanding of the task.

3.2.2 Experimental procedure

Upon completion of the instruction and training part, snack food items were presented one by one to the participant. For each item, participants first indicated the maximum amount of money between 0 and 4 euros they would be willing to pay to receive this item after the experiment. We used a Becker, DeGroot and Marschak auction (BDM, Becker et al., 1964) and each willingness to pay stated by the participants was potentially relevant for their earnings.

The BDM auction was implemented as follows: At the end of the experiment one of the 44 items was randomly selected by letting the participant draw a numbered card out of a deck of 44 cards. For the selected snack food item a selling price between 0 and 3.99 euro was randomly determined by the participant by rolling three dice (one four sided die and two ten-sided dice). If the stated willingness to pay was below the randomly determined selling price, the participant did not buy the good. If the stated maximum willingness to pay was equal to or above the selling price, then the participant bought the good from us at the randomly determined selling price.

With this mechanism, it is in the participants' best interest to state their true maximum willingness to pay. This is because with their stated willingness to pay participants could only influence the chances to buy the item but not the selling price. Stating a higher willingness to pay than the true willingness to pay makes it more likely to buy the item, but it also increases the likelihood that the selling price is above the true willingness to pay. In this case the participants have to buy although they do not want to buy. In a similar manner understating the true willingness to pay may lead to situations where the item is not bought, despite the selling price being lower than the true willingness to pay.

The BDM mechanism is widely used in decision making research (Plassmann et al., 2007; Johnson et al., 2007). It resembles a real buying decision by measuring how much money participants are willing to give up to attain an item. The willingness to pay is our main measure of interest.

Participants also provided psychological measures of subjective value, namely liking ("Please imagine you would eat this item right now. How much do you think you would enjoy it?"), and wanting ("How much do you want to receive this item at the end of the experiment?") ratings on a four point Likert scale. It has been suggested that liking and wanting are closely related, but neurobiologically dissociable constructs (Berridge and Robinson, 2003; Berridge, 2009). We were therefore interested to see whether they might be differentially affected by the weight manipulation. Note that, even if this is the case, willingness to pay, liking and wanting should all reflect the subjective value of an item, and we expect them to be highly correlated. In addition, participants indicated the familiarity ("How well do you know this item?") of each item on a four point Likert scale. Familiarity was measured to be able to control for the fact that subjects might exhibit a higher valuation for more familiar items. Participants saw each item only once, and always made their willingness to pay decision first, followed by the three remaining questions in random order. Subjects were encouraged to answer spontaneously, but there were no time restrictions and subjects could complete the task in their own pace.

After the evaluation of all 44 snack food items, participants provided demographic information, answered questionnaires on food craving and impulsiveness for exploratory purposes, and were asked to indicate what they thought was the purpose of the weights and the hypothesis of the experiment. Before leaving the laboratory, subjects were debriefed about the purpose of the experiment.

3.2.3 Experimental manipulations

Physical effort. We manipulated the effort associated with making a reaching movement by attaching wristbands around the lower arms of our participants (see Figure 3.1). No explanation was given about the nature or the purpose of the wristbands. Each wristband could hold 10 metal bars, weighing 4.5 kilograms (9.9 pounds) in total (see Figure 3.1). For half of the items, each participant wore the wristbands with the bars (weight condition), for the other half of the items they wore the wristbands without the bars (no weight condition). Participants were instructed to sit such that the elbows were supported by the armrests of the chair and the wrists rested on their thighs. This was done to ensure that they were aware of the weight, but did not exert effort to hold the weight. The wristbands with weight would have made it significantly more demanding for the participants to grasp an item placed in front of them, but at no point during the experiment did they actually reach for an item, nor were they instructed to imagine any movements. Participants were provided with a small numeric keypad they held in their hands, which allowed them to enter their willingness to pay, liking, wanting and familiarity ratings without moving their arms.

The order of the weight conditions was counterbalanced across participants. To reduce the influence of strong preferences for certain types of snack food (like chocolate, licorice or fruit gums), the allocation of the 44 snack food items to the weight and no weight condition was not fully randomized. Instead, we formed pairs of comparable snack foods (like two types of chips or two different chocolate bars) and never presented both elements of a pair in the same weight condition (see Figure 3.S1 and Table 3.S1 in the Supporting Information). This ensured that the same amount of each type of snack foods was presented in both conditions to the participant. Which particular chocolate

Figure 3.1: Wristbands used in the experiment. In the weight condition each wristband held ten metal bars, which add up to 4.5 kilograms (9.9 pounds) of weight. In the no weight condition all metal bars were removed.

bar or chips bag was presented in the weight or no weight condition was randomized across participants.

Reachability. Between subjects we manipulated the reachability of the items. Participants in the physical condition had each food item placed physically in front of them at the time of decision making. Thus, during value judgements, the item was reachable with a simple arm movement. In the control condition, which we call the computer condition, items were not physically present and thus not physically reachable, but were presented to the participants as text on the computer screen. Participants in the computer condition were instructed that snack food items were normal package sizes, as available in the supermarket. The same items were used in the computer condition and the physical condition, with the exception of five items that were not available in stores when the computer condition was run. These were replaced with similar items. Note that our primary interest lies in the within-subject comparison.

When the snack food is physically reachable, it is possible to physically obtain the item with an arm movement. This arm movement would be more effortful when wearing weights. If these anticipated effort costs influence the valuation process, subjective valuation and willingness to pay in the physical condition will be lower in the weight condition compared to the no weight condition. In the computer condition items are not reachable by arm movements and thus not physically obtainable. Therefore, the increased effort of making an arm movement by having heavy wristbands around the lower arm should not influence the valuation process, and we expect no difference of our physical effort manipulation on valuation in the computer condition.

While it might seem a more natural choice for the control condition to present pictures of the snack food items on the computer screen, we chose not to do so because pictures of physical stimuli have some spatial properties, and have been found to affect cognitive processes based on the depicted physical properties (e.g. di Pellegrino et al., 2005), whereas this seems not to be the case for words (Helbig et al., 2006). Also note that any confounding effect of the weight on value judgements, such as stress or discomfort should still be present in the computer condition, while an effect that is specific to anticipated effort of reaching for the snack food should not.

3.3 Results

Our final data set contained decisions from 24 participants (12 female, 12 male, mean age 22.6, SD = 3.0 years) in the physical condition and from 26 participants (13 female, 13 male, mean age 22.4, SD = 2.7 years) in the computer condition. Two participants in the physical condition had to be excluded from the analysis because of a computer malfunction. From the computer condition two participants were excluded, one who bid zero for all items, indicating that he was not motivated to buy any snack food, the other participant reported that he changed his bidding strategy halfway through the experiment. Note that including these participants in the data analysis does not change any significance level of our hypothesis tests reported below.

In the physical condition, when snack food items were physically in front of subjects, participants reported an average willingness to pay of 1.33 euro in the no weight condition. However, the same participants decreased their average willingness to pay by 10 cents when they had heavy weight on their arms (paired samples t-test, $t(23) = 3.74$, $p < 0.01$, two-sided, mean Cohen's $d = 0.26$, see Figure 3.2). We did not observe this difference in valuation between the weight and the no weight condition when the item was not physically present but instead presented as text on a computer screen. Here, participants slightly increased their willingness to pay on average by 4 cents when wearing heavy weight (paired samples t-test, $t(25) = -0.81$, $p = 0.43$, two-sided).

From each subject, we obtained responses for 44 snack food items. Thus, our data set is hierarchically structured: Item evaluations are nested within subjects. To account for this, we fitted a random intercept regression model separately for each of the dependent variables (willingness to pay, liking, wanting). In this hierarchical linear model, each subject contributes 44 data points (1 response for each item), and the

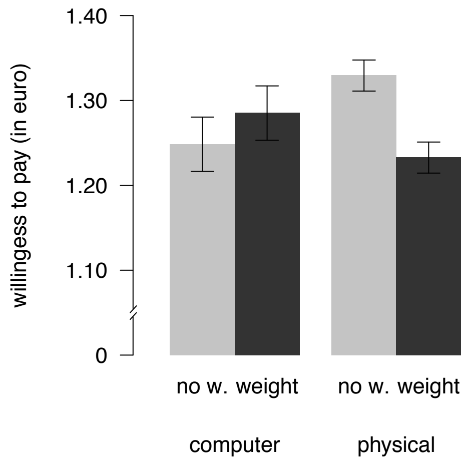


Figure 3.2: Effect of weight on willingness to pay separately for physical and computer condition. Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008) and are therefore only informative for evaluating within-subject differences between wearing wristbands with weight (black) vs. no weight (grey).

model accounts for these repeated measures by estimating a subject-specific intercept (Cohen et al., 2002; Gelman and Hill, 2007). By refraining from aggregating the data on the subject level, we are able to include control variables on the item level (in our case the familiarity of the item). In the regression models we also control for possible order effects of the weight manipulation (see Table 3.1 for willingness to pay). As a measure of effect size we report d, computed by dividing the regression coefficient by the item-level standard deviation obtained in the respective regression model (Schagen and Elliot, 2004). Aggregating the data on the subject level, as it is often done in ANOVA analysis revealed similar results (see Table 3.S2 in the Supporting Information).

In line with the reported differences above, participants in the physical condition showed a significant negative slope in willingness to pay across our weight manipulation (Physical $\times$ Weight coefficient = -0.15, p = 0.02, two-sided; d = 0.23), while there was no significant change in valuation in the computer condition across the anticipated effort manipulation (Weight coefficient = 0.02, p = 0.76, two-sided, see Table 3.1). This interaction was due to a significant decrease of willingness to pay in the physical condition when wearing heavy wristbands (Wald Test, chi2(1) = 22.21 $p < 0.01$; d = 0.20). When items were physically in front of participants, wearing heavy wristbands around their arms therefore significantly decreased willingness to pay, while this was not the case when the items were presented on a computer screen.

As can be seen in Figure 3.3, wanting and liking ratings were similarly affected by the weight manipulation (see Tables 3.S3 and 3.S4 in the Supporting Information for the full regression results on liking and wanting). The interaction of effort and presentation mode was statistically significant for wanting (random intercept regression controlling for familiarity and potential order effects, Physical $\times$ Weight coefficient = -0.21, p = 0.03, two-sided; d = 0.24) and marginally significant for liking ratings (Physical $\times$ Weight coefficient = -0.13, p = 0.08, two-sided; d = 0.15), however, Wald Tests directly comparing weight vs. no weight in the physical condition did not reach significance for liking and wanting ratings.

To test whether willingness to pay, liking, and wanting measure closely related constructs, we computed how strongly they were correlated for each subject. Correlations were generally high (willingness to pay and wanting: mean r = 0.60 , $p < 0.01$, range = [-0.12, 0.96]; willingness to pay and liking: mean r = 0.57, $p < 0.01$, range = [-0.03, 0.97]), and highest for wanting and liking (mean r = 0.81, $p < 0.01$, range = [0.21, 0.97]).

Lastly, we explored whether items in the physical condition were affected differently by the weights manipulation, based on how much subjects liked the items. For that, we performed a median-split of the data based on liking ratings, and again com-

Table 3.1

Random intercept regression with control variables.

Dependent variable: willingness to pay.

	Coef.	95% CI	p
constant (computer condition, no weight)	1.16	[0.78, 1.54]	< 0.01
physical condition	0.00	[-0.50, 0.50]	0.99
weight	0.02	[-0.10, 0.13]	0.76
physical $\times$ weight	-0.15	[-0.28, -0.03]	0.02
familiarity	0.13	[0.09, 0.18]	< 0.01
order of weight condition	-0.35	[-0.79, 0.10]	0.12
order $\times$ weight	0.02	[-0.15, 0.19]	0.81
order $\times$ physical	0.20	[-0.39, 0.80]	0.50
order $\times$ physical $\times$ weight	0.07	[-0.13, 0.26]	0.50

Note. 2200 trials, nested within 50 subjects. Standard errors are corrected for potential heteroscedasticity and autocorrelations at the subject level. All p values are two-sided. Order indicates whether participant started with the weight (Order = 1) or no weight condition (Order = 0).

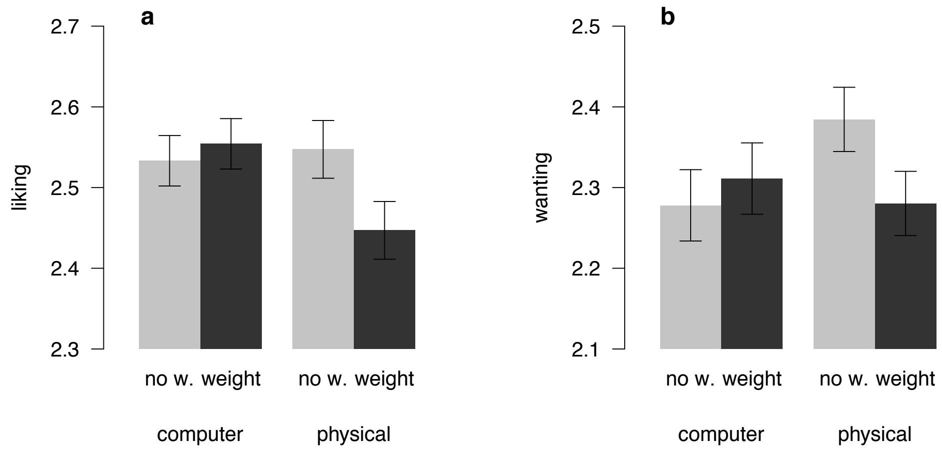


Figure 3.3: Effect of weight on liking (a) and wanting ratings (b) separately for physical and computer condition. Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008) and are therefore only informative for evaluating within-subject differences between wearing wristbands with weight (black) vs. without weight (grey).

pared willingness to pay across the two weight conditions. Note that liking ratings were also affected by the weight manipulation. To account for this, the median split was applied separately for each subject and each weight-condition, so that we compare the most liked half of the items in one condition with the most liked half in the other condition. Interestingly, the effect of the weights was much more pronounced for the better liked items, than for the less liked items (see Figure 3.4). As can be expected based on the high correlation between liking and wanting, a similar effect emerged when data was median-splitted based on wanting ratings (see Figure 3.S4 in the Supporting Information).

After the experiment, participants were asked to indicate what they thought was the purpose and the hypothesis of the experiment. In the physical condition, six subjects indicated a directed hypothesis, four of which hypothesized an effect in the opposite direction of our hypothesis. In the computer condition, four subjects indicated a directed hypothesis, two in each direction. Thus, we consider a demand effect resulting from correctly guessing the hypothesis unlikely.

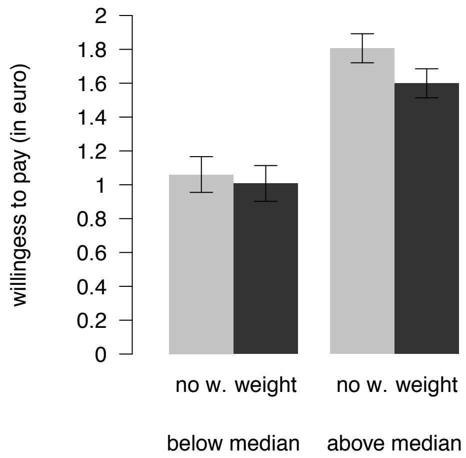


Figure 3.4: Effect of weight on willingness to pay for items with liking ratings below (left) or above (right) median-liking within its weight-condition. Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008) and are therefore only informative for evaluating within-subject differences between wearing wristbands with weight (black) vs. without weight (grey).

3.4 Discussion

Here we showed that physical constraints influence subjective value and actual purchasing decisions. Attaching heavy wristbands to participants' arms decreased willingness to pay and subjective valuation for different snack food items, even though participants in the experiment never actually had to reach out for an item, nor were they asked to imagine any arm movements. Our findings suggest that the anticipated metabolic costs of obtaining an item are automatically integrated in the value computation process when the item is physically reachable. Liking and wanting ratings were similarly affected by the weight manipulation, and we did not find evidence that liking and wanting are dissociable constructs in this task.

Importantly, the wristbands affected willingness to pay only when goods were physically in front of the subject, and not when goods were merely presented on a computer screen. Thus, our data suggest that the change in metabolic costs or effort due to the weights for making an arm movement only influence valuation when items are in fact reachable. In addition, we observed that the effect of weights on monetary valuation was most pronounced for items that were more liked.

From a normative perspective on decision making this finding is puzzling. Grasping the object was not required at any point in the experiment, and subjects knew that obtaining an item was not dependent on physical effort, but only on their stated willingness to pay. Consequently, such irrelevant aspects as having weights attached to the lower arms, should not influence the decision. In the following, we seek to explain why it might be biologically plausible that the physical constraints do influence value perception nevertheless.

In this experimental setting, as is often the case in modern daily life, choices occurred dissociated from physical action. In a supermarket, whether we can afford a product does not depend on physical constraints but on the amount of money we carry in our pocket. Similarly, we can order food online, making the decision even less dependent on physical action. However, apart from modern human society, whether something is easy to obtain or not is highly relevant. For example, any foraging decision in animals involves a trade-off between the value that can be obtained and the associated metabolic costs (Krebs and Stephens, 1986). Imagine a bear looking at a beehive high above in a tree or a group of lions tracking the movement of a gazelle. Whether it is optimal to climb the tree or to hunt the gazelle does not solely depend on the value of the goal, but also on the costs incurred by exerting the effort. When two actions yield similarly valued outcomes, the action that is less costly is the natural choice (Walton et al., 2006).

Because in natural environments, anticipated effort is so central to any decision, processes underlying the computation of goal value and effort costs could be closely intertwined. As a result, computing the value of a good may be automatically influenced by the anticipated effort of obtaining the good. Something that is easy to reach would then be attributed a higher value and perceived as more valuable because it is easy to grasp, even in choice situations in which grasping the object is not necessary to obtain it, like in a purchasing decision.

Some direct support for this hypothesis comes from an experiment by Beilock and Holt (2007). In this experiment, both skilled and novice typists made a series of binary choices between different printed letter pairs, each time choosing the letter pair they preferred. They found that only skilled typists showed a preference for letter pairs that can be easily typed, such as 'FJ', over pairs that are more difficult to type (e.g. 'FV'), whereas novices had random preferences. This effect of physical effort on preference vanished when skilled typists were instructed to hold a typing pattern in memory, presumably diminishing the capacity to compute the effort associated with typing the letter pair. Interestingly, this was only the case if the typing pattern that subjects needed to remember actually engaged the same fingers as required to type the letter pairs. This suggests that effects of physical effort on valuation and preference are indeed mediated by simulation of the relevant action.

Also, a substantial body of neurophysiological and neuroimaging data points at a strong interaction of value computation and motor-related processes (Cisek and Kalaska, 2010). The brain constantly keeps track of whether objects are within reach (Gallivan et al., 2009), and simulates movements that are feasible to manipulate the immediate environment, even in the absence of any intention or explicit goals (Tucker and Ellis, 1998; Cisek and Kalaska, 2005; Cisek, 2007; Baumann et al., 2009). Throughout the decision making process, decision-variables such as subjective value are represented not only in brain regions that have been directly linked to value computation and decision making, such as the ventromedial prefrontal cortex (Plassmann et al., 2007; Wallis, 2007; Padoa-Schioppa, 2011), but are also found in brain areas that are linked to motor preparation and execution (Platt and Glimcher, 1999; Gold and Shadlen, 2007; Louie et al., 2011; Klein-Flügge and Bestmann, 2012; Gluth et al., 2013). A recent neuroimaging study lends support to our interpretation that effort is automatically integrated with value (Kurniawan et al., 2013). When subjects received a positive outcome after exerting either high or low effort, activity in the ventral striatum, commonly interpreted as an evaluation of the outcome in the form of a reward prediction error, was higher after low effort. Thus, effort exerted to obtain an outcome was immediately integrated with the outcome value into a net-value. Our results suggest that such integration occurs to some extent also when assessing the value of the reward before obtaining it and even when physical actions are actually not required.

It seems plausible that anticipated effort of obtaining a good is most relevant if this good is highly liked, since it is more likely that the person will actually want to engage in a physical action to obtain the good. If, on the other hand, a good is not liked, anticipated effort should be less relevant and automatic motor processes might be triggered to a lesser extent. In line with this, we found that the more liked items were more strongly affected by the weights manipulation. However, since liking evaluations were obtained during our weight manipulation, this result should be interpreted with caution. Future research should aim to confirm this finding, ideally measuring liking independently of the experimental manipulation.

In summary, we show that irrelevant motor constraints affect value based decisions. Items that were more difficult to grasp were perceived as less valuable. This suggests that processes underlying the computation of value and effort are closely intertwined to an extent that anticipated effort for physically obtaining a good is automatically integrated in the value computation process, explaining why goods appear more valuable to subjects when physically present (Reb and Connolly, 2007; Knetsch and Wong, 2009), but not if reaching for them is obstructed by a barrier (Bushong et al., 2010). This mechanism is unexplored to this date and could help to explain maladaptive human behavior that occurs when reward is easy to reach, such as overeating or impulsive purchases.

References

Barsalou, L. W. (2008). Grounded Cognition. *Annual review of psychology*, 59(1):617–645.

Baumann, M. A., Fluet, M.-C., and Scherberger, H. (2009). Context-specific grasp movement representation in the macaque anterior intraparietal area. *The Journal of Neuroscience*, 29(20):6436–6448.

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

Beilock, S. L. and Holt, L. E. (2007). Embodied Preference Judgments Can Likeability Be Driven by the Motor System? *Psychological Science*, 18(1):51–57.

Berridge, K. C. (2009). 'Liking' and 'wanting' food rewards: Brain substrates and roles in eating disorders. *Physiology & Behavior*, 97(5):537–550.

Berridge, K. C. and Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, 26(9):507–513.

Bhalla, M. and Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of experimental psychology. Human perception and performance*, 25(4):1076–1096.

Brockmole, J. R., Davoli, C. C., Abrams, R. A., and Witt, J. K. (2013). The World Within Reach Effects of Hand Posture and Tool Use on Visual Cognition. *Current directions in psychological science*, 22(1):38–44.

Bushong, B., King, L. M., Camerer, C. F., and Rangel, A. (2010). Pavlovian Processes in Consumer Choice: The Physical Presence of a Good Increases Willingness-to-Pay. *American Economic Review*, 100(4):1556–1571.

Cacioppo, J. T., Priester, J. R., and Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65(1):5–17.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 362(1485):1585–1599.

Cisek, P. and Kalaska, J. F. (2005). Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. *Neuron*, 45(5):801–814.

Cisek, P. and Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience*, 33(1):269–298.

Cohen, J., West, S. G., and Aiken, L. (2002). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum Assoc Inc., New Jersey.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1):42–45.

Debreu, G. (1972). *Theory of Value. An Axiomatic Analysis of Economic Equilibrium*. Yale Univ. Press, New Haven, CT.

di Pellegrino, G. G., Rafal, R. R., and Tipper, S. P. S. (2005). Implicitly Evoked Actions Modulate Visual Selection: Evidence from Parietal Extinction. *Current Biology*, 15(16):1469–1472.

Förster, J. (2004). How Body Feedback Influences Consumers' Evaluation of Products. *Journal of Consumer Psychology*, 14(4):416–426.

Gallivan, J. P., Cavina-Pratesi, C., and Culham, J. C. (2009). Is That within Reach? fMRI Reveals That the Human Superior Parieto-Occipital Cortex Encodes Objects Reachable by the Hand. *The Journal of Neuroscience*, 29(14):4381–4391.

Gelman, A. and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge.

Gibson, J. J. (1986). The Ecological Approach To Visual Perception. Lawrence Erlbaum, New Jersey.

Gluth, S., Rieskamp, J., and Büchel, C. (2013). Classic EEG motor potentials track the emergence of value-based decisions. *NeuroImage*, 79:394–403.

Gold, J. and Shadlen, M. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30(1):535–574.

Greiner, B. (2004). An Online Recruitment System for Economic Experiments. In Kremer, K. and Macho, V., editors, *Forschung und wissenschaftliches Rechnen. GWDG Bericht*, pages 79–93. Göttingen.

Halász, V. and Cunnington, R. (2012). Unconscious Effects of Action on Perception. *Brain Sciences*, 2(4):130–146.

Helbig, H. B., Graf, M., and Kiefer, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, 174(2):221–228.

Johnson, E. J., Häubl, G., and Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3):461–474.

Klein-Flügge, M. C. and Bestmann, S. (2012). Time-Dependent Changes in Human Corticospinal Excitability Reveal Value-Based Competition for Action during Decision Processing. *The Journal of Neuroscience*, 32(24):8373–8382.

Knetsch and Wong (2009). The endowment effect and the reference state: Evidence and manipulations. *Journal of Economic Behavior and Organization*, 71(2):407–413.

Krebs, J. R. and Stephens, D. W. (1986). *Foraging Theory*. Princeton University Press, Princeton.

Kurniawan, I. T., Guitart-Masip, M., Dayan, P., and Dolan, R. J. (2013). Effort and Valuation in the Brain: The Effects of Anticipation and Execution. *The Journal of Neuroscience*, 33(14):6160–6169.

Labroo, A. A. and Nielsen, J. H. (2010). Half the thrill is in the chase: Twisted inferences from embodied cognitions and brand evaluation. *Journal of Consumer Research*, 37(1):143–158.

Levy, D. J. and Glimcher, P. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, 22:1027–1038.

Longo, M. R. and Lourenco, S. F. (2006). On the nature of near space: Effects of tool use and the transition to far space. *Neuropsychologia*, 44(6):977–981.

Louie, K. K., Grattan, L. E. L., and Glimcher, P. (2011). Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex. *The Journal of Neuroscience*, 31(29):10627–10639.

Lourenco, S. F. and Longo, M. R. (2009). The plasticity of near space: Evidence for contraction. *Cognition*, 112(3):451–456.

Montague, P. R. and Berns, G. S. (2002). Neural Economics and the Biological Substrates of Valuation. *Neuron*, 36(2):265–284.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2):61–64.

Morgado, N., Gentaz, É., Guinet, É., Osiurak, F., and Palluel-Germain, R. (2013). Within reach but not so reachable: Obstacles matter in visual perception of distances. *Psychonomic Bulletin & Review*, 20(3):462–467.

Niedenthal, P. M. (2007). Embodying emotion. *Science*, 316(5827):1002–1005.

Padoa-Schioppa, C. (2011). Neurobiology of Economic Choice: A Good-Based Model. *Annual Review of Neuroscience*, 34(1):333–359.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience*, 27(37):9984–9988.

Platt, M. L. and Glimcher, P. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400:233–238.

Priester, J. R., Cacioppo, J. T., and Petty, R. E. (1996). The influence of motor processes on attitudes toward novel versus familiar semantic stimuli. *Personality and Social Psychology Bulletin*, 22(5):442–447.

Proffitt, D. R. (2006). Embodied Perception and the Economy of Action. *Perspectives on Psychological Science*, 1(2):110–122.

Proffitt, D. R., Bhalla, M., Gossweiler, R., and Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin & Review*, 2(4):409–428.

Ramenzoni, V. C., Riley, M. A., Shockley, K., and Davis, T. (2008). Carrying the height of the world on your ankles: encumbering observers reduces estimates of how high an actor can jump. *Quarterly journal of experimental psychology (2006)*, 61(10):1487–1495.

Reb, J. and Connolly, T. (2007). Possession, feelings of ownership and the endowment effect. *Judgment and Decision Making*, 2(2):107–114.

Schagen, I. P. and Elliot, K. (2004). *But What Does It Mean? The Use of Effect Sizes in Educational Research.* NFER, Slough.

Strack, F., Martin, L. L., and Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 54(5):768–777.

Sugovic, M. and Witt, J. K. (2013). An older view on distance perception: older adults perceive walkable extents as farther. *Experimental Brain Research*, 226(3):383–391.

Tom, G., Pettersen, P., Lau, T., Burton, T., and Cook, J. (1991). The Role of Overt Head Movement in the Formation of Affect. *Basic and Applied Social Psychology*, 12(3):281–289.

Tucker, M. and Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of experimental psychology. Human perception and performance*, 24(3):830–846.

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

Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.

Wallis, J. D. (2007). Orbitofrontal Cortex and Its Contribution to Decision-Making. *Annual Review of Neuroscience*, 30(1):31–56.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., and Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks*, 19(8):1302–1314.

Wells, G. L. and Petty, R. E. (1980). The Effects of Over Head Movements on Persuasion: Compatibility and Incompatibility of Responses. *Basic and Applied Social Psychology*, 1(3):219–230.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4):625–636.

Witt, J. K., Proffitt, D. R., and Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception*, 33:577–590.

Woods, A. J., Philbeck, J. W., and Danoff, J. V. (2009). The various perceptions of distance: An alternative view of how effort affects distance judgments. *Journal of experimental psychology. Human perception and performance*, 35(4):1104–1117.

Supporting Information

Stimuli

Physical condition

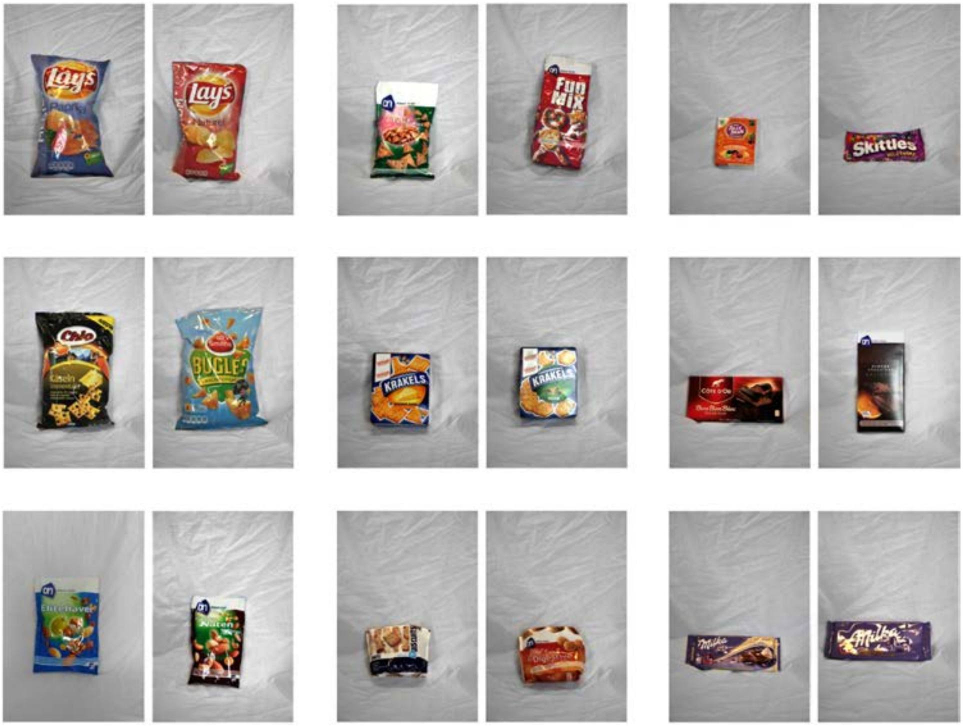


Figure 3.S1: Pictures of the snack food items used in the physical condition. Snack foods near to each other comprise pairs. Elements of a pair were never presented in the same condition within one participant. One was presented in the heavy weight, the other one in the low weight condition. Which item was presented in which condition was randomized across participants.

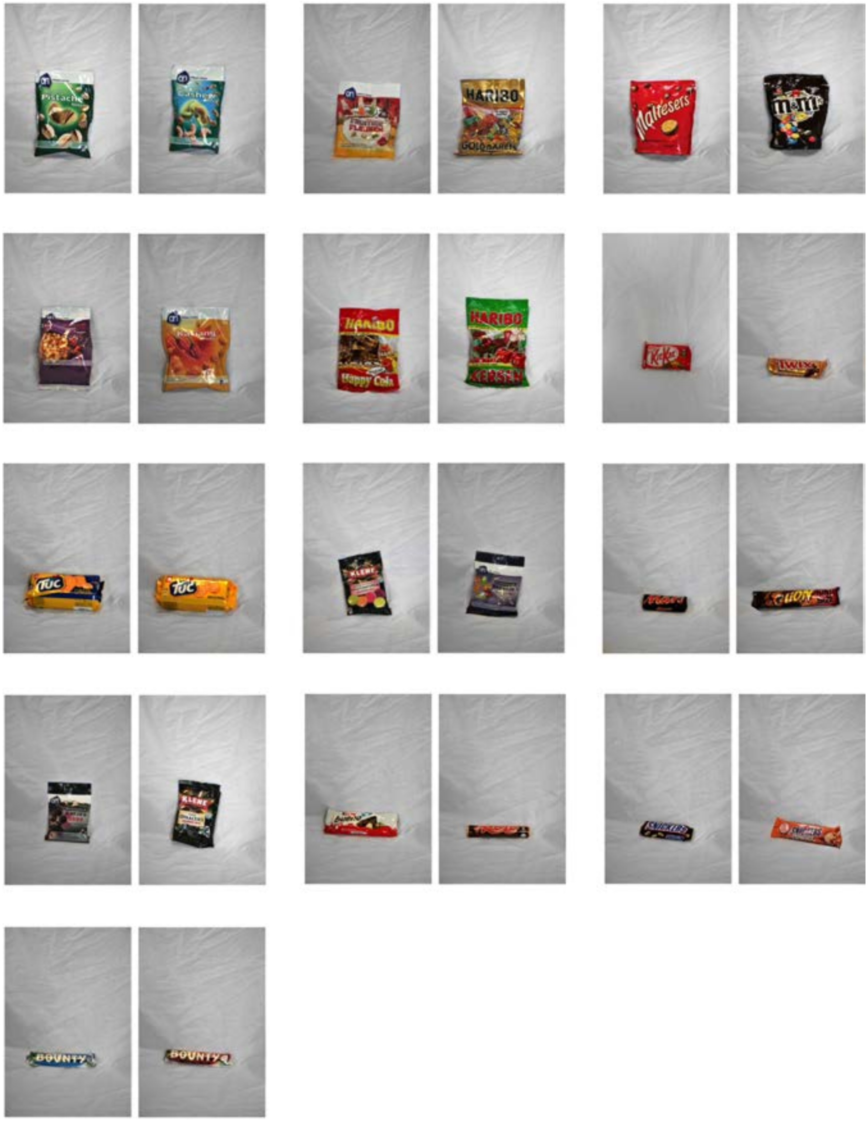


Figure 3.S1 (continued): Pictures of the snack food items used in the physical condition. Snack foods near to each other comprise pairs. Elements of a pair were never presented in the same condition within one participant. One was presented in the heavy weight, the other one in the low weight condition. Which item was presented in which condition was randomized across participants.

Computer condition

Table 3.S1

Stimuli used in the computer condition. Names and descriptions (in parentheses) were presented together on the computer screen. Snack foods in the same row comprise pairs.

Name	Name
Digestive Mini's Milk Chocolate* (Cookies)	Choc Chip Mini's* (Cookies)
Salty Liquorice (Liquorice)	Sweet Liquorice (Liquorice)
Soft Fruitbears (Fruitgum)	Fruity Winegums (Fruitgum)
Haribo Happy Cola (Fruitgum)	Haribo Happy Cherries (Fruitgum)
Haribo Gummibears (Fruitgum)	Fruity Bottles (Fruitgum)
Jelly Beans (Fruit Candy)	Skittles (Fruit Candy)
Bounty (Chocolate Bar)	Bounty Dark (Chocolate Bar)
Cote d'or BonBonBloc Praline (Chocolate)	Bitter Chocolate with Orange (Chocolate)
Kinder Bueno (Chocolate Bar)	Rolo (Chocolates with Caramel)
KitKat (Chocolate Bar)	KitKat Chunky* (Chocolate Bar)
Choco M&Ms (Chocolate Candy)	Maltesers (Chocolate Candy)

Note. *These items were not identical but close substitutes to items in the physical condition, all other items were identical across the two conditions. Snack foods in the same line comprise pairs.

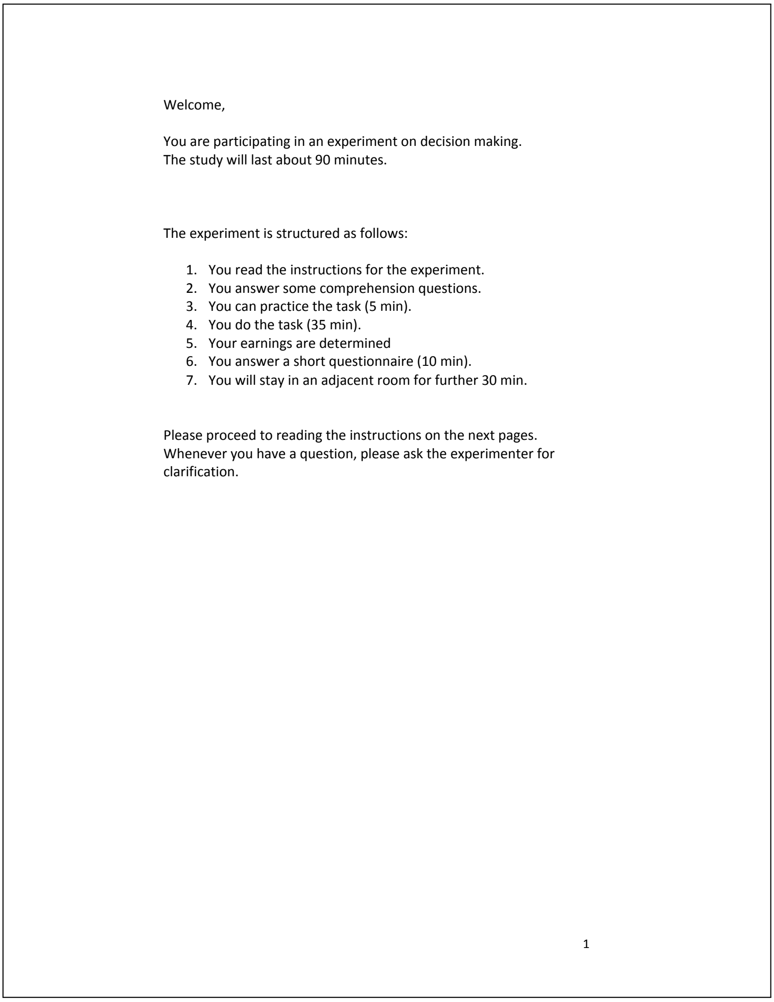
Table 3.S1 (continued)

Stimuli used in the computer condition. Names and descriptions (in parantheses) were presented together on the computer screen. Snack foods in the same row comprise pairs.

Name	**Name**
Mars (Chocolate Bar)	Snickers (Chocolate Bar)
Milka Milk (Chocolate)	Milka Cream (Chocolate)
Twix (Chocolate Bar)	Lion (Chocolate Bar)
Bugles Cheese (Nacho Chips)	Chio Cheese (Potato Chips)
Lay's Paprika (Potato Chips)	Lay's Natural (Potato Chips)
Fun Mix (Salty Crackers)	Trio Zoutjes (Salty Crackers)
Kaas Zoutjes* (Cheese Crackers)	Snack Zoutjes* (Salty Crackers)
Tuc Cheese (Crackers)	Tuc Naturel (Crackers)
Thai Sweet Chili Nuts (Spicy Peanuts)	Katjang Pedis (Spicy Peanuts)
Cashews (Nuts)	Pistachios (Nuts)
Nut and Raisins Mix (Nuts)	Nut Mix (Nuts)

Note. *These items were not identical but close substitutes to items in the physical condition, all other items were identical across the two conditions. Snack foods in the same line comprise pairs.

Experimental instructions



Welcome,

You are participating in an experiment on decision making.
The study will last about 90 minutes.

The experiment is structured as follows:

1. You read the instructions for the experiment.
2. You answer some comprehension questions.
3. You can practice the task (5 min).
4. You do the task (35 min).
5. Your earnings are determined
6. You answer a short questionnaire (10 min).
7. You will stay in an adjacent room for further 30 min.

Please proceed to reading the instructions on the next pages.
Whenever you have a question, please ask the experimenter for clarification.

1

Figure 3.S2: Instructions given to the participants in the physical condition. Page 1/6.

Instructions

Task
In this experiment, you will have an opportunity to buy a snack food from our store using €14.50 that you receive from us. You will receive the €14.50 once you have read the instructions and answered the comprehension questions correctly.

At the end of the experiment, you will be asked to stay in an adjacent room for 30 minutes. During this time, the only food that you will be allowed to eat is whatever snack you bought from us during the experiment.

The snack food items that are available in this experiment are real and recently purchased for the purpose of this experiment. You will see all of them during the experiment.

Your task in this experiment is to decide and tell us the maximum amount that you would currently be willing to pay for each of these items.

Round structure
The experiment consists of many rounds, all of which have a similar structure. In each round:

1. First, you will see which item is on offer in this round
2. Then you will enter the maximum amount that you are willing to pay for this item (in €). You will need to enter a number between €0 and €4. You can also enter decimal numbers like €3.47 using the dot as a decimal point.
3. You will answer a few questions about the item.

There is no strict time limit for giving an answer. Nevertheless, try to answer spontaneously, without thinking too much. After 22 of these rounds there will be a break, then there will be another 22 rounds. You will see each item only once. The items will be right in front of you so you can see them well. Please do not touch the items.

During this task a wristband will be attached to each of your arms.

2

Figure 3.S2 (continued): Instructions given to the participants in the physical condition. Page 2/6.

Your earnings in this experiment
At the beginning of the experiment, you receive €14.50. You can use this money to buy one item of snack food from us. Whatever you do not spend remains yours, just like in everyday transactions. Although you will place bids on 44 items, you will only be allowed to buy one of them.

How do we determine whether you bought or not?
At the end of the experiment, **one** of the rounds will be randomly chosen. You will be asked to draw a card from a stack of cards numbered from 1 to 44. The round with the number that you draw will be the one that counts. Note that every round has the chance to be selected, and only one round will be selected. Therefore, you don't need to worry about spreading your €14.50 budget. In fact, you can treat every round as if it were the only round. Each time you have to bid on an item, it is in your best interest to report exactly your maximum willingness to pay for being allowed to eat this item at the end of the experiment.

After you picked a card to randomly select the round that counts, you can see which item was presented in this round and what was your bid on this item. Then, the actual price for the item will be determined randomly. Therefore, you will be asked to throw 3 dice that will determine the price. The price shown is the actual price for which you can buy the item from us. Please note that your bid does not influence the actual price of the item.

Whether you indeed buy the item from us depends on your bid and the actual, randomly determined price. If your bid is higher than or equal to the actual price (so you would be willing to pay the actual price) you will buy the item at the actual price and keep the rest of the €14.50.
On the other hand, if your bid is lower than the actual price(so you would not be willing to pay this actual price), then you do not buy the item and keep the €14.50.

Note that if you buy you never pay more than the actual, randomly determined price! If your bid is higher than the actual price you do not have to pay your bid, but just the actual price! Therefore, the best you can do in each round is simply to estimate what the item is worth to you (the maximum you would be willing to pay for it) and bid exactly this amount.

3

Figure 3.S2 (continued): Instructions given to the participants in the physical condition. Page 3/6.

You should not bid more money on an item than you actually are willing to pay. Stating higher bids increases the chance that you will buy the item. However, the downside of this is that this involves the risk of buying the item at a price that is higher than what you are willing to pay for it.

For example: Suppose that the most you would like to pay for a bag of biscuits is €3, but in order to increase the chances of getting the biscuits you decide to bid €4. The actual price is randomly determined at €3.60. Then, you have to purchase the biscuits for €3.60, a price that is higher than what the biscuits are actually worth to you (€3).

You might think that your best strategy is to bid lower than your actual valuation for the item. This is incorrect. The price that you pay is determined by the numbers you throw with the dies and not by your bid. Bidding lower than your true value you would not affect the price that you pay, but you run the risk of not buying although the price is acceptable to you.

For example: Suppose that the maximum you would like to pay for a chocolate bar is €3.50, but in order to keep more money you decide to bid only €1. The actual price turns out to be €2. You will not buy the chocolate bar because you bid only €1. Had your bid been your true value of €3.50, you would have purchased the chocolate bar for €2 and kept €12.50 in cash.

To sum up, the best you can do in your own interest is to bid exactly the amount which you are maximally willing to pay for the item at stake.

Here are a few examples, to make this mechanism clear:

Example 1: In the round selected for payment, Manuel was bidding on chips. Manuel's bid in this round was €2.65. The randomly selected actual price turns out to be €2.00.
Manuel buys the chips because the actual price (€2.00) is lower than his maximum willingness to pay (€2.65). Manuel gets the chips and pays €2.00. He keeps €12.50 from his initial €14.50.

Example 2: In the selected round, Manuel was bidding on popcorn. Manuel's bid was €1.35. The randomly selected actual price turns out to be €4.00.

4

Figure 3.S2 (continued): Instructions given to the participants in the physical condition. Page 4/6.

Manuel doesn't buy the popcorn because the store's price (€4) is higher than Manuel's maximum willingness to pay (€1.35). He keeps the €14.50.

Example 3: In the selected round, Manuel bid €2.80 on gummibears. The actual price for the gummibears turns out to be €2.50. Manuel buys the gummibears for €2.50 because his bid is greater than the actual price. He keeps €12.00 and gets the gummibears.

Please ask questions NOW if anything remains unclear.

You will have a chance to practice this task for 5 rounds that will not count towards determining your final earnings.

Please let the experimenter know that you finished reading.

5

Figure 3.S2 (continued): Instructions given to the participants in the physical condition. Page 5/6.

Comprehension questions: (Choose all that apply)

1. *Suppose that in the round that is selected for payment you entered a **bid of €4** for a bag of chips. The randomly selected actual price turns out to be **€3**. What happens? Choose all that apply.*

   a) *I buy the chips for €4*
   b) *I buy the chips for €3*
   c) *My earnings in cash are €10.50*
   d) *My earnings in cash are €14.50*
   e) *My earnings in cash are €11.50*
   f) *I don't buy the chips*

2. *Suppose that in the selected round you had **bid €1** for a pack of crackers**.** The randomly selected actual **price** turns out to be **€2.75**. What happens? Choose all that apply.*

   a) *I buy the crackers for €2.75*
   b) *I buy the crackers for €1*
   c) *My earnings in cash are €14.50*
   d) *My earnings in cash are €11.25*
   e) *My earnings in cash are €13.50*
   f) *I don't buy the crackers*

3. *Suppose that in the round that is selected for payment, **you bid €3.20** for a bag of peanuts. **The actual price turns out to be €3.20**. What happens? Choose all that apply.*

   a. *I buy the peanuts for €3.20*
   b. *I buy the peanuts for €2.50*
   c. *My earnings in cash are €11.30*
   d. *My earnings in cash are €12.00*
   e. *My earnings in cash are €14.50*
   f. *I don't buy the peanuts*

6

Figure 3.S2 (continued): Instructions given to the participants in the physical condition. Page 6/6.

Instructions

Task
In this experiment, you will have an opportunity to buy a snack food from our store using €14.50 that you receive from us. You will receive the €14.50 once you have read the instructions and answered the comprehension questions correctly.

At the end of the experiment, you will be asked to stay in an adjacent room for 30 minutes. During this time, the only food that you will be allowed to eat is whatever snack you bought from us during the experiment.

Please note: All of the snack food items that are available in this experiment are regular size, as available in the supermarket. They have been recently purchased for the purpose of this experiment.

Your task in this experiment is to decide and tell us the maximum amount that you would currently be willing to pay for each of these items.

Round structure
The experiment consists of many rounds, all of which have a similar structure. In each round:

1. On the screen you will see the name and a description of the food item that is on offer in this round.
2. Then you will enter the maximum amount that you are willing to pay for this item (in €). You will need to enter a number between €0 and €4. You can also enter decimal numbers like €3.47 using the dot as a decimal point.
3. You will answer a few questions about the item.

There is no strict time limit for giving an answer. Nevertheless, try to answer spontaneously, without thinking too much. After 22 of these rounds there will be a break, then there will be another 22 rounds. You will see each item only once.

During this task a wristband will be attached to each of your arms.

2

Figure 3.S3: Page 2 of the instruction given to the participants in the computer condition. The other pages of the instructions were identical to the physical condition.

Regression models

Table 3.S2
Repeated measures analysis of variance.
Dependent variable: willingness to pay.

	df	**SS**	**MS**	**F**	**p**
within subjects					
anticipated effort	1	0.01	0.01	0.01	0.93
familiarity	1	1.59	1.59	2.81	0.10
order	1	0.98	0.98	1.74	0.19
anticipated effort × order	1	0.43	0.43	0.76	0.39
error	45	25.40	0.56		
between subjects					
presentation mode	1	0.02	0.01	1.10	0.30
familiarity	1	0.09	0.09	5.56	0.02
anticipated effort × presentation mode	1	0.08	0.08	4.75	0.03
presentation mode × order	1	0.02	0.02	0.88	0.35
anticipated effort × presentation mode × order	1	0.01	0.01	0.39	0.54
error	45	0.76	0.02		

Note. 50 subjects. Data aggregated over items, two data points for each subject. Within-subject factor: Anticipated effort (no weights vs. weights). Between-subject factor: Presentation mode (physical vs. computer screen). All p values are two-sided.

Table 3.S3
Random intercept regression model with control variables.
Dependent variable: wanting ratings.

	Coef.	95% CI	p
constant (computer condition, no weight)	0.45	[0.30, 0.61]	< 0.01
physical condition	0.10	[-0.11, 0.32]	0.35
weight	0.14	[-0.00, 0.28]	0.06
physical $\times$ weight	-0.21	[-0.41, -0.02]	0.03
familiarity	0.35	[0.28 ,0.41]	< 0.01
order of weight condition	0.29	[0.02, 0.57]	0.04
order $\times$ weight	-0.26	[-0.47, -0.05]	0.01
order $\times$ physical	0.11	[-0.30, 0.53]	0.59
order $\times$ physical $\times$ weight	0.22	[-0.08, 0.52]	0.15

Note. 2200 trials, nested within 50 subjects. Standard errors are corrected for potential heteroscedasticity and autocorrelations at the subject level. All p values are two-sided. Order indicates whether participant started with the weight (Order = 1) or no weight condition (Order = 0).

Table 3.S4
Random intercept regression model with control variables.
Dependent variable: liking ratings.

	Coef.	95% CI	p
constant (computer condition, no weight)	0.81	[0.65, 0.98]	< 0.01
physical condition	-0.07	[-0.32, 0.19]	0.60
weight	0.09	[-0.02, 0.20]	0.11
physical $\times$ weight	-0.13	[-0.29, 0.02]	0.08
familiarity	0.35	[0.30, 0.41]	< 0.01
order of weight condition	0.06	[-0.23, 0.34]	0.70
order $\times$ weight	-0.19	[-0.33, -0.04]	0.01
order $\times$ physical	0.27	[-0.13, 0.67]	0.19
order $\times$ physical $\times$ weight	0.09	[-0.14, 0.33]	0.44

Note. 2200 trials, nested within 50 subjects. Standard errors are corrected for potential heteroscedasticity and autocorrelations at the subject level. All p values are two-sided. Order indicates whether participant started with the weight (Order = 1) or no weight condition (Order = 0).

Supplemental figures

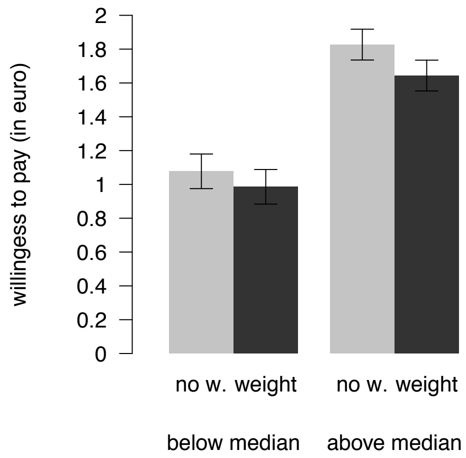


Figure 3.S4: Effect of weight on willingness to pay for items with wanting ratings below (left) or above (right) median-liking within its weight-condition. Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008) and are therefore only informative for evaluating within-subject differences between wearing wristbands with weight (black) vs. without weight (grey).

Part II

Social Decisions

Chapter 4

Be Nice if You Have to – The Neurobiological Roots of Strategic Fairness

Abstract. Social norms, such as treating others fairly regardless of kin relations, are essential for the functioning of human societies. Their existence may explain why humans, among all species, show unique patterns of prosocial behavior. The maintenance of social norms often depends on external enforcement, as in the absence of credible sanctioning mechanisms prosocial behavior deteriorates quickly. This sanction-dependent prosocial behavior suggests that humans strategically adapt their behavior and act selfishly if possible but control selfish impulses if necessary. Recent studies point at the role of the dorsolateral prefrontal cortex (DLPFC) in controlling selfish impulses. We test whether the DLPFC is indeed involved in the control of selfish impulses as well as the strategic acquisition of this control mechanism. Using repetitive transcranial magnetic stimulation, we provide evidence for the causal role of the right DLPFC in strategic fairness. Because the DLPFC is phylogenetically one of the latest developed neocortical regions, this could explain why complex norm systems exist in humans but not in other social animals.

Based on Sabrina Strang*, Jörg Gross*, Teresa Schuhmann, Arno Riedl, Bernd Weber, Alexander Sack (in press). Social Cognitive and Affective Neuroscience. *shared first authorship

4.1 Introduction

Humans among all animals are unique in their ability to establish highly complex social norm systems (Sethi and Somanathan, 1996; Ostrom, 2000; Gintis, 2003; Tomasello and Rakoczy, 2003; Fehr and Rockenbach, 2004; Fehr and Fischbacher, 2004). Fairness and cooperation norms often demand to restrict immediate self-interest in favor of benefits of the group, the society or another individual in need. The widespread prevalence of such norms in human societies is puzzling from an evolutionary perspective (Fehr and Fischbacher, 2004; Melis and Semmann, 2010), as they are informal, often vaguely defined and, as such, should be easy to circumvent. Especially in large groups with anonymous interactions, free-riding should dominate (Bowles and Gintis, 2003). Indeed, experiments have shown that without credible punishment threats, fair and cooperative behavior, like sharing with others or contributing to a group project, can deteriorate quickly (Fehr and Gächter, 2002; Egas and Riedl, 2008; Gächter et al., 2008; Riedl et al., 2009). On the other hand, there is convincing evidence that fair and cooperative behavior can emerge and be maintained when there is the threat that free-riding will be sanctioned (Fehr and Gächter, 2000; Fehr and Gächter, 2002; Fehr and Fischbacher, 2004; Fehr et al., 2007; Egas and Riedl, 2008; Gächter et al., 2008; Riedl et al., 2009). This indicates that humans are sensitive to punishment threats, enabling them to act selfishly when they can and to act strategically fairly when they have to.

The neurobiological basis of this ability to adapt behavior strategically and thereby controlling immediate selfish impulses has been explored in recent studies. These suggest that activity in the right prefrontal cortex is associated with the control of selfish impulses (Wout et al., 2005; Knoch et al., 2006; Knoch and Fehr, 2007; Knoch et al., 2009) and the ability to adapt behavior strategically (Fehr et al., 2007; Ruff et al., 2013). Knoch et al. (2006) found that the disruption of the right dorsolateral prefrontal cortex (DLPFC), using transcranial magnetic stimulation (TMS), led participants to accept an offer that yielded a higher financial payoff for themselves much more frequently than to reject it in favour of a financially less attractive but fair outcome. A functional magnetic resonance imaging (fMRI) study by Fehr et al. (2007) showed that acting fairly because of strategic reasons is correlated with increased activity in the right DLPFC. Most recently, using transcranial direct current stimulation (tDCS), Ruff et al. (2013) demonstrated that suppressing neural excitability (with cathodal tDCS) of the right lateral prefrontal cortex (LPFC) led to a lower degree of strategic fairness, whereas enhancing neural excitability (with anodal tDCS) of the right LPFC increased strategically fair behavior. Interestingly, cathodal tDCS over right LPFC also decreased immediate selfish responses, while enhancing the right LPFC using anodal tDCS led to a higher degree of immediate selfishness.

This result pattern is intriguing and yet puzzling at the same time. The fact that suppressing neural activity of the right LPFC with cathodal tDCS decreases immediate selfishness is in conflict with an earlier result of Knoch et al. (2006) who found an increase of immediate selfishness when disrupting the dorsal part of the right LPFC using repetitive TMS (rTMS). However, these two studies differ in various methodological aspects, which might account for this apparent contradiction. Ruff et al. (2013) applied a different intervention method and used a broader target region. The strength of the study by Ruff et al. (2013) clearly lies in the differential effects revealed on strategic fairness as well as immediate selfishness contrasting anodal (enhancing) vs cathodal (suppressing) tDCS over the same brain region, i.e. the right LPFC. This focus on the right LPFC naturally neglected the possible contribution of and comparison with left LPFC, as tested, e.g. in Knoch et al. (2006). Moreover, all mentioned studies used between-subject experimental designs, which allow inferences on the population level but do not allow investigating individual differences in these effects across stimulation conditions. Hence, although these studies strongly suggest that the right LPFC is functionally relevant for decisions that involve trade-offs between immediate selfish goals on the one hand and fair and cooperative behavior on other hand, it remains necessary to further investigate to what extent the right, and not the left, LPFC is involved in controlling immediate selfish impulses and strategically fair behavior. In the current study we used a within-subject design applying rTMS (or sham) either to the right or left DLPFC of male participants in order to test on the individual level whether the right and/or left DLPFC are causally linked to, first, the control of immediate selfishness and, second, the strategic acquisition of this control mechanism when the threat of norm enforcement demands it. Moreover, to explore whether a shift in beliefs or perception could explain the results, we elicited participants' beliefs about norm enforcement and perception of the fairness of behaviors while under the different TMS and sham conditions.

4.2 Materials and methods

4.2.1 Subjects

We studied 17 male participants (mean age 23.5 years, ranging between 20 and 41 years) with normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. None of the participants had taken part in a TMS experiment before. They received medical approval for participation and gave written informed consent after being instructed about the procedure. The study was approved by the local Medical Ethical Commission.

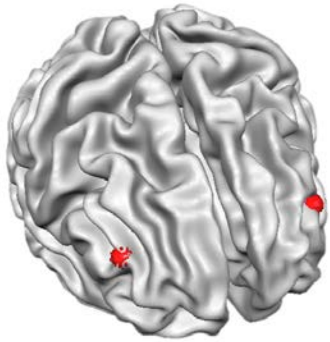


Figure 4.1: Target area for the magnetic brain stimulation. Each target was selected based on the individual anatomical image obtained in a separate MRI measurement. The red dots represent the two target points in Talairach space: x = ±39, y = ±37, z = ± 22.

4.2.2 TMS-procedure

Participants were tested in four sessions separated by at least 1 week. In the first session a T1-weighted magnetic resonance image was acquired. The other three sessions were TMS sessions. Each participant took part in each of the three TMS conditions (left DLPFC, right DLPFC and sham). The condition order was counterbalanced across participants.

A surface reconstruction on the MRI images was made to recover the spatial surface of the cortical sheet based on the white-grey matter boundary using Brain Voyager QX 2.4 (BrainInnovation, Maastricht, The Netherlands). We then identified the right and left DLPFC based on the coordinates established by Aronson et al. (2003) and Knoch et al. (2006); x = ±39, y = ±37, z = ±22, radius = 6 (Figure 4.1). The coordinates, given in Talairach space, were transformed to each participant's individual brain space.

Biphasic TMS pulses were applied using the MagVenture R30 stimulator (MagVenture A/S, Farum, Denmark) and a figure-of-eight coil (MC-B70, inner radius 10mm, outer radius 50mm). The maximum output of this coil and stimulator combination is 1.9 T and 150 A/μS. At the beginning of the first TMS session, individual resting motor thresholds (RMTs) were determined. The mean RMT was 34.6% (SD = 3.7), ranging from 28 to 40% of maximal stimulator output (MSO). Stimulation intensity was applied at 110% RMT.

For sham stimulation, a figure-of-eight placebo coil (MC-BP70) was used. This coil produced the same acoustic stimulation as the active coil while not inducing a magnetic field. The coil was manually held tangentially to the skull over the right/left DLPFC using the online visualization function of Brain Voyager TMS Neuronavigation. Participants received 15 min, 1 Hz rTMS (900 pulses) offline stimulation over the left or right DLPFC. Sham stimulation was applied either over the left DLPFC or right DLPFC, balanced over participants. Participants were told beforehand that intensity of the TMS stimulation could vary across sessions. In the debriefing, we asked participants how the different TMS conditions might have affected their behavior. None of the participants indicated any directed hypotheses.

4.2.3 Task

Resembling the task used in Fehr et al. (2007), two different games were used, a standard Dictator Game (DG) and a Dictator Game with punishment option (DGp). In both games, two players, a dictator and a recipient, interact with each other. Each player receives an initial endowment of 25 monetary units (MUs, 1 MU = 0.16 euro). Additionally, the dictator receives 100MUs and can distribute these freely between himself and the recipient. In DG, the recipient is passive and the game ends after the dictator has made a decision. In DGp, the recipient can punish the dictator after being informed about the distribution. To punish the dictator, the recipient has to spend his own MUs. For every MU the recipient uses for punishment, the dictator's payoff is reduced by 5 MUs. Thus, in case the dictator does not transfer any MUs and the recipient applies maximum punishment by spending his 25 MUs, both participants end up with 0 MUs.

4.2.4 Experimental procedure

Dictators and recipients were invited separately. First, 60 recipients were invited to the BEElab (Behavioral & Experimental Economics laboratory, Maastricht). They received written instructions about the rules of both games (DG and DGp). In the instructions, dictators and recipients were neutrally labelled as player A and B. To maximize the number of observations per recipient, we implemented the so-called strategy method (Selten, 1967). Recipients were asked how many of their MUs they would spend for punishment for every possible transfer the dictator could make. That is, we gathered a punishment response for each possible dictator transfer. Importantly, this method does not imply that punishment choices of recipients are hypothetical. The chosen punishment is real, as it has real monetary consequences for the recipient as well as the dictator. Specifically, depending on how much the dictator later actually chose to transfer to the recipient, the corresponding punishment decision was selected and final payoffs were calculated and paid accordingly (Brandts and Charness, 2011). Additionally a photo was taken of every recipient.

The 17 participants invited to the TMS sessions were allocated the role of dictators. They received written instructions about the rules of both games (DG and DGp) and were asked to answer a set of comprehension questions before the TMS stimulation. At the beginning of each session, after TMS stimulation, they saw a group picture of the 20 recipients who they would be paired with in the following 20 rounds. This was done to emphasize that in each round they interact with a different real person and that each decision bears real consequences. In half of the rounds, participants were paired with recipients who could punish (DGp condition), and in the other half, they were paired with recipients who could not punish (DG condition). These conditions were randomized over rounds and participants were informed about the condition by a symbol on the computer screen. Hence, in the DG condition, participants knew that they would not face any consequences for acting selfishly, whereas in the DGp condition, punishment by the recipient could decrease their payoffs substantially. In each round, participants were asked how much, if any, of the 100 MUs they wanted to transfer to the other participant. To avoid learning effects, dictators received no feedback about the punishment decisions of recipients until the end of the experiment. We opted for not giving feedback because providing feedback would have had the problematic downside that potential TMS effects on learning could not have been disentangled from the effects we are interested in. Experienced punishment often has a larger impact on behavior than imagined punishment. In repeated public goods game experiments with punishment, for example, first round cooperation (imagined punishment) is often smaller than cooperation in later rounds (experienced punishment; Fehr and Gächter, 2000, Fehr and Gächter, 2002, Egas and Riedl, 2008). Thus, by giving no feedback we are likely observing a lower bound of the effect of (potential) punishment on dictator transfers. This means that the incentive for acting strategically fairly is on the lower side and, therefore, inhibiting effects of our TMS intervention are likely to be on the conservative side too.

To test whether fairness perception or punishment beliefs were systematically affected by our TMS manipulation, participants saw five hypothetical transfers (from 0 to 50 MUs in steps of 10) from a hypothetical dictator and were asked to make fairness judgments for each transfer. Furthermore they were asked about their punishment expectation and own punishment expenses, were they in the role of the recipient.

Decisions in both games and the elicitation of fairness perceptions, punishment beliefs and hypothetical own punishment expenses were completed within 5–6 minutes after the rTMS stimulation. For each session, 1 of the 20 rounds was randomly selected and paid out in cash. Participants knew about this procedure upfront and were informed about the selected rounds and the associated earnings after the last session.

4.2.5 Analysis

Dictator transfers are censored below by zero. We therefore fitted random-intercept Tobit regressions (Gelman and Hill, 2007) to the data using R and JAGS (see Lunn et al., 2009). In each regression model, variables, coding the session number as well as the sequence of the conditions were added to control for potential learning and order effects (see Supporting Information). To test whether participants behave more selfishly when TMS is applied over the right DLPFC compared with sham and TMS over the left DLPFC, transfer decisions in the DG without punishment were regressed on dummy predictors coding the three TMS condition (sham condition as baseline).

To test the causal involvement of the right and/or left DLPFC in the ability to act strategically fairly, we first classified participants into ‘adapters’ and ‘non-adapters’; those who gave more in the DGp during sham were classified as ‘adapters’ and those who gave less or equal were classified as ‘non-adapters’. For each participant the transfer difference across DG and DGp as a measure for strategic adaption was calculated and regressed on dummy predictors coding the three TMS condition (sham condition as baseline) and a dummy variable indicating whether a dictator was a non-adapter to test for changes in strategic fairness across TMS conditions for adapters and non-adapters separately.

4.3 Results

On average, dictators transferred relatively little, although significantly more than zero, to recipients in the DG during sham (average transfer = 6.6, one sample t-test, $t(16) = 2.7$, $p < 0.05$, two-sided). Transfer rates are smaller than observed in some other DG experiments but similar to experiments with large social distance between dictators and recipients (Hoffman et al., 1996; Camerer, 2003). The low transfers in sham give little room for observing the hypothesized effect of increased selfishness when inhibiting the right or left DLPFC. Nevertheless, we find that, on average, participants gave significantly less (almost 50%; 3 MUs on average) to recipients when TMS was applied to the right DLPFC compared with TMS over the left DLPFC and sham, respectively (see Figure 4.2; random-intercept Tobit regression, rDLPFC dummy, 95% confidence interval (CI) = [-14.4, -3.9], $p < 0.05$). There was no significant difference in transfers between sham and TMS over the left DLPFC (random-intercept Tobit regression, lDLPFC dummy, 95% CI = [-6.9, 3.1]).

To test our second hypothesis, whether the right DLPFC is also causally involved in the ability to act strategically fairly, we analyzed the change in strategic adaption over

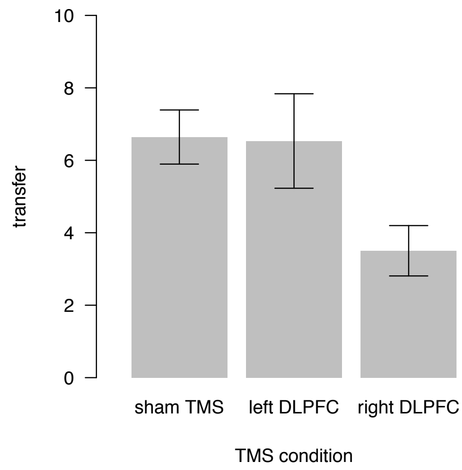


Figure 4.2: Average transfers in the DG condition. Mean transfers to recipients without the ability to punish (DG condition) for each TMS session. Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008).

the TMS conditions. Not all participants showed strategic adaption during sham. Four participants gave constantly nothing, regardless of DG and TMS condition. These participants also reported that they did not expect any punishment from the recipients for unfair transfers. One participant offered very low amounts to the recipients (between 0 and 15 MUs) and did not change offers across DG and DGp and one participant actually gave more to recipients without punishment power (DG) than to recipients with punishment power (DGp). However, a majority of the dictators (11 of 17) adapted strategically during sham and were classified as 'adapters'. On average, with sham TMS, adapters transferred 6MUs in DG but a 5-fold of it (32 MUs) in DGp.

Figure 4.3 shows how this strategic adaption was affected by the disruption of the right and left DLPFC by plotting the mean transfers of DGp and DG for the three TMS conditions. When facing a recipient with punishment ability, participants who adapted strategically during sham did so significantly less when TMS was used to disrupt the right DLPFC compared with sham (random-intercept regression, rDLPFC dummy, 95% CI = [-8.0, -4.6], $p < 0.05$). We observed the highest strategic adaption when the left DLPFC was disrupted. However, the difference to sham stimulation was not significant (random-intercept regression, lDLPFC dummy, 95% CI = [0.6, 4.3]).

Most recipients were willing to use their MUs to punish unfair behavior by the dictators. Consequently, by giving less to recipients with punishment ability during TMS

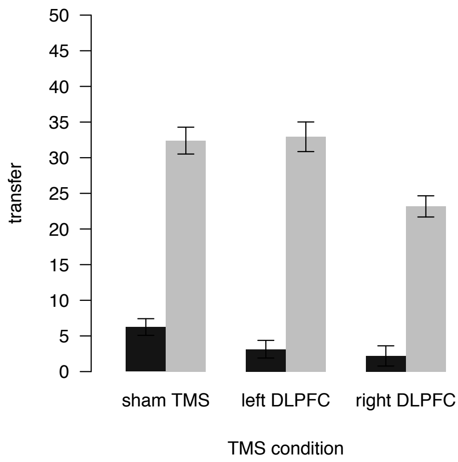


Figure 4.3: Strategic adaption across TMS conditions. Mean transfers of adapters in DG (black) and DGp (grey) across the TMS conditions. Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008).

over the right DLPFC, dictators earned 4.1MUs less on average in each interaction compared with sham (random-intercept regression, rDLPFC dummy, 95% CI = [-4.7, -3.3], $p < 0.05$).

In principle, it could be possible that TMS stimulation alters the judgement of how (un)fair an offer is and/or the belief about how likely it is that unfair offers will be punished. To test for this we explored whether fairness judgements or beliefs about punishment were different across the three TMS conditions. As expected, we find that adapting dictators judge offers to be fairer the more is offered to the recipient (random-intercept regression, offer predictor, 95% CI = [0.10, 0.15], $p < 0.05$). These fairness judgements did not significantly change across the TMS conditions (see Figure 4.4; random-intercept regression, offer $\times$ left TMS, 95% CI = [-0.05, 0.02]; offer $\times$ right TMS, 95% CI = [-0.04, 0.03]. Regarding punishment, during sham, adapting dictators expected that unfair offers would be punished more severely (see Figure 4.5; random-intercept regression, offer predictor, CI = [-0.56, -0.37], $p < 0.05$). Importantly, this expectation did not change significantly across the TMS conditions (random-intercept regression, offer $\times$ right, CI = [-0.05, 0.22]; offer $\times$ left, CI = [-0.04, 0.22]).

Hence, neither fairness judgements nor expected punishment of adapters were significantly affected by the TMS conditions. Interestingly, in comparison with adapters,

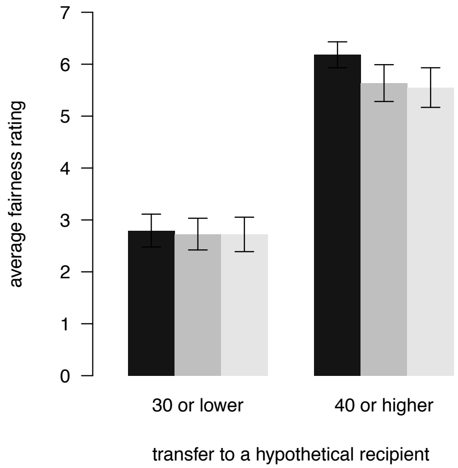


Figure 4.4: Fairness judgments. Fairness judgments of adapters (from 1 = 'very unfair' to 7 = 'very fair') for low (≤30) and high (≥40) hypothetical transfers after each TMS session (black: sham; dark grey: TMS over the left DLPFC; light grey: TMS over the right DLPFC). Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008).

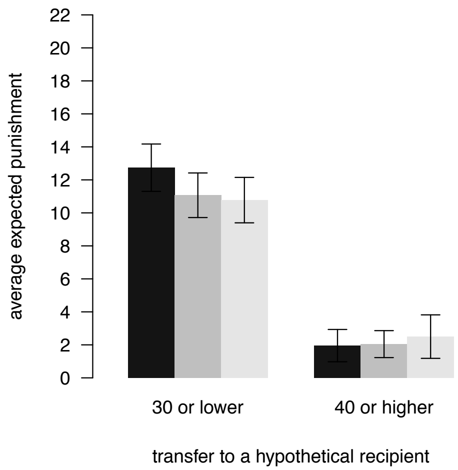


Figure 4.5: Expected punishment. Expected punishment of adapters (from 0 to 25 punishment points) for low (≤30) and high (≥40) hypothetical transfers after each TMS session (black: sham; dark grey: TMS over the left DLPFC; light grey: TMS over the right DLPFC). Error bars show the within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008).

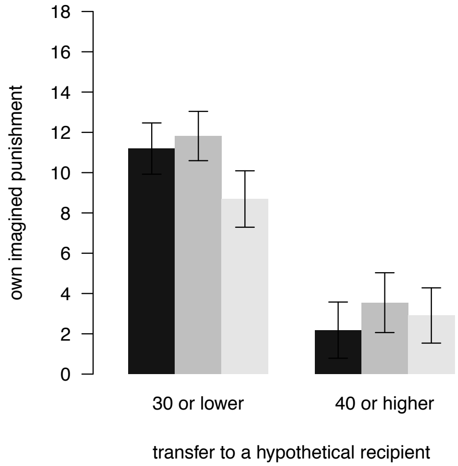


Figure 4.6: Own imagined punishment. Own imagined punishment of adapters (from 0 to 25 punishment points) for low (≤30) and high (≥40) hypothetical transfers after each TMS session (black: sham; dark grey: TMS over the lDLPFC; light grey: TMS over the rDLPFC). Error bars show the within- subject standard errors of the mean (Cousineau, 2005; Morey, 2008).

non-adapters reported different fairness judgements as well as punishment beliefs. They showed a significantly smaller increase of rated fairness for increasing offers (random-intercept regression, non-adapter × offer, CI = [-0.18, -0.10], $p < 0.05$), which did not change significantly across TMS conditions (random-intercept regression, non-adapter × offer × left, CI = [-0.01, 0.10]; non-adapter × offer × right, CI = [-0.03, 0.08]. Compared with adapters, they also believed that dictators are punished significantly less severely in general (random-intercept regression, non-adapter dummy, CI = [-43.9, -15.3], $p < 0.05$). Thus, fairness judgements of non-adapters were less sensitive to changes in transfers and they generally expected less punishment.

When dictators were asked to imagine to be in the role of the receiver confronted with unfair transfers by a hypothetical dictator, adapters reported that they would be less willing to spend their MUs for punishment while the right DLPFC was disrupted by TMS, as compared with sham and TMS over the left DLPFC (Figure 4.6, random-intercept Tobit regression, rDLPFC dummy, 95% CI = [-7.8, -1.3], $p < 0.05$). Non-adapters indicated that they would punish significantly less compared with adapters in general (random-intercept regression, non-adapter dummy, CI = [-60.1, -4.9], $p < 0.05$).

4.4 Discussion

Our results reveal that experimental perturbation of the right, but not the left, DLPFC systematically altered, first, the degree of prosocial behavior and, second, the ability to act strategically fairly. The first result is consistent with previous findings by Knoch et al. (2006) who studied responses to fair or selfish behavior. We extent their finding and show that even when subjects can actively decide to behave selfishly or fairly they are on average more selfish when the rDLPFC is disrupted.

The average transfers during sham in the DG were relatively low compared with other studies. Behavior in the DG is known to be sensitive to framing (List (2007); Bardsley, 2008). Especially social distance has been shown to influence transfer rates (Hoffman et al., 1996; Leider et al., 2009). Leider et al. (2009) showed that dictators give significantly more to receivers that are socially close. Hoffman et al. (1996) used a double-blind procedure that maximizes social distance and find transfer rates of 6–8%, similar to the rates we observed in the DG with sham TMS. In our experiment, receivers and dictators were invited separately to the experiment, and receivers were, thus, not present when dictators made their transfer decisions. This certainly increased social distance between dictators and receivers and might explain the relatively low transfers observed in our study. Further, it has been shown that when participants have to exert effort to earn their endowment, transfer rates decrease (Cherry et al., 2002; Oxoby and Spraggon, 2008). In our study, dictators had to come to the lab four times (three times more often than the receivers) and had to undergo the TMS procedure three times. Thus, compared with the receivers, the effort they invested was much higher, and dictators might have thought that they deserved to keep more.

The relatively low offers of dictators in the DG with sham TMS left little room for a decrease of offers due to disruption of the right or left DLPFC. The fact that we nevertheless find a significant decrease in offers after TMS over the right DLPFC in comparison with sham TMS and TMS over the left DLPFC is reassuring for the interpretation that the rDLPFC controls selfish impulses not only when responding to offers (Knoch et al., 2006) but also when actively making offers.

During sham and TMS over the left DLPFC, most dictators adapted their behavior strategically by transferring more of their money to recipients who were able to punish them. Similar to DG transfers, our DGp transfers were lower than in comparable studies (Fehr et al., 2007; Ruff et al., 2013), and similar reasons as discussed above might play a role. Moreover, in the DGp, participants did not receive immediate feedback about punishment and, hence, could not experience but only anticipate punishment. It is conceivable that participants thought that punishment threats are not credible and without feedback they could not update their beliefs. Some participants indeed reported that they did not expect any punishment from the recipients for unfair transfers. These participants mostly responded rationally to their beliefs and did not adapt their behavior to the punishment threat. The majority of participants, however, reported that they expected to be punished for unfair offers and also that punishment would increase with the unfairness of the offer. These participants also adapted their behavior accordingly in the sham TMS condition. However, when disrupting the right DLPFC, the same dictators not only shared their money less generously with recipients but also showed significantly less strategic adaption in their unfair behavior in case a punishment threat was present. Thus, participants who consistently adapted their behavior strategically during sham and TMS over the left DLPFC were less capable of doing so when TMS was applied to the right DLPFC. Our study therefore provides causal evidence for the functional role of the right DLPFC not only in overriding immediate selfish impulses but also in acting strategically fairly, an ability paramount for obeying fairness norms.

Interestingly, and in contrast to Knoch et al. (2006), where the disruption of the right DLPFC led to an increase in earnings, this lack of adaption was maladaptive because recipients were willing to spend their money to punish unfair behavior, which decreased the payoff of the dictators substantially. Hence, by disrupting the right DPLFC participants not only failed to comply with widely shared fairness norms but thereby also failed to maximize their own payoff.

Our results also indicate that the failure to adapt strategically is neither explained by altered fairness perception nor by changes in beliefs about recipients' punishment behavior as suggested by Sanfey et al. (2014). In line with previous findings (Knoch et al., 2006; Ruff et al., 2013), fairness judgments were not significantly affected by TMS, indicating that disrupting the right DLPFC impaired the control of selfish impulses without altering fairness perception. This suggests that fairness perception and decisions on complying with a fairness norm are to some degree independently represented in the brain, enabling us to know what is right, but do otherwise. There is also no evidence that beliefs about recipients' punishment behavior were affected by our TMS intervention. Across all three TMS conditions, participants either believed that there would be no or little punishment (non-adapters) or that there would be punishment and that it would increase with the unfairness of offers (adapters). These results indicate that perturbation of the right DLPFC can alter strategic behavior, but not the underlying motive or belief system that led to strategic fairness in the first place. Importantly, participants under TMS of the right DLPFC reported that they themselves would use less money for punishing unfair transfers. That is, although perceiving small transfers as unfair, participants in the right DLPFC TMS condition indicated not to be willing to spend money to punish unfair behavior of others. Consistent with the results of Knoch et al. (2006), this suggests that even in the role of the recipient, participants would act more selfishly by withholding costly punishment. This implies that less social norm violations would be punished and more selfish behavior would be tolerated, pointing to a possible involvement of the right DLPFC not only in norm compliance but also in the enforcement of norms that demand the restriction of selfish behavior. Further research needs to be conducted to investigate the role of the DLPFC in norm enforcement directly.

An inherent characteristic of a within-subject design is that participants have to engage in a task repeatedly. This may lead to memory effects or habit formation, which may influence behavior in later sessions. We controlled for this by counterbalancing the order of TMS conditions and also controlled for it in our statistical analysis. Evidence that order effects may only be of limited importance also comes from experiments showing that when participants restart an experimental task, behavior is similar to the one in the previous task. This so-called restart effect was first observed by Andreoni (1988) and has been replicated numerous times (Andreoni and Croson, 2008). Hence, memory effects or habit formation are unlikely to confound our results.

Ruff et al. (2013) employed the same paradigm (DG and DGp) and investigated differences in strategic adaption across groups of female participants, while decreasing and increasing the neural excitability of the right LPFC using cathodal and anodal tDCS. Using a different intervention method and male participants, our findings are mostly in agreement with their results. They show that strategic adaption is significantly lower when decreasing excitability of the right LPFC, while fairness judgements are unaffected. Transfers in the DG condition were generally higher in their sample but adaption rates were comparable in size. Ruff et al. (2013) had an additional non-social control condition, showing that strategic adaption is only altered in a social context. In contrast to Ruff et al. (2013) who report an increase in transfers in the DG when disrupting the right LPFC with cathodal tDCS, we find that transfers to recipients who cannot punish (DG condition) significantly decrease when disrupting the right DLPFC. This finding is also in line with previous findings of increased selfishness (Wout et al., 2005; Knoch et al., 2006; Fehr et al., 2007) and the subsequent interpretation of a causal role of the right DLPFC in controlling selfish impulses. One possible explanation for the observed differences to Ruff et al. (2013) may be that the spontaneous reaction that is controlled by a secondary process is in fact not always selfish but sometimes prosocial to some extent (Rand et al., 2013; Schulz et al., 2014), and that this is different between male (current study) and female participants (Ruff et al., 2013). Future research specifically designed to explore these questions is needed to identify the exact role of the right DLPFC in voluntary giving (DG). Another possible explanation for the observed differences lies in the different techniques used to stimulate/suppress neural activity (TMS vs tDCS). In addition to the possible differences in neurophysiological effects induced by both techniques, they also differ in spatial resolution, thereby potentially affecting different (sub) regions within the LPFC (Priori et al., 2009).

While our study demonstrated a crucial involvement of the right DLPFC in strategic fairness, the specific interplay of the right DLPFC and other brain areas is not resolvable with our study design. A recent study, combining fMRI with TMS, suggests that the DLPFC and the ventromedial prefrontal cortex (VMPFC) are part of a frontal cortex network, responsible for orchestrating normative choice (Baumgartner et al., 2011). From this perspective, and in line with other research linking the VMPFC to the computation of abstract value signals guiding simple choice (see e.g. Chib et al., 2009; Levy and Glimcher, 2011; McNamee et al., 2013; Gross et al., 2014 – chapter 2), one possible interpretation is that the VMPFC is computing the expected value, integrating selfish goals with expected punishment for violating cooperation norms, whereas the DLPFC is linked to the execution of actions based on this valuation.

The complex norm system we observe in human societies, which is not present in other social animals, might be related to the fact that phylogenetically the DLPFC is one of the latest neocortical regions (Fuster, 2001). In a similar vein, the DLPFC is also ontogenetically one of the latest developing brain regions (Gogtay et al., 2004; Shaw et al., 2008), which in the context of our current findings could help to explain why young children up to the age of 3–8 years are not able to fully implement social rules like sharing resources with others (Fehr et al., 2008). Taken together, our study provides strong evidence for a direct neurobiological basis of social norm compliance, a cornerstone for the functioning of human society.

References

Andreoni, J. (1988). Why free ride?: Strategies and learning in public goods experiments. *Journal of Public Economics*, 37(3):291–304.

Andreoni, J. and Croson, R. (2008). Partners Versus Strangers – The Effect of Random Rematching in Public Goods Experiments. In Plott, C. and Smith, V., editors, *Handbook of Experimental Economics Results, Vol. 1*, pages 776–783. North-Holland, Amsterdam.

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

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, 11(2):122–133.

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., and Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, 14(11):1468–1476.

Bowles, S. and Gintis, H. (2003). Origins of human cooperation. In Hammerstein, P., editor, *Genetic and Cultural Evolution of Cooperation*. MIT Press, Cambridge.

Brandts, J. and Charness, G. (2011). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics*, 14(3):375–398.

Camerer, C. F. (2003). *Behavioral Game Theory*. Princeton University Press, Princeton.

Cherry, T. L., Frykblom, P., and Shogren, J. F. (2002). Hardnose the Dictator. *American Economic Review*, 92(4):1218–1221.

Chib, V. S., Rangel, A., Shimojo, S., and 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.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1):42–45.

Egas, M. and Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the National Academy of Sciences / Biological Sciences*, 275:871–878.

Fehr, E., Bernhard, H., and Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, 454:1079–1084.

Fehr, E. and Fischbacher, U. (2004). Social norms and human cooperation. *TRENDS in Cognitive Sciences*, 8(4):185–190.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415:137–140.

Fehr, E. and Rockenbach, B. (2004). Human altruism: Economic, neural, and evolutionary perspectives. *Current Opinion in Neurobiology*, 14:784–790.

Fehr, E., Spitzer, M., Fischbacher, U., Herrnberger, B., and Gron, G. (2007). The neural signature of social norm compliance. *Neuron*, 56:185–196.

Fuster, J. M. (2001). The prefrontal cortex – An update. *Neuron*, 30:319–333.

Gächter, S., Renner, E., and Sefton, M. (2008). The long-run benefits of punishment. *Science*, 322:1510.

Gelman, A. and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge.

Gintis, H. (2003). The Hitchhiker's Guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of theoretical biology*, 220(4):407–418.

Gogtay, N. N., Giedd, J. N. J., Lusk, L. L., Hayashi, K. M. K., Greenstein, D. D., Vaituzis, A. C. A., Nugent, T. F. T., Herman, D. H. D., Clasen, L. S. L., Toga, A. W. A., Rapoport, J. L. J., and Thompson, P. M. P. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, 101(21):8174–8179.

Gross, J., Woelbert, E., Zimmermann, J., Okamoto-Barth, S., Riedl, A., and Goebel, R. (2014). Value Signals in the Prefrontal Cortex Predict Individual Preferences across Reward Categories. *The Journal of Neuroscience*, 34(22):7580–7586.

Hoffman, E., McCabe, K., and Smith, V. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 86(3):653–660.

Knoch, D. and Fehr, E. (2007). Resisting the Power of Temptations. *Annals of the New York Academy of Sciences*, 1104(1):123–134.

Knoch, D., Fehr, E., Schneider, F., Schunk, D., and Hohmann, M. (2009). Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proceedings of the National Academy of Sciences*, 106:20895–20899.

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

Leider, S., Möbius, M. M., Rosenblat, T., and Do, Q.-A. (2009). Directed Altruism and Enforced Reciprocity in Social Networks. *The Quarterly Journal of Economics*, 124(4):1815–1851.

Levy, D. J. and Glimcher, P. (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.

List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *The Journal of Political Economy*, 115(3):482–493.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in medicine*, 28(25):3049–3067.

McNamee, D., Rangel, A., and O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, 16(4):479–485.

Melis, A. P. and Semmann, D. (2010). How is human cooperation different? *Philosophical transactions of the Royal Society, Biological sciences*, 365(1553):2663–2674.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2):61–64.

Ostrom, E. (2000). Collective action and the evolution of social norms. *The Journal of Economic Perspectives*, 14(3):137–158.

Oxoby, R. J. and Spraggon, J. (2008). Mine and yours: Property rights in dictator games. *Journal of Economic Behavior and Organization*, 65(3):703–713.

Priori, A., Hallett, M., and Rothwell, J. C. (2009). Repetitive transcranial magnetic stimulation or transcranial direct current stimulation? *Brain Stimulation*, 2(4):241–245.

Rand, D. G., Greene, J. D., and Nowak, M. A. (2013). Spontaneous giving and calculated greed. *Nature*, 489(7416):427–430.

Riedl, A., Ule, A., Schram, A., and Cason, T. (2009). Indirect punishment and generosity toward strangers. *Science*, 326:1701–1704.

Ruff, C. C., Ugazio, G., and Fehr, E. (2013). Changing Social Norm Compliance with Non-invasive Brain Stimulation. *Science*, 342:482–484.

Sanfey, A. G. A., Stallen, M., and Chang, L. J. (2014). Norms and expectations in social decision-making. *TRENDS in Cognitive Sciences*, 18(4):172–174.

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

Sethi, R. and Somanathan, E. (1996). The evolution of social norms in common property resource use. *The American Economic Review*, 86(4):766–788.

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J. L., Giedd, J. N., and Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience*, 28(14):3586–3594.

Tomasello, M. and Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind & Language*, 18(2):121–147.

Wout, M., Kahn, R., Sanfey, A. G. A., and Aleman, A. (2005). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *NeuroReport*, 16(16):1849–1852.

Supporting Information

Experimental instructions

INSTRUCTIONS

In the following experiment we want to investigate how people behave in an interactive situation. In addition to your show up fee you will earn money based on the decisions you make during the experiment. It is important that you understand the procedure of the experiment. Please read the following instructions carefully and do not hesitate to ask questions if anything remains unclear!

During the experiment you will not earn Euros but money units (MU):

100 MUs = 16€
Thus,
1 MU = 0,16 €

There are 20 rounds. In each round you will play with a different participant. So you will never meet the same person again during the experiment. Of course you can see that right now there are no other people in this room to interact with. Instead of bringing 20 different people to the lab now, they have already taken part in the experiment and have given a hypothetical response to every possible decision you can make during this experiment. Attached to the instruction you can find photos of all participants you will interact with during this experiment. The payoffs of these participants depend on your decisions, so they did not receive any money yet, but will receive it after this experiment.

How does one round work?

In every round you will interact with one other participant, we call you participant A and we call the person you interact with participant B. In each round you will interact with a different participant B.

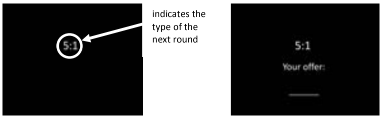
At the beginning of each round you and participant B each receive an endowment of 25 MUs.

You receive an additional endowment of 100 MUs. You can transfer as much of these MUs as you want to participant B (from 0 MUs to 100 MUs in steps of 1 MU). In each round, you simply type in how many MU's you want to transfer to participant B. You can change what you typed in pressing "←" on the keyboard. If you are confident about your decision press "ENTER" to confirm it and the next round will begin.

Figure 4.S1: Instructions given to the participants. Page 1/4.

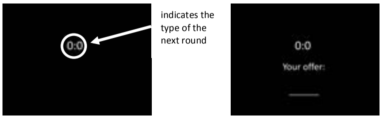
There are two types of rounds, the 5:1 and the 0:0 rounds.

5:1 rounds:



In the 5:1 rounds, he or she can spend his or her MUs to decrease your payoff in a ratio of 5:1. That means that for each MU participant B spends, 5 MU will be subtracted from your final payoff. Since all player B's have given a hypothetical response to every possible decision you can make during this experiment, you can imagine that player B observes your decision and then makes his decision.

0:0 rounds:



In the 0:0 rounds participant B cannot spend any MU's and therefore cannot decrease your payoff. The round is therefore finished after you have decided on an allocation of the 100 MUs and participant B will be informed about this allocation.

After finishing all 20 rounds you will be asked to fill in a short questionnaire.

At the end of the experiment:

At the end of the experiment one out of all rounds will be selected randomly and you and participant B will be paid based on this selected round after the last session.

2

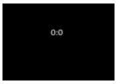
Figure 4.S1 (continued): Instructions given to the participants. Page 2/4.

Trainings questions:

In order to make sure that you understand the procedure of this experiment, we would like you to answer the following questions:

1. How much endowment do you receive at the beginning of each round? ______________________________
2. How much endowment does participant B receive at the beginning of each round? ______________________________
3. In the 5:1 rounds, by how many MUs can participant B decrease your payoff using one of his or her MUs? ______________________________
4. In the 0:0 rounds, by how many MUs can participant B decrease your payoff using one of his or her MUs? ______________________________
5. Does the following picture indicate a round in which participant B can decrease your payoff? ______________________________

6. Does the following picture indicate a round in which participant B can decrease your payoff? ______________________________



7. Has participant B received any payoff yet? ______________________________
8. How many Euros are 100 MUs? ______________________________
9. How many MUs can you maximally transfer to Player B? ______________________________
10. How many MUs can you minimally transfer to Player B? ______________________________
11. With how many other people will you interact throughout 5 rounds in this experiment? ______________________________

3

Figure 4.S1 (continued): Instructions given to the participants. Page 3/4.

Trainingstrial:

0:0

your offer

____ ⟵ Please indicate a hypothetical offer

The following two questions are related to the offer you have made.

How many MUs would you earn additionally to your 25 MUs endowment? _________

How many MUs would Player B earn additionally to his/her 25 MUs endowment? _________

5:1

your offer

____ ⟵ Please indicate a hypothetical offer

The following five questions are related to the offer you have made.

How many MUs would you earn if Player B would use 25 MUs to decrease your payoff?

____

How many MUs would Player B earn if he/she would use 25 MUs to decrease your payoff?

____

How many MUs would you earn if Player B would use 0 MUs to decrease your payoff?

____

How many MUs would Player B earn if he/she would use 0 MUs to decrease your payoff?

____

4

Figure 4.S1 (continued): Instructions given to the participants. Page 4/4.

Regression models

First Model: behavior in the Dictator Game (DG)

To test (i) whether participants act more selfishly when TMS is applied over the right DLPFC compared to sham and TMS over the left DLPFC, transfer decisions of the dictator game without punishment (DG) were regressed on a dummy predictor coding the three TMS conditions (sham condition as baseline). We controlled for the order in which participants experienced the three TMS conditions in two ways. For each observation a session variable coding for the session number was introduced, as well as a dummy variable coding for the condition order (e.g. sham-left-right).

A substantial fraction of the dictator transfers was zero. We therefore treated the data as left censored and fitted a Bayesian random-intercept Tobit regression model to the data using R and JAGS. Non-informative Gaussian priors (m = 0, SD = 100) were used for each predictor and non-informative uniform distributions (range 0 to 100) for the level-1 and level-2 error term. For the fitting we used three chains. $\hat{R}$ was below 1.1 for all parameters, indicating good mixing of the three chains and thus high convergence (Brooks and Gelman, 1998). Table 4.S1 shows estimated beta coefficients together with the 95% highest density interval (HDI, also called Bayesian confidence interval) for each predictor. Note that, since non-informative priors were used, a 95% HDI that only contains negative or positive values can be interpreted as significant at a $p = 0.05$ two-sided threshold in a frequentist framework. However, since the predictors are treated as random variables in the Bayesian framework, the HDI can further be interpreted as the probability distribution of this parameter and the estimate as the point with the highest likelihood.

Participants gave significantly less when the right DLPFC was disrupted by TMS compared to sham (TMS right coefficient, see Table 4.S1). When the left DLPFC was disrupted, we did not find evidence for a significant change in transfer rates compared to sham (TMS left coefficient, see Table 4.S1). Examining the posterior distributions of the TMS left and TMS right parameter revealed an estimated difference of -7.3 with a 95% HDI ranging from -10.3 to -4.3. Thus, participants not only gave significantly less to recipients under right TMS compared to sham but also compared to left TMS. Figure 4.S2 shows residual diagnostic plots for the fitted model and a posterior predictive check comparing the actual observed transfer frequencies and the frequencies expected by the fitted model based on 10,000 simulations (Gelman, Meng and Stern, 1996). As can be seen in Figure 4.S2a, transfers above 40 MUs were slightly overestimated by the model.

Table 4.S1

Coefficients and 95% interval for the Bayesian random-intercept Tobit regression model testing hypothesis (i).

		Coef.	**95% CI**
Fixed part			
intercept	Sham TMS	-35.14	[-114.41, 46.00]
TMS	left DLPFC	-1.83	[-6.86, 3.14]
TMS	right DLPFC	-9.15	[-14.37, -3.92]
order	session	4.17	[1.66, 6.68]
order	left-right-sham (n = 2)	19.38	[-80.21, 120.99]
order	left-sham-right (n = 3)	-24.34	[-128.65, 73.70]
order	right-left-sham (n = 4)	-0.80	[-98.65, 93.48]
order	right-sham-left (n = 2)	-10.32	[-117.01, 97.41]
order	sham-left-right (n = 3)	-21.64	[-118.88, 73.43]
order	sham-right-left (n = 3)	34.12	[-58.34, 127.47]
Random part			
error term	intercept	58.67	[30.05, 95.54]
error term	y	15.67	[14.02, 17.61]

Note. Dependent variable: Transfer decisions in the dictator game without punishment (DG).

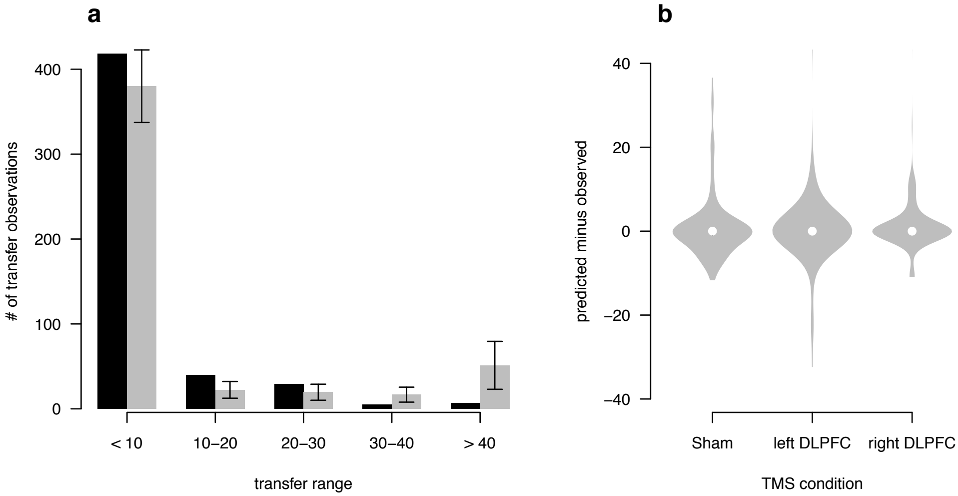


Figure 4.S2: Posterior predictive check and residual plots of the fitted model. (a) Posterior predictive simulation of transfer frequencies (±1 SD, grey bars) compared to actual observed frequencies (black bars) and (b) residual distributions separated by TMS condition. White point shows the median.

Second Model: strategic adaption across Dictator Games with (DGp) and without punishment (DG)

To test hypothesis (ii), whether participants show less strategic fairness when TMS is applied over the right DLPFC compared to sham and TMS over the left DLPFC we first looked at the behavior of each dictator during sham TMS and classified participants into 'adapters' and 'non-adapters'. Those who gave more to recipients with punishment opportunity over the 20 dictator game rounds during sham were classified as 'adapters' and those who gave less or equal to recipients with punishment opportunity were classified as 'non-adapters'. For each participant we calculated the transfer difference across DG and DGp as a measure for strategic adaption and regressed it on the dummy predictor coding the three TMS conditions (sham condition as baseline) as well as the non-adaption-dummy predictor, which takes value 1 for participant classified as non-adapter and 0 otherwise. Like in the first model we controlled for the order of the TMS treatments.

A random intercept regression was fitted to the data using R and JAGS. Non-informative Gaussian priors (m = 0, SD = 100) were used for each predictor and non-informative uniform distributions (range 0 to 100) for the level-1 and level-2 error term. For the fitting we used three chains. $\hat{R}$ was below 1.1 for all parameters, indicating good mixing of the chains and thus high convergence (Brooks and Gelman, 1998). Table 4.S2 shows estimated beta coefficients together with the 95% confidence interval for each predictor. Note that predictor coefficients that are not interactions with the non-adapter dummy (fixed part, adapters, see Table 4.S2) can be interpreted as the behavior of adapters, while the predictor coefficients that do interact with the non-adapter dummy (fixed part, non-adapters, see Table 4.S2) is the change in behavior of non-adapters compared to adapters.

Participants who adapted during sham did so significantly less when the right DLPFC was disrupted (TMS right, see Table 4.S2). We did not observe a significant change in strategic adaption of adapters during the disruption of the left DLPFC (TMS left, see Table 4.S2). Examining the posterior distributions of the TMS left and TMS right parameter revealed an estimated difference of -8.91 with a 95% HDI ranging from -9.90 to -7.93. Thus, participants who adapted strategically during sham did so significantly less during TMS over the right DLPFC compared to TMS over the left DLPFC.

Figure 4.S3 shows residual diagnostic plots for the fitted model and a posterior predictive check comparing the actual observed transfer change frequencies and the frequencies expected by the fitted model based on 10,000 simulations (Gelman, Meng and Stern, 1996). As can be seen in Figure 4.S3a, there is no systematic over- or underestimation of transfer change frequencies by the model.

Third Model: transfer change across Dictator Games with (DGp) and without punishment (DG)

We also fitted a model on the individual trial level by using the transfers in each round and regressed it on all TMS condition $\times$ punishment condition $\times$ non-adaption interaction terms. We controlled for the order of the TMS treatments and treated the data as left censored. In this model, a change in strategic adaption (TMS condition $\times$ punishment condition predictor) due to TMS can be analyzed while controlling for a possible change in selfishness in the DG due to TMS (TMS condition predictor).

Non-informative Gaussian priors (m = 0, SD = 100) were used for each predictor and non-informative uniform distributions (range 0 to 100) for the level-1 and level-2 error term. For the fitting we used three chains. $\hat{R}$ was below 1.1 for all parameters, indicating good mixing of the three chains and thus high convergence (Brooks and Gelman, 1998). Table 4.S3 shows estimated beta coefficients together with the 95% confidence interval for each predictor. Note that predictor coefficients that are not interactions with the non-adapter dummy (fixed part, adapters, see Table 4.S3) can be interpreted as the behavior of adapters, while the predictor coefficients that do interact with the non-

Table 4.S2
Coefficients and 95% interval for the Bayesian random-intercept Poisson regression model testing hypothesis (ii).

		Coef.	95% CI
Fixed part (adapters)			
intercept	Sham TMS, no punishment, adapters	26.69	[-50.16, 103.45]
TMS	left DLPFC	2.61	[-0.64, 4.31]
TMS	right DLPFC	-6.30	[-8.04, -4.59]
Fixed part (non-adapters)			
NA	non-adapters	-26.45	[-42.73, -9.89]
NA × TMS	non-adapters, left DLPFC	2.85	[-0.04, 5.73]
NA × TMS	non-adapters, right DLPFC	14.58	[11.79, 17.40]
order	session	0.01	[-0.68, 0.69]
order	left-right-sham (n = 2)	4.18	[-74.62, 82.88]
order	left-sham-right (n = 3)	-14.22	[-92.36, 63.83]
order	right-left-sham (n = 4)	1.01	[-76.70, 79.51]
order	right-sham-left (n = 2)	9.06	[-69.21, 87.65]
order	sham-left-right (n = 3)	1.94	[-76.01, 80.78]
order	sham-right-left (n = 3)	-5.17	[-84.76, 73.28]
Random part			
error term	intercept	13.78	[8.63, 22.99]
error term	y	10.88	[10.49, 11.29]

Note. Dependent variable: Transfer difference.

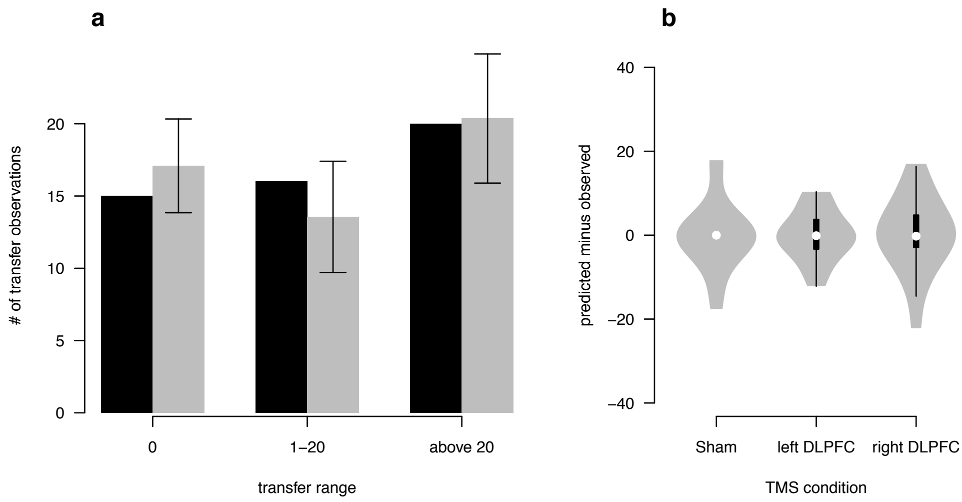


Figure 4.S3: Posterior predictive check and residual plots of the fitted model. (a) Posterior predictive simulation of transfer change frequencies (± 1 SD, grey bars) compared to actual observed frequencies (black bars) and (b) residual distributions separated by TMS condition. White point shows the median.

adapter dummy (fixed part, non-adapters, see Table 4.S3) is the change in behavior of non-adapters compared to adapters.

Participants who adapted during sham did so significantly less when the right DLPFC was disrupted (TMS right $\times$ DG coefficient, see Table 4.S3). We did not observe a significant change in strategic adaption of adapters during the disruption of the left DLPFC (TMS left $\times$ DG coefficient, see Table 4.S3). Examining the posterior distributions of the TMS left $\times$ Punishment and TMS right $\times$ Punishment parameter revealed an estimated difference of -8.90 with a 95% HDI ranging from -13.40 to -6.49. Thus, participants who adapted strategically during sham did so significantly less during TMS over the right DLPFC compared to TMS over the left DLPFC. Figure 4.S4 shows residual diagnostic plots for the fitted model and a posterior predictive check comparing the actual observed transfer frequencies and the frequencies expected by the fitted model based on 10,000 simulations (Gelman, Meng and Stern, 1996). As can be seen in Figure 4.S4a, there is no systematic over- or underestimation of transfer-frequencies by the model.

Table 4.S3
Coefficients and 95% interval for the Bayesian random-intercept Tobit regression model testing hypothesis (ii).

		Coef.	95% CI
Fixed part (adapters)			
intercept	Sham TMS, no punishment, adapters	-0.27	[-67.34, 62.73]
TMS	left DLPFC	-2.86	[-5.86, 1.13]
TMS	right DLPFC	-3.86	[-6.87, -0.87]
DG	with punishment	25.97	[22.95, 28.94]
TMS × DG	with punishment, left DLPFC	3.53	[-0.65, 7.77]
TMS × DG	with punishment, right DLPFC	-5.37	[-9.56, -1.12]
Fixed part (non-adapters)			
NA	non-adapters	0.66	[-10.30, 11.67]
NA × TMS	non-adapters, left DLPFC	7.87	[2.89, 12.90]
NA × TMS	non-adapters, right DLPFC	2.25	[-2.72, 7.26]
NA × DG	non-adapters, with punishment	-29.36	[-34.29, -24.30]
NA × TMS × DG	non-adapters, with punishment, left DLPFC	-3.83	[-10.96, 3.13]
NA × TMS × DG	non-adapters, with punishment, right DLPFC	13.62	[6.54, 20.62]
order	session	0.71	[-0.13, 1.56]
order	left-right-sham (n = 2)	-2.23	[-74.25, 46.86]
order	left-sham-right (n = 3)	-15.96	[-86.60, 32.12]
order	right-left-sham (n = 4)	-0.81	[-72.26, 47.73]
order	right-sham-left (n = 2)	-4.19	[-76.12, 45.32]
order	sham-left-right (n = 3)	-6.98	[-78.98, 41.24]
order	sham-right-left (n = 3)	-2.72	[-74.36, 46.96]
Random part			
error term	intercept	8.88	[5.49, 14.84]
error term	y	11.11	[10.63, 11.62]

Note. Dependent variable: Transfer decisions.

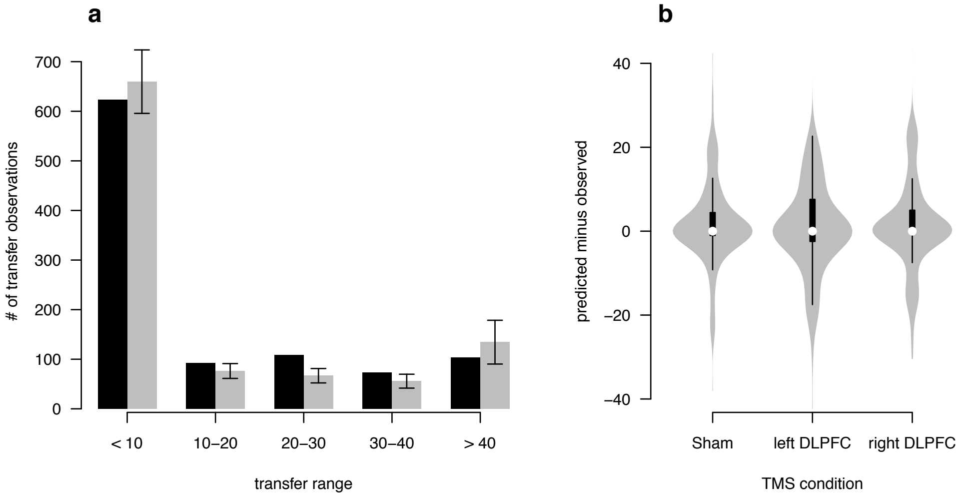


Figure 4.S4: Posterior predictive check and residual plots of the fitted model. (a) Posterior predictive simulation of transfer frequencies ($\pm$1 SD, grey bars) compared to actual observed frequencies (black bars) and (b) residual distributions separated by TMS condition. White point shows the median.

Fairness judgements

After each dictator game in each TMS session, participants made fairness judgements about five hypothetical transfers (from 0 to 50 MUs in steps of 10) on a scale from 1 ("very unfair") to 7 ("very fair").

To test whether fairness judgements were systematically influenced by TMS we regressed the responses on a dummy predictor coding the three TMS conditions (sham condition as baseline), as well as a predictor coding for the different hypothetical offers in increasing order, the non-adaption dummy already used in the above model, and control variables for the order. Non-informative Gaussian priors (m = 0, SD = 100) were used for each predictor and non-informative uniform distributions (range 0 to 100) for the level-1 and level-2 error term. For the fitting we used three chains. $\hat{R}$ was below 1.1 for all parameters, indicating good mixing of the three chains and thus high convergence (Brooks and Gelman, 1998). Table 4.S4 shows estimated beta coefficients together with the 95% confidence interval for each predictor. Note again, that predictor coefficients that are not interactions with the non-adapter dummy (fixed part, adapters, see Table 4.S4) can be interpreted as the behavior of adapters, while the predictor coefficients that do interact with the non-adapter dummy (fixed part, non-adapters, see Table 4.S4) is the change in behavior compared to adapters.

As can be expected, fairness judgements significantly increased with the hypothetical offer during sham for adapters (offer coefficient, see Table 4.S4). For adapters, this increase in fairness judgements was not significantly altered by the TMS manipulations (TMS left $\times$ offer coefficient and TMS right $\times$ offer coefficient, see Table 4.S4). Interestingly, non-adapters rated offers to be more fair compared to adapters (non-adaption coefficient, see Table 4.S4) but with increased offer showed a significantly lower slope in rating higher offers as more fair (non-adaption $\times$ offer, see Table 4.S4). This indicates that their fairness judgements were less influenced by the size of the offer. This relative insensitivity to changes in the offer could explain why they did not adapt in the first place. Since they did not perceive higher offers as fairer (at least not as much as observed for adapters), it could be that they did not feel an obligation to make higher offers when under the threat of punishment.

Figure 4.S5 shows residual diagnostic plots for the fitted model and a posterior predictive check comparing the actual observed transfer frequencies and the frequencies expected by the fitted model based on 10,000 simulations (Gelman, Meng and Stern, 1996). As can be seen in Figure 4.S5a, fairness judgements in the center were slightly over- or underestimated, while judgements at the end of the scale were neither systematically over- nor underestimated by the model. Overall, the model captured the general frequency trend of judgements.

Expected punishment

Next to fairness evaluations, participants were asked about how many MUs (from 0 to 25) they believed the receivers would on average spent on punishment for a given hypothetical offer (from 0 to 50 MUs in steps of 10).

To test whether punishment expectations were systematically influenced by TMS we regressed the responses to dummy predictors coding the three TMS conditions (sham condition as baseline), as well as a predictor coding for the hypothetical offer, the non-adaption dummy already used in the above models, and predictors coding for the session and TMS order.

Again, the data was treated as left-censored and non-informative Gaussian priors (m = 0, SD = 100) were used for each predictor and non-informative uniform distributions (range 0 to 100) for the level-1 and level-2 error term. For the fitting we used three chains. $\hat{R}$ was below 1.1 for all parameters, indicating good mixing of the three chains and thus high convergence (Brooks and Gelman, 1998). Table 4.S5 shows estimated beta coefficients together with the 95% confidence interval for each predictor. Note again, that predictor coefficients that are not interactions with the non-adapter dummy (fixed

Table 4.S4
Coefficients and 95% interval for the Bayesian random-intercept regression model testing the influence of TMS on fairness judgements.

		Coef.	95% CI
Fixed part (adapters)			
intercept	Sham TMS, adapters, zero offer	4.18	[-38.86, 33.45]
TMS	left DLPFC	0.23	[-0.86, 1.35]
TMS	right DLPFC	-0.18	[-1.31, 0.95]
offer	hypothetical offer	0.13	[0.10, 0.15]
offer × TMS	offer, left DLPFC	-0.02	[-0.05, 0.02]
offer × TMS	offer, right DLPFC	0.00	[-0.04, 0.03]
Fixed part (non-adapters)			
NA	non-adapters	3.25	[1.61, 4.87]
NA × offer	non-adapters, offer	-0.14	[-0.18, -0.10]
TMS × NA	left DLPFC, non-adapters	-0.84	[-2.61, 0.91]
TMS × NA	right DLPFC, non-adapters	-0.74	[-2.52, 1.03]
TMS × offer × NA	hypothetical offer, left DLPFC, non-adapters	0.05	[-0.01, 0.10]
TMS × offer × NA	hypothetical offer, right DLPFC, non-adapters	0.03	[-0.03, 0.08]
order	session	-0.01	[-0.24, 0.23]
order	left-right-sham (n = 2)	-2.73	[-32.16, 40.27]
order	left-sham-right (n = 3)	-3.98	[-33.25, 39.20]
order	right-left-sham (n = 4)	-3.68	[-32.95, 39.51]
order	right-sham-left (n = 2)	-3.62	[-32.90, 39.60]
order	sham-left-right (n = 3)	-4.56	[-34.06, 38.48]
order	sham-right-left (n = 3)	-3.43	[-32.78, 39.66]
Random part			
error term	intercept	0.87	[0.33, 1.52]
error term	y	1.63	[1.48, 1.79]

Note. Dependent variable: fairness judgements.

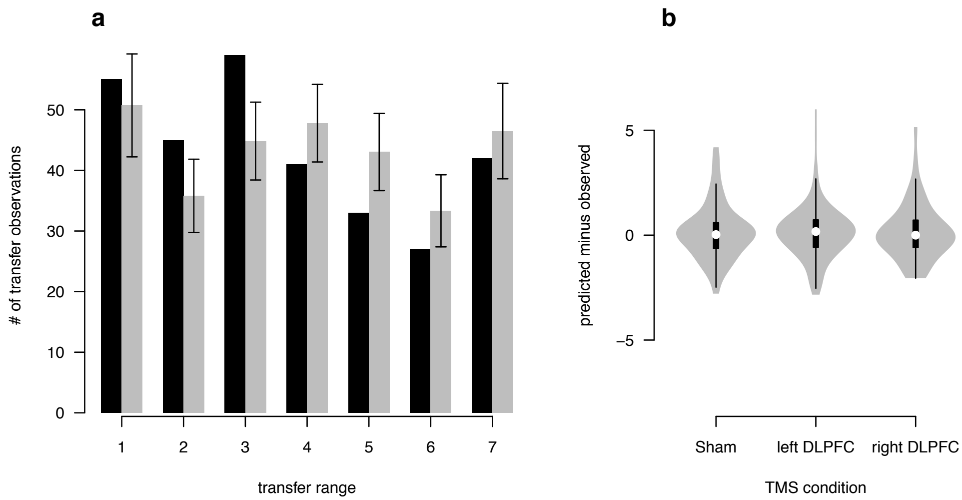


Figure 4.S5: Posterior predictive check and residual plots of the fitted model. (a) Posterior predictive simulation of transfer frequencies (±1 SD, grey bars) compared to actual observed frequencies (black bars) and (b) residual distributions separated by TMS condition. White point shows the median.

part, adapters, see Table 4.S5) can be interpreted as the behavior of adapters, while the predictor coefficients that do interact with the non-adapter dummy (fixed part, non-adapters, see Table 4.S5) is the change in behavior of non-adaptors compared to adapters.

As can be expected, with increased offer, adapting participants expected the receivers to punish less. This expectation was not significantly altered by TMS in adapters. Interestingly, non-adapters not only expected significantly less punishment in general (NA coefficient, see Table 4.S5) but also expected that punishment does not increase with increasingly unfair offers (slope of 0.06 for increasing offers; coefficient NA × offer plus offer). This could explain why they did not adapt to the punishment threat in the first place. It could also mean that they did not believe that punishment would actually take place.

Figure 4.S6 shows residual diagnostic plots for the fitted model and a posterior predictive check. As can be seen in Figure 4.S6a, expected punishment in the range from 0-25 MUs was slightly overestimated by the model.

Table 4.S5

Coefficients and 95% interval for the Bayesian random-intercept regression model testing the influence of TMS on expected punishment.

		Coef.	95% CI
Fixed part (adapters)			
intercept	Sham TMS, adapters, zero offer	11.72	[-46.92, 77.46]
TMS	left DLPFC	-3.45	[-7.12, 0.20]
TMS	right DLPFC	-3.55	[-7.22, 0.14]
offer	hypothetical offer	-0.46	[-0.56, -0.37]
TMS × offer	left DLPFC, offer	0.09	[-0.04, 0.22]
TMS × offer	right DLPFC, offer	0.08	[-0.05, 0.22]
Fixed part (non-adapters)			
NA	non-adapters	-28.66	[-43.87, -15.29]
NA × offer	non-adapters, offer	0.52	[0.34, 0.70]
TMS × NA	left DLPFC, non-adapters	14.88	[7.69, 22.13]
TMS × NA	right DLPFC, non-adapters	8.21	[-0.98, 15.68]
TMS × offer × NA	hypothetical offer, left DLPFC, non-adapters	-0.33	[-0.57, 0.09]
TMS × offer × NA	hypothetical offer, right DLPFC, non-adapters	-0.17	[-0.42, 0.07]
order	session	0.85	[-0.06, 1.77]
order	left-right-sham (n = 2)	7.04	[-60.81, 67.75]
order	left-sham-right (n = 3)	6.43	[-61.93, 66.21]
order	right-left-sham (n = 4)	4.64	[-61.41, 63.96]
order	right-sham-left (n = 2)	6.98	[-60.57, 66.89]
order	sham-left-right (n = 3)	5.50	[-61.16, 65.16]
order	sham-right-left (n = 3)	13.21	[-55.39, 74.16]
Random part			
error term	intercept	10.32	[5.64, 19.37]
error term	y	5.78	[5.20, 6.44]

Note. Dependent variable: expected punishment.

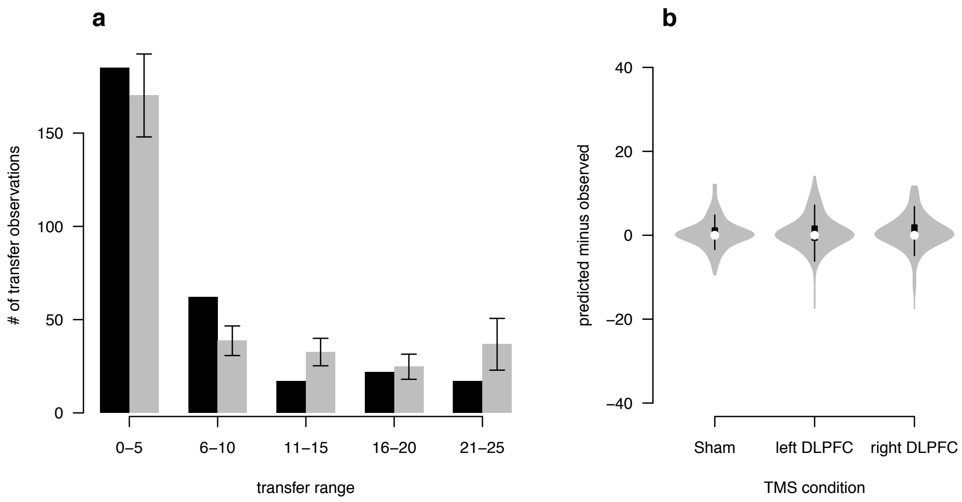


Figure 4.S6: Posterior predictive check and residual plots of the fitted model. (a) Posterior predictive simulation of transfer frequencies (±1 SD, grey bars) compared to actual observed frequencies (black bars) and (b) residual distributions separated by TMS condition. White point shows the median.

Own imagined punishment

Lastly, participants were asked how many MUs (from 0 to 25) they would spent on punishment were they in the shoes of a receiver receiving hypothetical offers from 0 to 50 MUs in steps of 10.

As above, own imagined punishment expenses were regressed on a dummy predictor coding the three TMS conditions (sham condition as baseline), as well as a predictor coding for the hypothetical offer, the adaption dummy already used in the above models, and predictors coding for the session and TMS order. The data was treated as left-censored and non-informative Gaussian priors (m = 0, SD = 100) were used for each predictor and non-informative uniform distributions (range 0 to 100) for the level-1 and level-2 error term. For the fitting we used three chains. $\hat{R}$ was again below 1.1 for all parameters, indicating good mixing of the three chains and thus high convergence (Brooks and Gelman, 1998). Table 4.S6 shows estimated beta coefficients together with the 95% confidence interval for each predictor.

Adapting participants indicated that they would spent less MUs on punishment the fairer the offer is (offer coefficient, see Table 4.S6) and non-adapting participants did not significantly differ (NA $\times$ offer coefficient, see Table 4.S6). However, non-adapters indicated to spent significantly less on punishment (non adapters coefficient, see Table 4.S6). Adapting participants reported that they would spent less MUs on punishment while under the influence of TMS over the right DLPFC (TMS right coefficient, see Table 4.S6).

Figure 4.S7 shows residual diagnostic plots for the fitted model and a posterior predictive check. As can be seen in Figure 4.S7a, expected punishment in the range from 0–25 MUs was slightly overestimated by the model.

Table 4.S6

Coefficients and 95% interval for the Bayesian random-intercept regression model testing the influence of TMS on imagined punishment.

		Coef.	95% CI
Fixed part (adapters)			
intercept	Sham TMS, adapters, zero offer	3.15	[-49.32, 60.14]
TMS	left DLPFC	0.23	[-3.70, 4.18]
TMS	right DLPFC	-3.71	[-7.84, -1.30]
offer	hypothetical offer	-0.47	[-0.58, -0.37]
offer × TMS	offer, left DLPFC	0.04	[-0.10, 0.18]
offer × TMS	offer, right DLPFC	0.08	[-0.06, 0.22]
Fixed part (non-adapters)			
NA	non-adapters	-31.05	[-60.11, -4.92]
NA × offer	non-adapters, offer	0.21	[-0.02, 0.43]
TMS × NA	left DLPFC, non-adapters	-2.15	[-10.66, 6.35]
TMS × NA	right DLPFC, non-adapters	6.97	[-1.32, 15.38]
TMS × offer × NA	hypothetical offer, left DLPFC, non-adapters	0.15	[-0.15, 0.45]
TMS × offer × NA	hypothetical offer, right DLPFC, non-adapters	-0.02	[-0.32, 0.29]
order	session	1.45	[0.40, 2.51]
order	left-right-sham (n = 2)	13.72	[-47.47, 68.92]
order	left-sham-right (n = 3)	2.35	[-58.11, 62.34]
order	right-left-sham (n = 4)	14.78	[-45.69, 73.35]
order	right-sham-left (n = 2)	-8.15	[-74.45, 49.58]
order	sham-left-right (n = 3)	22.18	[-37.28, 79.96]
order	sham-right-left (n = 3)	17.44	[-42.22, 76.78]
Random part			
error term	intercept	19.06	[9.36, 40.11]
error term	y	5.75	[5.08, 6.53]

Note. Dependent variable: own imagined punishment expenses.

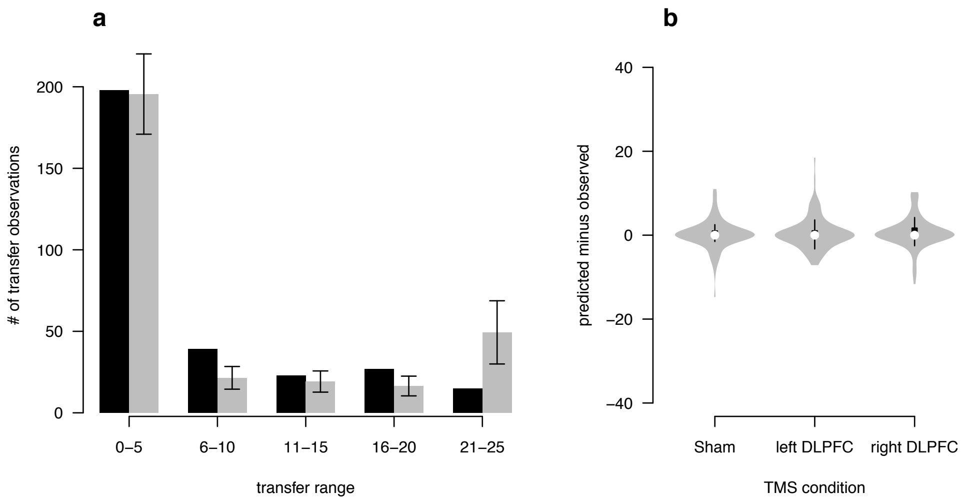


Figure 4.S7: Posterior predictive check and residual plots of the fitted model. (a) Posterior predictive simulation of transfer frequencies (± 1 SD, grey bars) compared to actual observed frequencies (black bars) and (b) residual distributions separated by TMS condition. White point shows the median.

Chapter 5

Building the Leviathan – Voluntary Centralization of Power Sustains Cooperation

Abstract. The prevalence of cooperation among humans is puzzling because cooperators can be exploited by free riders. Peer punishment has been suggested as a key to this puzzle, but cumulating evidence questions its effectiveness in sustaining cooperation. Punishment has to be powerful, meaning that the effect of punishment has to be sufficiently larger than its cost. Since group members can refrain from punishing non-cooperators, peer punishment poses a social dilemma in itself. By exploring a novel experimental setup, we show that the voluntary transfer of punishment power enables groups to overcome the problem of free riding when peer punishment can not. Participants are willing to empower individuals who act in the interest of the group. The endogenous establishment of power centralization solves the social dilemma inherent to peer punishment, sustains cooperation, and increases welfare. Our results could explain why hierarchical power structures are widespread among animals and humans.

Based on Jörg Gross, Zsombor Méder, Sanae Okamoto-Barth, Arno Riedl (submitted).

5.1 Introduction

The scale of cooperation observed among humans constitutes a puzzle for the social and biological sciences. Cooperative efforts bear the risk of exploitation as selfish agents can reap their benefits without contributing to the common good. Nevertheless, cooperation is frequently sustained in human societies (Trivers, 1971, Alexander, 1987, Fehr and Fischbacher, 2003, Nowak and Sigmund, 2005). Peer punishment has been proposed as a possible means to sustain cooperation (Güth et al., 1982, Yamagishi, 1986, Boyd and Richerson, 1992, Ostrom et al., 1992, Fehr and Gächter, 2000, Fehr and Gächter, 2002, de Quervain et al., 2004, Walker and Halloran, 2004, Falk et al., 2005). Experiments on public goods dilemmas show that cooperation deteriorates quickly in the absence of sanctioning mechanisms, but can stabilize when the punishment of free riders is possible (Yamagishi, 1986, Yamagishi, 1988, Fehr and Gächter, 2000, Fehr and Gächter, 2002, Gächter et al., 2008).

However, peer punishment does not guarantee cooperation. Recent research documented crucial limitations of peer punishment in its ability to sustain cooperation and foster welfare. The cost of punishment has to be sufficiently low relative to its effect on the punished (Masclet et al., 2003, Casari, 2005, Falk et al., 2005, Egas and Riedl, 2008, Nikiforakis and Normann, 2008, Ohtsuki et al., 2009). Further, non-cooperators sometimes punish cooperators out of spite or retribution, thereby undermining cooperation (Dreber et al., 2008, Herrmann et al., 2008, Janssen and Bushman, 2008, Nikiforakis, 2008, Gächter and Herrmann, 2009, Rand et al., 2010, Rand and Nowak, 2011, Dreber and Rand, 2012). The excessive use of punishment can stabilize cooperation but leads to adverse results for group welfare (Masclet et al., 2003, Dreber et al., 2008, Gächter et al., 2008, Janssen et al., 2010, Engelmann and Nikiforakis, ress). Finally, not all members of a group are willing to participate in the punishment of non-cooperators. It is frequently observed that some, while willing to cooperate, are not willing to punish non-cooperators. Hence, peer punishment generates a second-order social dilemma in which cooperators not willing to punish can second-order free ride on those who do engage in costly punishment (Boyd et al., 2003, Panchanathan and Boyd, 2004, Fowler, 2005, Kiyonari and Barclay, 2008, Perc, 2012).

Social institutions provide an alternative for upholding cooperation through centralized punishment mechanisms (Yamagishi, 1986, Ostrom, 1990, Casari and Plott, 2003, Rockenbach et al., 2006, Sefton et al., 2007, Kosfeld et al., 2009, Boyd et al., 2010, Sigmund et al., 2010, Guala, 2012). Laws are issued to tackle 'tragedies of the commons', like problems of over-fishing, littering or air pollution. Contracts are made between individuals to prevent exploitation in mutual agreements like rentals, insurances or investments. Authorities, like courts or the police force, enforce these formal institutions.

An essential characteristic of these executive institutions is that they embody a centralized power to punish (Flew, 1954).

Centralized power is also an important feature of groups that are not fully governed by laws or contracts, ranging from hunter-gatherer chiefdoms to modern internet communities. For example, Wikipedia provides a global public good to which everybody can contribute, but only a small share of its editors holds the right to enforce policy and sanction antisocial behaviour.

We propose that this centralization is created and maintained through the voluntary transfer of power to a few trusted individuals. In particular, the transfer of punishment power, fundamental to executive institutions, could solve the problems associated with decentralized peer punishment. We test this hypothesis experimentally by introducing a new experimental setup, which we call the 'power transfer game'.

5.2 Materials and methods

5.2.1 The power transfer game

The power transfer game consists of three stages: the power transfer, the contribution, and the punishment stage. In our experiment, participants played the game in groups of five.

In the power transfer stage, group members can give up and transfer punishment power to other group members at no cost. Each group member has a power of 1 at her disposal. Power can be transferred in units of 0.1 and can be distributed among multiple group members. The sum of power kept and received from others determines one's punishment effectiveness (see below). After all power transfer decisions are made, everyone is informed about how much punishment power each group member has. Power transfer does not change the total punishment power in the group but (may) change its distribution among group members.

In the contribution stage, representing the classic public goods game, participants receive an endowment of 20 monetary units (20 MUs = 0.50 euro) and decide how much to contribute to a 'group project'. Group members keep any MUs not contributed. The sum of MUs contributed to the group project is then multiplied by 1.5, and distributed equally among all five group members, regardless of how much each individual contributed. This poses a social dilemma: The return of each contributed MU is 1.5 MU for the group, but only 0.3 MU for the individual. Therefore, if participants are selfish payoff-maximisers, they should not contribute at all. However, if everyone contributes their entire endowment to the group project, participants will earn 30 MUs (20 MUs $\times$ 5 group members $\times$ 1.5 multiplier / 5 group members), whereas if no one contributes, participants will only earn 20 MUs. Contribution decisions are made simultaneously. Afterwards, everyone is informed about how many MUs each of the other group members contributed and how many they kept for themselves.

Finally, in the punishment stage, group members are able to punish their peers. Punishment is dealt out by assigning between 0 and 10 punishment points. Punishment decisions are made simultaneously. Subsequently, group members see how many punishment points each group member assigned, and to whom. Punishment reduces the earnings of both the punisher and the punished. For each assigned punishment point, the punisher pays 1 MU. Importantly, the amount of MUs that are deducted from the punished is determined by the power of the punisher. For example, assume that in the power transfer stage, group member A decided to transfer all of her power to group member B, and no one else transferred any power. B would now have a power of 2, whereas A would have a power of 0. Suppose both A and B decide to punish group member C. Because of her increase in power, each point B uses to punish C will lead to a reduction of 2 MUs in earnings for C. On the other hand, A, who transferred all her power to B, and did not receive any power from others, would have a punishment effectiveness of 0. Thus, A would still pay 1 MU for each punishment point that she assigns to C, but those would not decrease the earnings of C. A game theoretic description and a detailed presentation of the computer interface can be found in the Supporting Information.

5.2.2 Experimental implementation

Participants were recruited from the subject pool of the behavioural and experimental economics lab (BEElab) at Maastricht University and were invited via e-mail. Each experimental session comprised at least 3 and at most 5 groups. Participants (n = 350, mean age = 21.1, SD = 2.6, 204 female) were allocated to one of three conditions that differed in how punishment power was determined. The experiment lasted for a total of 20 rounds in fixed groups of five. In the endogenous condition (n = 135) participants played the power transfer game described above. Voluntary transfer of power was only possible in this condition.

Power transfer decisions made in the previous round served as the default option for the current round. However, participants could freely change their allocation each round.

The remaining participants were allocated to one of two control conditions: the fixed condition (n = 80) or the exogenous condition (n = 135). In the fixed condition, each participant had a punishment power of 1 and participants were not able to transfer any power. Thus, each group member had a 1:1 effectiveness-to-cost ratio of punishment during the whole experiment. In the exogenous condition, power transfer was not voluntary. Instead, for each group in the endogenous condition we created a twin group in the exogenous condition, in which the history of power transfers and, hence, punishment effectiveness, was mirrored at the individual level. Thus, each group member followed the same change in punishment power across rounds as its twin. The purpose of the exogenous condition was to test whether it is important that power transfer is voluntarily. In the endogenous condition, power transfers can make some group members more effective punishers. However, group members are also free to determine who these effective punishers are going to be. The exogenous condition therefore disentangles the effects of increasing punishment effectiveness and voluntarily selecting group members to hold punishment power.

In all three conditions, the different stages of the game were introduced sequentially to the participants (Figure 5.1). The experiment started with a round consisting of only a contribution stage (public goods game). The second round consisted of a contribution and a punishment stage (public goods game with punishment). In the third round, the power mechanism was introduced to the experiment according to the condition. Subsequent rounds followed the structure of the third round.

5.3 Results

Participants in all three conditions transferred roughly half of their endowment to the group project in the first round. In the fixed condition with decentralized 1:1 punishment, cooperation decreased steadily (Figure 5.2a; mixed effect regression, round coefficient = -0.28, 95% CI = [-0.51, -0.05]). In contrast, in the endogenous condition, with voluntary transfer of power, initial cooperation was not only sustained, but even increased slightly over time (Figure 5.2a; mixed effect regression, round $\times$ endogenous condition coefficient = 0.46, 95% CI = [0.16, 0.74]). This was not the case for groups in the exogenous condition. Lacking the freedom to decide whom to transfer power to, these groups showed a decline in cooperation that was not significantly different from that in the fixed condition (Figure 5.2a; mixed effect regression, round $\times$ exogenous condition coefficient = 0.22, 95% CI = [-0.06, 0.51]). Thus, only the voluntary transfer of power could solve the social dilemma.

The punishment histories for all three conditions are displayed in Figure 5.2b. Overall, average MUs assigned for punishment declined over the course of the experiment. This decline was the strongest in the endogenous condition (mixed effect regression, round $\times$ endogenous condition coefficient = -0.07, 95% CI = [-0.13, -0.00]). Higher lev-

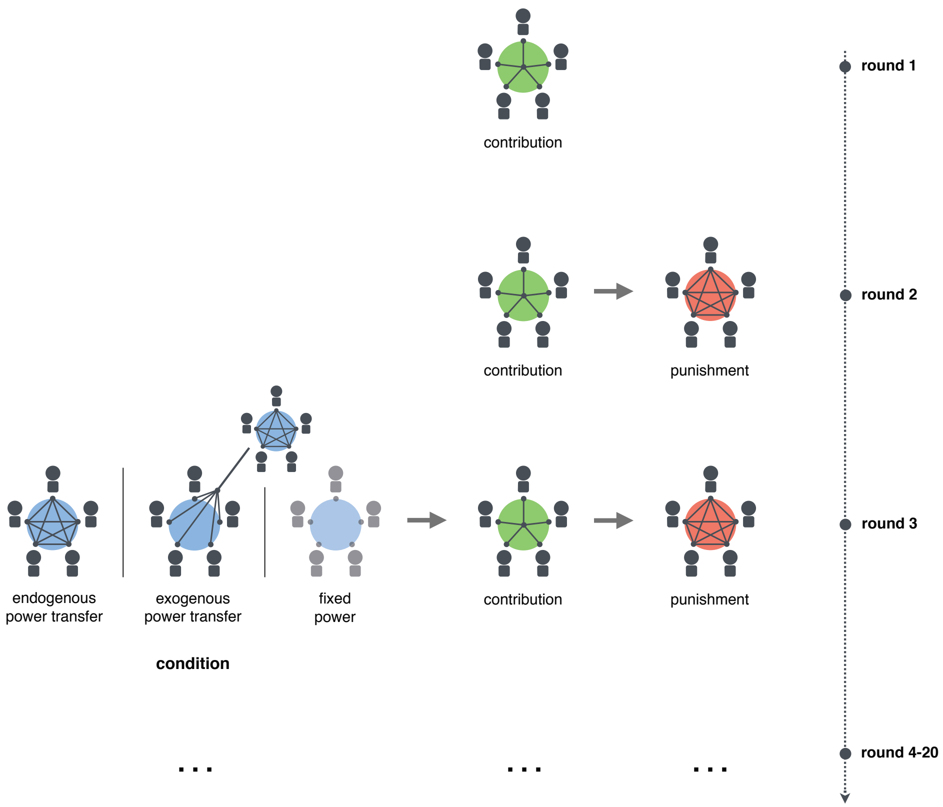


Figure 5.1: Timeline of the experiment. Groups in all three conditions started with one round of a contribution stage, followed by round 2, consisting of a contribution stage and a punishment stage. In round 3, the experimental manipulation was introduced. In the endogenous condition, representing the power transfer game, group members were able to transfer power to other group members before the contribution and punishment stages. Each exogenous condition group mirrored the power transfers of one endogenous condition group and thus group members were not able to transfer power voluntarily. In the fixed condition, power transfers were not possible, and everyone's power was fixed to 1. Rounds 4 to 20 followed the structure of round 3, according to the condition.

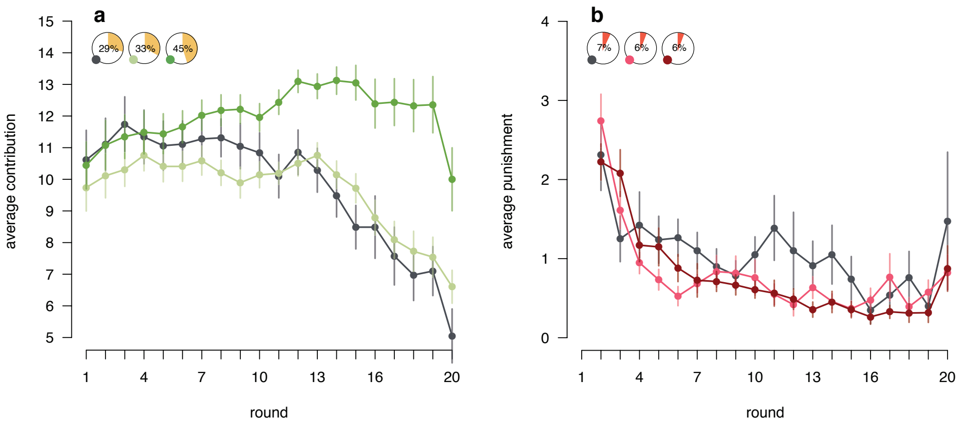


Figure 5.2: Cooperation and punishment over rounds. (a) Mean contributions to the group project for endogenous (dark green), exogenous (light green) and fixed (grey) conditions. Yellow pie charts show overall earnings as a percentage of the social optimum (maximum cooperation without punishment, 30 MUs per group member = 100%), compared to the selfish outcome (minimal cooperation without punishment, 20 MUs per group member = 0%) for each condition. (b) Average amount of MUs spent on punishment in the endogenous (dark red), exogenous (light red) and fixed (grey) conditions. Red pie charts show the average amount of MUs lost due to punishment dealt and received as a percentage of the total earnings for each condition. Error bars show the within-subject standard errors of the mean.

els of cooperation, and the more pronounced decline in punishment in the endogenous condition paid off for participants in terms of earnings. Participants with the ability to transfer power earned progressively more compared to participants in the two control conditions (Figure 5.S13, mixed effect regression, round $\times$ endogenous condition coefficient = 1.47, 95% CI = [0.56, 2.41]; difference between round $\times$ endogenous and round $\times$ exogenous condition coefficient = 0.79, 95% CI = [0.32, 1.26]). In contrast, there was no significant difference in earnings over rounds between the fixed and exogenous condition (mixed effect regression, round $\times$ exogenous condition coefficient = 0.68, 95% CI = [-0.24, 1.59]). Thus, voluntary power transfer enabled participants to achieve an outcome much closer to the social optimum, thereby enhancing group welfare.

To understand how groups in the endogenous condition solved the cooperation dilemma, we first looked at the power centralization that emerged over time. Power was already transferred in the first round with a power transfer stage (round 3, see Figure 5.1). A substantial fraction of participants (37%) were willing to transfer power in this round. The amount of power held by the most powerful group member increased significantly over rounds (Figure 5.3a, mixed effect regression, round coefficient = 0.02, 95% CI = [0.00, 0.04]). Thus, power became more and more centralized. Centralization of power was positively related to cooperation in the endogenous condition, but not in the exogenous condition. For each group, we computed the correlation across rounds between power held by the most powerful group member and average cooperation. For groups who could transfer power voluntarily, higher power centralization was associated with higher average group cooperation (Figure 5.3b, mean Pearson's r = 0.24, one-sample t-test, $t(25) = 2.9$, $p < .01$, two-sided). In contrast, groups in the exogenous condition, experiencing exactly the same power centralization but without the ability to transfer power voluntarily, correlations between power centralization and cooperation were not significantly different from zero (Figure 5.3b, mean Pearson's r = 0.08; one-sample t-test, $t(26) = 1.3$, $p = 0.21$, two-sided).

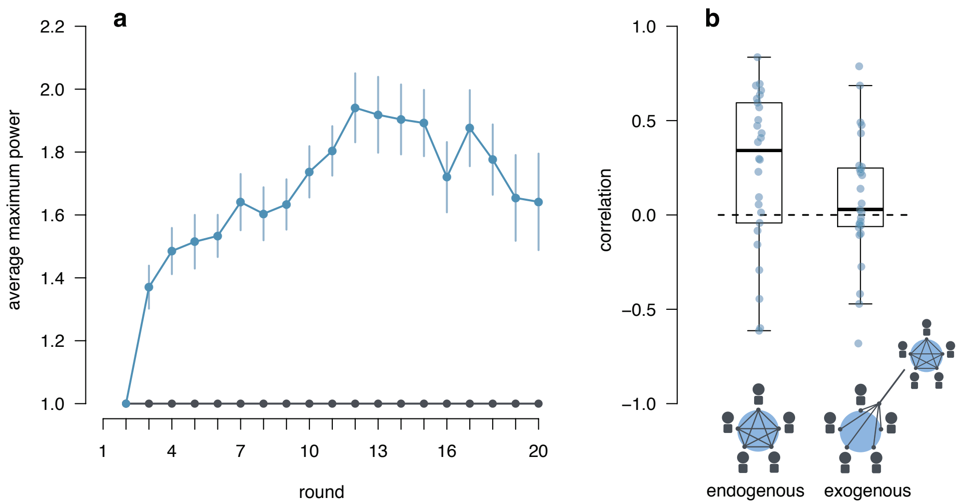


Figure 5.3: Power and cooperation. (a) Change of average power of the most powerful group member over rounds in the endogenous condition (blue). In the exogenous condition, power transfers were identical to the endogenous condition, and thus, the average power of the most powerful group member was the same, too. In the fixed condition, power was fixed to 1 (grey). Error bars show the within-subject standard errors of the mean. (b) Distribution of correlations across rounds between maximum power and cooperation for each group in the endogenous and exogenous condition. Thick horizontal bars represent the medians.

Next, we analysed decisions in the endogenous condition on the individual level in order to understand who transferred and who received power, how it was used and what effect it had on group members. Interestingly, even though participants were unaware of the subsequent introduction of the power transfer mechanism, behaviour in the first two rounds reliably predicted the average power status of a group member. Initial cooperators, those who contributed at or above the group average in the first round, received significantly more power over the course of the experiment than initial free riders, defined as group members who contributed less than the group average (Mann-Whitney U-test, $U = 2847.5$, $p < .01$, two-sided). Similarly, group members who punished free riders in the first punishment stage (round 2) received significantly more power from other group members than those who were not willing to punish (Mann-Whitney U-test, $U = 2294$, $p = 0.02$, two-sided).

Looking at power transfers from round to round, power transfers were mostly initiated by non-punishers. Group members with a lower than average punishment expenditure had a significant higher likelihood to give up power (mixed effect logistic regression, t-1 punishment difference coefficient = 0.51, 95% CI = [0.17, 0.84]). In line with the first two rounds, the likelihood of receiving power was significantly increased by being a cooperator or spending MUs on punishing free riders in the previous round (mixed effect logistic regression, t-1 cooperator coefficient = 0.52, 95% CI = [0.18, 0.86]; t-1 punishing free rider coefficient = 0.78, 95% CI = [0.39, 1.19]). In turn, gaining power further increased the odds of punishing free riders (mixed effect logistic regression, power coefficient = 1.55, 95% CI = [0.85, 2.23]). Since those willing to engage in costly punishment and cooperating above group average were more likely to gain power, and, in turn, gaining power further increased the likelihood of spending own MUs on punishment, powerful group members earned less than the group average (correlation of power and earnings, Spearman's rank correlation $r = -0.24$, $p < .01$, Figure 5.S15). This indicates that the behaviour of powerful group members was not driven by selfish payoff-maximisation.

Group members increased their contributions in response to punishment and power changes. The more MUs someone lost due to receiving punishment in the previous round, the more she increased her contribution to the group project (mixed effect regression, earning reduction coefficient = 0.31, 95% CI = [0.25, 0.37]). Furthermore, the higher the increase in power centralization from the previous round, the more group members increased their contributions compared to the previous round (mixed effect regression, power change coefficient = 4.76, 95% CI = [3.06, 6.48]). Thus, group members reacted not merely to actual punishment, but also to the threat of powerful punishment.

However, not all groups in the endogenous condition were able to solve the social dilemma. Out of 27 groups, cooperation increased steadily over time in 17 (cooperative groups), while it decreased in 10 (non-cooperative groups). This increase or decrease in cooperation is not explained by initial propensities to cooperate: Cooperation in the first round was not significantly different between cooperative and non-cooperative groups (t-test, $t(24) = 0.8$, $p = 0.43$, two-sided). To reach a solution to the free rider problem from a state of weak and decentralized punishment, a punisher needs to be sufficiently empowered by her peers. Therefore, some group members have to be willing to give up power and confide in that their power will be used in their interest. We measured this willingness by the total amount of power transferred within the group. Further, they have to select the right person to shift power to, somebody who will use this power for the benefit of the group. To evaluate whether this selection was successful, we calculated the share of rounds in which the group member most willing to punish free riders in the past became the most powerful.

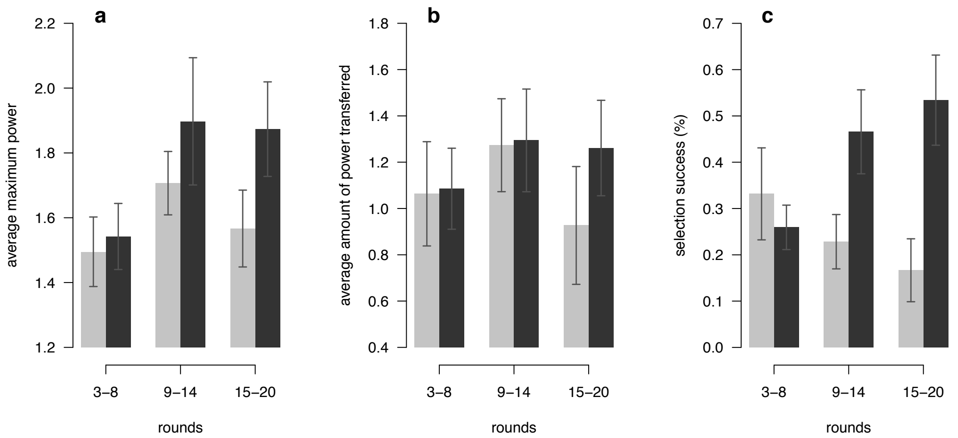


Figure 5.4: Characteristics of cooperative and non-cooperative groups across time intervals. Bars depict groups in which cooperation declined (light grey), or increased (dark grey). (a) Power centralization, measured by the power of the most powerful group member; (b) average amount of power transferred; (c) selection success, measured by the share of rounds in which the most active punisher of non-cooperators of past rounds was the most powerful.

Power centralization, the willingness to give up power as well as selection success were not different in the first third of the experiment across cooperative and non-cooperative groups (Figure 5.4). However, power centralization increased more sharply in cooperative groups and remained stable towards the end of the experiment (Figure 5.4a). This result supports the findings reported above (see Figure 5.2b). In non-cooperative groups, power centralization decreased towards the end of the experiment.

This observed difference was not driven by willingness to give up power. The average amount of power transferred was similar in the first two thirds of the experiment (Figure 5.4b). Instead, cooperative and non-cooperative groups diverged in their success to centralize power in the hands of a group member who reliably punished free riders over past rounds (Figure 5.4c). Thus, groups that transferred sufficient power to the right group member could maintain cooperation. Figure 5.5 shows the power transfer networks. Although the initial networks had a similar structure, non-cooperative groups diverted more power away from the centre, and also transferred it along circles, leading to less power centralization. On the other hand, cooperative groups directed more and more power to one group member over time.

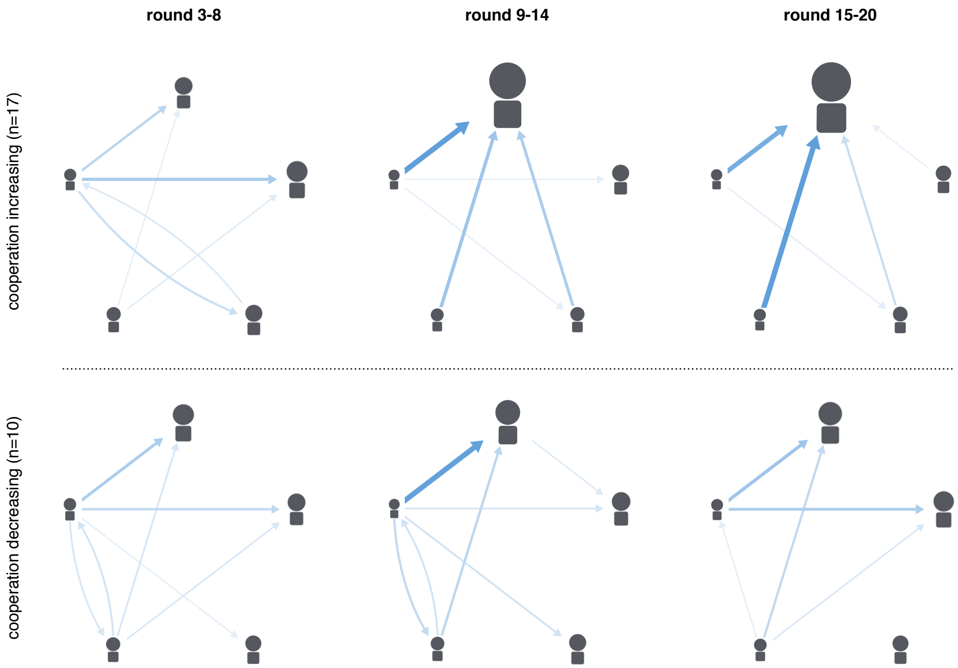


Figure 5.5: Power networks, by time interval and cooperation success. Each network shows the average power transfers (blue arrows) of groups in which either cooperation increased (top) or declined (bottom) in a given third of the experiment. The thickness of the line is proportional to the amount transferred. The size of the group members (nodes) is proportional to the amount of accumulated power.

5.4 Discussion

Power hierarchies help foster cooperation and lead to a better social outcome in environments where decentralized peer punishment is unable to sustain cooperation. The transfer of power solves the social dilemma by enabling group members who do not punish (second-order free riders) to empower cooperators who are willing to sacrifice private resources to bring free riders in line. Free riders anticipate this behaviour and raise their cooperation when they observe that a powerful group member is emerging.

Our work demonstrates the emergence of centralized punishment out of a ‘state of nature’, characterized by weak and decentralized punishment. The resulting power hierarchy overcomes the problems of peer punishment. As pro-social punishers gain power, anti-social punishment becomes more risky. Interestingly, the most powerful group members earned the least, indicating that their behaviour was not driven by financial incentives. They were instead willing to use their power for the sake of the group. Conversely, it is essential that power is concentrated in the right hands. When groups did not have the freedom to decide whom to direct power to, or failed to select the right group member, cooperation could not be sustained.

Social structures that are characterized by an unequal distribution of power are not only prevalent in human societies but also in other social animals (Ellis, 1995). For example, many nonhuman primates live in complex social groups organized in dominance hierarchies (de Waal, 1987, Cheney and Seyfarth, 1990). The emergence of social structures in which some group members have more power than others to enforce shared goals may thus be a crucial step in the evolution of cooperation. In human societies, institutions such as elected representative bodies, legal courts and law enforcement agencies govern much of social life. These institutions are built upon, and embody the centralization of power. The willingness to give up, transfer and centralize power, demonstrated here, can be seen as an important intermediary step and prerequisite to the constitution of such complex institutions.

References

Alexander, R. D. (1987). *The biology of moral systems*. Transaction Publishers, New Jersey.

Boyd, R., Gintis, H., and Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328(5978):617–620.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*, 100(6):3531–3535.

Boyd, R. and Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13(3):171–195.

Casari, M. (2005). On the design of peer punishment experiments. *Experimental Economics*, 8(2):107–115.

Casari, M. and Plott, C. R. (2003). Decentralized management of common property resources: Experiments with a centuries-old institution. *Journal of Economic Behavior and Organization*, 51(2):217–247.

Cheney, D. L. and Seyfarth, R. M. (1990). The representation of social relations by monkeys. *Cognition*, 37(1-2):167–196.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305:1254–1258.

de Waal, F. (1987). Dynamics of social relationships. In Smuts, R., Cheney, D. L., Seyfarth, R., Wrangham, R. W., and Struhsaker, T. T., editors, *Dynamics of social relationships*, pages 421–429. University of Chicago Press, Chicago.

Dreber, A., Rand, D., Fudenberg, D., and Nowak, M. A. (2008). Winners don't punish. *Nature*, 452:348–351.

Dreber, A. and Rand, D. G. (2012). Retaliation and antisocial punishment are overlooked in many theoretical models as well as behavioral experiments. *Behavioral and Brain Sciences*, 35(1):24.

Egas, M. and Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 275:871–878.

Ellis, L. (1995). Dominance and reproductive success among nonhuman animals: A cross-species comparison. *Ethology and Sociobiology*, 16(4):257–333.

Engelmann, D. and Nikiforakis, N. (in press). In the long-run we are all dead: On the benefits of peer punishment in rich environments. *Social Choice and Welfare*.

Falk, A., Fehr, E., and Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, 73(6):2017–2030.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425:785–791.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415:137–140.

Flew, A. (1954). The justification of punishment. *Philosophy*, 29(111):291–307.

Fowler, J. H. (2005). Second-order free-riding problem solved? *Nature*, 437(7058):E8.

Gächter, S. and Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1518):791–806.

Gächter, S., Renner, E., and Sefton, M. (2008). The long-run benefits of punishment. *Science*, 322:1510.

Guala, F. F. (2012). Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, 35(1):1–15.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3:367–388.

Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319:1362–1367.

Janssen, M. A. and Bushman, C. (2008). Evolution of cooperation and altruistic punishment when retaliation is possible. *Journal of Theoretical Biology*, 254(3):541–545.

Janssen, M. A., Holahan, R., Lee, A., and Ostrom, E. (2010). Lab experiments for the study of social-ecological systems. *Science*, 328(5978):613–617.

Kiyonari, T. and Barclay, P. (2008). Cooperation in social dilemmas: Free riding may be thwarted by second-order reward rather than by punishment. *Journal of Personality and Social Psychology*, 95(4):826–842.

Kosfeld, M., Okada, A., and Riedl, A. (2009). Institution formation in public goods games. *American Economic Review*, 99(4):1335–1355.

Masclet, D., Noussair, C., Tucker, S., and Villeval, M. C. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review*, 93(1):366–380.

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics*, 92(1-2):91–112.

Nikiforakis, N. and Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, 11(4):358–369.

Nowak, M. A. and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063):1291–1298.

Ohtsuki, H., Iwasa, Y., and Nowak, M. A. (2009). Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature*, 457(7225):79–82.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press, Cambridge.

Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, 86(2):404–417.

Panchanathan, K. and Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, 432(7016):499–502.

Perc, M. (2012). Sustainable institutionalized punishment requires elimination of second-order free-riders. *Scientific reports*, 2:344–344.

Rand, D. G., Armao IV, J. J., Nakamaru, M., and Ohtsuki, H. (2010). Anti-social punishment can prevent the co-evolution of punishment and cooperation. *Journal of Theoretical Biology*, 265(4):624–632.

Rand, D. G. and Nowak, M. A. (2011). The evolution of antisocial punishment in optional public goods games. *Nature Communications*, 2:1–7.

Rockenbach, B., Guererk, O., and Irlenbusch, B. (2006). The competitive advantage of sanctioning institutions. *Science*, 312(5770):108–111.

Sefton, M., Shupp, R., and Walker, J. M. (2007). The effect of rewards and sanctions in provision of public goods. *Economic Inquiry*, 45(4):671–690.

Sigmund, K., De Silva, H., Traulsen, A., and Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature*, 466(7308):861–863.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1):35–57.

Walker, J. M. and Halloran, M. A. (2004). Rewards and sanctions and the provision of public goods in one-shot settings. *Experimental Economics*, 7(3):235–247.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51(1):110–116.

Yamagishi, T. (1988). Seriousness of social dilemmas and the provision of a sanctioning system. *Social Psychology Quarterly*, 51(1):32–42.

Supporting Information

Experiment description

Game description

We introduce a new game, the ‘power transfer game’. In the power transfer game, a punishment stage and a power transfer stage are added to the standard public goods game. The game-theoretic description given below first presents the standard public goods game, and then adds the punishment and the power transfer mechanisms sequentially. For equilibrium analysis, we focus on subgame perfect and trembling hand perfect equilibria.

Public goods game

The public goods game is played by $n > 2$ players. Each player has an initial endowment of E monetary units (MUs). Players decide on how much of the endowment to contribute to a public good, keeping the remaining amount. Contributions to the public good are multiplied by m, and then distributed equally among all players. We assume $m > 1$ and $\frac{m}{n} < 1$; $\frac{m}{n}$ is called the ‘marginal per capita return’. Decisions on contribution are made simultaneously. The contribution of player i is denoted by c_i, with $0 \leq c_i \leq E$. After the contribution decisions are made, each player is informed about the contributions of all players, and the level of the public good.

In the public goods game, player i’s payoff is given by:

$$\pi_i = E - c_i + \frac{1}{n} m \sum_{j=1}^{n} c_j.$$

Selfish payoff-maximization prescribes a contribution of $c_i = 0$ for each individual i, as $\frac{m}{n} < 1$, and thus the private return on each invested MU is negative: $\frac{m}{n} - 1 < 0$. However, since $m > 1$, the social return of contributions is positive: $m - 1 > 0$. Therefore, social optimum, defined as the sum of all monetary payoffs, is reached when everybody contributes their full endowment, i.e. at $c_i = E$. With these contributions, the earnings of everyone will be mE. However, at individually rational contributions of zero, each player earns just their endowment E. Thus, everyone contributing their full endowment constitutes a Pareto-improvement over zero contributions. As selfish motives clash with group interests, the public goods game is a classical example of a social dilemma.

Because contributing zero is a strictly dominant strategy for every player, $c_i = 0$ for all players constitutes the only Nash-equilibrium of the game. When the public goods game is played for a finite number of periods, through a backward induction argument we get that contributions are zero for every period in the only subgame perfect equilibrium of the game.

In our experiment, the game was played with $n = 5$ players, and the contribution multiplier was set to $m = 1.5$. The initial endowment was $E = 20$, and only integer contributions were allowed.

Public goods game with punishment

A punishment option is added to the simple public goods game. Following contributions, as described in the previous section, and after being informed about their peers' contributions, players simultaneously make a punishment decision. Player i can assign $0 \leq d_{ij} \leq d_{\text{max}}$ deduction points (DP) to each other player $j \neq i$. For each DP that player i assigns to player j, player i's own payoff is reduced by the cost of punishment, pc, while the punished player j's payoff is reduced by the effectiveness of punishment pe. After punishment decisions are made, players are informed about who punished whom, and by how much (i.e. all DP assignments d_{ij}).

In the public goods game with punishment, player i's payoff is given by:

$$\pi_i = E - c_i + \frac{1}{n} m \sum_{j=1}^{n} c_j - pc \sum_{\substack{j=1 \\ j \neq i}}^{n} d_{ij} - pe \sum_{\substack{j=1 \\ j \neq i}}^{n} d_{ji}.$$

In the only subgame perfect equilibrium of the one-shot public goods game with punishment, punishments and contributions are zero. The reasoning is as follows: Once the contribution decisions are made, no selfish payoff-maximiser has any incentive to punish, because doing so would only reduce her payoffs. Therefore, it can be known from the start that $d_{ij} = 0$. Therefore, later punishment is not credible, and cannot raise contributions above $c_i = 0$. In the finitely repeated public goods game with punishment, the same backwards induction argument can be used, starting with the last period, to show that in a subgame perfect equilibrium, punishment and contribution is zero in all periods.

Almost ubiquitously in the literature, the cost of punishment is set at $pc = 1$. In the majority of experiments, the effectiveness of punishment is higher than its cost, with the most commonly used value of $pe = 3$. In our experiment, we set the *initial* effectiveness to the same level as its cost at $pe = pc = 1$ (see the next subsection). We limited the number of deduction points that could be assigned to $d_{\text{max}} = 10$. We imposed no lower constraint on the earnings of players. Thus, players could punish and be punished below zero income.

Power transfer game

We modify the public goods game with punishment by adding an additional decision *before* the contribution decision where players may transfer their power. Power determines the effectiveness of punishment. Players control an equal amount of power $\widehat{pe}$, and can decide to transfer it to other players. Thus, player i decides on how much power pt_{ij} to transfer to each player j (allowing for $j = i$), with $pt_{ij} \geq 0$ and $\sum_{j=1}^{n} pt_{ij} = \widehat{pe}$. Power transfer is free. Moreover, transferring power has no impact on the cost of punishment pc. The power pe_i of player i will be the total power she keeps for herself and receives from others: $pe_i = \sum_{j=1}^{5} pt_{ji}$. Players are informed about every player's power before they decide on their subsequent contributions.

In the power transfer game, player i's payoff is given by:

$$\pi_i = E - c_i + \frac{1}{n} m \sum_{j=1}^{n} c_j - pc \sum_{\substack{j=1 \\ j \neq i}}^{n} d_{ij} - pe_j \sum_{\substack{j=1 \\ j \neq i}}^{n} d_{ji}$$

To recapitulate, the items of the sum are, in turn, the initial endowment, the contribution to the public good, the individual return from the public good, the cost of punishment dealt to others, and the punishment received from others.

The cost of punishment pc is not affected by power transfers. Thus, the same argument showing that both punishments and contributions are zero in the only equilibrium of the public goods game with punishment can be extended to the power transfer game straightforwardly. Therefore, $c_i = 0$ and $d_{ij} = 0$ in the subgame perfect equilibria of the finitely repeated power transfer game. However, this equilibrium is not unique any more, because any level of power transfers is compatible with these choices. We can get a unique equilibrium by focusing instead on trembling hand perfection. In this case, 'slips of hand' – strategies that assign a positive probability to every pure strategy – should be taken into account. Thus, a player needs to consider the possibility that she might get punished by another by mistake. If the player has transferred any power to whomever punishes her, her payoff will be lower than if she had chosen not to transfer power. Because of this potentially harmful effect of power transfers, players should not transfer power to any other player. Therefore, in the only trembling hand perfect equilibrium, we get $pe_{ij} = 0$ for each player i, and other player $j \neq i$.

In our experiment, each player controlled a total power of $\widehat{pe} = 1$, so the constraint on power transfers was $\sum_{j=1}^{5} pt_{ij} = 1$.

Timeline and parameter setup

We assign each simultaneously made decision, together with the feedback/information provided after the decision to a *stage*. The timeline of the power transfer game is thus the following.

1. Power transfer stage: power transfer decision;
   players are informed about every player's effectiveness.

2. Contribution stage: contribution decision;
   players are informed about every player's contribution.

3. Punishment stage: DP assignment;
   players are informed about all DP assignments.

To summarize, our choice of parameters for the experiment was as follows:

- $n = 5$ players;
- contribution multiplier $m = 1.5$;
- initial endowment $E = 20$ MUs;
- maximum number of deduction points $d_{\text{max}} = 10$;
- punishment cost $pc = 1$;
- own power $\widehat{pe} = 1$.

Participants

Participants were recruited from the subject pool of the behavioural and experimental economics lab (BEElab) at Maastricht University and were invited via e-mail. Participants were randomly assigned to groups of five. Each experimental session comprised at least 3 and at most 5 groups.

Participants were paid the sum of monetary units (MUs) earned over the 20 rounds (10 MUs = 0.25 euro) plus a show-up fee of 3 euro and a small sum based on an incentivised social value orientation questionnaire. In three groups we had to abort the experiment because of computer malfunction. In these groups data is available for at least 14 rounds and used as such in the analysis.

Experimental setup

Participants were seated in separate cubicles and stayed there for the whole experiment. Each participant had a notepad and a pen to make notes.

The experiment started with one round of a PG. Participants read instructions explaining the rules of the PG on the computer screen. Instructions used neutral labels for describing the social dilemma: Participants were told that they would receive 20 MUs each round and that they have to decide how many MUs to contribute to a 'project'. The output of the project would then be distributed equally among all group members, irrespective of how much each member contributed. After reading the instructions, participants had to answer a set of comprehension questions about the rules of the PG.

First round: contribution

At the beginning of the first round each participant was asked to indicate how much of the 20 MUs to contribute to the project. When entering a number, the participant saw a graphical representation of the share of MUs she would contribute. After each participant in the group made a contribution decision, they saw how much each group member contributed. Group members were associated with a unique symbol by which they could be identified throughout the whole experiment. Participants saw a summary of the earnings of each group member and the outcome of the group project for this round. This summary was provided at the end of every round.

Second round: punishment

In the second round, punishment was introduced. Participants received instructions on the computer screen and answered a set of comprehension questions about the punishment rules. Instead of 'punishment' we used the neutral label 'deduction' and 'deduction points (DP)' in the instructions and throughout the experiment. Upon answering all questions correctly, participants entered the contribution stage of round 2. After getting informed about the contribution of each group member, participants simultaneously assigned between 0 and 10 deduction points to each other group member. When entering a number, participants saw the DP costs as well as the effect the punishment would have on the punished.

The deduction stage outcome summary showed which participants were punished and by whom using a graphical matrix representation. By going through the matrix by columns, participants could see how much a group member contributed in the contribution stage, how many DPs in total this group member spent on punishing others, and how many DPs were assigned to her by the others as well as the effect it had on her earnings. By going through the matrix row-wise, participants could see how much the corresponding group member punished the other group members. On this screen, participants also had the possibility to look at previous round behaviour. Thus, it was possible to review past contribution and punishment decisions of each group member. The round ended with the outcome screen.

In the second round, for each assigned DP, the punisher had to pay 1 MU and the punished would lose 1 MU. With the power transfer mechanism introduced in the third round and explained below, this effectiveness-to-cost ratio of punishment could change from round to round.

Third round: power transfer

Before entering the next round, participants received the third and final set of instructions. In the fixed condition, participants were told that contribution stage and punishment stage would now be repeated for another 18 rounds. In the endogenous condition, the power transfer stage was explained to the participants. Neutral labels for power was used. Instead of 'power', we used 'deduction effectiveness' and the stage was called 'shifting stage'. After answering a set of comprehension questions about the power transfer mechanism, participants started the third round with the power transfer stage. In this stage, participants had the possibility to transfer power to other group members. Each participant had a power of 1 and could transfer power in steps of 0.1 to other group members. Power could also be distributed among multiple group members (e.g. it was possible to transfer 0.5 to one participant, 0.2 to another participant and keep 0.3 to oneself).

After every participant made their power transfer decisions, they were shown the power each group member had for this round, based on the five transfer decisions. Note that participants could not see the individual transfer decisions of the other group member but only the outcome of these decisions, i.e. the total power of each group member for this round.

In the exogenous condition, instead of transferring power by themselves, power changed exogenously based on the power transfer decisions that participants made in one of the endogenous groups. In the instructions set, participants were told that "deduction effectiveness of you and the other group members can change" from round 3 to 20. Therefore, they could not transfer power voluntarily, but only saw the change in power at the beginning of each round, starting from round three.

A change in power changed the effectiveness of punishment. In the fixed condition, power was fixed to 1 and transferring power was not possible, just like in round two for all conditions. In the other two conditions power could change as explained above.

Rounds 4–20

In the exogenous and endogenous condition, power transfer, contribution and punishment stage were repeated for the consecutive 17 rounds, with the difference that in the exogenous condition, participants saw the change in power from round to round, without being able to influence it themselves.

Each round began with the power transfer stage. The transfer decisions made in the previous round served as the status quo for the current round. When entering round 4, participants would see the power status each group member had in the previous round together with the transfer decisions made by the participant in the previous power transfer stage. Thus, by default, the participant would make the same power allocation she did previous round. However, the participant could also decide to reverse or change the previous decision by changing the numbers below each bar accordingly (with the only constraint that the total amount of power transferred could never exceed 1).

By introducing punishment and power transfer stage round by round to the participants, we were able to measure baseline contribution and punishment rates across conditions. There should be no significant differences in average contributions in the first and average contributions and punishment in the second round between conditions, since the actual experimental manipulation started in the third round.

Computer interface

Figures 5.S1 to 5.S11 show what a hypothetical group member would see on the computer screen in the different stages for the hypothetical round 3 to the beginning of round 4 in the endogenous treatment. What is shown is meant as an example to explain the computer interface and is not real data.

At the beginning of round 3, our group member sees the first power transfer stage screen (Figure 5.S1). In this example, our group member decides to transfer 0.5 of her power to group member 3 and 0.2 of her power to group member 4 (Figure 5.S2).

After all group members made their power transfer decision, the power transfer outcome screen is shown. Due to power transfers of our and the other group members, group member 3 now holds the most power, while our group member is the least powerful in the group (Figure 5.S3).

The next screens show the contribution stage (Figure 5.S4). Our hypothetical group member decides to contribute 15 MUs to the group project (Figure 5.S5). After all group members made their contribution decision, the contribution outcome screen is shown (Figure 5.S6).

After the contribution stage, the group enters the deduction stage (punishment). Figure 5.S7 shows the input screen of the deduction stage. In the first row, the contributions of each group member of the previous contribution stage is shown. By entering numbers between 1 and 10 below each column, our group member is able to assign deduction points. This is shown in Figure 5.S8. Our group member decides to assign 2 DPs to group member 2 and 5 DPs to group member 5.

After all group members made their punishment decisions, the deduction outcome screen is shown. In this example, both our group member and group member 3 decides to punish group member 5 by 5 DPs for free riding in this round's contribution stage (Figure 5.S9). Since our group member only has a power of 0.3, her punishment is rather ineffective. The 5 DPs costs her 5 MUs but will only reduce the earnings of group member 5 by 1.5 MUs (0.3 effectiveness times 5 DPs). In contrast, group member 3 received power from other group members and has a total power of 2. The 5 DPs assigned to group member 5 also reduces her earnings by 5 (costs of punishment), but will reduce the earnings of group member 5 more substantially. Group member 5 would lose 10 MUs due to this punishment (2.0 effectiveness times 5 DPs). The same logic applies to the punishment of group member 2 by our group member and group member 3 (Figure 5.S9).

On the next screen the outcome of this round is shown (Figure 5.S10). Group members see the earnings of each group member as well as the outcome of the group project.

By pressing the button at the bottom of the screen, our group member can enter the next round, starting with the power transfer stage. Previously, she decided to transfer some of her power to group member 3 and 4. This decision serves as the default option in this round's power transfer stage (Figure 5.S11). By pressing 'accept & proceed', she would again transfer 0.5 and 0.2 of her power to these group members, respectively.

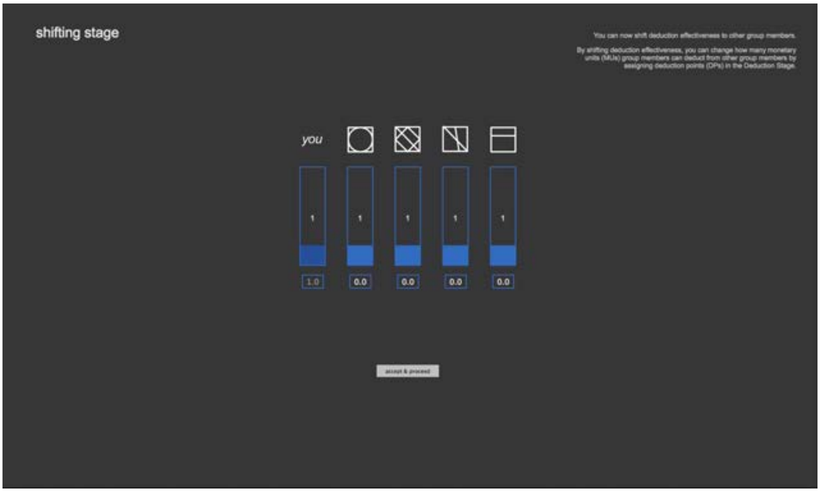


Figure 5.S1: Power transfer stage before input, round 3. Since this is the first round with power transfer, the status quo option, visualised here, is not to transfer any power to others. Note that this screen did not appear in the fixed and exogenous condition.

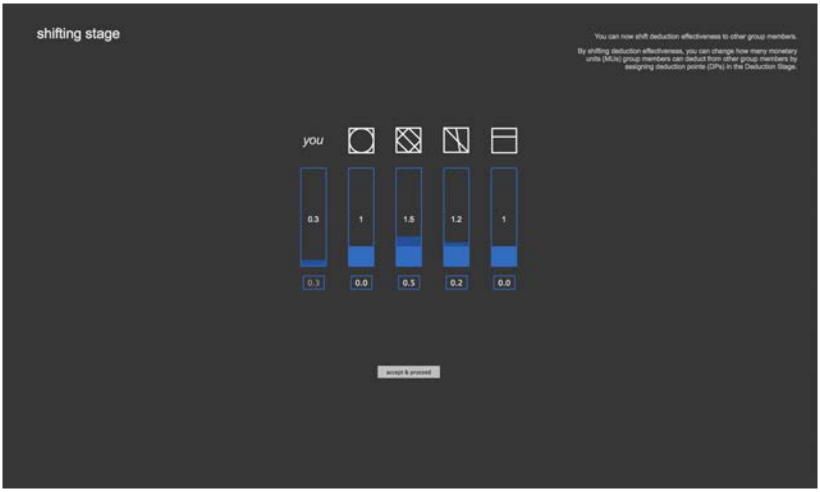


Figure 5.S2: Power transfer stage screen, round 3. Input to each field is followed by a graphical representation of how the power status of the respective group member will change (if nobody else transfers any power). Own power is shown in dark blue. Note that this screen did not appear in the fixed and exogenous condition.

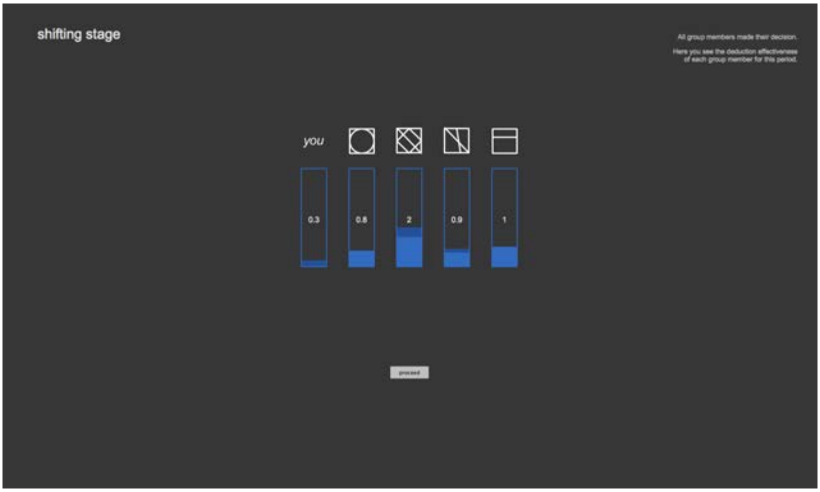


Figure 5.S3: Power transfer outcome screen, round 3. After every group member made her power transfer decision, the power transfer outcome screen is shown. Bars represent the power of each group member for this round. Power changes due to transferring own power are shown in dark blue. Note that this screen did not appear in the fixed condition.

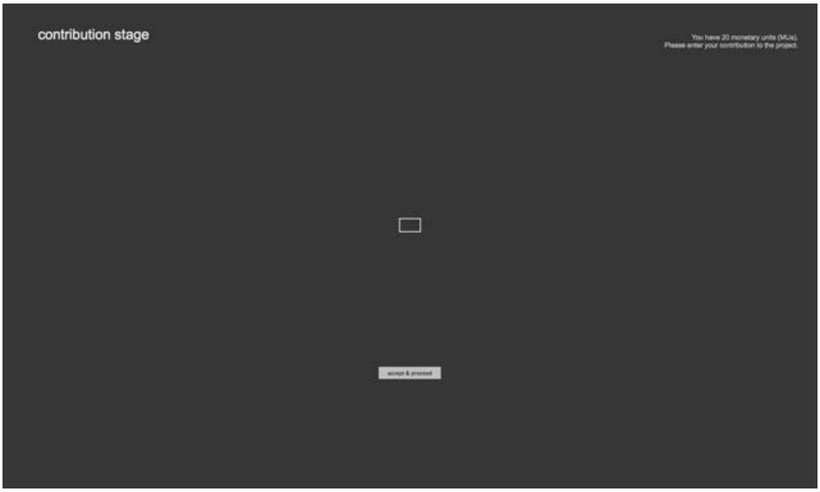


Figure 5.S4: Contribution stage before input, round 3. Participants decide simultaneously how much of their endowment of 20 monetary units (MUs) to contribute to the group project.

Figure 5.S5: Contribution stage input screen, round 3. After inputting a number, the participant sees a graphical representation of the fraction of her endowment she would contribute.

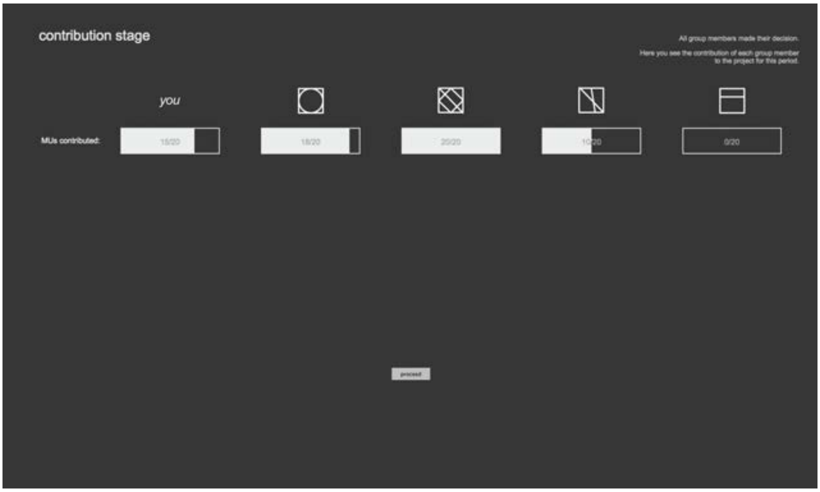


Figure 5.S6: Contribution stage outcome screen, round 3. After every participant makes her contribution decision, the outcome screen shows how much each group member contributed.

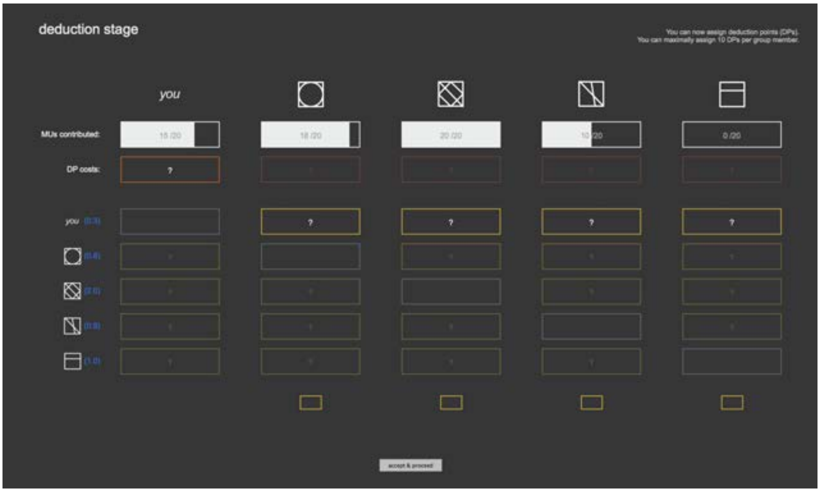


Figure 5.S7: Punishment stage screen before input, round 3. Each participant sees a 5×5 punishment matrix (yellow and grey rectangles) and assigns deduction points (DPs). The power of each group member is shown by the blue number. In the first row, the contribution of each group member of this round's contribution stage is shown.

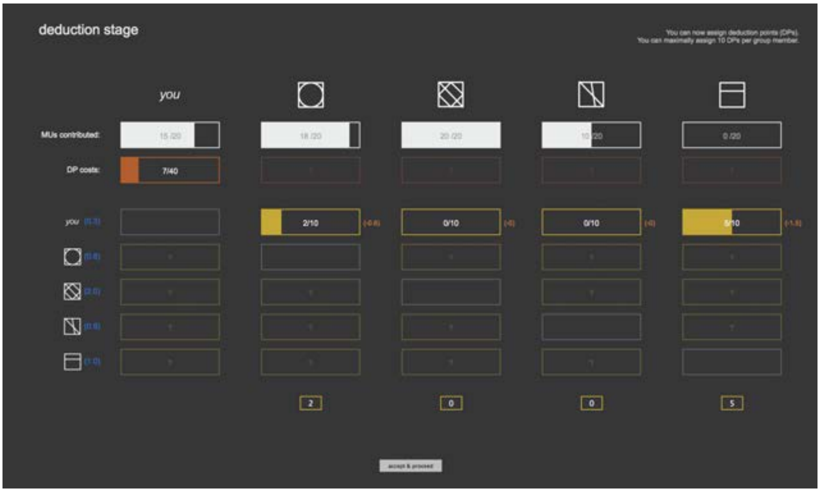


Figure 5.S8: Punishment stage screen after input, round 3. The participant sees the total cost (orange bar on the top left), the current DP assignment (yellow), as well as the effect on the group member (orange numbers) of her DP assignment.

Figure 5.S9: Punishment stage outcome screen. Participants see who was punished by whom. Each column indicates by how much a group member was punished, while the rows indicate by whom these deduction points were assigned by.

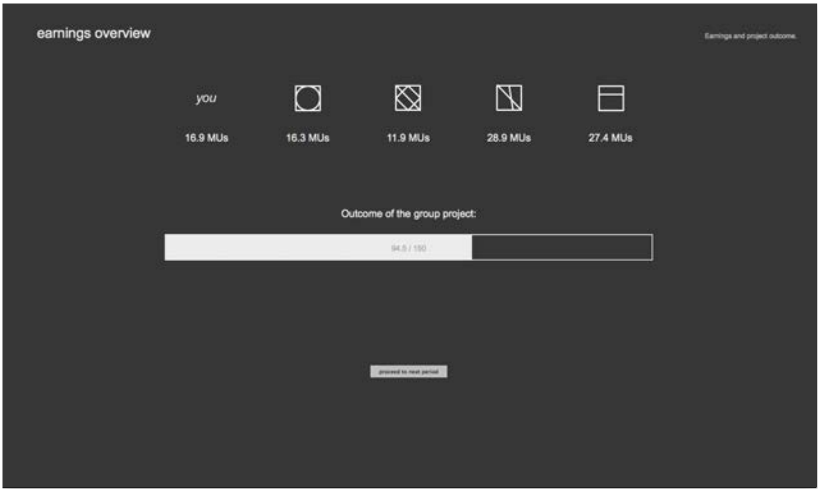


Figure 5.S10: Earnings screen, round 3. This screen is shown at the end of each round summarising earnings and the outcome of the group project (sum of contributions $\times 1.5$). The sum of MUs kept, MUs received from the group project, losses due to assigning DPs, and losses due to receiving punishment resulted in the payoffs seen on the screen.

Figure 5.S11: Power transfer stage screen before input, round 4. Each bar represents the power of each group member in the previous round. The power transfers made in the previous round serve as the status quo allocation for this round. Dark blue bar segments indicate the power allocated by this group member. Thus by pressing 'accept & proceed', the participant will again allocate 0.5 of her power to group member 3 and 0.2 of her power to group member 4.

Statistical analysis

Measurements

For the statistical analysis, we defined proxies and compound measures based on the behavioural data. Each of these measures are defined below.

Punishment behaviour

For the receiving power and using power models (see below), we were interested in how power was used with regards to punishment. We wanted to differentiate the punishment of free riders from the punishment of cooperators (antisocial punishment). Moreover, we opted for a measure that assigned higher value for punishing non-cooperators who deviated more from average contributions. Punishment behaviour was defined as:

$$
\begin{aligned}
\text{punishment behaviour}_{it} &= \sum_{\substack{j=1\\ i\neq j}}^{5} \left(\frac{\bar{c}_t - c_{jt}}{\sigma_{ct}} \times d_{ijt}\right) \\
\text{where } c_{jt} &= \text{contribution of player } j \text{ in round } t, \\
\bar{c}_t \text{ and } \sigma_{ct} &= \text{mean and standard deviation of contributions,} \\
d_{ijt} &= \text{deduction points assigned by } i \text{ to } j, \\
&\quad \text{in round } t, \text{ with } 2 \leq t \leq 20.
\end{aligned}
\tag{5.1}
$$

Thus, for each group member in each round, assigned deduction points were weighted by the standardised cooperation of the punished group member and summed up. Negative numbers therefore indicated, that punishment was predominantly used to punish cooperators (antisocial punishment), while positive numbers indicated that punishment was predominantly used to punish free riders (e.g. punishment behaviour value of 1 can be interpreted as using one deduction point to punish a group member with a contribution that was 1 standard deviation below group average in this round's contribution stage). Figure 5.S12 shows the percentage of punishment behaviour types across treatments. Antisocial punishment (free riders punishing cooperators) was observed less frequently than no punishment and the punishment of free riders.

Power centralisation

We summarized power centralisation through a single indicator: the power of the most powerful group member. In the fixed condition every group member had a fixed power of 1. For the endogenous and exogenous conditions, power centralization was defined by:

$$
\begin{aligned}
\text{power centralisation}_t &= \max_{i \in 1,\dots,5} pe_{it} \\
\text{where } pe_{it} &= \sum_{j=1}^{5} pt_{jit}, \text{ power (i.e. punishment effectiveness) of } i, \\
&\quad \text{in round } t, \text{ with } 3 \leq t \leq 20.
\end{aligned}
\tag{5.2}
$$

A minimal power centralisation of 1 is achieved by, for instance, no power being transferred in a given round. A full power centralisation value of 5 was achieved whenever all group members would transfer their power to a single individual.

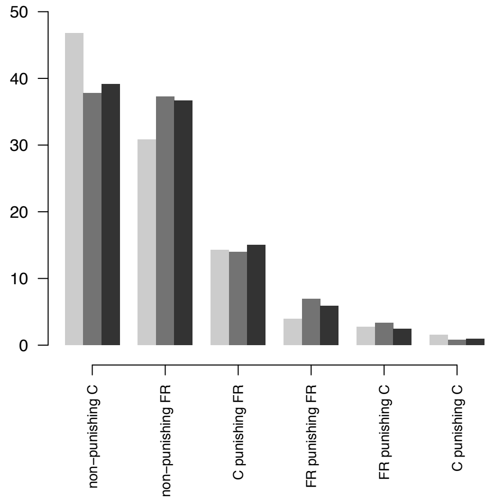


Figure 5.S12: Cooperation and punishment decisions. Percentage of punishment decisions observed across treatments (light grey = endogenous condition, grey = exogenous condition, dark grey = fixed condition). C = cooperator, FR = free rider.

Willingness to transfer power

We defined our measure for the willingness to transfer power as the sum of all power transferred in a certain round.

$$\begin{aligned} \text{willingness to transfer}_t &= \sum_{i=1}^{5} \sum_{\substack{j=1 \\ j \neq i}}^{5} pt_{jit} \\ \text{where } pt_{ijt} &= \text{power transferred from } j \text{ to } i, \\ & \quad \text{in round } t \text{, with } 3 \leq t \leq 20. \end{aligned} \tag{5.3}$$

In rounds when group members did not transfer any power between them, our willingness to transfer measure would take a value of 0. A willingness to transfer of 5 could be achieved by everyone transferring all their power.

Selection success

We introduced a measure to gauge a group's ability to select the most active prosocial punishers as the recipient of power transfers. First, building on Definition 5.1, we defined the aggregate (past) punishment behaviour as:

$$\text{aggregate punishment behaviour}_{it} = \sum_{j=2}^{t-1} \Big(\text{punishment behaviour}_{ij}\Big) \qquad (5.4)$$
$$\text{in round } t, \text{ with } 3 \leq t \leq 20.$$

This variable summarised information about the past behaviour of group members, as potential leaders of the in-group power hierarchy. Within a group, larger values for a certain member i indicated both the amount of resource i sacrificed for punishment, as well as her propensity to pick less cooperative group members, as a target for her punishment. If the past is a reliable predictor of an individual's future behaviour, those who showed a willingness to punish non-cooperators should receive the most power. Thus, for groups in the endogenous condition, we defined our selection success in the following manner:

$$\text{selection success}_t = \begin{cases} \frac{1}{|\operatorname*{arg\,max}_{i=1,\dots,5} pe_{it}|} & \text{if agg. punishment behaviour}_{it} \in \operatorname*{arg\,max}_{i=1,\dots,5} pe_{it}, \\ 0 & \text{otherwise,} \end{cases}$$
$$\text{in round } t, \text{ with } 3 \leq t \leq 20. \qquad (5.5)$$

Specifically, if in a certain round individual i, having the highest aggregate punishment behaviour, was strictly more powerful than all other group members, then for that round the selection success of the group would be 1. If, however, i was less powerful than some other group member, we would assign it a value of 0. In case of a tie, the value would be 1 divided by the number of group members that are equally (most) powerful. In such a case, the group has partially solved the selection problem, and we would assign $\frac{1}{2}$, $\frac{1}{3}$, etc. to the selection success variable. This is a rather rough estimate of selection success, as it only takes into account whether the most reliable punisher of non-cooperators gains the most power.

Statistical models

Because of the hierarchical structure of the data (participants clustered in groups and repeated measures over rounds), we fitted Bayesian mixed effects models to the data using R and JAGS.

Non-informative Gaussian priors (m=0, sd=100) were used for each predictor and non-informative uniform priors (range 0 to 100) for the error terms. In every model, random intercepts and slopes were allowed to covary. Therefore, the variance-covariance matrix was estimated alongside the fixed and random coefficients. For the correlations between random effects, non-informative uniform priors (range -1 to 1) were used. We used three parallel chains. For every estimated coefficient, the potential scale reduction factor (Gelman and Rubin Diagnostic) was below 1.05, indicating good mixing of the three chains and thus high convergence. Regression tables reported below show estimated coefficients together with the 95% confidence interval (CI, also called highest density interval in the Bayesian framework). Note that, since non-informative priors were used, a 95% CI that only contains negative or positive values can be interpreted as significant at a $p = .05$ two-sided threshold from frequentist perspective. Fitting the models using restricted maximum likelihood (REML) as implemented in the lme4 package in R revealed similar estimates and the same statistical inferences. However, models on the individual subject level failed to converge and also the censoring in the data could not be accounted for in these models.

Group level analysis

The aim of the group level analysis was, to compare cooperation (i.e. contribution to the group project), punishment and earnings across the three different conditions, as well as to analyse the increase of maximum power over rounds in the endogenous condition (and thus also the exogenous condition). For this, we aggregated the data by group members, such that for each group we had one data point for each round (e.g. average contribution).

Contribution. The fixed part of the contribution model contained two dummy variables coding the three experimental conditions (with the fixed condition as baseline), a continuous round predictor and the round $\times$ condition interactions. The random part contained a random intercept as well as a random slope for the round predictor for each group. Thus, for each group a separate baseline cooperation rate in round 1, and a separate slope of cooperation over rounds was estimated (see Equations 5.6). Since average group contribution could not exceed 20 and fall below 0, the data was treated as left and right censored.

$$
\begin{aligned}
y_i &\sim N(\mu_y, \sigma_y^2), \text{for } i = 1, \dots, n \\
\mu_y &= \alpha_{1j} + \beta_{1j}\text{round} + \alpha_2 + \beta_2\text{round} \\
&\quad + \beta_3\text{exogenous} + \beta_4\text{endogenous} \\
&\quad + \beta_5\text{exogenous} \times \text{round} + \beta_6\text{endogenous} \times \text{round} \\
\begin{pmatrix} \alpha_{1j} \\ \beta_{1j} \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha_1}^2 & \rho\sigma_{\alpha_1}^2\sigma_{\beta_1}^2 \\ \rho\sigma_{\alpha_1}^2\sigma_{\beta_1}^2 & \sigma_{\beta_1}^2 \end{pmatrix} \right), \text{for } j = 1, \dots, J
\end{aligned}
\tag{5.6}
$$

$$
\begin{aligned}
\text{where } n &= \text{ number of observations} \\
J &= \text{ number of groups}
\end{aligned}
$$

Table 5.S1 shows the estimated coefficients together with the 95% CI. First round cooperation was predicted to be 12.6 for the fixed condition. Groups in the exogenous and endogenous condition did not significantly differ from this initial level of cooperation (β_3 and β_4). In the fixed condition, there was a significant drop in cooperation over rounds (β_2). The drop in cooperation was not significantly different in the exogenous condition (β_5). In contrast, in the endogenous condition, cooperation over time was significantly higher (β_6).

Examining the posterior distributions of the exogenous $\times$ round and endogenous $\times$ round parameter revealed an estimated difference of 0.24 with a 95% CI ranging from 0.08 to 0.39. Thus, also compared to the exogenous condition, cooperation over time was significantly higher in the endogenous condition.

Punishment. The punishment model followed the same structure as the contribution model (see Equations 5.6), except for average deduction points spent as the dependent variable. Since average group punishment could not fall below 0, the data was treated as left censored.

Table 5.S2 shows the estimated coefficients together with the 95% CI. Punishment expenses in the second round (first round with punishment) did not significantly differ across conditions (α_2, β_2 and β_3). In the fixed condition, the use of punishment dropped over rounds (β_2). Groups in the exogenous condition did not deviate significantly from this trend (β_5). In the endogenous condition, the drop in punishment expenses over rounds was significantly higher compared to the fixed condition (β_6).

Table 5.S1

Contribution regression model.

		Coef.	95% CI
α_2	intercept (fixed condition, round 1)	12.60	[10.19, 14.95]
β_2	round	-0.28	[-0.51, -0.05]
β_3	exogenous condition dummy	-1.84	[-4.86, 1.01]
β_4	endogenous condition dummy	-1.47	[-4.56, 1.46]
β_5	exogenous condition $\times$ round	0.22	[-0.06, 0.51]
β_6	endogenous condition $\times$ round	0.46	[0.16, 0.74]
$\sigma^2_{\alpha_1}$	error-term random intercepts	0.47	[0.38, 0.56]
$\sigma^2_{\beta_1}$	error-term round slopes	2.41	[2.31, 2.51]
σ^2_y	error-term y	4.55	[3.76, 5.42]
ρ	correlation between random effects	-0.15	[-0.40, 0.09]

Note. Dependent variable: Contribution group average clustered by group.

Table 5.S2

Punishment regression model.

		Coef.	95% CI
α_2	intercept (fixed condition, round 2)	1.32	[0.70, 1.94]
β_2	round	-0.08	[-0.13, -0.03]
β_3	exogenous condition dummy	-0.19	[-0.97, 0.58]
β_4	endogenous condition dummy	0.10	[-0.66, 0.89]
β_5	exogenous condition $\times$ round	-0.01	[-0.08, 0.05]
β_6	endogenous condition $\times$ round	-0.07	[-0.13, -0.002]
$\sigma^2_{\alpha_1}$	error-term random intercepts	0.08	[0.05, 0.10]
$\sigma^2_{\beta_1}$	error-term round slopes	1.47	[1.39, 1.55]
σ^2_y	error-term y	1.04	[0.78, 1.32]
ρ	correlation between random effects	-0.60	[-0.82, -0.35]

Note. Dependent variable: Average punishment expense clustered by group.

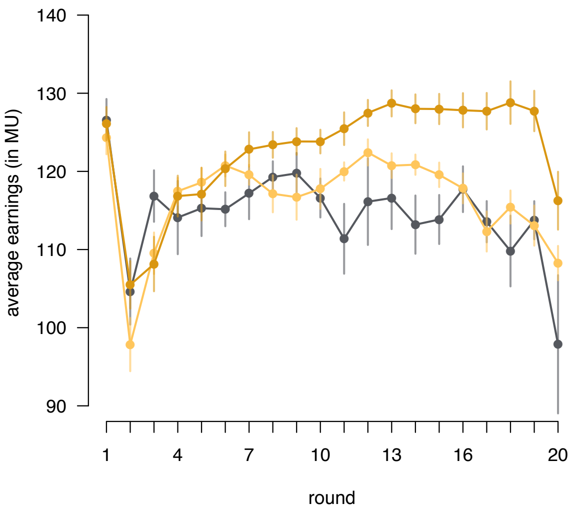


Figure 5.S13: Average earnings over rounds. Average earnings in the fixed condition (grey), exogenous condition (light yellow) and endogenous condition (dark yellow). Error bars show the within-subject standard errors of the mean.

Group earnings. Figure 5.S13 shows the average earnings over rounds by condition. There was a substantial drop in earnings in all three conditions after the introduction punishment to the experiment in round 2, as it is often observed in PG experiments with punishment.

The group earnings regression model followed the same structure as the contribution model (see Equations 5.6), except for aggregated earnings (in MUs) as the dependent variable. Since earnings for one round could not exceed 150 MUs (20 MUs $\times$ 5 participants $\times$ 1.5), the data was treated as right censored.

Table 5.S3 shows the estimated coefficients together with the 95% CI. Earnings in the first round did not significantly differ between fixed and exogenous condition, as well as fixed and endogenous condition (β_3 and β_4). Earnings were significantly higher over rounds in the endogenous condition compared to the fixed condition (β_6).

Examining the posterior distributions of the exogenous $\times$ round and endogenous $\times$ round parameter revealed an estimated difference of 0.79 with a 95% CI ranging from 0.32 to 1.26. Thus, also compared to the exogenous condition, earnings over time were significantly higher in the endogenous condition.

Table 5.S3
Earnings regression model.

		Coef.	**95% CI**
α_2	intercept (fixed condition, round 1)	118.72	[111.12, 127.07]
β_2	round	-0.36	[-1.08, 0.36]
β_3	exogenous condition dummy	-3.31	[-13.29, 6.79]
β_4	endogenous condition dummy	-3.30	[-13.30, 6.64]
β_5	exogenous condition $\times$ round	0.68	[-0.24, 1.59]
β_6	endogenous condition $\times$ round	1.47	[0.56, 2.41]
$\sigma^2_{\alpha_1}$	error-term random intercepts	1.36	[1.06, 1.68]
$\sigma^2_{\beta_1}$	error-term round slopes	13.58	[13.01, 14.15]
σ^2_y	error-term y	14.91	[12.01, 18.04]
ρ	correlation between random effects	-0.24	[-0.50, 0.02]

Note. Dependent variable: Average earnings clustered by group.

Maximum power. The maximum power regression model followed the same random structure as the contribution model (see Equations 5.6), but since power transfers only happened in the endogenous condition and exogenous condition groups merely mimicked these transfers over time, the model used only data of the endogenous condition and therefore contained only one predictor coding the round. The dependent variable was the amount of power of the most powerful group member in a particular round. Since maximum power could not fall below 1, the data was treated as left censored.

Table 5.S4 shows the estimated coefficients together with the 95% CI. Already in the first power transfer stage (round 3), the most power a participant accumulated on average was predicted to be 1.5 by the model (α_2). Power accumulation increased over rounds. In each round, the maximum power was estimated to increase by 0.02 on average (β_2).

Individual level analysis

The aim of the the individual level analysis was to analyse who received power, who was willing to give up power and how power affected contributions in the endogenous condition on a round by round basis. Therefore, the data was not aggregated on a group

Table 5.S4

Power change regression model.

		Coef.	95% CI
α_2	intercept (round 3)	1.455	[1.237, 1.672]
β_2	round	0.024	[0.004, 0.043]
$\sigma^2_{\alpha_1}$	error-term random intercepts	0.041	[0.023, 0.060]
$\sigma^2_{\beta_1}$	error-term round slopes	0.579	[0.536, 0.624]
σ^2_y	error-term y	0.499	[0.314, 0.699]
ρ	correlation between random effects	-0.063	[-0.574, 0.465]

Note. Dependent variable: Dependent variable: maximum power clustered by group.

level and thus was clustered by groups and by individuals (over time). The regression models accounted for that by having two grouping levels (see below).

In the receiving, giving and using power model, the distributions of the dependent variables were non-normal and highly restricted. We therefore transformed the dependent variable in these models into a dichotomous variable and fitted logistic regressions.

Also some predictors were transformed to dichotomous 'type' variables. This has the downside of losing some statistical power, as well as the ability to make more detailed quantitative statements (e.g. with a one point increase in the independent variable, the probability of the dependent variable being 1 changes by x). On the other hand, coefficients can be interpreted more easily. For example, by converting contributions to a binary variable (at or above/below group average), participants are classified into free riders (those who contributed less than group average) and cooperators (those who contributed at least or above group average).

Receiving rower. The dependent variable of the receiving power model was power received by other group members (0 = no power received, 1 = power received). The fixed part of the model contained the continuous round predictor, a dummy predictor indicating free riding or cooperation (0 = below average contribution, 1 = equal or above average contribution) in previous round's contribution stage, and a dummy predictor coding the punishment behaviour (see Equation 5.1; 0 = antisocial or no punishment, 1 = punishment of free riders) of previous round's punishment stage as predictors.

The random part of the model contained a random intercept and a random slope for the round predictor for each group, as well as a random intercept and a random slope for the round predictor for each participant. Thus, for each group a separate baseline of the likelihood of power transferring in round 3 (first power transfer stage), and a separate slope in how the likelihood of transferring power changed over rounds was estimated. Separately, the model estimated the likelihood of power being transferred in round 3, as well as the change in the likelihood of receiving power over rounds for each participant (see Equations 5.7).

$$
\begin{aligned}
Pr(y_i = 1) &\sim \text{logit}^{-1}(\mu_y), \text{for } i = 1, \dots, n \\
\mu_y &= \alpha_{1j} + \beta_{1j}\text{round} + \alpha_{2k} + \beta_{2k}\text{round} \\
&\quad + \alpha_3 + \beta_3\text{round} + \beta_4\text{contribution type}_{t-1} \\
&\quad + \beta_5\text{punishment type}_{t-1}
\end{aligned}
$$

$$
\begin{aligned}
\begin{pmatrix} \alpha_{1j} \\ \beta_{1j} \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\alpha_1} & \rho_1\sigma^2_{\alpha_1}\sigma^2_{\beta_1} \\ \rho_1\sigma^2_{\alpha_1}\sigma^2_{\beta_1} & \sigma^2_{\beta_1} \end{pmatrix} \right), \text{for } j = 1, \dots, J \\
\begin{pmatrix} \alpha_{2k} \\ \beta_{2k} \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\alpha_2} & \rho_2\sigma^2_{\alpha_2}\sigma^2_{\beta_2} \\ \rho_2\sigma^2_{\alpha_2}\sigma^2_{\beta_2} & \sigma^2_{\beta_2} \end{pmatrix} \right), \text{for } k = 1, \dots, K
\end{aligned}
\tag{5.7}
$$

$$
\begin{aligned}
\text{where } n &= \text{ number of observations} \\
J &= \text{ number of groups} \\
K &= \text{ number of subjects}
\end{aligned}
$$

Table 5.S5 shows the estimated coefficients together with the 95% CI as well as the odds ratio (exponential of coefficient). According to the model, cooperators had a 68% increase in odds of receiving power (β_4). The odds of receiving power more than doubled for participants who punished free riders in the previous punishment stage (β_5).

Giving away power. The regression model for giving away power followed the same random structure described in Equations 5.7. The dependent variable coded the transfer of power to other group member (0 indicated no transfer of power and 1 indicated transfer of power).

In this analysis, we were interested in whether the willingness to spend points on punishment predicted the likelihood of giving away power. Therefore, we used the dif-

Table 5.S5

Receiving power regression model.

		Coef.	odds ratio	95% CI
α_3	intercept (round 3)	0.16		[-1.13 1.53]
β_3	round	-0.14	0.87	[-0.25 -0.03]
β_4	contribution type$_{t-1}$	0.52	1.68	[0.18 0.86]
β_5	punishment type$_{t-1}$	0.78	2.18	[0.39 1.19]
$\sigma^2_{\alpha_1}$	error-term random intercepts (group level)	3.23		[2.08, 4.53]
$\sigma^2_{\beta_1}$	error-term round slopes (group level)	2.54		[1.87, 3.26]
$\sigma^2_{\alpha_2}$	error-term random intercepts (individual level)	0.23		[0.13, 0.33]
$\sigma^2_{\beta_2}$	error-term round slopes (individual level)	0.27		[0.19, 0.34]
σ^2_y	error-term y	49.97		[1.39, 96.31]
ρ_1	correlation random effects (group level)	-0.55		[-0.89, -0.14]
ρ_2	correlation random effects (individual level)	-0.46		[-0.72, -0.18]

Note. Dependent variable: power received (0 = no, 1 = yes) clustered by group and participant.

ference in the amount of points spent on punishment compared to group average as a dummy predictor (0 = punishment expense equal or above group average, 1 = punishment expense below group average). The fixed part of the model also contained the continuous round predictor and a dummy predictor indicating free riding or cooperation (0 = below average contribution, 1 = equal or above average contribution) in previous round's contribution stage.

Table 5.S6 shows the estimated coefficients together with the 95% CI as well as the odds ratio. According to the model, punishing below group average increased the odds of transferring power to other group members in the next round by 67% (β_5).

Table 5.S6

Giving away power regression model.

		Coef.	**odds ratio**	**95% CI**
α_3	intercept (round 3)	-1.42		[-2.52, -0.41]
β_3	round	-0.04	0.96	[-0.13, 0.04]
β_4	contribution difference$_{t-1}$	0.03	1.03	[-0.32, 0.38]
β_5	punishment difference$_{t-1}$	0.51	1.67	[0.17, 0.84]
$\sigma^2_{\alpha_1}$	error-term random intercepts (group level)	1.08		[0.04, 2.17]
$\sigma^2_{\beta_1}$	error-term round slopes (group level)	4.55		[3.48, 5.67]
$\sigma^2_{\alpha_2}$	error-term random intercepts (individual level)	0.07		[0.00, 0.16]
$\sigma^2_{\beta_2}$	error-term round slopes (individual level)	0.34		[0.26, 0.42]
σ^2_y	error-term y	49.96		[0.39, 95.35]
ρ_1	correlation random effects (group level)	-0.19		[-1.00, 0.83]
ρ_2	correlation random effects (individual level)	-0.64		[-0.83, -0.44]

Note. Dependent variable: Power giving (0 = no power was transferred, 1 = power was transferred) clustered by group and participant.

Usage of power. The regression model for using power followed the same random structure described in Equations 5.7. The dependent variable coded the punishment behaviour (0 indicated no punishment or antisocial punishment and 1 indicated punishment of free riders).

As predictors we used the continuous round predictor, the group contribution of the present round and a dummy variable indicating whether the participant had power above 1 in the present round (0 = power below or equal 1, 1 = power more than 1), as well as the interaction of this dummy with the group contribution.

Table 5.S7 shows the estimated coefficients together with the 95% CI as well as the odds ratio. According to the model, having a power greater than 1 increased the odds to punish free riders in the consecutive punishment stage nearly fivefold (β_5). Higher group contributions decreased these odds slightly. For each additional MU invested by the group, the odds for a powerful group member to punish decreased by 1% (β_6).

Table 5.S7

Using power regression model.

		Coef.	odds ratio	95% CI
α_3	intercept (round 3)	-1.03		[-1.78, -0.28]
β_3	round	-0.12	0.89	[-0.18, -0.06]
β_4	group contribution	-0.01	0.99	[-0.02, 0.00]
β_5	power	1.55	4.71	[0.85, 2.23]
β_6	group contribution $\times$ power	-0.01	0.99	[-0.02, -0.001]
$\sigma^2_{\alpha_1}$	error-term random intercepts (group level)	0.62		[0.02, 1.17]
$\sigma^2_{\beta_1}$	error-term round slopes (group level)	1.66		[1.19, 2.12]
$\sigma^2_{\alpha_2}$	error-term random intercepts (individual level)	0.09		[0.03, 0.16]
$\sigma^2_{\beta_2}$	error-term round slopes (individual level)	0.13		[0.08, 0.18]
σ^2_{y}	error-term y	49.97		[0.04, 95.00]
ρ_1	correlation random effects (group level)	-0.30		[-1.00, 0.68]
ρ_2	correlation random effects (individual level)	-0.64		[-0.88, -0.35]

Note. Dependent variable: punishment behaviour (0 = no punishment or antisocial punishment, 1 = punishment of free riders) clustered by group and participant.

Effect of power and punishment. The regression model for the effect of power and punishment followed the same random structure described in Equations 5.7. The dependent variable coded the change in cooperation from previous' round. Since the individual change in cooperation could not exceed 20 and fall below -20, the data was treated as left and right censored.

As predictors we used the continuous round predictor, the reduction in earnings due to punishment in the last round and the change in maximum power from last round, as well as the interaction of change in power and earning reduction due to punishment.

Table 5.S8 shows the estimated coefficients together with the 95% CI. According to the model, contribution decisions were influenced by actual punished, as well as changes in power of the most powerful group member (threat of getting punished). For every MU a participant lost due to getting punished in the last round increased her contribution by 0.3 MUs (β_4). A change in power of 0.1 of the most powerful group member, increased contributions by 0.5 MUs (β_5). Additionally, getting punished, followed by an increase in power, also increased contributions significantly (β_6).

Table 5.S8

Effect of power and punishment regression model.

		Coef.	**95% CI**
α_3	intercept (change to round 3)	-0.12	[-0.49, 0.25]
β_3	round	-0.03	[-0.08, 0.01]
β_4	earnings reduction (from punishment in round $t-1$)	0.31	[0.25, 0.37]
β_5	change in maximum power (from round $t-1$)	4.76	[3.06, 6.48]
β_6	earnings reduction $\times$ change in power	0.54	[0.08, 0.99]
$\sigma^2_{\alpha_1}$	error-term random intercepts (group level)	0.46	[0.00, 0.89]
$\sigma^2_{\beta_1}$	error-term round slopes (group level)	0.11	[0.00, 0.28]
$\sigma^2_{\alpha_2}$	error-term random intercepts (individual level)	0.07	[0.01, 0.12]
$\sigma^2_{\beta_2}$	error-term round slopes (individual level)	0.01	[0.00, 0.03]
σ^2_y	error-term y	3.73	[3.62, 3.84]
ρ_1	correlation random effects (group level)	-0.73	[-1.00, 0.23]
ρ_2	correlation random effects (individual level)	-0.27	[-1.00, 0.83]

Note. Dependent variable: change in cooperation from round$_{t-1}$ clustered by group and participant.

Individual level analysis for the exogenous condition

To analyse how gaining power and a change in maximum power affected individual decisions in the exogenous condition in which power transfers were not possible, we fitted the 'usage of power' and the 'effect of power and punishment' models described above also to the data of the exogenous condition groups.

Usage of power (exogenous condition). The regression model for using power followed the same structure as described above for the endogenous condition.

Table 5.S9 shows the estimated coefficients together with the 95% CI as well as the odds ratio. Like in the endogenous condition, having a power greater than 1 increased the odds to punish free riders in the consecutive punishment stage.

Table 5.S9

Using power regression model for the exogenous condition.

		Coef.	odds ratio	95% CI
α_3	intercept (round 3)	-2.57		[-3.33, -1.78]
β_3	round	-0.01	0.99	[-0.05, 0.03]
β_4	group contribution	0.00	1.00	[-0.01, 0.02]
β_5	power	1.69	5.42	[1.03, 2.34]
β_6	group contribution $\times$ power	-0.01	0.99	[-0.02, 0.00]
$\sigma^2_{\alpha_1}$	error-term random intercepts (group level)	0.79		[0.30, 1.30]
$\sigma^2_{\beta_1}$	error-term round slopes (group level)	1.67		[1.26, 2.07]
$\sigma^2_{\alpha_2}$	error-term random intercepts (individual level)	0.06		[0.01, 0.10]
$\sigma^2_{\beta_2}$	error-term round slopes (individual level)	0.09		[0.05, 0.13]
σ^2_y	error-term y	50.01		[0.18, 95.02]
ρ_1	correlation random effects (group level)	-0.18		[-0.84, 0.80]
ρ_2	correlation random effects (individual level)	-0.79		[-0.97, -0.59]

Note. Dependent variable: punishment behaviour (0 = no punishment or antisocial punishment, 1 = punishment of free riders) clustered by group and participant.

Effect of power and punishment (exogenous condition). The regression model for the effect of power and punishment followed the same structure as described above for the endogenous condition.

Table 5.S10 shows the estimated coefficients together with the 95% CI. Like in the endogenous condition, contribution decisions were influenced by actual punishment (β_4). However, contrary to the endogenous condition, a change in maximum power from last round (threat of punishment) did not affect contribution decisions significantly (β_5).

Table 5.S10

Effect of power and punishment regression model for the exogenous condition.

		Coef.	**95% CI**
α_3	intercept (change to round 3)	-0.25	[-0.73 0.22]
β_3	round	-0.03	[-0.08 0.02]
β_4	earnings reduction (from punishment in round $t-1$)	0.38	[0.30 0.46]
β_5	change in maximum power (from round $t-1$)	-0.13	[-0.54 0.29]
β_6	earnings reduction $\times$ change in power	0.19	[0.08 0.30]
$\sigma^2_{\alpha_1}$	error-term random intercepts (group level)	0.15	[0.00, 0.38]
$\sigma^2_{\beta_1}$	error-term round slopes (group level)	0.23	[0.00, 0.66]
$\sigma^2_{\alpha_2}$	error-term random intercepts (individual level)	0.02	[0.00, 0.04]
$\sigma^2_{\beta_2}$	error-term round slopes (individual level)	0.03	[0.00, 0.08]
σ^2_y	error-term y	4.56	[4.42, 4.69]
ρ_1	correlation random effects (group level)	-0.26	[-1.00, 0.83]
ρ_2	correlation random effects (individual level)	-0.25	[-1.00, 0.82]

Note. Dependent variable: change in cooperation from round$_{t-1}$ clustered by group and participant.

Other remarks

Power and cooperation. In Figure 5.3b, we reported the correlation of maximum power and average cooperation across rounds for each group. For the calculation of these correlations we omitted round 20 from each group. This was done because of the sharp drop of cooperation in the last round (known as the endgame effect). As can be seen in Figure 5.S14, with increased maximum power, mean contribution increased. Round 20 can be identified as an outlier.

Including round 20 in the reported analysis did not change the statistical inferences. Quantitatively, the average correlation dropped from $r = 0.24$ to $r = .21$ for the exogenous condition groups.

Power and earnings. As reported, receiving power was correlated with lower earnings compared to the other group members. Figure 5.S15 shows how earnings decreased, the more power a participant received in the experiment.

First and second round types. When examining the connection between behavior in the first two rounds, and later power status, we omitted an analysis of antisocial punishers, because antisocial punishment was very rare. The distribution of initial punishment behaviour was as follows: 75 participants decided to not punish in the second round, 49 participants punished primarily free riders, and only 11 participants punished antisocially.

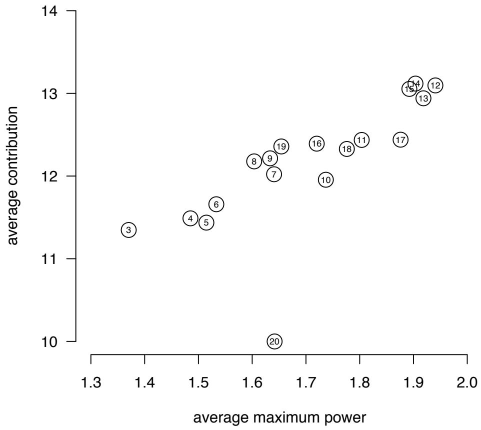


Figure 5.S14: Average cooperation and average maximum power in each round of the exogenous condition. The numbers in each circle indicate the round.

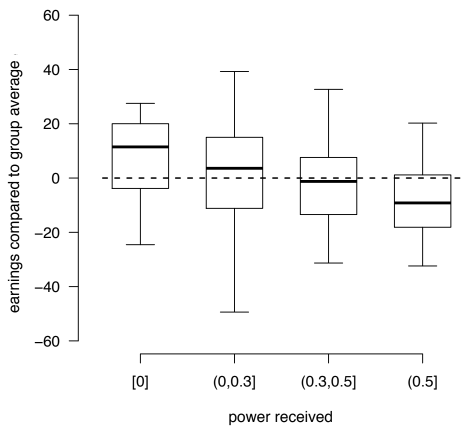


Figure 5.S15: Earnings and power. Average earnings of group members compared to the group average for different average amounts of power received over the experiment. Horizontal bar depicts the median.

Discussion

Concluding Remarks

In my thesis I combined the methodological repertoire of behavioral economics, neuroscience and psychology to gain insights into the complex universe of human decision making. In the sphere of decision making, psychologists can provide the conceptual framework for the inner architecture of our perceptual and cognitive systems, economists developed a rigorous mathematical and normative theory about how decisions should be made and under what circumstances rationality is violated, biologists lay out the genetic and evolutionary ground rules of development and with modern neuroscience, for the first time, the biological underpinnings of latent psychological constructs can be investigated. Similarly to a lot of social scientists, I share the hope that with the convergence of different disciplines, interdisciplinary theories will evolve that can describe human behavior more accurately and at the same time can be used for prediction, eventually guiding policy making.

Crossing the boundaries can also help to balance the perspective on human decision making. In economics, there is a strong emphasis on rational and optimal behavior[1], whereas in modern psychology, humans are often characterized as ill-adapted to their environment or as constrained in their ability of processing information, leading to distorted cognitions and sub-optimal behavior. The combination of the biological, psychological and economic perspective can help to understand under what circumstances human behavior does not meet the rational benchmark, what psychological and neurological constraints are responsible for this and why these constraints evolved in the history of natural selection.

By broadening the perspective on human decision making and carefully considering the biological, economic, psychological but also cultural perspective, consilience, the attempt to unite scientific disciplines, has another important justification: It can outline the boundaries of theories that were generated in the encapsulated tradition of a single discipline. All too often in the last two centuries we have seen reductionist theories, born out of the radicalization of few ideas specific to one subfield of the social or

[1]One argument why human behavior should approximate rationality is that at least in competitive environments, decision making biases should be easily exploitable by more rational behavior and should therefore be driven out in the long run (Alchian, 1950).

the natural sciences, in which simplicity was confused with predictive power. The danger of reductionism looms when scientists delude themselves into believing that they found simple and powerful, yet descriptively accurate accounts of complex phenomena like human decision making, as can be illustrated by three major concepts of biology, psychology and economics: evolutionary theory, reinforcement learning theory and rational choice theory.

In the late 18th century, scholars were convinced that facial features like the shape of the nose or the size of the ears can be used to predict personality traits and cognitive abilities, and that eventually, the study of physiognomic characteristics would be able to identify criminals or soon to be criminals. Phrenologists believed that measuring the shape of the skull would reveal the manifestation of 'mental faculties', determining intelligence or emotionality. The idea that psychology can be reduced to physiognomy, that the latent can be traced back to the manifest (the shape of the head or the body), was ultimately a perverted generalization of theories of heredity, and later, Darwinian evolution. Eventually, these biological principles, were extended beyond their justifiable application and used for the scientific legitimization of discrimination against women, homosexuals or black people up until the early 20th century (see Figure 6.1). What is considered pseudoscience today was not a practice of some untalented researchers. Instead, well respected scientists, like the French physiologist Paul Broca who pioneered neuroanatomy, or the Italian physician Cesare Lombroso, four times nominated for the Nobel prize, were lured by the Siren's voice of reductionist theories into believing that complex behavior can be reduced to superficial and simple principles.

In psychology, reinforcement learning had to serve as the one concept that would rule them all, eventually predicting and controlling the most complex human behavior. Still a powerful concept in learning theory today, the generalization of reinforcement learning beyond its descriptively valid boundaries fueled the hubris of behaviorists like John Watson: "Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select – doctor, lawyer, artist, merchant-chief, and, yes, even beggarman and thief [...]" (Watson, 1930, p. 82). Phenomena that were hardly reducible to reinforcement learning, like mental abilities, were simply denied in radical behaviourism: "Mental life and the world in which it is lived are inventions. They have been invented on the analogy of external behavior occurring under external contingencies. Thinking is behavior. The mistake is in allocating the behavior to the mind" (Skinner, 1971, p. 50).

As the mind and the cognitive abilities of humans marked the boundaries and threats for the predictive power of reinforcement learning models, the observation of irrational

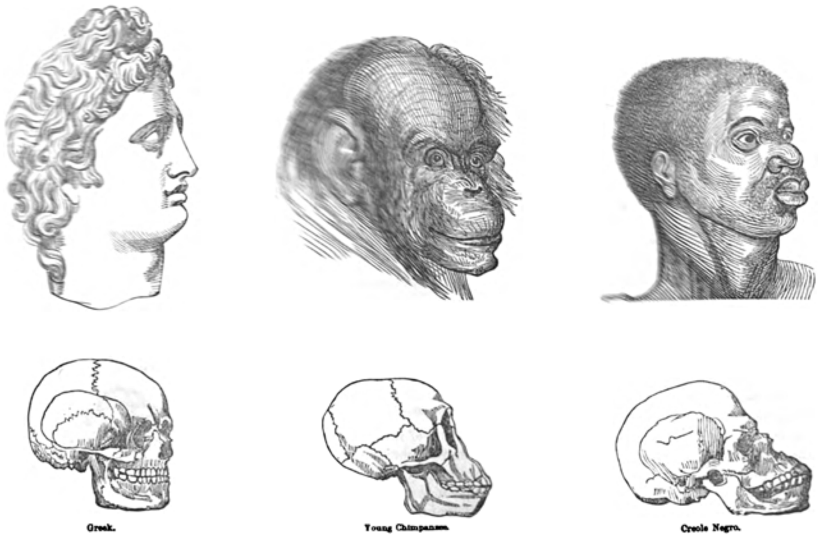


Figure 6.1: Distorted drawings of skulls, trying to raise doubts that "the Orang-Outan and Chimpanzee are more widely separated from certain African and Oceanic Negroes than are the latter from the Teutonic or Pelasgic types" (Nott, 1855, p. 458, picture adapted from Nott, 1855).

behavior is a threat to the predictive power of neoclassical rational choice theory. Although most economists see rational choice theory as a simplified, yet useful misrepresentation of human behavior, attempts were made to reframe the irrational as rational, like in the work of the Nobel prize winning economist Gary Becker. Becker proposed that drug addiction is ultimately driven by rational decision making: "addictions, even strong ones, are usually rational in the sense of involving forward-looking maximization with stable preferences" (Becker and Murphy, 1988, p. 675). What psychologists would characterize as highly maladaptive, Becker tried to impute with rationality, in the hope that the useful misrepresentation of humans as rational selfish agents is in fact the 'real deal', an accurate description of human nature. Becker's teacher Milton Friedman, perfectly aware of the reductionist nature of rational choice theory, nevertheless popularized free market policies that are based on the assumption of rationality through public speeches and his own television series.

Evolution, reinforcement learning and rational choice are important frameworks to think about human decision making. In certain domains they are powerful in prediction and at the same time descriptively accurate. However, generalizing these concepts to the extent that descriptive validity is lost and the illusion of control and predictive power still maintained, will not advance the understanding of human decision making. Interdisciplinary communication and cooperation, I believe, is therefore particularly well suited to define the justifiable application of broad concepts to broad behaviour (e.g., whether the rational choice framework can be extended to substance addiction and if so, to what extent). The cognitive revolution in psychology for example, that defined the boundaries of reinforcement learning, did not gain momentum until researchers from neighboring disciplines started to cross the discipline border. Although the work of Edward Tolman that I outlined in the introduction, has already pointed to the descriptive weaknesses of behaviorist theories, Tolman's work was very much ignored by his colleagues, and his cognitive concepts like latent learning went mostly unnoticed.[2] It was the combination of young psychologists, computer scientists and linguists that formulated the limits of behaviorist theories and came up with new, more complex theories for behaviour for which the reinforcement learning framework was ill-suited. For example, when Skinner published his book 'Verbal Behavior' (Skinner, 1957), proposing that language acquisition can be solely explained by reinforcement mechanisms, he caught the attention of the young linguist Noam Chomsky. Just as Watson defined the boundaries for psychologists in his speech at Columbia University in 1913, about 50 years later, Chomsky's famous critique of Skinner's book outlined the weaknesses and explanatory boundaries of the behaviorist program from a linguist's perspective (Chomsky, 1959). Today, the emergence of behavioral economics and neuroeconomics, in which economists work together with neuroscientists, biologists and psychologists, could ultimately serve the same purpose in economics. Just like computer scientists and linguists in the late 1960s and 1970s did it for reinforcement learning models, behavioral economists, neuroscientists, and psychologists start to investigate the limits of the rational choice framework.

Therefore, my hope for what consilience can provide is not only the formulation of theories that can increase predictive power. Instead, I believe consilience could help prevent that simple theories, usually characterized by the narrow generalization of concepts that emerged in one discipline, go rogue. Ultimately, the convergence of different disciplines can remind researchers of the boundaries of their own discipline and to be more cautious about what theories of human psychology can really provide to the social sphere.

In this spirit, I would like to conclude by pointing out the limitations of the studies I presented in my thesis.

[2] How important Tolman's concepts still are today shows the 2014 Noble prize in Physiology or Medicine that was awarded to John O'Keefe and May-Britt and Edvard Moser who worked on how the brain learns to spatially represent the environment, eventually constructing a latent map.

In the first study, we showed how value signals in the frontal cortex can be used to predict preferences across distinct reward categories: snack foods and activities. Together with numerous other studies (see Levy and Glimcher, 2012 for a review and Bartra et al., 2013 for a meta study), this result points to the existence of a utility calculation machinery in our brain, very much in line with the Benthamite utility concept. However, in our experiment, we do not investigate how these value signals are affected by different framings of choices. In fact, we do not investigate actual decisions but only the construction of subjective value, which we believe underlies decisions, but might be affected by various situational variables. Further, we only use two reward categories. In John Stuart Mill's famous critique of Bentham's utility idea, he proposed that there exist qualitatively different forms of value that are not represented on the same scale and are thus not comparable. It remains to be seen if the ventromedial prefrontal cortex really represents subjective value across all decision domains, from simple food choice to, for example complex moral considerations.

In the second study we demonstrated how the arbitrary restriction of arm movements can influence purchasing decisions. Although we propose an evolutionary theory, more evidence is needed to understand the concrete mechanism of this phenomenon. Can it also be observed in animals? Can we find evidence that indeed activation in the motor cortex underlies this decision making bias?

In the third study, we show how the disruption of the dorsolateral prefrontal cortex can lead to selfish and maladaptive behaviour. In the literature and also in our paper, the prefrontal cortex is often characterized as a Freudian super-ego that enables us to control selfish impulses (Knoch et al., 2006). Fundamentally, this research addresses the question of human nature itself, by attributing selfishness to the subcortical and evolutionary old parts of our brain and fairness and prosocial behaviour to the neocortical regions that only developed recently in evolutionary history. Although the simplicity of this theory is tempting, recent research suggests that spontaneous behaviour, less controlled by reasoning and higher order cognitive processes, does not have to be selfish but can be equally generous and prosocial (Rand et al., 2013, Ferguson et al., 2014, Schulz et al., 2014). Further, recent experimental work suggests that the framing of the situation very much influences whether our first impulse is to be nice or greedy (Bednar et al., 2012, Schulz et al., 2014).

In the last study, we investigated how voluntary transfer of power can solve the free rider problem in the public goods dilemma. In this paper we share the perspective of early social contract philosophers like Thomas Hobbes, who believed that a chaotic decentralized state is undermining cooperation and social structures are needed to fight exploitation and stabilize cooperation (Hobbes, 1651). However, does power centralization always have such beneficial characteristics? What if power is misused and exploited? By varying the employed power transfer game such that, for example power cannot be taken back as soon as it is transferred to another individuals, we should be able to explore the boundaries of the benefits of power centralization and investigate when legitimate power turns into force and tyranny.

References

Alchian, A. A. (1950). Uncertainty, evolution, and economic theory. *The Journal of Political Economy*, 58(3):211–221.

Bartra, O., McGuire, J. T., and 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.

Becker, G. S. and Murphy, K. M. (1988). A Theory of Rational Addiction. *The Journal of Political Economy*, 96(4):675–700.

Bednar, J., Chen, Y., Liu, T. X., and Scott, E. (2012). Behavioral spillovers and cognitive load in multiple games: An experimental study. *Games and Economic Behavior*, 74:12–31.

Chomsky, N. (1959). A Review of B. F. Skinner's Verbal Behavior. *Language*, 35(1):26–58.

Ferguson, E., Maltby, J., Bibby, P. A., and Lawrence, C. (2014). Fast to Forgive, Slow to Retaliate: Intuitive Responses in the Ultimatum Game Depend on the Degree of Unfairness. *PLoS ONE*, 9(5):e96344.

Hobbes, T. (1651). *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiasticall and Civil*. Andrew Crooke, London.

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

Levy, D. J. and Glimcher, P. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, 22:1027–1038.

Nott, J. C. (1855). Comparative Anatomy of Races. In Morton, S. G., editor, *Types of Mankind*, pages 411–465. Lippincott, Grambo & Co., Philadelphia.

Rand, D. G., Greene, J. D., and Nowak, M. A. (2013). Spontaneous giving and calculated greed. *Nature*, 489(7416):427–430.

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

Skinner, B. F. (1957). *Verbal Behavior*. Copley Publishing Group, Acton.

Skinner, B. F. (1971). *Beyond Freedom and Dignity*. Hackett Publishing, Indianapolis.

Watson, J. B. (1930). *Behaviorism*. University of Chicago Press, Chicago.