

Motivating forces of human actions Neuroimaging reward and social interaction

Henrik Walter^{a,b,*}, Birgit Abler^b, Angela Ciaramidaro^c, Susanne Erk^a

^a *Department of Psychiatry, Laboratory for Neuroimaging and Neurophysiology, Johann Wolfgang Goethe University Frankfurt, Heinrich-Hoffmann-Str. 10, 60598 Frankfurt, Germany*

^b *Department of Psychiatry, University of Ulm, Leimgrubenweg 12-14, 89075 Ulm, Germany*

^c *Center for Cognitive Science, Department of Psychology, University of Turin, via Po, 14–10123 Turin, Italy*

Available online 25 July 2005

Abstract

In neuroeconomics, reward and social interaction are central concepts to understand what motivates human behaviour. Both concepts are investigated in humans using neuroimaging methods. In this paper, we provide an overview about these results and discuss their relevance for economic behaviour. For reward it has been shown that a system exists in humans that is involved in predicting rewards and thus guides behaviour, involving a circuit including the striatum, the orbitofrontal cortex and the amygdala. Recent studies on social interaction revealed a mentalizing system representing the mental states of others. A central part of this system is the medial prefrontal cortex, in particular the anterior paracingulate cortex. The reward as well as the mentalizing system is engaged in economic decision-making. We will discuss implications of this study for neuromarketing as well as general implications of these results that may help to provide deeper insights into the motivating forces of human behaviour.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Neuroeconomics; Ventral striatum; Medial prefrontal cortex; Theory of Mind; Mentalizing

1. Introduction

Neuroeconomics can broadly be defined as the interdisciplinary enterprise to investigate economic-related behaviour by using neuroscientific methods. Thus, neuroeconomists are interested in the neural correlates of the motivating forces of behaviour and decision-making. They study mechanisms of behaviour in game theory, e.g. [67,90–92,96,99] or the processing of financial rewards, e.g. [19,33,38,59–62,81,87]. The study of consumer products on information processing is also known as neuromarketing, e.g. [18,35,40,69]. The main idea is that neuroscientific knowledge about the factual decision-making processes in human agents will inform and stimulate theories in economics. But why should that be necessary?

As described by Camerer [25], early economic theory was built on a psychological model of humans, which

considered human behaviour as the result of a process of decision-making, weighting costs and benefits of actions to maximize utility. A problem of this model was that feelings and preferences do influence decision-making but could not be measured directly. This problem was solved by introducing the concepts of ordinal utility and revealed preference in the 1940s. Revealed preference theory simply equates unobserved preferences with observed choices. Circularity is avoided by assuming that people behave consistently. Ordinal utility and revealed preferences as well as later extensions – discounted, expected and subjective expected utility and Bayesian updating – gave economists an easy way to avoid the messy reality of the psychology underlying utility. Economists then spent decades developing mathematical techniques to make economic predictions without having to measure thoughts or feelings directly.

However, with the advent of cognitive neuroscience in the 1980s, techniques and devices have been developed to measure thoughts and feelings at least indirectly, i.e. by measuring

* Corresponding author. Tel.: +49 69 6301 5897; fax: +49 69 6301 3833.
E-mail address: h.walter@med.uni-frankfurt.de (H. Walter).

their neural underpinnings [107]. It was only a matter of time that problems and questions of economics reached the neuroscience labs and that researchers began to cooperate in order to make progress on them.

What can cognitive neuroscience teach economics? Cognitive neuroscience deconstructs the picture of perfectly rational humans, which are deliberating their choices by weighting costs and benefits until a deliberative equilibrium is reached. Although humans are definitely capable of conscious deliberation, many, if not most economically relevant decision processes are characterized by certain other features: first, they rely on automatic, fast and effective cognitive processes, which are not under direct volitional control [8]. Second, they are under the influence of unrecognized and finely tuned affective mechanisms, which often play a decisive role in action [26,29,83]. Third, many of these processes have been shaped by evolution in order to serve social purposes [2,21,22]. Thus, decision-making and evaluation in economic contexts will be influenced by mechanisms dedicated to social interaction [18].

The purpose of this paper is to give a selective overview over recent research in the cognitive neurosciences that deal with some of these aspects. In particular, we will discuss reward and motivation (first part) and social interaction as studied by game theory (second part).

2. Reward and motivation

2.1. The concept of reward

From an evolutionary point of view, the capacity to seek rewards as goals is essential for the survival and reproduction of mobile organisms. Generally, rewards can be defined as those stimuli which positively reinforce the frequency or intensity of a behaviour pattern. Food, water and sexual stimuli are called *primary rewards* as they reinforce behaviour without being learned. It can safely be assumed that the reward mechanisms pertaining to such stimuli are mostly innate because they are essential for survival and reproduction. Other stimuli, such as cultural goods or money, are called secondary rewards, as they reinforce behaviour only after having been learned. It is assumed that they acquire reward value through association with primary rewards. Indeed, almost any stimulus can do so through appropriate learning.

According to Schultz [97], three functions of reward can be distinguished: first, rewards induce learning, as they make an organism come back for more (positive reinforcement). Second, they induce approach and consuming behaviour for acquiring the reward object and third, they induce hedonic (positive) feelings. Rewards serve as goals of behaviour, if reward, action and the contingency between both are represented in the brain during the action. By contrast, punishers induce avoidance learning, withdrawal behaviour and negative emotions.

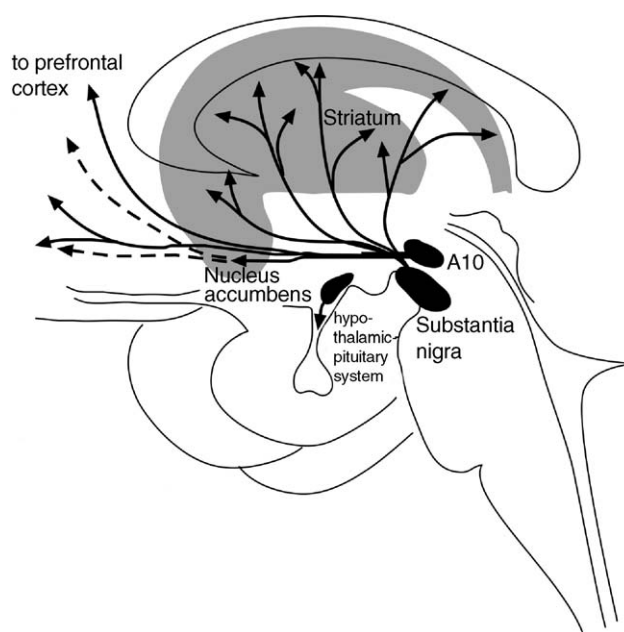


Fig. 1. Schema of the dopaminergic mesolimbic–mesocortical reward system. The dopaminergic neurons innervating the ventral striatum (mesolimbic system) and the prefrontal cortex (mesocortical system) have their origin in the ventral tegmental area (adapted from [103]).

2.2. The mesolimbic–mesocortical dopamine system

In the 1950s, it was discovered that rats will self-stimulate themselves excessively if they have free access to electrodes, which are implanted in the medial forebrain bundle. This bundle has its origin in the ventral tegmental area (VTA) A10 and projects to the ventral striatum including the nucleus accumbens (NAc) and the forebrain. In humans, a similar system exists, the so-called mesolimbic–mesocortical dopamine system (Fig. 1).

For a long time, it has been assumed that dopamine mediates the hedonic feeling associated with receiving rewards, i.e. feelings of pleasure and lust. This assumption was based on experiments with blocking or inducing dopamine transmission in animals and men. Also, hedonic drugs, like cocaine, induce dopamine release and have been found to activate the ventral striatum in humans addicted to cocaine in neuroimaging studies [20].

However, there is now accumulating evidence that dopamine itself does not mediate the hedonic component (“liking”) but rather the motivational component (“wanting” or “incentive salience”) in rewarded behaviour. Although the destruction of the dopamine cells in rats stops all motivated behaviour (wanting), the animals still do show signs of liking and disliking liquid foods when fed compulsory [14,13,93,103].

2.3. Attractive sport cars activate the reward circuitry

As an example for an early neuroeconomic study involving clearly secondary rewards, we will shortly describe a

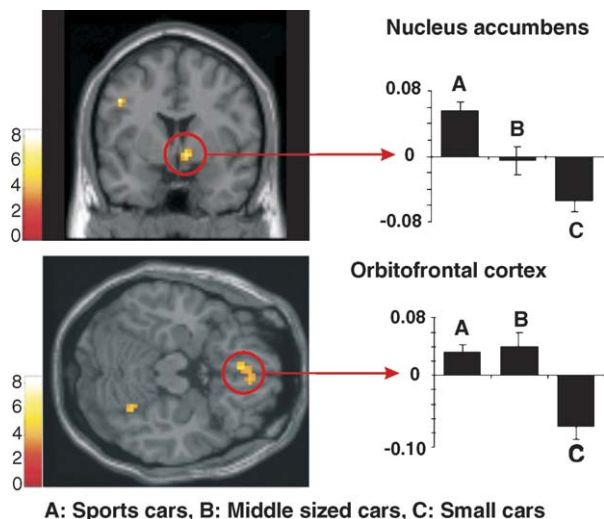


Fig. 2. Judging attractive sport cars activates the reward circuitry. The figure shows the effects sizes of the hemodynamics response while different types of cars were evaluated for their attractiveness (adapted from [40]).

study of our own group [40]. In this study, men were presented with different types of grey-scaled photographs of cars with all brand labels removed: sport cars, middle-sized cars and small cars. Their task was to assess the attractiveness of the cars on a five-point scale, irrespective of the costs or practical considerations, for example space for children. Behavioural results showed that sports cars were judged significantly more and small cars significantly less attractive than medium-sized cars. More attractive cars elicited stronger activations in several brain regions including areas implicated in reward processing. The contrast sports cars > small cars showed significantly more activation in right ventral striatum and medial orbitofrontal gyrus (Fig. 1), as well as in left anterior cingulate cortex, bilateral dorsolateral prefrontal cortex, right fusiform gyrus and left occipital cortex (Fig. 2).

Why should sports cars activate the reward circuit? A preliminary answer might be, because people simply *like* sports cars. However, this explanation is too simple. According to the concepts described above, we should not ask why people *like* sports cars but rather why they *want* them. More precisely, we may ask to what kind of primary reinforcer sport cars may be linked. Certainly, sport cars symbolize speed, power and independence. Apart from that, however, we propose that sports cars do signal social dominance. Social relationships are a very important part in the life of higher mammals. Higher mammals are well able to recognize signals related to social hierarchy and dominance. Our explanation is supported by a study of Morgan et al. in monkeys [74]. These authors have shown that the establishment of a social hierarchy has measurable consequences on the sensitivity of the brain reward circuitry. After 3 months of social housing macaque monkeys differ in amount and availability of dopamine D₂ receptors in relation to their acquired social rank: in contrast to subordinate monkeys dominant monkeys showed an increase of the amount and availability of

D₂ receptors in the ventral striatum and less cocaine self-administration.

But what kind of relation do sports cars bear to social dominance? Dominance and social rank are mediated by individual attributes signaling wealth and superfluity. A classical example is the peacock's elaborate tail: the peacock's tail has no apparent survival value and might actually hinder survival in making its owner more conspicuous to its predators and too clumsy to escape [16,28]. The tail thus fulfils no meaningful function except for signaling that its owner is obviously strong enough to be able to invest energy in such a superfluous structure. In fact, peacocks that are able to produce the most fancy and ornamental tails are "fitter" in the Darwinian sense, which has been shown in a study reporting that the offspring of those peacocks grew faster and had better survival rates [84]. In human societies, it is known that the demonstration of wealth and superfluity is also a strong signal for social dominance. This has been shown already in rural societies like the Indian tribe of the Kwakiutl in British Columbia, where the aim of a feast called Potlatch is to give away or to destroy more goods than other competing chieftains. By this behaviour the respective chieftain demonstrates his power and wealth [52]. Sports cars also differ from natural reward stimuli like food, sex and faces (mates) and do not have an intrinsic reward value. In contrast, they are neither economically nor ecologically appropriate, certainly not spacious for more than two persons and often dangerous for their owners or drivers. However, they signal high social rank, social dominance and wealth, e.g. can be regarded as the human equivalent to the peacocks tail.

2.4. Studying reward in the monkey brain

A lot of what we know about reward processing is based on animal research. Single cell recordings in monkeys have demonstrated the involvement of several brain structures in reward processing. Apart from the mesolimbic–mesocortical system the dorsolateral prefrontal cortex (DLPFC), the anterior cingulate cortex (ACC) and the posterior cingulate cortex, the frontal eye fields, the parietal cortex and the thalamus have been implicated in reward processing. We will shortly describe the main forms of reward coding in monkeys according to Schultz [97] as they provide the conceptual and empirical background of many human neuroimaging studies.

Behavioural experiments in monkeys often are choice tasks. Often, monkeys are trained with neutral cues that announce a reward, sometimes the reward is omitted or comes unexpected. In the simplest form of reward processing dopaminergic neurons detect the receipt of reward by an increase of activity after the reward has been obtained (Fig. 3). If a monkey learns that a cue predicts a reward by Pavlovian conditioning, there are neurons that after conditioning respond to these reward-predicting stimuli (Fig. 3). If a predicted reward is omitted, dopaminergic neurons decrease their firing rate at the time of reward (not shown in Fig. 3). Reward predictions may trigger actions. Therefore, it is use-

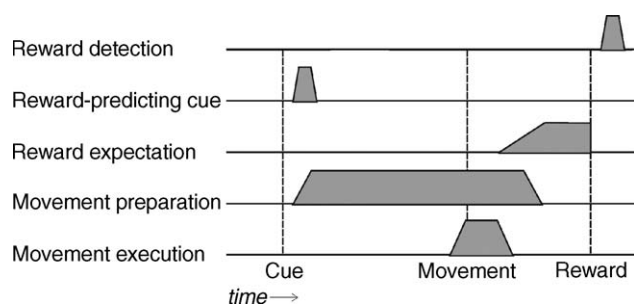


Fig. 3. Forms of reward coding in the monkey (adapted from [97]).

ful that the representation of a reward (which now serves as a goal) is maintained until the goal is obtained. Therefore, there are also neurons which are active during that time (Fig. 3). Compared to the phasic signal during obtaining a reward or at the time of cue representation, the firing of these neurons is more extended and shows a ramp-like shape, i.e. increases as the delivery of the reward approaches. Rewards serve as goals guiding actions and consequently there are also neurons found that fire during preparation and execution of movements directed at rewards (Fig. 3). To summarize very briefly, we can say that neurons of the reward systems in monkeys fire when reward is obtained, predicted, expected and acted for. We will now turn to the human reward system and neuroimaging studies.

2.5. Reward in human neuroimaging studies

Since the experiment of Breiter et al. [20] with direct manipulation of the reward system with cocaine, reward has been studied in humans with several neuroimaging methods, especially fMRI. As a note of caution in interpreting these studies, one should bear in mind that the BOLD-signal of fMRI has a delay of about 6 s after onset of neural activity, that it cannot distinguish between excitation and inhibition and that the most interesting regions of the reward system are in regions which are highly susceptible to artefacts. Nevertheless, numerous studies have found a common set of structures activated in tasks using a wide range of rewarding stimuli. The regions activated vary with the task and are dependent on methodological limitations, but the most consistently activated structures include the orbitofrontal cortex (OFC), the amygdala and the ventral striatum/nucleus accumbens (compare Fig. 4).

Other regions have been also identified like the dorsolateral and medial prefrontal cortex or the anterior cingulate. They are assumed to integrate reward information in action and decision-making and we will discuss one of these areas (the medial prefrontal cortex) later on. For the following selective overview we will restrict the focus to the brain areas related to detecting, predicting and valuing rewards [70].

The human OFC–amygdala–NAc reward system has been shown to respond to primary rewards, e.g. sugar water [82], tomato juice and chocolate milk [63], appetitive smell [43],

chocolate [102] and sexual stimuli [7] as well as to secondary rewards, e.g. money [19,59,62] (compare Fig. 5), music [17] or cars [40]. Activation of the reward circuitry has also been found in studies using social stimuli like beautiful faces [3], social interactions [90], affect-laden words [43] and pleasant touch [95].

Like in animals, it has been shown that responses in the reward circuit shift from the reward delivery itself to the (arbitrary) cue announcing the reward [79]. At the beginning of a conditioning experiment with sugar water the ventral striatum and the OFC were active upon receiving the sugar water. After conditioning, these regions were only active when the cue appeared that announced the sugar water. In terms of learning theory: the reward system reacts first to the primary, unconditioned reinforcer (the sugar water) and then, after learning, to the secondary, learned reinforcer (the cue). Fig. 5 shows activation of the NAc when different levels of financial reward are expected [1].

Although, it is safe to assume that reward components are processed not by separate regions but rather by interaction of different brain structures it is useful to associate the core structures with core functions in order to get an idea of the working of the reward system. We will now discuss three core structures of the reward circuit in more detail [70].

2.5.1. Orbitofrontal cortex: reward valence

The orbitofrontal cortex is a region with dense afferent connections from the primary taste and olfactory cortex as well as from higher visual areas and the somatosensory cortex [94]. In macaques, the OFC response to an apple is strong, if the alternative is a less tasty cereal, but weak, when the alternative is a very tasty raisin [105]. In line with these findings, human neuroimaging studies have shown that OFC activity is correlated with the reward value of stimuli. The OFC responds to the level of monetary reward [38] as well as expected monetary reward [19,60], to the valence of odours indicated by abstract cues [44] and to the level of attractiveness of cars [40].

In a recent study, it has been shown that activity in the ventromedial PFC is correlated positively with the subjective preferences for cultural familiar drinks (Pepsi and Coke) [69]. There is some evidence for a medial–lateral dissociation in the sense that the medial OFC is active for rewarding stimuli (gaining money) whereas the lateral OFC is active for punishing stimuli (losing money) [81]. More generally, the lateral OFC seems to be active when learned contingencies have to be inverted (reversal learning) [64,78].

2.5.2. Amygdala: reward intensity

The amygdala is well known for its role in emotional information processing, especially in fear conditioning [65] and in the perception of emotional face expressions [2]. Traditionally, it has been linked with negative emotions [23]. However, the notion that the amygdala only responds to negative stimuli

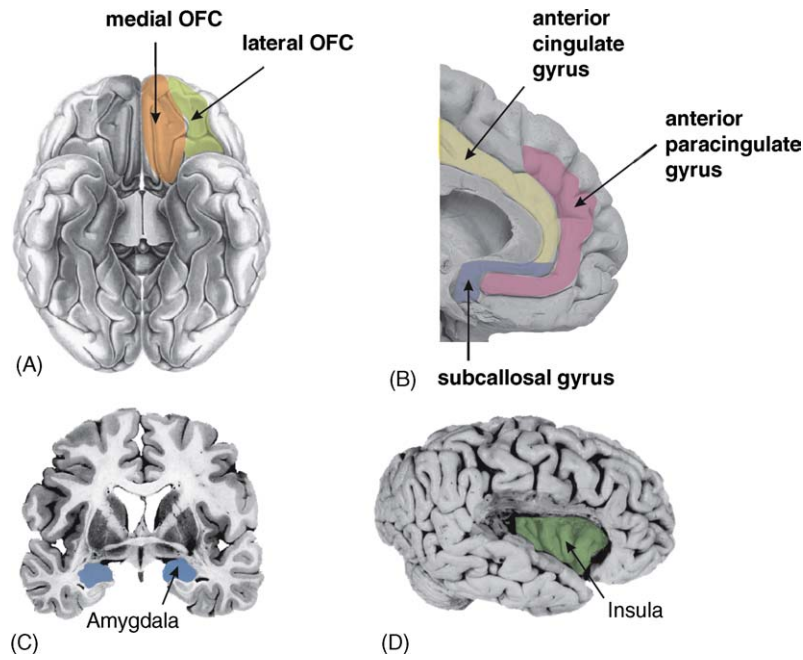


Fig. 4. Neuroanatomy: (A) shows the brain upside down, (B) shows a medial view of the prefrontal cortex and (C) a coronal section. In (D) part of the cortex has been removed to show the insula.

has been questioned recently. With fMRI, it has been shown that the left amygdala reacts to negative as well as positive stimuli [51] with some evidence for a left (positive)–right (negative)-asymmetry [51,110]. The amygdala has also been found reacting to the prediction of financial reward [56] and to the devaluation of a previously rewarded stimulus [44].

However, valence and intensity are not easy to disentangle, as negative stimuli are generally more arousing. A recent fMRI study explicitly addressed this question using different odours, which allowed to manipulate valence and intensity independently [5]. It was found that only the OFC reacts to valence, whereas the amygdala reacts to intensity indepen-

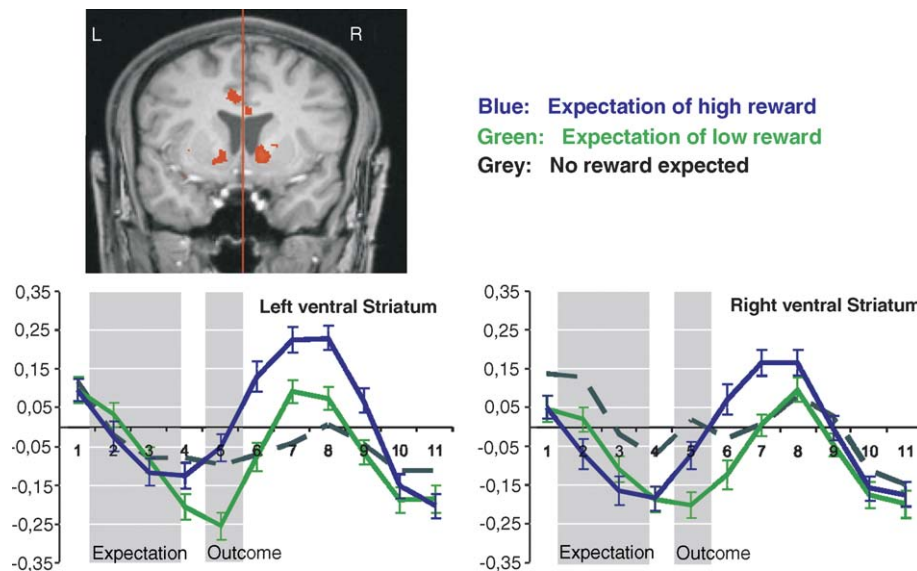


Fig. 5. Reward value is represented in the N. Accumbens (NAc) in an incentive monetary delay task. The result of an experiment of a group with four subjects is shown (for the whole group compare [1]). Three different abstract cues announced that there was a chance of 60% to win either €1, €0.20 or nothing if a simple perceptual decision was performed correctly after the expectation period. The figure shows the result of contrast high reward versus low reward for the expectation period (for winning and losing trials) with the NAc active on both sides. For the outcome period there was no accumbens activity. Note, that the maximum of activation is about 6 s after the end of the expectation period corresponding to the hemodynamic delay (x-axis, time in scans; TR (time of repetition), 1.5 s; y-axis, effect size, statistical threshold on the voxel level $p < 0.005$, uncorrected).

dent of valence (see also [101]). Also, in animal experiments the amygdala is activated strongly when positive reinforcement is involved [9]. In the context of reward processing the activation of the amygdala thus can be interpreted as a structure coding for the intensity of rewards.

2.5.3. Ventral striatum: reward prediction and salience

The ventral striatum/nucleus accumbens has been found to be activated in human neuroimaging experiments when a reward is received [33] as well as when a reward is expected [82]. Although ventral striatal activity correlates with the amount of money and thus indicates reward value [59], the predictability of the reward seems to be the most important information the NAc carries. In humans, this was clearly shown for the first time by Berns et al. [12]. Mildly pleasurable stimuli (juice or water) only elicited activity in the NAc, when their occurrence was unpredictable but not when it was predictable. The prediction error theory of NAc activation (for a recent review see [70]) has been developed from animal studies in monkeys, see for example [55,98]. The prediction error is the discrepancy between the value of probability (or magnitude or timing) with which a reward is predicted and the actual outcome of this value. A positive prediction error means that the prediction was too pessimistic, i.e. the reward is higher than expected. A negative prediction error means that the prediction was too optimistic, i.e. the reward was less than expected. By comparing representations of expected and actual outcome an organism can adjust its behaviour, i.e. it can learn. The response of the NAc resembles the teaching signals that have been employed in temporal difference learning in computational theory [104]. Therefore, the signal in the NAc carries information about the prediction error. This has now also been confirmed in human subjects with fMRI experiments using appetitive conditioning with a pleasant taste reward [68,79] or conditioning with fruit juice [68].

Other authors have proposed that it is not the prediction error but salience in general that is coded by the NAc [113,112]. Salience can be defined as the property of a stimulus that is both unexpected and elicits an attentional-behavioural switch. In the study of Zink et al. [113], the ventral striatum was active when money could be gained by acting in a specific way but not when money was gained passively. This shows that reward processing has to be understood in the context of motivated behaviour. Moreover, feedback after correct trials without rewarding character did not elicit activation in the ventral striatum. Thus, motivation and incentives for actions seem to be important aspects, which lead to an activation of the reward system. Another piece of evidence for the salience theory is that negatively rated stimuli have been found to elicit activation in the NAc. This has been shown by Berrera et al. [10] for noxious thermal stimuli. Although it has been argued that this might be a pain relief effect, a recent study did not support this interpretation. An activation of the NAc was found already in the expectation phase of slight electrical shocks [57].

2.6. Social interaction

We will now turn to another important subject of neuroeconomic studies, namely to the study of social interaction. Game theory has formalized social interactions in which the players may profit or lose [24]. The standard approach in cognitive neuroscience is individualistic. The individualistic approach can easily be deduced from methodological constraints: normally, cognitive neuroscientists put people into a scanner and confront them with simple stimuli. Due to methodological technological limitations, most studies on social interaction until now also use offline designs, in which scanned subjects do not actually take part in social interaction, but do represent social interaction. In contrast, in an ‘online-design’ subjects under investigation actually interact with con-specifics—or at least are told so [42,67,90–92,96]. ‘Hyperscanning’ is a term that has been introduced for a design in which two actually interacting subjects are located each in a scanner while the brain activities of both are recorded [73].

Researchers have begun to investigate what happens in the brain of subjects when they are involved in games like the prisoner’s dilemma [90,99], the Ultimatum Game [96] or trust games [67]. Furthermore, studies have been conducted that investigate social interaction in a moral context, e.g. social exclusion [37], the consequences of suppressing prejudice [89], moral decision-making [43,47,48,53], social norm violation [15] or punishing people for unfair actions [31]. We will describe results of game theoretic inspired studies on cooperation and punishment, focusing on online-designs.

2.7. Cooperation and competition

In the first fMRI study on game theory volunteers played a standard trust and reciprocity game each with a human or a computer counterpart for cash rewards [67]. By varying payoff numbers, different incentives for cooperation could be studied. After playing, subjects were divided into cooperative and non-cooperative subjects according to behavioural results. In a group analysis, activation during waiting for the result screen was compared when playing against a computer or when playing against a human: only cooperators showed activation of the medial prefrontal cortex. This was interpreted by the authors as an indication for the existence of a convergence zone that binds attention for mutual gains during cooperation.

In another fMRI study, women played an iterated Prisoner’s dilemma game either with human partners or with a computer [90]. In the prisoner’s dilemma, subjects will profit most, if they defect and their partners cooperate (D/C), and least, if it is the other way round (C/D)—with the total payoff matrix (subject/partner) being $D/C > C/C > D/D > C/D$. Mutual cooperation with a human partner was associated with consistent activation in brain areas that have been linked with reward processing: NAc, OFC and the rostral anterior cingulate cortex. OFC activation was also found when cooperating with the computer. However, a direct comparison between

playing humans versus playing the computer was not calculated in this study (but see [91]).

In a recent fMRI experiment, a simple game was played (building patterns in a 5×5 array) alone or interactive with a partner outside the scanner [32]. In the partner situation, the subject was instructed to play either cooperatively (helping to build a pattern) or competitively (blocking the partner). For the social > alone game the authors found activation in a fronto-parietal network, which they attributed to the higher executive load associated with an interactive game. Furthermore, both conditions activated the anterior insula (see Fig. 4). This region is rich in vegetative input and has been related to mainly negative emotions like disgust [27,86] or pain [30,85]. Comparing cooperation and competition against each other distinct regions of the brain were active. Notably, cooperation was associated with an activation of the left medial OFC. This was interpreted by the authors as evidence for positive reward value of cooperation. In contrast, competition was associated with increased activation in the inferior parietal cortex and the (dorsal) medial prefrontal cortex. According to the authors the inferior parietal cortex can be attributed to the sense of agency, or rather the distinction between self and others. The medial PFC activation was linked to similar Theory of Mind (ToM) studies in which the mental state of others had to be represented [41]. However, it has to be noted that activation was around 3 cm more dorsal than the medial PFC activation found by typical TOM studies or the cooperation studies by McCabe et al. [67].

All three studies so far described used “online” designs. Singer et al. used a different, but very clever design [99]. Subjects first played the Prisoner’s dilemma with human partners acting as defectors or cooperators—at least so they were told. Later, subjects were presented with photos of the faces of their game partners and with other (neutral) faces while they were scanned. The task was to judge gender. The hypothesis was that former social interaction partners would elicit different activations according to (i) their status as defectors or cooperators and (ii) their degree of moral responsibility. Moral responsibility was manipulated by introducing half of the partners playing on their own thus being responsible for their moves whereas the other half was said to follow a given response sequence thus being not responsible. There was more activation for cooperator than for defector faces (compared to neutral faces) with the most robust findings in the amygdala, the NAc, the lateral OFC, the insula, the fusiform gyrus and the superior temporal sulcus (STS). These findings are well consistent with general models of social cognition [2]. Notably, they include all central reward areas as described above as well as the insula. An effect of moral responsibility (intentional > non-intentional—this comparison controls for the receipt of money) was found in the ventral striatum and the bilateral lateral OFC for cooperators. Defector faces elicited more activation only in one region, namely the medial OFC. The pattern of OFC activation is thus in contrast to the above mentioned idea that medial OFC processes reward and lateral OFC rather punishment. In general, activation to

defectors were weaker than to cooperators, like in Rilling et al. [90].

In a follow up study with the Prisoner’s dilemma, Rilling et al. also could demonstrate a relation of the reward system with social interaction: in both ventromedial prefrontal cortex and ventral striatum, reciprocated and unreciprocated cooperation were associated with positive and negative BOLD responses, respectively. The authors interpret their results as consistent with the hypothesis that mesencephalic dopamine projection sites carry information about errors in reward prediction that allow to learn who can and cannot be trusted to reciprocate favours [92].

The comparison human versus computer was explicitly addressed in a positron emission tomography (PET) study with a computerized version of the competitive game ‘stone, paper, scissors’ in an online-design [42]. The game was played under three different experimental conditions. In the mentalizing condition subjects believed they were playing against a human and thus adopted what Dennett describes as ‘the intentional stance’ [34], which is to treat a system as an agent, attributing to it beliefs and goals. In the comparison condition, volunteers believed they were playing against a computer using a predetermined, rule-based strategy, and thus treated their opponent not as an agent but as a machine. In fact, in both instances during the critical scanning window they played against a random sequence. The only difference between the conditions was the attitude, or ‘stance’, adopted by the volunteer. In a third condition, the volunteers were told they were playing against a random sequence that acted as a low-level control. The main comparison of the mentalizing condition versus rule solving showed only one region of significant activation—the anterior paracingulate cortex bilaterally. The location of this region closely matched the activation found in the other on-line study by [67].

To summarize, in several studies it has been shown that cooperative social interaction does activate the core structures of the reward circuitry whereas non-cooperative behaviour did not. Furthermore, the insula has been found to be activated in an interactive game. The most consistently activated structure in social interaction paradigms was the medial prefrontal cortex, which was found mainly to be active when playing against humans but not against computers. We will now discuss this structure in more detail.

2.8. Medial prefrontal cortex and mentalizing

The medial prefrontal cortex has been found to be activated in a variety of paradigms, like Theory of Mind (review in [41]), moral decision-making (reviews in [46,72]), different kinds of self-monitoring and self-reflection (review in [77]) as well as in the experience of emotion and during autonomic arousal (reviewed in [41,77]). The medial PFC belongs – together with the posterior cingulate and the temporal cortices – to a set of regions that has been dubbed the “default mode” of brain function [49,50]. The medial prefrontal cortex is ‘deactivated’ – more precisely shows a signal

decrease relative to the control or resting condition – when a demanding task is performed. In a typical Theory of Mind task, for example the physical control condition will induce a signal decrease in the medial prefrontal cortex relative to rest whereas the Theory of Mind condition will induce no signal change and thus “activates” relative to the control condition.

Gallagher and Frith [41,42] have argued that the anterior paracingulate cortex (aPCC) – a subregion of the medial PFC – is specifically involved in the process of mentalizing. The aPCC is the most anterior part of the paracingulate cortex, where it lies anterior to the genu of the corpus callosum and the anterior cingulate cortex proper (Fig. 4). The paracingulate cortex (approximately corresponding to BA 32) is often considered to be part of the ACC that incorporates the cytoarchitecturally defined Brodmann areas 24, 25 and 33. There is some evidence that the ACC has undergone some recent evolutionary change. First, it contains an unusual type of projection neuron (spindle cell) found in human and in some other higher primates (pongids and hominids) but not in monkeys [76]. Furthermore, in humans these cells are not present at birth, but first appear at around 4 months of age [4]. In the human brain BA32 often extends anteriorly into the paracingulate gyrus [4]. This is dependent on the presence or absence of a paracingulate sulcus, the incidence of which is approximately 50% [4]. The presence of a paracingulate sulcus indicates that this frontal region is likely to be located on a gyral crown and this increased cortical folding might be indicative of a progressive evolution of this region in humans [111]. However, BA32 has been described cytoarchitecturally as a cingulofrontal transition area [36] and therefore anatomically (and speculatively functionally) distinct from the ACC proper. As Gallagher and Frith [41] note, it remains to be seen whether the recent evolutionary changes observed in the ACC are relevant to the more anterior region of the medial frontal lobe, where activations associated with mentalizing are observed.

It is of notable interest that activation of the medial prefrontal cortex has been described also in studies investigating moral decision-making [47,48,53] and evaluating social transgression [15]. If we ask for the reason, why a capacity of mentalizing should evolve, we have to take into account the social nature of humans. Mentalizing or the ability to read the mental state of others will only develop if subjects commu-

nicate with each other and transfer meaning. Based on such evolutionary considerations, we therefore hypothesized that it is not mentalizing per se, but mentalizing in the context of social interaction, which activates the aPCC. From this hypothesis it follows that aPCC activation should correlate with the degree of social interaction in ToM tasks. Recently, we were able to demonstrate this in an off-line design using cartoons [109] (Fig. 6). In this experiment, cartoons were used. In order to find the correct ending of the story subjects had to read the mind of characters in the mentalizing task (control condition was a cartoon story involving physical causality). The manipulation of social interaction in a mentalizing context was achieved by constructing stories involving simple intentions directed at objects without any social interaction (A), prospective intentions that were directed at future social interaction (B) and communicative intentions (C) that is intentions in social interaction involving communication. As hypothesized, we found the aPCC activated in $C > B > A$ compared to the physical causality cartoons.

That communicative intentions do activate the anterior paracingulate cortex (and the temporal poles) has also been shown in a recent experiment involving two different sensory channels (eye gaze and calling one’s name) [58]. Following the ideas of Leslie of a decoupling mechanism, we further propose an extension of her decoupling hypothesis. Leslie argued that mental representations are decoupled from the physical world so that they are no longer subject to normal input–output relations [66]. We suggest a more specific decoupling function of the anterior PCC, namely that it is active when the task is to keep two or more interacting minds separate in one’s own mind. Thus, decoupling may also play an important role in situations where mental self–other distinctions are required. This can explain why the aPCC is also found in self-monitoring or reflecting tasks.

As already described, medial PFC activation is mainly found in interaction with humans but not with computers. However, a recent fMRI study comparing two types of games—prisoner’s dilemma (see above) and Ultimatum Game (see below) describe some perplexing findings [91]. For both games, activations were detected in two of the three classic ToM areas, i.e. in the anterior paracingulate cortex and posterior superior sulcus. Although both regions showed stronger responses to human partners in both

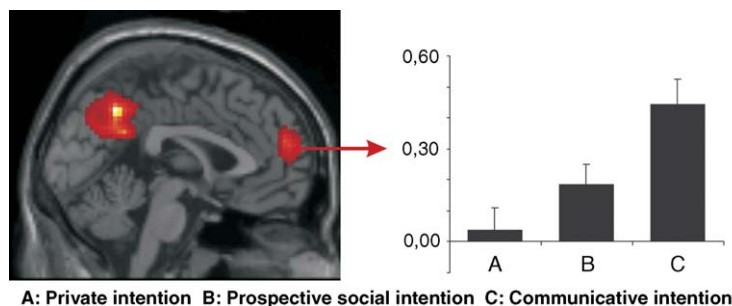


Fig. 6. Anterior paracingulate cortex activation varies with the degree of social interaction in a Theory of Mind design (adapted from [109]).

games, they were also activated when playing against the computer. This suggests that either this neural system can also be activated by reasoning about the unobservable states of non-human systems, or that participants imbue their computer partners with human attributes. The authors propose that subjects might have reasoned more about the ‘computer’s mind’ because in the computer trials they always encountered the same symbol for the computer whereas in the human condition they interacted with 10 different human partners. Thus, participants may perceive the computer trials as an iterated interaction with a single entity, as opposed to a series of single shot interactions, as for the human trials. The findings are consistent with the idea that an intentional stance can be held against nearly any object [34]. It further points to the difficulty of inferring conclusions about mental states from brain activations. People may think in patterns that are not stimulus bound.

In summary, a specific function of the anterior paracingulate cortex is associated with mentalizing, i.e. the attribution of mental states others and or oneself. The representation of social interaction does modulate its activity considerably.

We will now turn to the neural processes that are associated with processes that pertain when people do not cooperate.

2.9. Unfairness and punishment

In order to judge social interaction in real life different mechanisms must exist which evaluate actions of others quick and immediately in order to be able to react appropriately according to an ever changing environment. In a recent study based on game theory, Sanfey et al. investigated the neural correlates of experiencing unfairness in the Ultimatum Game (UG) [96]. In the UG two subjects are paired, and the first player (proposer) is provisionally allotted a divisible “pie” (usually money). The proposer then offers a portion of the pie to a second person (responder). The responder, knowing both the offer and the total amount of the pie, then has the opportunity to either accept or reject the proposer’s offer. If the responder accepts, he or she receives the amount offered and the proposer receives the remainder (the pie minus the offer). If the responder rejects the offer, then neither player receives anything. In either case, the game ends; the two subjects receive their winnings and depart. Players typically receive payments in cash and remain anonymous to other players. According to the canonical assumption of economics all participants should try to maximize their income. Responders, faced with a choice between zero and a positive payoff should accept any positive offer. Knowing this, proposers should offer the smallest non-zero amount possible. However, in every experiment yet conducted the vast majority of subjects violated this prediction, and this is also valid for non-western societies [54].

Sanfey et al. scanned responders while they were playing the Ultimatum Game against human partners or a computer with a pie of US\$ 10. Behavioural results were very similar to those typically found in Ultimatum Game experiments.

Participants accepted all fair offers (US\$ 5), with decreasing acceptance rates as the offers became less fair. Unfair offers of US\$ 2 and 1 made by human partners were rejected at a significantly higher rate than those offers made by a computer. Comparing unfair with fair offers in humans, activation of the insula, the anterior cingulate cortex and the dorsolateral PFC (DLPFC) was found. The insula is involved in the representation of vegetative states and the experience of pain [30,85] as well in the evaluation and representation of specific negative emotional states [23] in particular anger and disgust [27,86]. The activation in the right insula was correlated with the rejection rate. Furthermore, for unfair offers DLPFC activity was higher when unfair offers were accepted. The DLPFC has been linked to cognitive processes such as goal maintenance and executive control [71]. The authors interpret these findings as evidence for the insula and the DLPFC as the “twin demands” of the Ultimatum Game task, with the emotional goal of resisting unfairness and the cognitive goal of accumulating money, respectively.

In a recent study of our group, we investigated what happens when a subject has a 60% chance to win money, but does not receive it. Omission of reward in such a context was associated with activation of the right anterior insula and the right ventrolateral prefrontal cortex. We interpret these findings as neural correlates of the subjective feeling of frustration [1].

Another fMRI study showed that the insula does not only react to unfairness a person experiences itself but also to physical pain that is applied to a significant other. Singer et al. [100] scanned subjects while they could see their partner outside the scanner. Painful stimuli were applied either to the subject or the subject’s partner. Common activations for self-pain and other-pain were found in the anterior insula as well as the rostral ACC.

A study on social exclusion showed that the dorsal anterior cingulate cortex and another part of the prefrontal cortex, namely the ventrolateral PFC may serve as twins in regulating the experience of being socially separated by other people [37]. In an fMRI study, subjects were scanned while playing a virtual ball tossing game with people on the Internet—at least so they were told. There were two conditions in which they were excluded from the game: intentional (by the other players) or unintentional (by pretended technical difficulties). The right anterior insula and the anterior cingulate cortex were more active during exclusion than during inclusion. However, only ACC activity was correlated positively with self-reported distress. The right ventral prefrontal cortex (RVVFC) was active during exclusion and correlated negatively with self-reported distress. ACC changes mediated the RVVFC-distress correlation, suggesting that RVVFC regulates the distress of social exclusion by disrupting ACC activity.

From these studies it can be concluded that the anterior insula (and possibly the ACC) do serve important functions experiencing unfairness, frustration, moral disgust and separation. It will be interesting to see if further studies can corroborate the idea that the emotional (anterior insula) reaction

can be equated with a moral sense for unfairness whereas the cognitive areas of the DLPFC serve utility maximizing goals. In this respect, a recent study on moral decision-making is interesting. This study found that difficult decisions in moral dilemmas recruit the anterior prefrontal cortex (BA 10) when utilitarian decisions are made, i.e. decisions that maximize the number of lives saved, even if these decisions have to be made against intuitive moral feelings [47].

The two studies mentioned have investigated the responder of unfairness or pain. But what happens in the brain of the punisher? A recent PET-study by de Quervain et al. has investigated this question for the case of altruistic punishment [31]. In this study, a trust game was played. The general logic of the game is that player A gives money to player B, who is free to multiply this investment (and share the profit) or keeps the money for himself. Player B could be punished (loses money units) for non-cooperation in different ways: when player A decides to punish, punishment is intentional (non-intentional punishment is achieved by a random procedure). When player A has to pay for punishment it is costly (=not free): the more A pays for punishment, the more B will lose. If punishment is not effective, i.e. if B does not lose money, it is symbolic. Thus, there were four conditions of punishment: “intentional and costly” (IC), “intentional and free” (IF), “intentional and symbolic” (IS) and “non-intentional and costly” (NC).

The first main result was an activation of the dorsal striatum (caudate nucleus) only for intentional and effective punishment, i.e. only for IC and IF, but not IS and NC. Moreover, subjects with stronger activations in the dorsal striatum were willing to incur greater costs in order to punish, i.e. showed a positive correlation between activation in the dorsal striatum and the amount invested in punishment. The dorsal striatum has been linked to the processing of rewards that accrue as a result of goal-directed actions in single cell recordings in monkeys as well as in neuroimaging studies [80]. The authors interpret this activation as anticipated satisfaction for the punishing of norm violators. Furthermore, they found activation of the medial orbitofrontal (BA 11) and ventromedial (BA 10) PFC in the IC condition compared with the IF condition. BA 10 has been associated with the integration of two or more separate cognitive operations in the pursuit of higher behavioural goals [88]. The orbitomedial PFC is involved in difficult choices that require the coding of reward value [6,39]. Therefore, the authors conclude that these activations also provide indirect support for the hypothesis that punishing defectors involves satisfaction, because if that were not the case, no benefits would have to be weighed against the costs of punishing and no integration would have to take place. Although punishment in this experiment is altruistic from an evolutionary point of view, because it involves costs for the punisher, the neural activations give evidence that it is not psychologically altruistic.

In summary, three main results emerge from these studies. First, the anterior insula (and the anterior cingulate) are involved when people experience negatively rated social events like being treated unfair, being frustrated, being

socially excluded or seeing other people suffer. Thus, mental pain seems related to bodily pain in neuroanatomical terms. Second, these emotional reactions can be influenced and controlled by regions of the prefrontal cortex. Acting against one’s gut feelings is associated with the activation of prefrontal structures. Third, neuroscience points to the fact that executing punishment measures is related to experiencing reward.

2.10. Implications and conclusion

We hope to have shown convincingly that it has become possible today to measure thoughts and feelings at least in an indirect way in economic contexts. This will help to understand better what motivates people in their actions. Now what does follow from that? We might distinguish between possible practical applications and more general, theoretical implications. Let us first turn to the practical issues.

Public discussion of several of the aforementioned studies has boosted the idea of neuromarketing. The idea of neuromarketing is that understanding the human brain will help marketing in devising better methods to sell products. For example, it might be argued that it is now possible to distinguish whether people only like products or whether they want them. Furthermore, if people tend to buy products that they want, which seems a reasonable assumption, than it might indeed be possible to gain information about the likelihood that people will buy them. Recently, it has been shown that brand information does considerably influence brain activation. Comparing informed to blinded tasting of soft-drinks results in increased activation in the right DLPFC as well as the bilateral hippocampus for the Coke, but not the Pepsi condition. If only sensory cues are available (blinded tasting) the ventromedial prefrontal cortex activation correlated with subjective preference. The authors therefore conclude that brand information does influence subjects expressed preferences through the DLPFC and the hippocampus [69]. In another study, subjects had to choose a brand from two options (coffee brands for women, beer brands for men) [35]. It was found that the first choice brand was special: it was associated with a decreased activation in the working memory network, including the DLPFC and increased activation of the ventromedial prefrontal cortex. The authors interpret their findings such that for a first choice brand the reasoning chain is less important than the emotional chain for decision according to the model put forward by Damasio and Bechara [11].

Although the two studies mentioned do not converge on the exact role of the prefrontal regions in processing brand information, they demonstrate that neuroscientific methods can provide information about consumer preferences or, eventually, buying behaviour in real life. The latter, however, has not been proven yet. Nonetheless, if it will do so effectively, we might predict that these methods will be used in marketing research. This has raised worries about ethical issues. Do we have to fear manipulation of consumers? Or is

this a big chance for big business? Certainly, the public debate raises fears (and hopes) that generally are greater than justified. Although it is true that any technique can be abused we have to bear certain facts in mind. First, consumer behaviour is influenced by many factors, which cannot be manipulated in the lab. Second, it is unlikely that large cohorts of people will be studied with neuroimaging, which may be necessary to get information that goes beyond basic research and can be used for applied business research. Third, due to methodological limitations, it is not possible to measure brain activity of people if they do not want to. Fourth, the ethical problem is not specific to neuroscience but a question that pertains to all marketing research. The only way to soothe worries of manipulation is to discuss such research in public and not to hide marketing strategies. This seems to work already in standard marketing research. Would any ethically concerned person object if psychologist finds out that humour sells and that it is possible to investigate what people find funny? If we see funny ads, we all know that the fun associated with an ad is intended to influence our buying. Is it objectionable if marketing would find out with the help of brain science that people are more attracted to say cars with green lights? Well, some people may think so. It will be a matter of public debate if the use of neuromarketing methods raises too many worries so that companies will restrain from it. But then most of marketing research outside of the neuroscience domain itself will have to be objected, too.

What are the general implications of the research on reward and social interaction? As all research about mental phenomena it might deepen our understanding of ourselves. If we know how we function, we might better be able to control and influence ourselves. The basic message however is that affective and social mechanisms are always at work, like it or not. So, we have to learn to live with them but not against them. But not only practical implications are important. Some research might simply change our intuitive view of ourselves, even if our behaviour is not directly affected. For example, Greene and Cohen have argued that neuroscience might have a transformative effect on the law despite the fact that existing legal doctrine can, in principle, accommodate whatever neuroscience will tell us [45]. They argue that new neuroscience will change the law, not by undermining its current assumptions, but by transforming people's moral intuitions regarding free will and responsibility [106,108]. For example, much of what appears to be recognising eternal moral truths might be nothing else than the fast, automatic and affective appreciation of constellations that make us believe we just see right or wrong. Moreover, our Theory of Mind Mechanisms might contribute to the fact that we naturally assume that people are unmoved mover's with the ability to do otherwise whenever they want. This change in moral outlook will result not from the discovery of crucial new facts or clever new arguments, but from a new appreciation of old arguments, bolstered by vivid new illustrations provided by cognitive neuroscience. The authors predict, and recommend, a shift away from punishment aimed at retribution in favour of a more progressive,

consequentialist approach to the criminal law [45] (but compare [75]).

It is not surprising that many of the possible implications discussed do touch moral questions. Unrevealing the inner workings of the mind immediately leads to question of ethical relevance, as, for example social interaction is inherently morally coloured. Therefore, we should be careful not to abuse the authority neuroscience currently has in our society. This does not only mean to avoid harm. It also means that we should be aware of our responsibility not to abuse neuroscience by inducing inappropriate fears or hopes that will never be fulfilled.

References

- [1] B. Abler, H. Walter, S. Erk, Neural correlates of frustration, *Neuroreport* 16 (2005) 669–672.
- [2] R. Adolphs, Cognitive neuroscience of human social behaviour, *Nat. Rev. Neurosci.* 4 (2003) 165–178.
- [3] I. Aharon, N. Etcoff, D. Ariely, C.F. Chabris, E. O'Connor, H.C. Breiter, Beautiful faces have variable reward value: fMRI and behavioral evidence, *Neuron* 32 (2001) 537–551.
- [4] J.M. Allman, A. Hakeem, J.M. Erwin, E. Nimchinsky, P. Hof, The anterior cingulate cortex. The evolution of an interface between emotion and cognition, *Ann. N. Y. Acad. Sci.* 935 (2001) 107–117.
- [5] A.K. Anderson, K. Christoff, I. Stappen, D. Panitz, D.G. Ghahremani, G. Glover, J.D. Gabrieli, N. Sobel, Dissociated neural representations of intensity and valence in human olfaction, *Nat. Neurosci.* 6 (2003) 196–202.
- [6] F.S. Arana, J.A. Parkinson, E. Hinton, A.J. Holland, A.M. Owen, A.C. Roberts, Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection, *J. Neurosci.* 23 (2003) 9632–9638.
- [7] B.A. Arnow, J.E. Desmond, L.L. Banner, G.H. Glover, A. Solomon, M.L. Polan, T.F. Lue, S.W. Atlas, Brain activation and sexual arousal in healthy, heterosexual males, *Brain* 125 (2002) 1014–1023.
- [8] J.A. Bargh, T.L. Chartrand, The unbearable automaticity of being, *Am. Psychol.* 54 (1999) 462–479.
- [9] M.G. Baxter, E.A. Murray, The amygdala and reward, *Nat. Rev. Neurosci.* 3 (2002) 563–573.
- [10] L. Becerra, H.C. Breiter, R. Wise, R.G. Gonzalez, D. Borsook, Reward circuitry activation by noxious thermal stimuli, *Neuron* 32 (2001) 927–946.
- [11] A. Bechara, H. Damasio, D. Tranel, A.R. Damasio, Deciding advantageously before knowing the advantageous strategy, *Science* 275 (1997) 1293–1294.
- [12] G.S. Berns, S.M. McClure, G. Pagnoni, P.R. Montague, Predictability modulates human brain response to reward, *J. Neurosci.* 21 (2001) 2793–2798.
- [13] K.C. Berridge, T.E. Robinson, What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.* 28 (1998) 309–369.
- [14] K.C. Berridge, T.E. Robinson, Parsing reward, *Trends Neurosci.* 26 (2003) 507–513.
- [15] S. Berthoz, J.L. Armony, R.J. Blair, R.J. Dolan, An fMRI study of intentional and unintentional (embarrassing) violations of social norms, *Brain* 125 (2002) 1696–1708.
- [16] S. Blaffer-Hrdy, *Mother Nature: a History of Mothers, Infants and Natural Selection*, Pantheon, New York, 1999.
- [17] A.J. Blood, R.J. Zatorre, P. Bermudez, A.C. Evans, Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions, *Nat. Neurosci.* 2 (1999) 382–387.

- [18] S. Braeutigam, S.P. Rose, S.J. Swithenby, T. Ambler, The distributed neuronal systems supporting choice-making in real-life situations: differences between men and women when choosing groceries detected using magnetoencephalography, *Eur. J. Neurosci.* 20 (2004) 293–302.
- [19] H.C. Breiter, I. Aharon, D. Kahneman, A. Dale, P. Shizgal, Functional imaging of neural responses to expectancy and experience of monetary gains and losses, *Neuron* 30 (2001) 619–639.
- [20] H.C. Breiter, R.L. Gollub, R.M. Weisskoff, D.N. Kennedy, N. Makris, J.D. Berke, J.M. Goodman, H.L. Kantor, D.R. Gastfriend, J.P. Riorden, R.T. Mathew, B.R. Rosen, S.E. Hyman, Acute effects of cocaine on human brain activity and emotion, *Neuron* 19 (1997) 591–611.
- [21] L. Brothers, The social brain: a project for integrating primate behavior and neurophysiology in a new domain, *Concepts Neurosci.* 1 (1990) 27–51.
- [22] J.T. Cacioppo, G.G. Gerntson, R. Adolphs, C.S. Carter, R.J. Davidson, M.K. McClintock, B.S. McEwen, M.J. Meaney, D.L. Schacter, E.M. Sternberg, S.S. Suomi, S.E. Taylor (Eds.), *Foundations in Social Neuroscience*, MIT Press, Cambridge, MA, 2002.
- [23] A.J. Calder, A.D. Lawrence, A.W. Young, Neuropsychology of fear and loathing, *Nat. Rev. Neurosci.* 2 (2001) 352–363.
- [24] C.F. Camerer, *Behavioral Game Theory Experiments in Strategic Interaction*, Princeton University Press, 2003.
- [25] C.F. Camerer, G. Loewenstein, D. Prelec, Neuroeconomics: how neuroscience can inform economics, *J. Econ. Lit.* 43 (2005) 9–64.
- [26] A.R. Damasio, H. Damasio, Y. Christen, (Eds.) *Neurobiology of Decision-Making*, Springer, New York, 1996.
- [27] A.R. Damasio, T.J. Grabowski, A. Bechara, H. Damasio, L.L. Ponto, J. Parvizi, R.D. Hichwa, Subcortical and cortical brain activity during the feeling of self-generated emotions, *Nat. Neurosci.* 3 (2000) 1049–1056.
- [28] C. Darwin, *The Descent of Man, and Selection in Relation to Sex*, vol. 2, John Murray, London, 1871.
- [29] R.J. Davidson, H. Abercrombie, J.B. Nitschke, K. Putnam, Regional brain function, emotion and disorders of emotion, *Curr. Opin. Neurobiol.* 9 (1999) 228–234.
- [30] K.D. Davis, The neural circuitry of pain as explored with functional MRI, *Neurol. Res.* 22 (2000) 313–317.
- [31] D.J. de Quervain, U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, E. Fehr, The neural basis of altruistic punishment, *Science* 305 (2004) 1254–1258.
- [32] J. Decety, P.L. Jackson, J.A. Sommerville, T. Chaminade, A.N. Meltzoff, The neural bases of cooperation and competition: an fMRI investigation, *Neuroimage* 23 (2004) 744–751.
- [33] M.R. Delgado, L.E. Nystrom, C. Fissell, D.C. Noll, J.A. Fiez, Tracking the hemodynamic responses to reward and punishment in the striatum, *J. Neurophysiol.* 84 (2000) 3072–3077.
- [34] D.C. Dennett, *The Intentional Stance*, MIT, Cambridge, MA, 1987.
- [35] M. Deppe, W. Schwindt, H. Kugel, H. Plassmann, P. Kenning, Non-linear responses within the medial prefrontal cortex when specific implicit information influences economic decision-making, *J. Neuroimaging* 15 (2005) 171–182.
- [36] O. Devinsky, F. Putnam, J. Grafman, E. Bromfield, W.H. Theodore, Contributions of anterior cingulate cortex to behavior, *Brain* 118 (1995) 279–306.
- [37] N.I. Eisenberger, M.D. Lieberman, K.D. Williams, Does rejection hurt? An FMRI study of social exclusion, *Science* 302 (2003) 290–292.
- [38] R. Elliott, K.J. Friston, R.J. Dolan, Dissociable neural responses in human reward systems, *J. Neurosci.* 20 (2000) 6159–6165.
- [39] R. Elliott, J.L. Newman, O.A. Longe, J.F. Deakin, Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study, *J. Neurosci.* 23 (2003) 303–307.
- [40] S. Erk, M. Spitzer, A.P. Wunderlich, L. Galley, H. Walter, Cultural objects modulate reward circuitry, *Neuroreport* 13 (2002) 2499–2503.
- [41] H.L. Gallagher, C.D. Frith, Functional imaging of 'Theory of Mind', *Trends Cogn. Sci.* 7 (2003) 77–83.
- [42] H.L. Gallagher, A.I. Jack, A. Roepstorff, C.D. Frith, Imaging the intentional stance in a competitive game, *Neuroimage* 16 (2002) 814–821.
- [43] J.A. Gottfried, J. O'Doherty, R.J. Dolan, Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging, *J. Neurosci.* 22 (2002) 10829–10837.
- [44] J.A. Gottfried, J. O'Doherty, R.J. Dolan, Encoding predictive reward value in human amygdala and orbitofrontal cortex, *Science* 301 (2003) 1104–1107.
- [45] J. Greene, J. Cohen, For the law, neuroscience changes nothing and everything, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359 (2004) 1775–1785.
- [46] J. Greene, J. Haidt, How (and where) does moral judgment work? *Trends Cogn. Sci.* 6 (2002) 517–523.
- [47] J.D. Greene, L.E. Nystrom, A.D. Engell, J.M. Darley, J.D. Cohen, The neural bases of cognitive conflict and control in moral judgment, *Neuron* 44 (2004) 389–400.
- [48] J.D. Greene, R.B. Sommerville, L.E. Nystrom, J.M. Darley, J.D. Cohen, An fMRI investigation of emotional engagement in moral judgment, *Science* 293 (2001) 2105–2108.
- [49] D.A. Gusnard, E. Akbudak, G.L. Shulman, M.E. Raichle, Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function, *Proc. Natl. Acad. Sci. U.S.A.* 98 (2001) 4259–4264.
- [50] D.A. Gusnard, M.E. Raichle, Searching for a baseline: functional imaging and the resting human brain, *Nat. Rev. Neurosci.* 2 (2001) 685–694.
- [51] S. Hamann, H. Mao, Positive and negative emotional verbal stimuli elicit activity in the left amygdala, *Neuroreport* 13 (2002) 15–19.
- [52] M. Harris, *Cows, Pigs, Wars and Witches. The Riddles of Culture*, Random House, New York, 1974.
- [53] H.R. Heekeren, I. Wartenburger, H. Schmidt, H.P. Schwintowski, A. Villringer, An fMRI study of simple ethical decision-making, *Neuroreport* 14 (2003) 1215–1219.
- [54] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, K. Hill, F. Gil-White, M. Gurven, F. Marlowe, J.Q. Patton, N. Smith, D. Tracer, 'Economic man' in cross-cultural perspective: behavioral experiments in 15 small-scale societies, *Behav. Brain Sci.*, in press.
- [55] J.R. Hollerman, W. Schultz, Dopamine neurons report an error in the temporal prediction of reward during learning, *Nat. Neurosci.* 1 (1998) 304–309.
- [56] D.W. Hommer, B. Knutson, G.W. Fong, S. Bennett, C.M. Adams, J.L. Varner, Amygdalar recruitment during anticipation of monetary rewards: an event-related fMRI study, *Ann. N. Y. Acad. Sci.* 985 (2003) 476–478.
- [57] J. Jensen, A.R. McIntosh, A.P. Crawley, D.J. Mikulis, G. Remington, S. Kapur, Direct activation of the ventral striatum in anticipation of aversive stimuli, *Neuron* 40 (2003) 1251–1257.
- [58] K.K. Kampe, C.D. Frith, U. Frith, "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality, *J. Neurosci.* 23 (2003) 5258–5263.
- [59] B. Knutson, C.M. Adams, G.W. Fong, D. Hommer, Anticipation of increasing monetary reward selectively recruits nucleus accumbens, *J. Neurosci.* 21 (2001) RC159.
- [60] B. Knutson, G.W. Fong, C.M. Adams, J.L. Varner, D. Hommer, Dissociation of reward anticipation and outcome with event-related fMRI, *Neuroreport* 12 (2001) 3683–3687.
- [61] B. Knutson, G.W. Fong, S.M. Bennett, C.M. Adams, D. Hommer, A region of mesial prefrontal cortex tracks monetarily rewarding

- outcomes: characterization with rapid event-related fMRI, *Neuroimage* 18 (2003) 263–272.
- [62] B. Knutson, A. Westdorp, E. Kaiser, D. Hommer, fMRI visualization of brain activity during a monetary incentive delay task, *Neuroimage* 12 (2000) 20–27.
- [63] M.L. Kringelbach, J. O'Doherty, E.T. Rolls, C. Andrews, Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness, *Cereb. Cortex* 13 (2003) 1064–1071.
- [64] M.L. Kringelbach, E.T. Rolls, The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology, *Prog. Neurobiol.* 72 (2004) 341–372.
- [65] J. LeDoux, The emotional brain, fear, and the amygdala, *Cell Mol. Neurobiol.* 23 (2003) 727–738.
- [66] A.M. Leslie, Pretending and believing: issues in the theory of ToMM, *Cognition* 50 (1994) 211–238.
- [67] K. McCabe, D. Houser, L. Ryan, V. Smith, T. Trouard, A functional imaging study of cooperation in two-person reciprocal exchange, *Proc. Natl. Acad. Sci. U.S.A.* 98 (2001) 11832–11835.
- [68] S.M. McClure, G.S. Berns, P.R. Montague, Temporal prediction errors in a passive learning task activate human striatum, *Neuron* 38 (2003) 339–346.
- [69] S.M. McClure, J. Li, D. Tomlin, K.S. Cypert, L.M. Montague, P.R. Montague, Neural correlates of behavioral preference for culturally familiar drinks, *Neuron* 44 (2004) 379–387.
- [70] S.M. McClure, M.K. York, P.R. Montague, The neural substrates of reward processing in humans: the modern role of fMRI, *Neuroscientist* 10 (2004) 260–268.
- [71] E.K. Miller, J.D. Cohen, An integrative theory of prefrontal cortex function, *Annu. Rev. Neurosci.* 24 (2001) 167–202.
- [72] J. Moll, R. de Oliveira-Souza, P.J. Eslinger, Morals and the human brain: a working model, *Neuroreport* 14 (2003) 299–305.
- [73] P.R. Montague, G.S. Berns, J.D. Cohen, S.M. McClure, G. Pagnoni, M. Dhamala, M.C. Wiest, I. Karpov, R.D. King, N. Apple, R.E. Fisher, Hyperscanning: simultaneous fMRI during linked social interactions, *Neuroimage* 16 (2002) 1159–1164.
- [74] D. Morgan, K.A. Grant, H.D. Gage, R.H. Mach, J.R. Kaplan, O. Prioleau, S.H. Nader, N. Buchheimer, R.L. Ehrenkauf, M.A. Nader, Social dominance in monkeys: dopamine D2 receptors and cocaine self-administration, *Nat. Neurosci.* 5 (2002) 169–174.
- [75] S. Morse, New neuroscience, old problems. Neuroscience and the law, in: B.B. Garland, M.S. Frankel, K. Schaefer, C. Liu (Eds.), *Rain, Mind, and the Scales of Justice*, Dana Press, New York, in press.
- [76] E.A. Nimchinsky, E. Gilissen, J.M. Allman, D.P. Perl, J.M. Erwin, P.R. Hof, A neuronal morphologic type unique to humans and great apes, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 5268–5273.
- [77] G. Northoff, F. Bermpohl, Cortical midline structures and the self, *Trends Cogn. Sci.* 8 (2004) 102–107.
- [78] J. O'Doherty, H. Critchley, R. Deichmann, R.J. Dolan, Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices, *J. Neurosci.* 23 (2003) 7931–7939.
- [79] J.P. O'Doherty, P. Dayan, K. Friston, H. Critchley, R.J. Dolan, Temporal difference models and reward-related learning in the human brain, *Neuron* 38 (2003) 329–337.
- [80] J. O'Doherty, P. Dayan, J. Schultz, R. Deichmann, K. Friston, R.J. Dolan, Dissociable roles of ventral and dorsal striatum in instrumental conditioning, *Science* 304 (2004) 452–454.
- [81] J. O'Doherty, M.L. Kringelbach, E.T. Rolls, J. Hornak, C. Andrews, Abstract reward and punishment representations in the human orbitofrontal cortex, *Nat. Neurosci.* 4 (2001) 95–102.
- [82] J.P. O'Doherty, R. Deichmann, H.D. Critchley, R.J. Dolan, Neural responses during anticipation of a primary taste reward, *Neuron* 33 (2002) 815–826.
- [83] J. Panksepp, *Affective Neuroscience: the Foundations of Human and Animal Emotions*, Oxford University Press, New York, Oxford, 1998.
- [84] M. Petrie, Improved growth and survival of offspring of peacocks with more elaborate trains, *Nature* 371 (1994) 598–599.
- [85] R. Peyron, B. Laurent, L. Garcia-Larrea, Functional imaging of brain responses to pain. A review and meta-analysis, *Neurophysiol. Clin.* 30 (2000) 263–288.
- [86] M.L. Phillips, A.W. Young, C. Senior, M. Brammer, C. Andrew, A.J. Calder, E.T. Bullmore, D.I. Perrett, D. Rowland, S.C. Williams, J.A. Gray, A.S. David, A specific neural substrate for perceiving facial expressions of disgust, *Nature* 389 (1997) 495–498.
- [87] N. Ramnani, R.C. Miall, Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation, *Cereb. Cortex* 13 (2003) 318–327.
- [88] N. Ramnani, A.M. Owen, Anterior prefrontal cortex: insights into function from anatomy and neuroimaging, *Nat. Rev. Neurosci.* 5 (2004) 184–194.
- [89] J.A. Richeson, A.A. Baird, H.L. Gordon, T.F. Heatherton, C.L. Wyland, S. Trawalter, J.N. Shelton, An fMRI investigation of the impact of interracial contact on executive function, *Nat. Neurosci.* 6 (2003) 1323–1328.
- [90] J. Rilling, D. Gutman, T. Zeh, G. Pagnoni, G. Berns, C. Kilts, A neural basis for social cooperation, *Neuron* 35 (2002) 395–405.
- [91] J.K. Rilling, A.G. Sanfey, J.A. Aronson, L.E. Nystrom, J.D. Cohen, The neural correlates of Theory of Mind within interpersonal interactions, *Neuroimage* 22 (2004) 1694–1703.
- [92] J.K. Rilling, A.G. Sanfey, J.A. Aronson, L.E. Nystrom, J.D. Cohen, Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways, *Neuroreport* 15 (2004) 2539–2543.
- [93] T.E. Robinson, K.C. Berridge, The neural basis of drug craving: an incentive-sensitization theory of addiction, *Brain Res. Brain Res. Rev.* 18 (1993) 247–291.
- [94] E.T. Rolls, The orbitofrontal cortex and reward, *Cereb. Cortex* 10 (2000) 284–294.
- [95] E.T. Rolls, J. O'Doherty, M.L. Kringelbach, S. Francis, R. Bowtell, F. McGlone, Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices, *Cereb. Cortex* 13 (2003) 308–317.
- [96] A.G. Sanfey, J.K. Rilling, J.A. Aronson, L.E. Nystrom, J.D. Cohen, The neural basis of economic decision-making in the Ultimatum Game, *Science* 300 (2003) 1755–1758.
- [97] W. Schultz, Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology, *Curr. Opin. Neurobiol.* 14 (2004) 139–147.
- [98] W. Schultz, P. Apicella, E. Scarnati, T. Ljungberg, Neuronal activity in monkey ventral striatum related to the expectation of reward, *J. Neurosci.* 12 (1992) 4595–4610.
- [99] T. Singer, S.J. Kiebel, J.S. Winston, R.J. Dolan, C.D. Frith, Brain responses to the acquired moral status of faces, *Neuron* 41 (2004) 653–662.
- [100] T. Singer, B. Seymour, J. O'Doherty, H. Kaube, R.J. Dolan, C.D. Frith, Empathy for pain involves the affective but not sensory components of pain, *Science* 303 (2004) 1157–1162.
- [101] D.M. Small, M.D. Gregory, Y.E. Mak, D. Gitelman, M.M. Mesulam, T. Parrish, Dissociation of neural representation of intensity and affective valuation in human gustation, *Neuron* 39 (2003) 701–711.
- [102] D.M. Small, R.J. Zatorre, A. Dagher, A.C. Evans, M. Jones-Gotman, Changes in brain activity related to eating chocolate: from pleasure to aversion, *Brain* 124 (2001) 1720–1733.
- [103] M. Spitzer, Models of schizophrenia: from neuroplasticity and dopamine to psychopathology and clinical management, in: K. S. Y. Leclubier (Eds.), *Dopamine in the Pathophysiology and Treatment of Schizophrenia*, Martin Dunitz, Taylor & Francis Group, London/New York, 2003, pp. 155–175.
- [104] R.S. Sutton, A.G. Barto, Toward a modern theory of adaptive networks: expectation and prediction, *Psychol. Rev.* 88 (1981) 135–170.

- [105] L. Tremblay, W. Schultz, Relative reward preference in primate orbitofrontal cortex, *Nature* 398 (1999) 704–708.
- [106] H. Walter, *Neurophilosophy of Free Will*, MIT Press, Cambridge, MA, 2001.
- [107] H. Walter, *Funktionelle Bildgebung in Psychiatrie und Psychotherapie (Functional Neuroimaging in Psychiatry and Psychotherapy)*, Schattauer, Stuttgart, 2004.
- [108] H. Walter, Willensfreiheit, Verantwortlichkeit und Neurowissenschaft (Free will, responsibility and neuroscience), *Psychol. Rundschau* 55 (2004) 169–177.
- [109] H. Walter, M. Adenzato, A. Ciaramidaro, I. Enrici, L. Pia, B.G. Bara, Understanding intentions in social interaction: The role of the anterior paracingulate cortex, *J. Cogn. Neurosci.* 16 (2004) 1854–1863.
- [110] T. Zalla, E. Koechlin, P. Pietrini, G. Basso, P. Aquino, A. Sirigu, J. Grafman, Differential amygdala responses to winning and losing: a functional magnetic resonance imaging study in humans, *Eur. J. Neurosci.* 12 (2000) 1764–1770.
- [111] K. Zilles, E. Armstrong, A. Schleicher, H.J. Kretschmann, The human pattern of gyrification in the cerebral cortex, *Anat. Embryol. (Berl.)* 179 (1988) 173–179.
- [112] C.F. Zink, G. Pagnoni, M.E. Martin-Skurski, J.C. Chappelow, G.S. Berns, Human striatal responses to monetary reward depend on saliency, *Neuron* 42 (2004) 509–517.
- [113] C.F. Zink, G. Pagnoni, M.E. Martin, M. Dhamala, G.S. Berns, Human striatal response to salient non-rewarding stimuli, *J. Neurosci.* 23 (2003) 8092–8097.