The Neuroscience of Social Decision-Making

James K. Rilling[1,2,3,4] and Alan G. Sanfey[5,6,7]

[1]Department of Anthropology, [2]Department of Psychiatry and Behavioral Sciences,
[3]Center for Behavioral Neuroscience and [4]Yerkes National Primate Center, Emory
University, Atlanta, Georgia 30322; email: jrillin@emory.edu

[5]Donders Institute for Brain, Cognition & Behavior and [6]Behavioral Science Institute,
Radboud University Nijmegen, NL-6500 HB Nijmegen, Netherlands, and
[7]Department of Psychology, University of Arizona, Tucson, Arizona 85721

Key Words

trust, reciprocity, altruism, fairness, revenge, norms

Abstract

Given that we live in highly complex social environments, many of our
most important decisions are made in the context of social interactions.
Simple but sophisticated tasks from a branch of experimental economics
known as game theory have been used to study social decision-making
in the laboratory setting, and a variety of neuroscience methods have
been used to probe the underlying neural systems. This approach is in-
forming our knowledge of the neural mechanisms that support decisions
about trust, reciprocity, altruism, fairness, revenge, social punishment,
social norm conformity, social learning, and competition. Neural sys-
tems involved in reward and reinforcement, pain and punishment, men-
talizing, delaying gratification, and emotion regulation are commonly
recruited for social decisions. This review also highlights the role of the
prefrontal cortex in prudent social decision-making, at least when social
environments are relatively stable. In addition, recent progress has been
made in understanding the neural bases of individual variation in social
decision-making.

Contents

INTRODUCTION

The study of decision-making attempts to understand our fundamental ability to process multiple alternatives and to choose an optimal course of action. Historically, the majority of research on decision-making has examined individual decisions in which we must consider purely our own values and preferences in order to select an option. For example, experimental participants are often asked to choose between monetary gambles or to evaluate a choice-set described in terms of different attributes. However, given that we live in highly complex social environments, many of our most important decisions are made in the context of social interactions, with these decisions additionally dependent on the concomitant choices of others (Sanfey 2007). These social decisions can be defined as decisions that affect others as well as ourselves and are therefore typically informed by both self and other-regarding preferences (Fehr & Camerer 2007). We encounter these social decision scenarios on a daily basis: Should I trust this person? Should I reciprocate this person's trust? Should I treat this person fairly? Should I punish this person for treating me unfairly? Should I be deferential to this person? Whom should I learn from, model or imitate? And should I abide by social norms?

Though made in a unique context, social decisions appear to share key elements with individual decisions. Like many decisions, difficult social decisions involve psychological conflict, such as between self-interest and the interests of others as when we decide whether to help another at a personal cost. Social decisions can also involve conflict between short-term rewards and more distant, but potentially larger, rewards. Am I willing to endure the immediate costs of altruism in order to reap the long-term benefits of a sustained cooperative relationship? Finally, as with individual decisions, challenging social decisions can involve conflict between emotion and reason (Frith & Singer 2008, Sanfey et al. 2006). Indeed, both emotion and reason may provide wisdom in social decision-making. Social emotions often help us to reach more adaptive decisions than would be possible by reasoning alone (Damasio 1994, Frank 1988), as for example when guilt dissuades us from harming relationships with selfish behavior. Conversely, the ability to override social-emotional biases with cognitive control may also be prudent in some circumstances, as when suppressing indignation over unfair treatment by a more powerful other.

Though social decisions are undoubtedly important, the requisite interactive scenarios can be challenging to recreate in the laboratory. How then can we experimentally study the neuroscience of social decision-making? What kinds of tools are available? This article outlines the current methods that have been employed in understanding social decision-making and discusses the empirical findings that are emerging from this rapidly growing field. We focus in particular on neuroscientific investigations of these important questions and provide an

overview of what is currently understood regarding the neural basis of social decision-making. Though these research questions are often embedded within the larger field of social neuroscience, we limit our focus here to the neuroscience of decision-making within social interactions, mostly in the context of interactive games. Thus, we do not review a large body of important (and potentially relevant) work within the broader domain of social neuroscience, including the neuroscience of moral decision-making (Greene et al. 2004, Moll et al. 2005), theory of mind (Gallagher & Frith 2003, Saxe et al. 2004), face processing (Haxby et al. 2002, Todorov et al. 2008), attitudes toward outgroup members (Eberhardt 2005, Harris & Fiske 2006), and the role of the medial prefrontal cortex in social cognition in general (Lieberman 2007, Mitchell 2009).

TASKS

Simple but sophisticated tasks from experimental economics, using game theory as a framework, have been used to study social decision-making in the laboratory, and researchers have in turn employed a variety of neuroscience methods to investigate the underlying neural systems. Game theory is a collection of rigorous models attempting to understand and explain situations in which decision-makers must interact with one another (Neumann & Morgenstern 1947). It offers a rich source of both behavioral tasks and data in addition to well-specified models for the investigation of social interaction. The games used have the advantage of being easy for participants to understand, offer quite compelling social scenarios, and are relatively straightforward to adapt to neuroscientific study, all of which goes a long way toward explaining their extensive use in recent years. These tasks have been used to study several aspects of social decision-making, primarily reciprocal exchange, responses to fairness and equity, and altruism and punishment.

Reciprocal exchange has been extensively studied using the Prisoner's Dilemma (PD) and Trust games. In the Trust Game, a player (the investor) must decide how much of an endowment to invest with a partner (the trustee). Once transferred, this money is multiplied by some factor and then the trustee has the opportunity to return money to the investor but, importantly, need not return anything. If the trustee honors trust and returns money, both players end up with a higher monetary payoff than the original endowment. However, if the trustee abuses trust and keeps the entire amount, the investor takes a loss. Thus, the Trust Game models both decisions to trust and decisions to reciprocate trust. If the investor and trustee interact only once during the game, Game Theory predicts that a rational and selfish trustee will never honor the trust given by the investor. The investor, realizing this, should never place trust in the first place and so will invest zero in the transaction. Despite these grim theoretical predictions, a majority of investors do in fact send some amount of money to the trustee, often approximately half of their endowment, and this trust is generally reciprocated (Camerer 2003).

The standard PD game is similar to the Trust Game except that both players now simultaneously choose whether or not to trust each other, without knowledge of their partner's choice. In the PD game, payoffs depend on the interaction of the two choices. The largest payoff to the player occurs when he or she defects and the partner cooperates, with the worst outcome when the decisions are reversed (player cooperates while partner defects). Mutual cooperation yields a modest payoff to both players, whereas mutual defection provides a lesser amount to each. The predicted solution to the PD game is mutual defection, a worse outcome for both players than mutual cooperation, but again, in most iterations of the game, players exhibit more trust than expected, with mutual cooperation occurring about 50% of the time (Camerer 2003). Both the Trust Game and the PD game can also be played as iterated, multiple-round games, though these variants change both the optimal and actual game strategies due to the "shadow of the future" (Axelrod & Hamilton 1981), that is, the effect of potential future consequences on current choices.

PD: Prisoner's Dilemma

The Ultimatum Game (UG) is often used to examine responses to fairness. In the UG, two players must divide a sum of money, with the proposer specifying the division. The responder then has the option of accepting or rejecting this offer. If the offer is accepted, the sum is divided as proposed. If it is rejected, neither player receives anything. The UG therefore models decisions about resource allocation on the part of the proposer, as well as responses to fairness and inequity in the responder. If people are motivated purely by self-interest, the responder should accept any offer, and, knowing this, the proposer will offer the smallest nonzero amount. However, once again, this game theoretic prediction is at odds with observed behavior across a wide range of societies (Henrich et al. 2005), with both fair offers and rejections of unfair offers often observed. Thus, people's choices in the UG do not conform to a model in which decisions are driven by financial self-interest, and neuroscience has begun to offer clues as to the mechanisms underlying these decisions.

Altruism has been modeled using the Dictator Game (DG), essentially a simplified version of the UG, in which the second player is a passive recipient of the proposer's offer and therefore cannot reject it. With no material incentive to offer anything, a proposer who offers a nonzero amount is considered altruistic, and proposal magnitude reflects the degree of altruism toward the second player.

Both anonymous and nonanonymous versions of the above games have been studied with neuroimaging. Because of their interest in "pure" game play, economists have typically emphasized the importance of anonymous interactions to eliminate reputation effects or personal characteristics of partners that could bias choices. However, psychologists and neuroscientists are generally interested in these social factors and how they influence game decisions, and so they often include known partners as part of these experiments. It can also be argued that humans are evolutionarily unprepared for social interactions with completely anonymous partners, and therefore use of a more ecologically valid design is justified.

In addition to these classic game theory designs, a number of more recent studies have employed new and creative paradigms that model other aspects of social decision-making, such as social conformity (Klucharev et al. 2009), norm-abiding social behavior (Spitzer et al. 2007), revenge and altruistic punishment (de Quervain et al. 2004, Singer et al. 2006), and reputation management (Izuma et al. 2008). These approaches offer some interesting variants on the questions tackled by the standard tasks. Overall, the full complement of tasks outlined here is providing researchers with useful experimental scenarios with which to ask questions regarding the neural basis of social decision-making, and their results are discussed below.

NEUROSCIENCE METHODS

The methods that are being used to probe the neural bases of social decision-making include functional neuroimaging, the study of brain-damaged neurological patients, transcranial magnetic stimulation, pharmacologic manipulations, genetic association studies, and studies of psychiatric patients with pathological social decision-making, as well as lesion and single-cell recording studies in nonhuman primate models of human social decision-making.

The majority of current research on the neuroscience of social decision-making is derived from functional magnetic resonance imaging (fMRI) studies in which changes in cerebral blood flow are imaged as subjects play interactive social games inside the MRI scanner. Typically, computerized game paradigms are projected onto a screen in the scanner, and subjects make choices by pressing buttons in response to game scenarios. Compared with other functional neuroimaging techniques, fMRI is less invasive, less expensive, and has good spatial and temporal resolution. However, while effective in identifying brain regions that are involved in social decision-making, fMRI is less effective in

UG: Ultimatum Game

identifying brain regions that are essential for social decision-making. For this, studies of neurological patients are helpful, and tremendous insight into the neuroscience of social decision-making has been gleaned from the study of patients with damage to the ventromedial prefrontal cortex (VMPFC) (Bechara & Damasio 2005, Beer et al. 2003, Damasio 1994, Mah et al. 2004). However, these lesions often span large regions of cortex that likely involve multiple functions, which limits the specificity of structure-function mapping. Repetitive transcranial magnetic stimulation (rTMS), in which an oscillating magnetic field is used to induce electric current in the brain, enables temporary, directed disruption of cortical regions and is a useful complement to the study of neurological patients in understanding which brain regions are essential for normal social decision-making.

In addition to the above methods, pharmacological manipulations can inform our knowledge of the neurochemical basis of human social decision-making. Monoamine (e.g., serotonin), neuropeptide (e.g., oxytocin), and steroid hormone (e.g., testosterone) levels have all been experimentally manipulated and tested for effects on social decision-making in game theoretic paradigms. The density and distribution of neurochemical receptors can be imaged with positron emission tomography (PET), and individual variation in receptor patterns could in theory be linked with variation in social decision-making; however, PET ligands for many of the receptors of interest are currently unavailable. Nevertheless, individual variation in the genes that code for neuropeptide receptors such as oxytocin (OT) and vasopressin (AVP) has been linked with social decision-making (Israel et al. 2009, Knafo et al. 2008).

Many psychiatric conditions also involve deficits in social decision-making. Depressed patients often withdraw from social interactions; patients with social anxiety disorder, borderline personality disorder, and autism often incorrectly interpret social interactions; psychopaths persistently violate social norms and selfishly manipulate others; and patients with conduct disorder can exhibit inappropriate levels of aggression. Identifying the brain abnormalities underlying these disorders can therefore potentially shed light on the neural systems that mediate social decisions. Furthermore, use of these games can potentially play a valuable role in the assessment of, and intervention in, decision-making styles in these disorders.

Finally, a large body of research has examined decision-making at the cellular level in nonhuman primate models using single-cell recording (Kable & Glimcher 2009), and a subset of these studies has focused on social decision-making in particular (Klein et al. 2008, Seo et al. 2009). These studies, which normally cannot be performed in humans for obvious ethical reasons, are an important complement to the study of large-scale neural systems involved in social decision-making in humans.

Below we summarize what has been learned in applying these varied methods to the study of human social decision-making (**Figure 1a,b**).

RECIPROCAL EXCHANGE

From a comparative mammalian perspective, one remarkable feature of human social life is the extent to which we engage in the reciprocal exchange of aid with nonrelatives, since, in nonhuman animals, most altruism is directed toward genetic relatives. Although cooperation does occur among nonrelatives, particularly in social mammals such as lions, meerkats, and primates, most examples are best explained by mutualism, in which both partners gain immediate benefits from their cooperation (Clutton-Brock 2009b). For example, in wild dogs, cooperation between hunting partners can increase their per capita success in catching or defending prey (Creel & Creel 2001). Mutualism differs from reciprocal altruism, which encumbers net costs at the time assistance is provided, though these are then offset by later benefits (Trivers 1971). One significant consequence of a temporal delay between receiving and returning help is that natural selection can favor cheating (i.e., accepting but not reciprocating a favor). It may

VMPFC: ventromedial prefrontal cortex

rTMS: repetitive transcranial magnetic stimulation

Neuropeptide: a molecule composed of short chains of amino acids found in brain tissue that influences neural activity, such as oxytocin and vasopressin

Ligands: molecules that bind to another (e.g., a receptor). Radioactively tagged ligands can be used to label neurotransmitter receptors for imaging

OT: oxytocin

Borderline personality disorder: a personality disorder marked by a long-standing pattern of instability in interpersonal relationships, behavior, mood, and self-image

be this barrier to the evolution of reciprocal altruism that accounts for the limited number of documented cases among nonhuman animals. In contrast, reciprocal altruism is pervasive in human society (Clutton-Brock 2009a). Indeed, hunter-gatherers like the Kalahari Bushmen, who have been studied intensively by anthropologists because they may provide a glimpse of human nature unconfounded by recent dramatic environmental changes (Konner 2002), depend upon reciprocal food sharing for their very survival (Lee 1979).

Trust

Relationships based on reciprocal altruism are inherently unstable. Both parties may be tempted to act according to short-term self-interests by accepting but then not reciprocating a favor, and both parties may fear these same selfish impulses in their partner and can therefore be reluctant to risk placing trust. Given this instability, it has been theorized that the pervasiveness of reciprocal altruism in humans required the evolution of a suite of psychological specializations to support it (Trivers 1971). One such specialization is a willingness to take the social risk of helping another despite the possibility of nonreciprocation; in other words, a willingness to trust (#1 in **Figure 1***a*,***b***). Decisions to trust a previously unknown partner are strongly associated with general judgments of facial trustworthiness (van 't Wout & Sanfey 2008), and neuroimaging and neuropsychological studies have established that the amygdala is centrally involved in assessments of trust. Untrustworthy faces activate the amygdala in fMRI paradigms, even when the judgment is made implicitly (Winston et al. 2002), and people with amygdala lesions have deficits in the ability to appropriately judge facial trustworthiness (Adolphs et al. 1998). Thus, it is reasonable to suppose that dampening amygdala activity would increase behavioral expressions of trust. Several studies have now demonstrated that the neuropeptide oxytocin (OT) reduces amygdala activity in male subjects (Baumgartner et al. 2008,

Domes et al. 2007, Kirsch et al. 2005, Petrovic et al. 2008, Singer et al. 2008) although not in women (Domes et al. 2010), and others have shown that OT also increases behavioral expressions of trust (Baumgartner et al. 2008, Kosfeld et al. 2005). Thus, trusting another person may involve OT-mediated suppression of amygdala activity and dampening the accompanying fear of betrayal. The adult human pair-bond is a good example of a cooperative social relationship between nonrelatives that is based on trust. Therefore, it is of note that OT mediates pair-bonding in monogamous rodent species (Young et al. 2005). Similar mechanisms may be at play in humans, as intranasal OT increases positive communication during couple conflict (Ditzen et al. 2009). Thus, social attachment between unrelated adult humans may be mediated in part by oxytocin.

Of course, the neural substrate of the decision to trust is not confined to subcortical structures. Patients with lesions to VMPFC, which also include damage to the frontal pole and the anterior cingulate cortices, exhibit less trust in the Trust Game (Krajbich et al. 2009). Consistent with this finding, an fMRI study of the Trust Game reported greater activation in the frontal pole during the decision to trust as compared to the decision to reciprocate trust (Krueger et al. 2008). VMPFC patients have been said to exhibit "myopia for the future" (Bechara & Damasio 2005), and the frontal pole has been implicated in protecting long-term mental plans from immediate environmental demands (Koechlin & Hyafil 2007) and in valuing future rewards (Kable & Glimcher 2007). One hypothesis is that this region registers the long-term benefits that could emerge from a successful partnership, which can help to surmount the immediate fear of betrayal associated with deciding to trust.

Decisions to trust should also be based on inferences of others' trustworthiness—if partners do not appear trustworthy, whether assessed by facial features or by knowledge about their past behavior, we should be wary about interacting with them. Therefore, trust decisions could be expected to engage brain systems implicated in

theory of mind. Indeed, partner feedback in the PD game reliably activates several regions that have been implicated in theory of mind, including dorsomedial prefrontal cortex (DMPFC), posterior cingulate, and the temporo-parietal junction (TPJ), with each of these areas engaged more when playing with a human than a computer partner (Rilling et al. 2004a). Similarly, another study showed that DMPFC activity is high during the initial stages of building a trusting relationship but then subsides once trust has been established (Krueger et al. 2007), suggesting that this region may be involved more specifically in learning whether someone is trustworthy. This was the conclusion of a recent study that tracked social prediction errors when following a confederate's advice and found activity related to social prediction errors in DMPFC, superior temporal sulcus (STS), and TPJ (Behrens et al. 2008). Social anxiety disorder is associated with attenuated DMPFC activation during a trust game (Sripada et al. 2009), perhaps implying that social learning mechanisms are short-circuited by the limbic hyperactivity characteristic of this disorder (Etkin & Wager 2007), which may lead to prematurely judging a social stimulus as threatening.

Reciprocating Trust

As noted above, in relationships based on reciprocal altruism (#2 in **Figure 1a,b**) there is an obvious temptation to accept but then not reciprocate a favor. For example, asking for help when moving house but then not returning that favor may be beneficial in the short term but will quite likely incur long-term costs by discouraging the altruistic from granting future favors. In other words, reciprocity is important for the maintenance of relationships. Throughout the animal kingdom, the bias for immediate gratification is strong (Kagel et al. 1995), and in some cases, this bias prevents the establishment of stable, cooperative relationships with others. However, remarkably frequently, people overcome these biases. What are the motivations for doing so, and how does the human brain accomplish this? Part of the answer may be that the

short-term social reward associated with mutual cooperation can outweigh the short-term material rewards gained from cheating. That is, the subjective utility of mutual cooperation can exceed that of unilateral defection (Fehr & Camerer 2007). Functional MRI studies of human subjects engaged in Prisoner's Dilemma or related trust games have shown that reciprocated cooperation is associated with activation of two brain regions involved in reward processing, the caudate nucleus (Delgado et al. 2005; Rilling et al. 2002, 2004b) and the orbitofrontal cortex (OFC) (Rilling et al. 2002, 2004b). Moreover, the strength of response in the caudate predicts the degree of future cooperation (King-Casas et al. 2005, Rilling et al. 2002), suggesting that activation of this brain region can positively reinforce cooperation, either by rendering mutual cooperation immediately rewarding or by providing a learning signal after feedback. Either way, evolution may have effectively removed the need to delay gratification. Although the material payoff from mutual cooperation may be realized later, the social payoff can be immediate.

A recent study (Li et al. 2009) suggests that the act of being trusted may also be inherently rewarding. This study focused on activation in trustees' brains after they learned that the investor had transferred money to them. In one condition, investors were allowed to threaten trustees with a financial penalty for nonrepayment, whereas in the other condition this was not possible. VMPFC was more active when trustees received money from the investor without sanction threats as opposed to with threats, and the magnitude of this VMPFC activation predicted levels of trustee repayment (Li et al. 2009). One interpretation of these findings is that subjects anticipate keeping more money when they are not threatened with costly sanctions compared to when they are. Though there are obvious dangers in assuming that activation of reward regions implies that an act is inherently rewarding (Poldrack 2006), an alternative and more interesting possible interpretation is that the act of being trusted in the absence of sanction threats is rewarding in and of itself and

DMPFC:
dorsomedial prefrontal cortex

TPJ:
temporo-parietal junction

Tryptophan: the amino acid precursor of serotonin

Selective serotonin reuptake inhibitor (SSRI): a drug that increases serotonin transmission by blocking synaptic reuptake

Insula: an island of cortex buried within the sylvian fissure

ACC: anterior cingulate cortex

DLPFC: dorsolateral prefrontal cortex

thereby reinforces reciprocation. We presumably feel a greater social bond with someone who does not threaten retribution, and this may be a mechanism by which greed is overcome and may also explain why patients with VMPFC lesions are less trustworthy in the Trust Game (Krajbich et al. 2009).

Intriguingly, serotonin may modulate these value functions. In both humans and laboratory animals, experimental manipulation of the serotonin system demonstrates that low serotonin levels decrease the value of delayed rewards, steepening delayed reward discounting (Denk et al. 2005, Schweighofer et al. 2008). Lowering serotonin levels through tryptophan depletion decreases cooperation by second movers in the PD game, and elevating the levels, through selective serotonin reuptake inhibitor (SSRI) treatment, has the opposite effect (Tse & Bond 2002, Wood et al. 2006). These effects of serotonin may be mediated by VMPFC, as the human VMPFC is known to be modulated by serotonin (Robbins 2000), serotonin agonists increase VMPFC metabolism (Siever et al. 1999), and tryptophan depletion mimics the effects of VMPFC lesions on behavior in the UG (Crockett et al. 2008) (discussed below). Thus, serotonergic input to VMPFC may promote reciprocity through increasing the value of long-term benefits associated with mutual cooperation (Wood et al. 2006).

An alternative motivation for reciprocal behavior beyond a reward explanation is that it may be driven by the minimization of potential negative affect, primarily guilt. That is, the reason for cooperation is that we anticipate feeling guilty if we would not reciprocate generous behavior. Here again, VMPFC patients are relevant, as both qualitative observations of their social behavior (Koenigs & Tranel 2007) as well as more formal modeling based on their behavior in economic games (Krajbich et al. 2009) suggest that they have impairments in guilt. Thus, the expression of guilt, and perhaps more generally the elicitation of emotions based on imagined outcomes (Krajbich et al. 2009), can play an important role in social decision-making, and affective processing areas such as

VMPFC and insula (see below) may be involved in the neural instantiation of these processes.

If contemplation of defection elicits guilt, the decision to defect might be expected to involve conflict. Indeed, a recent study showed that breaking a promise to reciprocate trust, compared to honoring that same promise, was associated with activation in the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (DLPFC) (Baumgartner et al. 2009), regions consistently implicated in cognitive conflict and cognitive control, respectively (Botvinick et al. 2001, Miller & Cohen 2001, Pochon et al. 2008). These results suggest that for most people, breaking a promise to reciprocate requires cognitive effort, and conversely that honoring such a promise is our prepotent response bias. Thus, whether through innate, genetic predispositions or through socialization, the tendency to reciprocate altruism appears to become ingrained in our biology and overridden only with cognitive effort.

Of course, there is significant individual variation in the tendency to reciprocate altruism, and explaining this variation is an important challenge for social neuroscience. A recent study classified subjects based on social value orientation, a measure of the tendency to value the outcomes of others. The researchers found that prosocial participants, who valued the outcomes of others, exhibited more ventral striatal activity when choosing to reciprocate compared with when choosing to defect, whereas proself participants had the opposite profile (van den Bos et al. 2009). These findings suggest that reciprocating, despite its lower material payoff, may have a higher reward value for prosocials, whereas defecting may have a higher reward value for proselfs. Accordingly, prosocials did in fact reciprocate more often than proselfs did. This study also reported a difference in activation of the insula as a function of social value orientation, whereby prosocials showed stronger activation when defecting and proselfs showed stronger activation when reciprocating. Thus, defecting may be aversive to prosocials, whereas reciprocating may be aversive to proselfs (van den Bos et al. 2009).

Finally, the neural correlates of decisions about reciprocity can be altered in psychiatric disorders. For example, the anterior insula of trustees is more active in response to low as compared to high expressions of trust by investors. However, this differential insula response is lacking in patients with borderline personality disorder, which has been interpreted as suggesting that these patients, in contrast to normal controls, fail to register low levels of trust as a norm violation (King-Casas et al. 2008). In another study of the iterated PD game, the decision to defect was associated with activation in ACC and DLPFC, generally interpreted as reflecting conflict and exertion of cognitive effort, but not in individuals who scored high on a measure of psychopathic personality. These individuals also defected at higher rates (Rilling et al. 2007), and thus defection may only be difficult for those who did not score high on psychopathic personality.

Responding to Breaches of Trust

In the iterated PD game, as in life, unconditional cooperators are vulnerable to exploitation by nonreciprocators (Axelrod 1984) (#3 in **Figure 1a,b**). Thus, humans have likely evolved psychological mechanisms to detect and avoid cooperating with nonreciprocators (Cosmides & Tooby 2000), which may be related to a more generalized aversion to "free riders" (i.e., those who accept benefits without paying expected costs), as suggested by behavioral economics experiments in which people often choose to punish free riders, even when the punishment is personally costly (Fehr & Gachter 2002).

In the iterated PD game, cooperation in combination with a partner's defection (the worst outcome) is associated with activation of the anterior insular cortex, which may be a neural correlate of an aversive response to free riding (Rilling et al. 2008) or of a more generalized response to norm violations (Montague & Lohrenz 2007). The anterior insula is involved in sensing the state of the viscera (e.g., heart, lungs, gut) and is activated in response to a variety of negative social interactions, from

social exclusion (Eisenberger et al. 2003), to receiving an unfair offer in an UG (Sanfey et al. 2003), to watching a loved one receive a painful stimulus (Singer et al. 2004). Anterior insula is also responsive to physically painful stimuli, and its activity is correlated with skin conductance responses (Critchley et al. 2000). These results and others suggest that the anterior insula is involved in mapping physiological states of the body, including pain, touch, and visceral sensations of autonomic arousal (Craig 2002, Critchley 2005). The right anterior insula, in particular, is thought to be a cortical station for interoception that may play a role in decision-making by instantiating valenced subjective feeling states (Damasio 1994). Finally, recent fMRI data implicate right anterior insula in aversive conditioning (Seymour et al. 2004). Collectively, these findings suggest that the anterior insula may be involved in signaling that a social encounter has differed from expectations and consequently marking negative interactions as aversive to help in learning to avoid such interactions in the future. Although the magnitude of activation in anterior insula does not by itself predict subsequent defection by the player in future interactions with the same nonreciprocating partner, correlated activity (i.e., functional connectivity) between anterior insula and lateral OFC does. This finding is consistent with evidence that lateral OFC is involved in the evaluation of punishing stimuli that may lead to behavioral changes (Kringelbach & Rolls 2004).

Seeking Forgiveness

When cooperation has ruptured, it may often be beneficial in the long run to try to repair it. Indeed, the motivation to reconcile following conflict seems to be present in many primate species (de Waal 2000). In the iterated Trust Game, repair of ruptured cooperation is often initiated by a coaxing response involving hyper-reciprocation that encourages increasing expressions of trust in the investor (King-Casas et al. 2008). Interestingly, this response also appears to be deficient in borderline personality

Visceral: related to the internal organs of the body, including the heart, lungs, and gut

disorder (King-Casas et al. 2008). The insula response to low levels of trust, discussed above, may be part of the neural mechanism supporting a decision to seek forgiveness.

Finally, it is important to emphasize that the mechanisms discussed here that have evolved to support reciprocal altruism in humans are imperfect. Humans are often overcome by greed or fear such that sustained, mutually beneficial relationships cannot be realized, or worse, may degenerate into a bitter series of defections. One important challenge for social neuroscience is to begin to specify the situations under which cooperation does not occur. An interesting speculation in this regard is that trade, and later money, both of which immediately discharge debts of reciprocity by effectively removing the delay between giving help and receiving payback, may have developed to help circumvent our imperfections in this regard (Ridley 2010).

SHARING AND RESOURCE DISTRIBUTION

Another type of social decision that has been intensively studied from a behavioral, economic, and neuroscientific perspective is the decision of how to distribute limited resources among multiple individuals. Like reciprocal altruism, this is likely an evolutionarily ancient necessity, as both archeological and ethnographic evidence suggest that big game hunting or scavenging was important in human evolution (Stanford & Bunn 2001). Decisions about how to distribute large kills, the successful hunting of which often depended on cooperation among many individuals, have significant nutritional and social consequences for group members and have been intensively studied by anthropologists for decades (Hawkes 1993, Hill & Kaplan 1994). In modern Western societies, decisions about distribution of resources are equally important, as for example when deciding how to allocate salary raises from a fixed pool of money or how medical resources should be divided up among patients of differing need.

Deciding Whether to be Fair

Those in control of limited resources often face the decision of whether to distribute those resources equitably or efficiently (#4 in **Figure 1a,b**). For example, the salary raise pool can be equally divided among employees or more can be given to those who have performed well in the past. The UG is a useful experimental tool for examining both decisions about, and responses to, fairness and equity. In the Ultimatum Game, most people can be said to be fair in that they will offer a responder more than their own personal minimum acceptable offer. However, this is not the case in VMPFC patients (Krajbich et al. 2009), who instead show "negative generosity" by offering less than they themselves demand. Experimentally elevating testosterone levels in men decreases UG generosity and also makes them more than twice as likely to exhibit this negative generosity effect (Zak et al. 2009). These two findings may be related, as a recent fMRI study has linked testosterone with decreased VMPFC activity (Mehta & Beer 2010). If, as mentioned above, VMPFC patients have deficits in eliciting emotions based on imagined outcomes, then both VMPFC lesions and experimentally elevated testosterone may interfere with the ability to envision the partner's emotional reaction to potential offers; that is, they may impair empathy. In fact, the pathway from the amygdala to the VMPFC is hypothesized to be involved in perceiving distress in others and in learning to avoid behaviors that provoke such distress (Blair 2008), and testosterone may impair the functioning of this pathway (van Wingen et al. 2010). It should be noted, however, that in contrast to the effects of testosterone on men, testosterone administration actually increases UG offers in women (Eisenegger et al. 2010). A possible explanation is that the effects on generosity in women are actually mediated by estrogen, which testosterone may displace from its binding protein (sex hormone–binding globulin) owing to its greater binding affinity (Wallen 2001).

In contrast to VMPFC damage and testosterone administration to men, both of which

decrease generosity in the UG, intranasal OT increases generosity (Zak et al. 2007). This finding is consistent with the generally prosocial effects of OT in both humans and other animals [although OT also mediates maternal aggression in rodents and perhaps envy and gloating in humans (Shamay-Tsoory et al. 2009)]. The effects of OT and testosterone on UG offers, coupled with their lack of an effect on offers in the DG that are thought to represent "pure altruism", suggest that they may be modulating the ability to both empathize with and predict the behavior of one's partner, but in opposite ways (Hurlemann et al. 2010, Zak et al. 2009).

The UG presents a decision about how to allocate resources between oneself and another, but many important social and political decisions involve allocating resources among third parties, sometimes with ambiguity regarding what the morally optimal choice is. An ingenious recent study used fMRI as participants made decisions about how to allocate meals to a group of children living in an orphanage in northern Uganda (Hsu et al. 2008). In every decision, each of three children began by receiving 24 meals, with the participant then forced to choose between two different decision options, each of which involved one or more children losing meals. Choices involved deciding between equity, in which variance among children was minimized but the overall number of meals donated was lower, and efficiency, in which the overall number of meals was higher but with greater inequity among children. Efficiency was associated neurally with activation in the putamen. When participants chose the more equitable distribution, the amount of inequity reduced by their choice was correlated with activation in the anterior insula, bilaterally. Moreover, participants who showed stronger inequity aversion in their decision-making had stronger responses to inequity in the bilateral insula. As mentioned above, the anterior insula is responsive to aversive social interactions and norm violations. These data suggest that participants with a stronger aversive response to inequity may be more likely to choose equitable

divisions even at the cost of decreasing overall efficiency. Thus, a version of a somatic marker (Bechara & Damasio 2005) may be at the root of our decisions to promote equity.

Responding to Unfairness and Inequity

Moreso than any other social decision, the neural basis of the response to unfairness has been probed with a variety of neuroscience methods. An initial fMRI study showed that receiving an unfair compared to a fair offer in the UG was associated with activation in the anterior insula, that this activation was stronger for offers from putative human versus computer partners, and that it scaled to the magnitude of unfairness (#5 in **Figure 1a,b**). Moreover, unfair offers that were subsequently rejected were associated with a stronger insula response than those that were subsequently accepted, suggesting that the magnitude of anterior insula activation influences the decision to accept or reject (Sanfey et al. 2003). In a follow-up study, the magnitude of skin conductance responses to unfair offers was also found to predict the likelihood of rejection (van 't Wout et al. 2006). Given that anterior insula activity is known to correlate with skin conductance responses (Critchley et al. 2000), these findings suggest that visceral feedback from the body is driving the rejection of unfair offers. The anterior insula has also been implicated in empathy (de Vignemont & Singer 2006), perhaps by simulating the somatic state and accompanying feelings of another. Therefore, the insula response to potential inequitable distributions among Ugandan orphans (described above) may represent an empathic response that drives the decision to choose in favor of a more equitable distribution.

A later fMRI study (Tabibnia et al. 2008) replicated this insula finding and also showed that activation of the ventrolateral prefrontal cortex (VLPFC) was associated with accepting unfair offers. Given that VLPFC has been implicated in emotion regulation (Ochsner & Gross 2005), the results of a path analysis

VLPFC: ventrolateral prefrontal cortex

suggest that VLPFC contributes to acceptance of unfair offers by reducing anterior insula–based negative affect.

As in the domain of trust and reciprocity, people vary in their response to inequity, and social value orientation explains a portion of this variance. When asked to evaluate the desirability of pairs of rewards for both themselves and others, prosocials dislike large absolute differences in distributions, whereas proselfs, or individualists, do not. In prosocials, the magnitude of the differences in distributions is positively correlated with amygdala activation, but not in individualists. These decisions are not affected by cognitive load, implying that inequity aversion in prosocials is driven by an automatic, bottom-up aversive response represented in the amygdala (Haruno & Frith 2010).

In addition to fMRI, the neural basis of the response to unfairness has been investigated in brain-damaged patients. Damage to the VMPFC is associated with higher rejection rates of unfair UG offers (Koenigs & Tranel 2007, Krajbich et al. 2009), but apparently only if the payment is abstract and delayed (Moretti et al. 2009). When payment is immediate and concrete (i.e., visible cash is present), rejection rates of VMPFC patients do not differ from those of controls. Given a role for VMPFC and frontal pole in representing the value of future or abstract outcomes (Moretti et al. 2009), these results have been interpreted to suggest that increased rejection rates in VMPFC patients in the abstract case stem from reduced reward value placed on future payoffs following acceptance (Moretti et al. 2009). Regardless of the explanation for increased rejection rates in the abstract case, it seems clear that the VMPFC and/or the frontal pole are important for decision-making in this single-shot case where there can be no long-term benefit from rejecting an offer.

Pharmacologic manipulations have provided some additional evidence as to the mechanisms underlying responses to unfairness. Experimentally decreasing central serotonin levels through tryptophan depletion is associated with higher rejection rates of unfair UG offers (Crockett 2009). The insula receives dense innervation from the dorsal raphe serotonin projection system (Way et al. 2007), and SSRI treatment, which enhances serotonin transmission, is associated with reduced anterior insula responses to perception of emotional stimuli (Arce et al. 2008). Tryptophan depletion, on the other hand, is associated with an enhanced insula response to emotional stimuli (Roiser et al. 2008). Thus, tryptophan depletion may increase UG rejection rates by removing inhibitory influences of serotonin on the insular response to unfair offers. Of course, these effects could also be mediated by VMPFC, especially since, as mentioned above, they mimic the effects of VMPFC lesions (Crockett 2009). Withdrawal of serotonin from VMPFC might therefore contribute to increased rejection rates by impairing emotion regulation (Crockett 2009) or by decreasing the value of an abstract monetary reward (Moretti et al. 2009) relative to the immediate social reward of successfully punishing the unfair proposer (de Quervain et al. 2004, Singer et al. 2006) (see below).

Additional pharmacologic studies have examined the influence of sex-steroid hormones, such as testosterone, on UG game behavior. This has been done by examining the effects of naturally occurring variation in hormone levels and by specific pharmacological manipulations. Men who reject low offers in the UG have higher salivary testosterone than do men who accept these offers (Burnham 2007), and following exogenous testosterone administration, testosterone levels are positively correlated with UG rejection thresholds (i.e., men with higher testosterone reject more easily) (Zak et al. 2009). However, these effects of testosterone do not hold for either pre- (Eisenegger et al. 2010) or postmenopausal (Zethraeus et al. 2009) women. As mentioned above, testosterone has been linked with decreased VMPFC activity (Mehta & Beer 2010) as well as decreased amygdala-orbitofrontal coupling (van Wingen et al. 2010). Thus, testosterone may increase rejection rates in men through a mechanism similar to tryptophan depletion: by impairing emotion regulation or by modulating

the relative reward value of accepting versus rejecting.

UG responder behavior has also been investigated using rTMS. Disruption of right, but not left, dorsolateral prefrontal cortex with rTMS is associated with decreased rejection rates of unfair offers in the single-shot UG (Knoch et al. 2006, van 't Wout et al. 2005). These results suggest that right DLPFC may be involved in the implementation of fairness norms (Spitzer et al. 2007) and in general demonstrate that DLPFC also plays an important, causal, role in UG decisions.

Finally, clinical populations have also been investigated using this task. Autistic children are more likely to accept unfair UG offers than are controls (Sally & Hill 2006). The UG robustly activates a neural network involved in theory of mind processing (Rilling et al. 2004a), and autistic children have impaired theory of mind skills. Autistic children may therefore lack the ability to perceive the unfair intent behind a low offer (Sally & Hill 2006), perhaps due to insufficient activation of a theory of mind neural network (Frith 2003), and may therefore be less apt to reject. A similar pattern holds for clinically depressed patients, who also accept more unfair offers than do controls (Harlé et al. 2010), although the precise mechanism underlying this latter result is still unclear.

ALTRUISM

One remarkable facet of human social behavior is that people often decide to help others even when it comes at personal cost, and when there is no expectation of receiving any material returns (#6 in **Figure 1***a,b*). In experimental settings, "pure" altruism has been modeled with the DG. In this game, a player chooses how to allocate a sum of money between themselves and another player, in a similar manner to the UG, but in the DG the second player is simply a passive recipient of the offer and has no recourse to punish. Differences between a given proposer's Ultimatum and Dictator offers can therefore be instructive in examining the effect of both altruistic and strategic concerns.

VMPFC patients give less in the DG compared with normal or brain-damaged controls (Krajbich et al. 2009), and given the association between altruism and empathy (Batson & Powell 2003), this effect may stem from a deficit in empathy in VMPFC patients. Interestingly, in contrast to UG offers, neither testosterone supplementation nor OT self-administration affects DG offers in men (Zak et al. 2007, 2009), providing some evidence for a dissociation between generosity and altruism. In any event, if these pharmacologic manipulations do influence empathy, as suggested above, the effects do not translate into altered DG behavior.

Despite the lack of an effect of intranasal OT on altruism in the DG, genetic studies have implicated both the OT and AVP systems in DG behavior. In one study, three of fifteen single nucleotide polymorphisms across the OT receptor gene showed significant associations with DG offers (Israel et al. 2009). Another study showed that longer versions of the RS3 microsatellite repeat element of the vasopressin 1a receptor gene were associated with increased AVP mRNA expression as well as larger DG allocations (Knafo et al. 2008). Thus, significant variability in altruistic behavior in the DG is explained by genetic variation in the receptors that bind AVP and OT.

The ecological validity of anonymous DG allocations as a measure of altruism is of course questionable, as in most cases we know something about the people or causes that we support, and often they know something about us. Two recent clever fMRI studies have used more realistic paradigms and scanned participants as they decide whether to donate endowed money to a variety of real charitable organizations. In one study, both accepting a monetary reward and voluntarily donating money to a specific charity activated the ventral tegmental area and the ventral striatum, both components of the mesolimbic dopamine system (Moll et al. 2006). In another study, subjects showing stronger ventral striatum activation to mandatory charitable donations were more likely to then voluntarily give to the charity. On the other hand, subjects showing stronger ventral

Single nucleotide polymorphisms: genetic variants that differ in a single DNA base pair

Microsatellite: any of numerous short segments of DNA that are distributed throughout the genome, that consist of repeated sequences of usually two to five nucleotides, and that tend to vary from one individual to another

striatum activation to payments to themselves were less likely to later donate to the charity (Harbaugh et al. 2007). Thus, the ventral striatum, a mesencephalic dopamine system target involved in reward processing (O'Doherty et al. 2004, Schultz 1998), may motivate decisions to voluntarily donate money to charitable organizations. Other brain regions may also be involved. In the first of the studies just mentioned, one region—the subgenual cingulate cortex—was actually more active when donating to a cause compared to when accepting a personal monetary reward (Moll et al. 2006). One interpretation of this finding is that the subgenual cingulate cortex represents social attachment to a cause promoted by a specific charity. The same study found that sacrificing money to donate to a charity activated the VMPFC and frontal pole to a greater extent than did accepting a personal monetary reward (Moll et al. 2006), a result consistent with the observation that VMPFC lesions are associated with less altruism in the DG.

NORM-ABIDING DECISION-MAKING

The establishment of large-scale cooperation through social norms is a uniquely human phenomenon (#7 in **Figure 1a,b**). Social norms are effective in shaping behavior, presumably because humans are highly sensitive to the opinions and approval of others. Two recent studies suggest that approval of others is processed within the same ventral striatal regions that respond to a wide range of nonsocial rewards. One of these studies scanned subjects using fMRI as they made decisions about whether to donate money to charities, as in the above experiments, but with the interesting added manipulation that in some cases these donations were observed by peers. The presence of observers increased both donation rates, albeit minimally, as well as the ventral striatal response that preceded decisions to donate. That the same ventral striatum region was also active when choosing to keep money for oneself in the absence of peer observers suggests that this activation may

reflect an anticipated reward from presumed peer approval (Izuma et al. 2009). In the second study, subjects were scanned as they received feedback from others who had recently evaluated the subjects based on personality questionnaires and a videotaped self-introduction. Receiving unambiguously positive peer evaluations was associated with activation in the same region of the striatum that was also activated by monetary reward in a nonsocial task (Izuma et al. 2008). Thus, social approval seems to be rewarding and may motivate norm-abiding decisions, such as when called upon to help those in need.

Despite these positive incentives for norm-abiding behavior, both ethnographic evidence (Sober & Wilson 1998) and behavioral economics experiments (Fehr & Gachter 2002) show that some people will only abide by social norms under threat of punishment. Thus, sensitivity to the threat of punishment is also an important motive for norm-abiding behavior. In a recent fMRI study (Spitzer et al. 2007), subjects were imaged while playing two different games. In one game, which resembles a DG, subjects (player A) received a monetary endowment that they could distribute freely between themselves and another player (player B). In this game, player B is a passive recipient of player A's monetary transfer. In the other game, player B could choose to pay money to financially punish player A after having been informed of player A's decision. Player A transferred substantially more money to player B in the punishment compared with the nonpunishment condition. Those subjects who showed the largest change in monetary transfer from the nonpunishment to the punishment condition also showed the greatest increase in activation of both the lateral OFC and the right DLPFC across conditions. Lateral OFC is involved in the evaluation of punishing stimuli that may lead to behavioral changes (Kringelbach & Rolls 2004), so it may hold a subjective representation of the punishment threat that motivates norm-abiding behavior. Activation in DLPFC is consistent with the rTMS results described above that implicate it in the implementation of fairness norms.

DLPFC, known for its role in cognitive control (Miller & Cohen 2001), may be overriding a prepotent selfish impulse to send less of the endowment to player B.

This human sensitivity to social approval is underscored by the finding that a conflict with group opinion triggers a prediction error signal within putative reinforcement learning circuitry (Klucharev et al. 2009). In this fMRI study, female participants rated female faces for attractiveness, after which they were informed of the "average European rating" of the face. When participants learned that the group rating differed from their own, activation was observed in the rostral cingulate zone, along with deactivation in the nucleus accumbens. Subjects rated the same set of faces again after the fMRI session, and in some cases adjusted their ratings to more closely conform to the group average. This conformity was associated with stronger activation of the rostral cingulate zone as well as stronger deactivation of the nucleus accumbens, consistent with a larger error signal. Furthermore, subjects with a greater tendency to conform showed greater deactivation of the ventral striatum in response to initial nonconformity. Thus, error-related signals in the rostral cingulate and nucleus accumbens alert us when our decisions deviate from social norms and can motivate subsequent conformity.

Social norms, and specifically the expectations engendered by these norms, may also provide an explanation for cooperative behavior more generally. Research examining the neural basis of deviations from expectation in nonsocial contexts, such as oddball detection paradigms, has consistently shown activation in a network including anterior insula, ACC, and supplementary motor area (SMA). The same network has been shown to be active when there is conflict with a social norm (Klucharev et al. 2009), when conforming to a norm (Berns et al. 2010). This suggests that decisions involving both trust and reciprocation may involve norm compliance, namely the social norm that one should both trust others and reciprocate trust that has been placed in oneself. Similarly, rejecting unfair UG offers is associated with increased activity in the insula, ACC/SMA, and DLPFC (Sanfey et al. 2003), and dividing money under threat of sanction (and presumably promoting greater conformity to the norm) is also associated with insula activity (Spitzer et al. 2007). Finally, norm violation accounts may help explain how we remember partners with whom we interact. In contrast to cheater-detection theories, which posit that we have enhanced memory for partners who deceive us, a recent fMRI study of the UG showed that memory for partners was actually better explained by whether the partners had deviated from the subject's initial expectations, irrespective of whether subjects had high or low expectations of these partners (Chang & Sanfey 2009). Violations were also associated with activation in the insula/ACC/SMA network. Taken together, these results suggest that enforcing a social norm may be associated with this neural system and that this may play an important role in social decision-making.

ALTRUISTIC PUNISHMENT

As we have discussed above, the threat of punishment is an important motive for conforming to social norms. But to be effective, the threat must be credible. For most of human history, punishment has been meted out privately rather than by legal institutions, and the effectiveness of such a system in stabilizing social norms is dependent on some individuals being motivated to punish norm violators or free-riders despite any inherent costs (Fehr & Fischbacher 2003). How does the human brain mediate this so-called altruistic punishment? Two neuroimaging studies of trust and PD games have shown that brain reward regions, including the caudate nucleus and related structures in the ventral striatum, are activated when subjects successfully punish others who have previously treated them unfairly (de Quervain et al. 2004, Singer et al. 2006). In one study, the effect was observed for male but not female subjects, where activation in reward areas in response to punishment of a nonreciprocating partner was correlated with self-reported desire for revenge

(Singer et al. 2006). In the other study, subjects showing stronger activation of reward areas were willing to incur greater costs in order to punish the cheating partner (de Quervain et al. 2004). Thus, the motive to altruistically punish is correlated with, and perhaps causally related to, activation in brain reward systems.

In most modern societies, judicial institutions relieve ordinary citizens of the responsibility for retribution, with punishment meted out by impartial third parties. We might therefore expect third-party decisions about altruistic punishment to be more dispassionate than second-party punishment in the above studies. However, a recent fMRI study suggests otherwise (Buckholtz et al. 2008). Here, participants were presented with written scenarios in which they had to decide whether a protagonist should be punished, and if so, to what extent. Consistent with previous evidence that DLPFC is implicated in penalizing norm violations, DLPFC was indeed involved in deciding whether or not to punish based on an assessment of criminal responsibility. However, decisions about punishment magnitude were positively correlated with activity in regions linked with affective processing, such as the amygdala. These results are consistent with the hypothesis that third-party sanctions are fueled by negative emotions toward norm violators (Buckholtz et al. 2008).

SOCIAL LEARNING

To a much greater extent than other animals, human behavior is shaped by what we learn from others (Henrich & McElreath 2003, Tomasello 1999). We learn complex subsistence and occupational skills, social norms, and the specific features of our language from others. Indeed, social learning (#8 in **Figure 1*a,b***) often has more influence over our behaviors than individual learning does. For example, an fMRI study involving the Trust Game showed that the prior moral reputation of a social partner can outweigh direct experience in deciding whether or not to trust the partner (Delgado et al. 2005). Interestingly, when trust

was not reciprocated by partners with good reputations, the caudate prediction error signal of the scanned investors was blunted relative to nonreciprocation by trustees with neutral or bad reputations. Thus, socially learned information can lead to a suppression of neural mechanisms involved in individual reinforcement-based learning.

Imitation and social learning are likely dependent on a putative mirror neuron network including the STS, inferior parietal cortex, and inferior frontal cortex (Rizzolatti & Fogassi 2007). However, decisions about whether to rely on social stimuli for learning seem to involve the anterior cingulate gyrus (ACCg). For example, ACCg lesions in monkeys abolish social interest (Rudebeck et al. 2006). Additional evidence comes from an fMRI study designed to distinguish brain regions involved in social learning from those involved in individual reward-based learning. Participants were asked to choose which of two stimuli would yield a reward and were able to draw on their prior history of reinforcement as well as a confederate's advice in making their decision. Activation in the ACCg reflected the value placed on confederate advice when deciding (Behrens et al. 2008). Thus, the ACCg is involved in valuing information from others (Behrens et al. 2009).

Given the centrality of social learning to human behavior and the need for individuals to learn accurate and useful information from others, the decision of who to learn from or imitate is crucial (Henrich & McElreath 2003). Natural selection may favor cognitive capacities that bias individuals to learn preferentially from those who are more successful, and one way of inferring success is through the deference or social status individuals receive from others (Henrich & McElreath 2003). Access to high-status individuals may therefore be valuable, as suggested by the observation that monkeys will forego food to acquire information about dominant monkeys (Deaner et al. 2005). In an fMRI study in which participants played a reaction-time game with both more highly ranked players than themselves (individuals

who were declared to be better than them at the game) and less highly ranked players than themselves (individuals who were declared to be worse than them at the game), viewing the higher-compared with the lower-ranked player was associated with increased ventral striatum activation, interpreted as reflecting the greater salience of the higher-ranked player (Zink et al. 2008). Thus, the ventral striatum response to high-status individuals may reflect our motivation to attend to and learn from them.

COMPETITIVE SOCIAL INTERACTIONS

A major component of social life, though one that has received relatively little research attention, are decisions made in competitive social interactions (#9 in **Figure 1a,b**). A recent study (Hampton et al. 2008) compared three different computational models with respect to their ability to explain subjects' behavior during a competitive game in which employees can "work" or "shirk" and employers can "inspect" or "not inspect." The model that best fit the data was an "influence learning" model, in which players make decisions based on predictions of how their opponents will respond to their own prior decisions. In contrast to the other two models, this model unambiguously involves theory of mind processing, demonstrating that mentalizing guides decision-making in this game. Subjects who assumed greater influence over their partner's choices, and who were therefore more strategic, had stronger activation in DMPFC, consistent with a large body of studies implicating this region in mentalizing (see above). Activity in the posterior STS, another putative mentalizing region (Saxe & Kanwisher 2003), corresponded to an update signal that captured the difference between the degree of influence expected on a given trial and the actual influence exerted once the outcome had been revealed. Therefore, this region is involved in learning about the degree of influence one has over a partner's strategy (Hampton et al. 2008).

In fact, mentalizing-related neural activity may play a more significant role in competitive than in cooperative social decision-making. In a two-player fMRI study, participants engaged in a pattern-completion task either with the help of another player (cooperation) or with their interference (competition). Relative to cooperation trials, competition trials activated DMPFC. These results echo those described above in which DMPFC activation is high in early rounds of an interactive game before trust and cooperation have been established but then tapers off in later rounds once trust is firmly in place (Krueger et al. 2007).

Finally, competitive social behavior is often motivated by envy. A recent study (Takahashi et al. 2009) used hypothetical scenarios involving social comparison to successfully provoke self-reported envy, the degree of which was positively correlated with activation in the dorsal anterior cingulate cortex (dACC). The dACC has been linked with social and psychological pain, but also cognitive conflict, and the authors of this paper speculate that the dACC activation represents the conflict between a normally positive self-concept and the feedback that someone else is superior. Intriguingly, this activation is correlated across subjects with the magnitude of the ventral striatum response to a superior other's misfortune. So it would seem that those most prone to envy may also be most prone to take pleasure at another's bad luck (Takahashi et al. 2009).

SUMMARY

Although the neurobiological study of social decision-making is still in its infancy, there are currently enough findings to propose some tentative models of how the brain makes social decisions. Clearly, prefrontal cortex plays an essential role in social decision-making. The VMPFC/frontal pole region seems to be involved in valuing the long-term benefits associated with cooperative relationships and perhaps also abstract rewards such as helping

anonymous others through charity donations. This region also plays a role in regulating emotional reactions that could jeopardize valued relationships. Other prefrontal regions are involved in different components of social decision-making. DLPFC is involved in exerting cognitive effort to override selfish impulses, as when abiding by fairness norms, and VLPFC is involved in overriding aversive reactions to unfair treatment as well as in representing the threat of punishment from others that motivates norm-abiding behavior. Finally, DMPFC is involved in learning whether someone can be trusted as well as in strategizing during competitive interactions.

However, the neural basis of social decision-making is not confined to prefrontal cortex. The dACC seems to function as a social alarm signal (Eisenberger et al. 2003) that reacts to social norm violations, as when breaking a promise, deviating from group opinion, or being outperformed by others (i.e., envy). The anterior insula may also play a role in signaling norm violations and also mediates inequity aversion, aversive responses to unreciprocated altruism, and motivates decisions that restore equity. It is additionally involved in empathy and third-party reactions to inequity. The ventral striatum, a target of the mesolimbic dopamine system, is a final common pathway for both social and nonsocial reward stimuli (Montague & Berns 2002). It mediates rewards from mutual cooperation, altruism, and social approval and appears to motivate revenge seeking as well as attention to high-status others as models for social learning. The amygdala is also involved in aversive responses to inequity and also seems to mediate the fear of betrayal, thereby inhibiting trust.

These neural systems are modulated by a number of neurochemicals. Serotonin promotes prosocial behavior, perhaps by augmenting VMPFC function, which would in turn have several important consequences, such as placing greater value on the long-term benefits of cooperative relationships, improving emotion regulation, and perhaps increasing empathy. Testosterone may have an opposite effect on VMPFC function. OT promotes trust through decreasing amygdala activity (in men) and increases generosity and perhaps empathy through some as yet unspecified neural mechanism.

Though a full discussion is beyond the scope of this review, it is notable that many of the neural regions described here are also involved in aspects of more traditional decision-making, including valuation, risk assessment, and decision conflict, to name but a few. Future research might fruitfully explore the degree to which social decision-making overlaps with the more fundamental mechanisms employed in individual decision-making, in order to generate a more complete model of how people choose and decide.

CONCLUSION

This review highlights the importance of a network of brain regions in decisions that promote prosocial behaviors. Areas such as the amygdala and insula can provide important affective biases to social decisions, and the prefrontal cortex appears to be involved in overriding selfish impulses, valuing abstract and distant rewards (like altruism and cooperation), and in generating certain prosocial emotions (such as guilt and empathy). Interestingly, the human prefrontal cortex is larger than that of any other primate, in terms of both absolute size and relative to the size of the rest of the brain (Preuss 2004, Rilling 2006). This may account for one of the most distinctive aspects of human social life, namely our propensity to cooperate with nonrelatives (Boyd & Richerson 2006). Food sharing, reciprocity, and cooperation among nonrelatives are unusual among primates, but those species that exhibit these behaviors, such as capuchin monkeys and chimpanzees (Brosnan 2009), have unusually large brains (and prefrontal cortices) for their body size (Rilling & Insel 1999). Although the prosocial behaviors of chimpanzees and capuchin monkeys are noteworthy and should not be

underestimated (Brosnan 2009, de Waal 2008), humans are remarkable in their degree of reliance on reciprocal altruism (Clutton-Brock 2009a), their upholding of fairness norms (including advantageous inequity aversion) (Brosnan 2009), and especially their tendency to often cooperate with unfamiliar others (Silk 2009). The enlarged human prefrontal cortex may therefore explain our special abilities in this regard. Overall, it is clear that the study of the neurobiology of social decision-making is growing rapidly, and the current state of knowledge as described here seems sure to offer many interesting avenues for future research.

SUMMARY POINTS

1. Given that we live in highly complex social environments, many of our most important decisions are made in the context of social interactions.

2. Simple but sophisticated tasks from experimental economics have been used to study social decision-making in the laboratory setting, and a variety of neuroscience methods have been used to probe the underlying neural systems.

3. Prefrontal cortex plays a critical role in social decision-making, and different regions have different functions. The VMPFC/frontal pole region seems to be involved in valuing the long-term benefits associated with cooperative relationships, valuing abstract rewards such as helping anonymous others through charity donations, and regulating emotional reactions that could jeopardize valued relationships. DLPFC is involved in exerting cognitive effort to override selfish impulses, as when abiding by fairness norms. VLPFC is involved in overriding aversive reactions to unfair treatment as well as in representing the threat of punishment from others that motivates norm-abiding behavior. Finally, DMPFC is involved in learning whether someone can be trusted and also in strategizing during competitive interactions.

4. The dorsal anterior cingulate cortex seems to function as a social alarm signal that reacts to social norm violations, as when breaking a promise, deviating from group opinion, or being outperformed by others (i.e., envy).

5. The anterior insula mediates inequity aversion and aversive responses to unreciprocated altruism and motivates decisions that restore equity. It is also involved in empathy and third-party reactions to inequity.

6. The ventral striatum mediates rewards from mutual cooperation, altruism, and social approval and appears to motivate revenge seeking as well as attention to high-status others as models for social learning.

7. The amygdala is also involved in aversive responses to inequity and seems to mediate the fear of betrayal, thereby inhibiting trust.

8. These neural systems are modulated by a number of neurochemicals including serotonin, which promotes prosocial behavior, perhaps by augmenting VMPFC function; testosterone, which may suppress VMPFC function; and oxytocin, which promotes trust through decreasing amygdala activity (in men) and increases generosity and perhaps empathy through some as yet unspecified neural mechanism.

FUTURE ISSUES

1. The models proposed in this review are based on the relatively limited evidence published to date and naturally should be considered preliminary. Explicit testing will be required to fully evaluate these models and to assist in further revision and expansion.

2. Only very few studies have combined fMRI of interactive tasks with pharmacological manipulations (e.g., Baumgartner et al. 2008). This is an important avenue for future research that will shed light on neurochemical influences on social decision-making. Although oxytocin is currently being widely studied in this regard, a wide range of other neurochemicals can be fruitfully employed in this research. For example, neurotransmitters involved in basic learning and reward processes, such as dopamine, can be used to link higher-level accounts of social behavior to more fundamental cognitive processes.

3. A useful complement to fMRI studies of social decision-making is to make use of the increasingly rich variety of formal mathematical models of these behaviors that are emerging from behavioral economics. These models encourage more specific descriptions of the social processes and can facilitate neural investigations of these processes.

4. Defining the neural correlates of individual variation in social decision-making is an important challenge for the field of social neuroscience. This variation might be explained by variation in genetics, personality, developmental settings, hormone and neurotransmitter levels, or even hormone and neurotransmitter receptor density. For example, variation in partner preference among monogamous male prairie voles is explained by variation in expression of the V1a vasopressin receptor. As PET ligands become available for receptors of interest, it should become possible to relate variation in receptor density with social decision-making in humans.

5. Although imaging brain function in the context of interactive games has been an important step toward increasing ecological validity of social neuroscience paradigms, a further step is to link brain activity in the scanner to real-world social behavior. This can be facilitated by asking subjects to record social behaviors in diaries or through experience-sampling methods (e.g., Eisenberger et al. 2007).

6. The emerging discipline of cultural neuroscience raises questions about whether the models outlined here will generalize beyond the developed Western cultures from which research participants are typically drawn (Chiao 2009). It will be important to evaluate cross-cultural variability in these models (Henrich 2010).

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Adolphs R, Tranel D, Damasio AR. 1998. The human amygdala in social judgment. *Nature* 393:470–74

Arce E, Simmons AN, Lovero KL, Stein MB, Paulus MP. 2008. Escitalopram effects on insula and amygdala BOLD activation during emotional processing. *Psychopharmacology* (*Berl.*) 196:661–72

Axelrod R, Hamilton WD. 1981. The evolution of cooperation. *Science* 211:1390–96

Axelrod RM. 1984. *The Evolution of Cooperation*. New York: Basic Books. 241 pp.

Batson C, Powell AA. 2003. Altruism and prosocial behavior. In *Handbook of Psychology: Personality and Social Psychology*, Vol. 5, ed. T Millon, MJ Lerner, pp. 463–84. New York: Wiley

Baumgartner T, Fischbacher U, Feierabend A, Lutz K, Fehr E. 2009. The neural circuitry of a broken promise. *Neuron* 64:756–70

Baumgartner T, Heinrichs M, Vonlanthen A, Fischbacher U, Fehr E. 2008. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* 58:639–50

Bechara A, Damasio A. 2005. The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52:336–72

Beer JS, Heerey EA, Keltner D, Scabini D, Knight RT. 2003. The regulatory function of self-conscious emotion: insights from patients with orbitofrontal damage. *J. Personal. Soc. Psychol.* 85:594–604

Behrens TE, Hunt LT, Rushworth MF. 2009. The computation of social behavior. *Science* 324:1160–64

Behrens TE, Hunt LT, Woolrich MW, Rushworth MF. 2008. Associative learning of social value. *Nature* 456:245–49

Berns GS, Capra CM, Moore S, Noussair C. 2010. Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage* 49:2687–96

Blair RJ. 2008. The amygdala and ventromedial prefrontal cortex: functional contributions and dysfunction in psychopathy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363(1503):2557–65

Botvinick M, Braver T, Barch D, Carter C, Cohen J. 2001. Conflict monitoring and cognitive control. *Psychol. Rev.* 108:624–52

Boyd R, Richerson PJ. 2006. Solving the puzzle of human cooperation. In *Evolution and Culture*, ed. S Levinson, pp. 105–32. Cambridge, MA: MIT Press

Brosnan SF. 2009. Responses to inequity in non-human primates. In *Neuroeconomics: Decision Making and the Brain*, ed. PW Glimcher, CF Camerer, E Fehr, RA Poldrack, pp. 285–302. London: Academic

Buckholtz JW, Asplund CL, Dux PE, Zald DH, Gore JC, et al. 2008. The neural correlates of third-party punishment. *Neuron* 60:930–40

Burnham TC. 2007. High-testosterone men reject low ultimatum game offers. *Proc. Biol. Sci.* 274:2327–30

Camerer CF. 2003. *Behavioral Game Theory*. Princeton, NJ: Princeton Univ. Press. 544 pp.

Chang LJ, Sanfey AG. 2009. Unforgettable ultimatums? Expectation violations promote enhanced social memory following economic bargaining. *Front. Behav. Neurosci.* 3:36

Chiao JY. 2009. Cultural neuroscience: a once and future discipline. *Prog. Brain Res.* 178:287–304

Clutton-Brock T. 2009a. Cooperation between non-kin in animal societies. *Nature* 462:51–57

Clutton-Brock T. 2009b. Structure and function in mammalian societies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364:3229–42

Cosmides L, Tooby J. 2000. The cognitive neuroscience of social reasoning. In *The New Cognitive Neurosciences*, ed. MS Gazzaniga, pp. 1259–70. Cambridge, MA: MIT Press

Craig AD. 2002. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3:655–66

Creel S, Creel NM. 2001. *The African Wild Dog: Behavior, Ecology, and Conservation*. Princeton, NJ: Princeton Univ. Press

Critchley HD. 2005. Neural mechanisms of autonomic, affective, and cognitive integration. *J. Comp. Neurol.* 493:154–66

Critchley HD, Elliot R, Mathias C, Dolan R. 2000. Neural activity relating to the generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *J. Neurosci.* 20:3033–40

Crockett MJ. 2009. The neurochemistry of fairness: clarifying the link between serotonin and prosocial behavior. *Ann. N. Y. Acad. Sci.* 1167:76–86

Crockett MJ, Clark L, Tabibnia G, Lieberman MD, Robbins TW. 2008. Serotonin modulates behavioral reactions to unfairness. *Science* 320:1739

Damasio AR. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam. 312 pp.

de Quervain DJ, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, et al. 2004. The neural basis of altruistic punishment. *Science* 305:1254–58

de Vignemont F, Singer T. 2006. The empathic brain: how, when and why? *Trends Cogn. Sci.* 10:435–41

de Waal FB. 2000. Primates—a natural heritage of conflict resolution. *Science* 289:586–90

de Waal FB. 2008. Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59:279–300

Deaner RO, Khera AV, Platt ML. 2005. Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr. Biol.* 15:543–48

Delgado MR, Frank RH, Phelps EA. 2005. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8:1611–18

Denk F, Walton ME, Jennings KA, Sharp T, Rushworth MF, Bannerman DM. 2005. Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology (Berl.)* 179:587–96

Ditzen B, Schaer M, Gabriel B, Bodenmann G, Ehlert U, Heinrichs M. 2009. Intranasal oxytocin increases positive communication and reduces cortisol levels during couple conflict. *Biol. Psychiatry* 65:728–31

Domes G, Heinrichs M, Glascher J, Buchel C, Braus DF, Herpertz SC. 2007. Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol. Psychiatry* 62:1187–90

Domes G, Lischke A, Berger C, Grossmann A, Hauenstein K, et al. 2010. Effects of intranasal oxytocin on emotional face processing in women. *Psychoneuroendocrinology* 35:83–93

Eberhardt JL. 2005. Imaging race. *Am. Psychol.* 60:181–90

Eisenberger NI, Gable SL, Lieberman MD. 2007. Functional magnetic resonance imaging responses relate to differences in real-world social experience. *Emotion* 7:745–54

Eisenberger NI, Lieberman MD, Williams KD. 2003. Does rejection hurt? An FMRI study of social exclusion. *Science* 302:290–92

Eisenegger C, Naef M, Snozzi R, Heinrichs M, Fehr E. 2010. Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature* 463:356–59

Etkin A, Wager TD. 2007. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry* 164:1476–88

Fehr E, Camerer CF. 2007. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11(10):419–27

Fehr E, Fischbacher U. 2003. The nature of human altruism. *Nature* 425:785–91

Fehr E, Gachter S. 2002. Altruistic punishment in humans. *Nature* 415:137–40

Frank RH. 1988. *Passions Within Reason: The Strategic Role of the Emotions*. New York: Norton. 304 pp.

Frith C. 2003. What do imaging studies tell us about the neural basis of autism? *Novartis Found. Symp.* 251:149–66; discussion 166–76, 281–97

Frith CD, Singer T. 2008. The role of social cognition in decision making. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363:3875–86

Gallagher HL, Frith CD. 2003. Functional imaging of "theory of mind." *Trends Cogn. Sci.* 7:77–83

Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD. 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44:389–400

Hampton AN, Bossaerts P, O'Doherty JP. 2008. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. USA* 105:6741–46

Harbaugh WT, Mayr U, Burghart DR. 2007. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316:1622–25

Harlé KM, Allen JJB, Sanfey AG. 2010. The impact of depression on social economic decisions. *J. Abnorm. Psychol.* 119:440–46

Harris LT, Fiske ST. 2006. Dehumanizing the lowest of the low: neuroimaging responses to extreme outgroups. *Psychol. Sci.* 17:847–53

Haruno M, Frith CD. 2010. Activity in the amygdala elicited by unfair divisions predicts social value orientation. *Nat. Neurosci.* 13:160–61

Hawkes K. 1993. Why hunter-gatherers work: an ancient version of the problem of public goods. *Curr. Anthropol.* 34:341–61

Haxby JV, Hoffman EA, Gobbini MI. 2002. Human neural systems for face recognition and social communication. *Biol. Psychiatry* 51:59–67

Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, et al. 2005. "Economic Man" in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* 28:795–815

Henrich JS, Heine J, Norenzayan A. 2010. The weirdest people in the world? *Behav. Brain Sci.* 33(2-3):61–83; discussion 83–135

Henrich J, McElreath R. 2003. The evolution of cultural evolution. *Evol. Anthropol.* 12:123–35

Hill K, Kaplan H. 1994. On why male forgers hunt and share food. *Curr. Anthropol.* 34:701–6

Hsu M, Anen C, Quartz SR. 2008. The right and the good: distributive justice and neural encoding of equity and efficiency. *Science* 320:1092–95

Hurlemann R, Patin A, Onur OA, Cohen MX, Baumgartner T, et al. 2010. Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *J. Neurosci.* 30(14):4999–5007

Israel S, Lerer E, Shalev I, Uzefovsky F, Riebold M, et al. 2009. The oxytocin receptor (OXTR) contributes to prosocial fund allocations in the dictator game and the social value orientations task. *PLoS ONE* 4:e5535

Izuma K, Saito DN, Sadato N. 2008. Processing of social and monetary rewards in the human striatum. *Neuron* 58:284–94

Izuma K, Saito DN, Sadato N. 2009. Processing of the incentive for social approval in the ventral striatum during charitable donation. *J. Cogn. Neurosci.* 22:621–31

Kable JW, Glimcher PW. 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10:1625–33

Kable JW, Glimcher PW. 2009. The neurobiology of decision: consensus and controversy. *Neuron* 63:733–45

Kagel JH, Battalio RC, Green L. 1995. *Economic Choice Theory: An Experimental Analysis of Animal Behavior*. London: Cambridge Univ. Press

King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR. 2008. The rupture and repair of cooperation in borderline personality disorder. *Science* 321:806–10

King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR. 2005. Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308:78–83

Kirsch P, Esslinger C, Chen Q, Mier D, Lis S, et al. 2005. Oxytocin modulates neural circuitry for social cognition and fear in humans. *J. Neurosci.* 25:11489–93

Klein JT, Deaner RO, Platt ML. 2008. Neural correlates of social target value in macaque parietal cortex. *Curr. Biol.* 18:419–24

Klucharev V, Hytonen K, Rijpkema M, Smidts A, Fernandez G. 2009. Reinforcement learning signal predicts social conformity. *Neuron* 61:140–51

Knafo A, Israel S, Darvasi A, Bachner-Melman R, Uzefovsky F, et al. 2008. Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. *Genes Brain Behav.* 7:266–75

Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E. 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314:829–32

Koechlin E, Hyafil A. 2007. Anterior prefrontal function and the limits of human decision-making. *Science* 318:594–98

Koenigs M, Tranel D. 2007. Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *J. Neurosci.* 27:951–56

Konner MJ. 2002. *The Tangled Wing: Biological Constraints on the Human Spirit*. New York: Henry Holt. 540 pp.

Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E. 2005. Oxytocin increases trust in humans. *Nature* 435:673–76

Krajbich I, Adolphs R, Tranel D, Denburg NL, Camerer CF. 2009. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J. Neurosci.* 29:2188–92

Kringelbach ML, Rolls ET. 2004. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog. Neurobiol.* 72:341–72

Krueger F, Grafman J, McCabe K. 2008. Neural correlates of economic game playing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363:3859–74

Krueger F, McCabe K, Moll J, Kriegeskorte N, Zahn R, et al. 2007. Neural correlates of trust. *Proc. Natl. Acad. Sci. USA* 104:20084–89

Lee RB. 1979. *The !Kung San: Men, Women, and Work in a Foraging Society*. London: Cambridge Univ. Press

Li J, Xiao E, Houser D, Montague PR. 2009. Neural responses to sanction threats in two-party economic exchange. *Proc. Natl. Acad. Sci. USA* 106:16835–40

Lieberman MD. 2007. Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58:259–89

Mah L, Arnold MC, Grafman J. 2004. Impairment of social perception associated with lesions of the prefrontal cortex. *Am. J. Psychiatry* 161:1247–55

Mehta PH, Beer J. 2010. Neural mechanisms of the testosterone-aggression relation: the role of orbito-frontal cortex. *J. Cogn. Neurosci.* 22:2357–68

Miller E, Cohen J. 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24:167–202

Mitchell JP. 2009. Social psychology as a natural kind. *Trends Cogn. Sci.* 13:246–51

Moll J, Krueger F, Zahn R, Pardini M, de Oliveira-Souza R, Grafman J. 2006. Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. USA* 103:15623–28

Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J. 2005. Opinion: the neural basis of human moral cognition. *Nat. Rev. Neurosci.* 6:799–809

Montague PR, Berns GS. 2002. Neural economics and the biological substrates of valuation. *Neuron* 36:265–84

Montague PR, Lohrenz T. 2007. To detect and correct: norm violations and their enforcement. *Neuron* 56:14–18

Moretti L, Dragone D, di Pellegrino G. 2009. Reward and social valuation deficits following ventromedial prefrontal damage. *J. Cogn. Neurosci.* 21:128–40

Neumann JV, Morgenstern O. 1947. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton Univ. Press. 648 pp.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–54

Ochsner KN, Gross JJ. 2005. The cognitive control of emotion. *Trends Cogn. Sci.* 9:242–49

Petrovic P, Kalisch R, Singer T, Dolan RJ. 2008. Oxytocin attenuates affective evaluations of conditioned faces and amygdala activity. *J. Neurosci.* 28:6607–15

Pochon JB, Riis J, Sanfey AG, Nystrom LE, Cohen JD. 2008. Functional imaging of decision conflict. *J. Neurosci.* 28:3468–73

Poldrack RA. 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10:59–63

Preuss TM. 2004. What is it like to be a human? In *The Cognitive Neurosciences*, ed. MS Gazzaniga, pp. 5–22. Cambridge, MA: MIT Press

Ridley M. 2010. *The Rational Optimist: How Prosperity Evolves*. New York: HarperCollins. 448 pp.

Rilling JK. 2006. Human and non-human primate brains: Are they allometrically scaled versions of the same design? *Evol. Anthropol.* 15:65–77

Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, et al. 2007. Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* 61:1260–71

Rilling JK, Goldsmith DR, Glenn AL, Jairam MR, Elfenbein HA, et al. 2008. The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* 46:1256–66

Rilling JK, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C. 2002. A neural basis for social cooperation. *Neuron* 35:395–405

Rilling JK, Insel TR. 1999. The primate neocortex in comparative perspective using magnetic resonance imaging. *J. Hum. Evol.* 37:191–223

Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD. 2004a. The neural correlates of theory of mind within interpersonal interactions. *NeuroImage* 22:1694–703

Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD. 2004b. Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *NeuroReport* 15:2539–43

Rizzolatti G, Fogassi L. 2007. Mirror neurons and social cognition. In *The Oxford Handbook of Evolutionary Psychology*, ed. RIM Dunbar, L Barrett, pp. 179–96. London: Oxford Univ. Press

Robbins TW. 2000. Chemical neuromodulation of frontal-executive functions in humans and other animals. *Exp. Brain Res.* 133:130–38

Roiser JP, Levy J, Fromm SJ, Wang H, Hasler G, et al. 2008. The effect of acute tryptophan depletion on the neural correlates of emotional processing in healthy volunteers. *Neuropsychopharmacology* 33:1992–2006

Rudebeck PH, Buckley MJ, Walton ME, Rushworth MF. 2006. A role for the macaque anterior cingulate gyrus in social valuation. *Science* 313:1310–12

Sally D, Hill E. 2006. The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27:73–97

Sanfey AG. 2007. Social decision-making: insights from game theory and neuroscience. *Science* 318:598–602

Sanfey AG, Loewenstein G, McClure SM, Cohen JD. 2006. Neuroeconomics: cross-currents in research on decision-making. *Trends Cogn. Sci.* 10:108–16

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. 2003. The neural basis of economic decision-making in the Ultimatum Game. *Science* 300:1755–58

Saxe R, Carey S, Kanwisher N. 2004. Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 55:87–124

Saxe R, Kanwisher N. 2003. People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind." *Neuroimage* 19:1835–42

Schultz W. 1998. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80:1–27

Schweighofer N, Bertin M, Shishida K, Okamoto Y, Tanaka SC, et al. 2008. Low serotonin levels increase delayed reward discounting in humans. *J. Neurosci.* 28:4528–32

Seo H, Barraclough DJ, Lee D. 2009. Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J. Neurosci.* 29:7278–89

Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, et al. 2004. Temporal difference models describe higher-order learning in humans. *Nature* 429:664–67

Shamay-Tsoory SG, Fischer M, Dvash J, Harari H, Perach-Bloom N, Levkovitz Y. 2009. Intranasal administration of oxytocin increases envy and schadenfreude (gloating). *Biol. Psychiatry* 66:864–70

Siever LJ, Buchsbaum MS, New AS, Spiegel-Cohen J, Wei T, et al. 1999. d,l-fenfluramine response in impulsive personality disorder assessed with [18F]fluorodeoxyglucose positron emission tomography. *Neuropsychopharmacology* 20:413–23

Silk JB. 2009. Social preferences in primates. In *Neuroeconomics: Decision-Making and the Brain*, ed. PW Glimcher, CF Camerer, E Fehr, RA Poldrack. London: Academic. 556 pp.

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD. 2004. Empathy for pain involves the affective but not sensory components of pain. *Science* 303:1157–62

Singer T, Seymour B, O'Doherty JP, Stephan KE, Dolan RJ, Frith CD. 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439:466–69

Singer T, Snozzi R, Bird G, Petrovic P, Silani G, et al. 2008. Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain. *Emotion* 8:781–91

Sober E, Wilson DS. 1998. *Unto Others—The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard Univ. Press

Spitzer M, Fischbacher U, Herrnberger B, Gron G, Fehr E. 2007. The neural signature of social norm compliance. *Neuron* 56:185–96

Sripada CS, Angstadt M, Banks S, Nathan PJ, Liberzon I, Phan KL. 2009. Functional neuroimaging of mentalizing during the trust game in social anxiety disorder. *Neuroreport* 20:984–89

Stanford CB, Bunn HT, eds. 2001. *Meat-Eating and Human Evolution*. New York: Oxford Univ. Press. 384 pp.

Tabibnia G, Satpute AB, Lieberman MD. 2008. The sunny side of fairness: preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychol. Sci.* 19:339–47

Takahashi H, Kato M, Matsuura M, Mobbs D, Suhara T, Okubo Y. 2009. When your gain is my pain and your pain is my gain: neural correlates of envy and schadenfreude. *Science* 323:937–39

Todorov A, Said CP, Engell AD, Oosterhof NN. 2008. Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* 12:455–60

Tomasello M. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard Univ. Press

Trivers RL. 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46:35–57

Tse WS, Bond AJ. 2002. Serotonergic intervention affects both social dominance and affiliative behaviour. *Psychopharmacology* (*Berl.*) 161:324–30

van 't Wout M, Kahn RS, Sanfey AG, Aleman A. 2005. Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport* 16:1849–52

van 't Wout M, Kahn RS, Sanfey AG, Aleman A. 2006. Affective state and decision-making in the Ultimatum Game. *Exp. Brain Res.* 169:564–68

van 't Wout M, Sanfey AG. 2008. Friend or foe: the effect of implicit trustworthiness judgments in social decision-making. *Cognition* 108:796–803

Van Den Bos W, van Dijk E, Westenberg M, Rombouts SA, Crone EA. 2009. What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Soc. Cogn. Affect. Neurosci.* 4:294–304

van Wingen G, Mattern C, Verkes RJ, Buitelaar J, Fernandez G. 2010. Testosterone reduces amygdala-orbitofrontal cortex coupling. *Psychoneuroendocrinology* 35:105–13

Wallen K. 2001. Sex and context: hormones and primate sexual motivation. *Horm. Behav.* 40:339–57

Way BM, Lacan G, Fairbanks LA, Melega WP. 2007. Architectonic distribution of the serotonin transporter within the orbitofrontal cortex of the vervet monkey. *Neuroscience* 148:937–48

Winston JS, Strange BA, O'Doherty J, Dolan RJ. 2002. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* 5:277–82

Wood RM, Rilling JK, Sanfey AG, Bhagwagar Z, Rogers RD. 2006. Effects of tryptophan depletion on the performance of an iterated Prisoner's Dilemma game in healthy adults. *Neuropsychopharmacology* 31:1075–84

Young LJ, Murphy Young AZ, Hammock EA. 2005. Anatomy and neurochemistry of the pair bond. *J. Comp. Neurol.* 493:51–57

Zak PJ, Kurzban R, Ahmadi S, Swerdloff RS, Park J, et al. 2009. Testosterone administration decreases generosity in the ultimatum game. *PLoS ONE* 4:e8330

Zak PJ, Stanton AA, Ahmadi S. 2007. Oxytocin increases generosity in humans. *PLoS ONE* 2:e1128

Zethraeus N, Kocoska-Maras L, Ellingsen T, von Schoultz B, Hirschberg AL, Johannesson M. 2009. A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proc. Natl. Acad. Sci. USA* 106:6535–38

Zink CF, Tong Y, Chen Q, Bassett DS, Stein JL, Meyer-Lindenberg A. 2008. Know your place: neural processing of social hierarchy in humans. *Neuron* 58:273–83
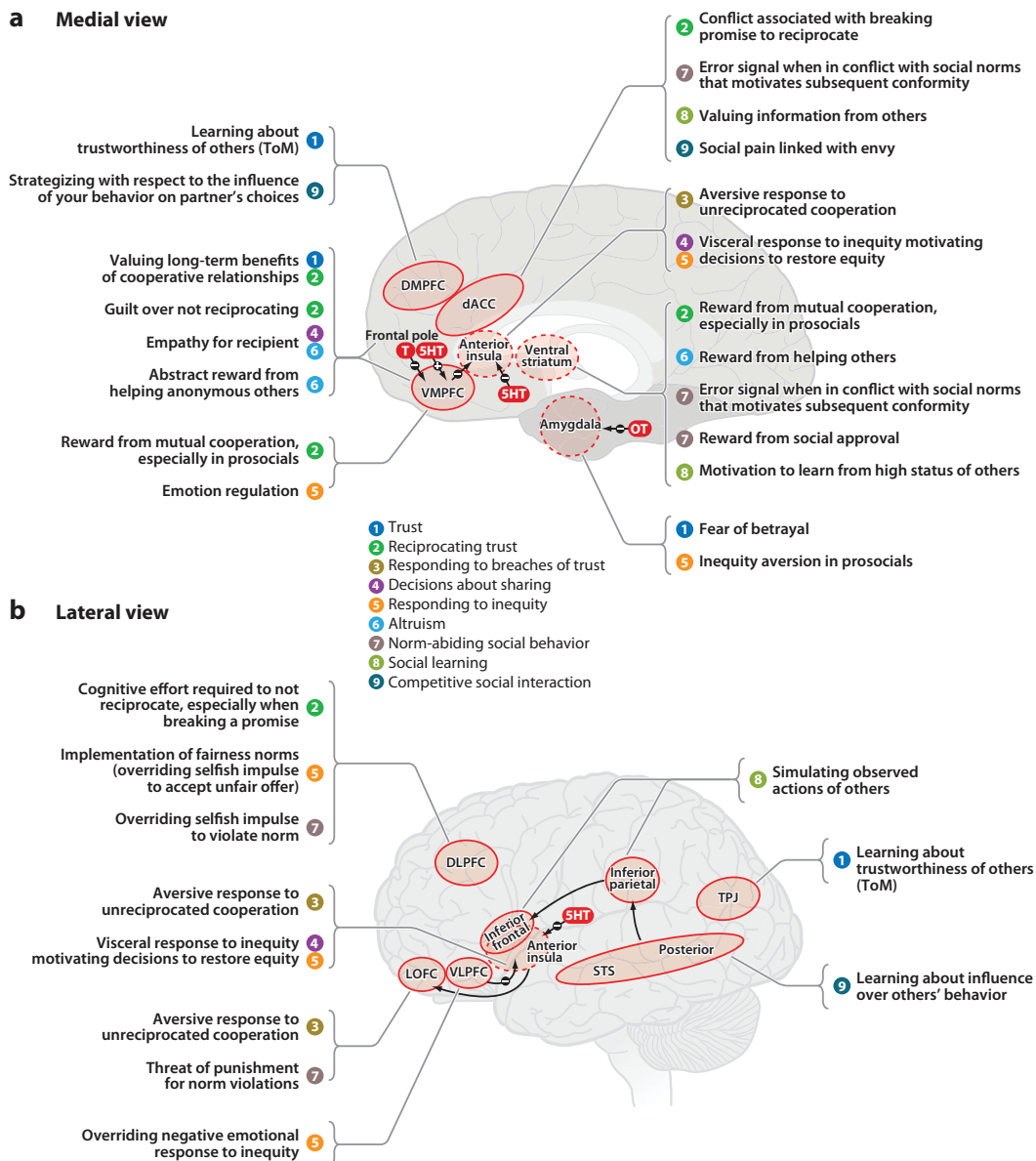
Figure 1

Model of the neural systems that mediate nine different types of social decisions, showing (*a*) medial and (*b*) lateral views of the human brain. Solid lines, surface structures; dashed lines, deep structures; −, inhibitory influences; +, stimulatory influences; arrows, white matter connections. DMPFC, dorsomedial prefrontal cortex; TPJ, temporo-parietal junction; VMPFC, ventromedial prefrontal cortex; dACC, dorsal anterior cingulate cortex; DLPFC, dorsolateral prefrontal cortex; VLPFC, ventrolateral prefrontal cortex; dACC, dorsal anterior cingulate cortex; LOFC, lateral orbitofrontal cortex; STS, superior temporal sulcus; 5-HT, serotonin; OT, oxytocin; T, testosterone.

**Annual Review of
Psychology**

Volume 62, 2011

Contents

**Psychophysiological Disorders and Psychological Dimensions
on Medical Disorders**

Research Methodology

Indexes

Errata

An online log of corrections to *Annual Review of Psychology* articles may be found at
http://psych.AnnualReviews.org/errata.shtml