Traditionally, studies of decision making have focused on situations in which people choose between outcomes that only affect themselves, such as playing a lottery where choices are described in terms of outcomes and probabilities and players have to pick the lottery that they prefer. However, many of our decisions in daily life are made in a social context in which the outcome of a decision has consequences not only for one’s self but for others as well, and where we need to consider the desires and values of others before deciding.

Game theory is a collection of models for understanding such instances of interactive decision making [1]. These models make clear mathematical predictions about behavior in social situations by describing what strategies decision makers converge on as they try to maximize their payoff. To study complex social processes, such as trust, cooperation, and reputation, game theory offers a rich variety of well-designed multiplayer games. Recently, a new interdisciplinary field has emerged, popularly known as neuroeconomics, which aims to better understand how people make decisions by integrating research from several different fields (for a detailed overview of the field see the handbook on neuroeconomics by Glimcher and colleagues [2]). As part of this neuroeconomic approach, researchers have begun to investigate the neural correlates of social decision making using some of these games.

In this article, we first provide a brief introduction as to how neuroeconomic studies are typically conducted. Following this, we describe the main findings of current neuroeconomic studies as they relate to game theory, and we conclude with an outline of potential future research directions in this field.

**NEUROECONOMICS**

Researchers in neuroeconomics have sought to investigate economic decision making in human subjects by combining a variety of methods from experimental economics, psychology, and neuroscience. These methods include (i) behavioral studies that examine how manipulation of the decision environment can affect choice, reaction time, skin conductance, or eye movements; (ii) studies with brain lesion patients that examine the consequences of abnormal brain function on decision making; (iii) experiments applying repetitive transcranial manipulation (rTMS) to temporarily disrupt activity within the brain; (iv) electroencephalography (EEG) or magnetoencephalography (MEG) studies, which measure at the scalp the electrical signals of neuronal firing; (v) pharmacological research to examine the effects of drug administration and neurotransmitters; (vi) genetic studies looking at the correlation between individual differences in the expression of certain genes and behavior; (vii) research using positron emission tomography (PET) to investigate the functioning of brain structures by the use of radioactive isotopes; and (viii) functional magnetic resonance imaging (fMRI) experiments that...
allow for the relatively direct measurement of real-time neural activity.

Over the last decade, the latter of these methods, fMRI, has emerged as the dominant technique in neuroeconomic studies. This method provides a noninvasive, indirect measure of the neural activity that occurs during decision making by assessing regional changes in the level of blood oxygenation. Typical fMRI experiments have a temporal resolution in the order of 1 s and a spatial resolution of 2–3 mm, which means that use of fMRI allows for the measurement of transient cognitive events, and also that relatively small structures within the brain can be imaged. A technical constraint of the method is that the blood oxygen level dependent response (BOLD) signal is rather noisy, requiring usually many trials to average the signal to a stable level. This has implications for the type of decision that can be studied, and as a consequence most game theoretic experiments undertaken with fMRI require many repetitions of the same game. This in turn means that genuine “single-shot” games (i.e., where players play only one game with only one partner) are typically not well suited to the MRI environment.

CURRENT RESEARCH

Recent research in neuroeconomics has begun to investigate the processes underlying social decision making using game theoretic paradigms. In contrast to standard behavioral studies in game theory, the neuroeconomic approach allows for the discrimination and modeling of processes that are hard to separate at the behavioral level. The combination of game theoretic models with the on-line measurement of brain activity during instances of decision making offers the promise of increasing our understanding of the processes and factors that are involved in social decision making.

An additional benefit of the neuroeconomic approach is that it has the potential to inform game theory models. This can be helpful, since actual decision behavior in game theory tasks often deviates from the models’ predictions. Game theory models assume that players in a social context interact rationally, with full information, and that they make purely self-interested decisions. However, several decades of laboratory experiments by psychologists and economists have found that players rarely behave according to these strict game theoretic strategies [3]. Instead, players often appear to care about the welfare of others, and in addition value factors like reciprocity and fairness. Identification of the neural mechanisms underlying social interaction in game theory paradigms may result in a more precise characterization of behavior and, subsequently, game theory models may be adapted to better fit decision-making behavior.

Although neuroeconomics is still a relatively young discipline, there have already been many studies conducted that have used game theoretic approaches to gain insight into the neural basis of social decision making [4,5]. The current findings can be usefully summarized in three themes: (i) reasoning about the intentions of others, (ii) social reward, and (iii) the role of emotions in decision making, each of which is expanded upon in the remainder of this section.

Reasoning about the Intentions of Others

One focus of game theory is to model reciprocal exchange, which has been studied extensively in the laboratory with games like the trust game [6–9] and the prisoner’s dilemma game [10–12]. A trust game is played with two players. One is termed the investor, who is endowed with a particular monetary amount, and the other is termed the trustee. The investor is instructed to choose how much of her endowment she would like to pass on to the trustee and how much she will keep for herself. Once this decision is made, the experimenter multiplies the transferred amount of money by some factor (usually 3 or 4) and passes it to the trustee. At this point the trustee has the option to return the money to the investor or to keep all the money to herself. If the trustee decides to return the money, thereby honoring the trust of the investor, both players finish the game with more money than they originally had. However, if the trustee decides to keep the money and thus does not repay the trust,
the investor ends up with a loss and the trustee is the only one who benefits.

As indicated above, game theory assumes that a decision maker acts in a rational and self-interested manner, and therefore the classical economic prediction would be that a trustee in the trust game would never repay the trust of the investor. The investor, in turn, will realize this and should therefore never invest in the trustee in the first place. However, although game theory predicts no transactions to occur, behavioral experiments have consistently shown that players in the trust game do in fact engage in social interaction: the majority of investors send money to the trustees and most of the time this trust is reciprocated with payments back to the investor [3].

In a standard laboratory prisoner’s dilemma game, game theoretic predictions are also usually violated. The prisoner’s dilemma game is quite similar to the trust game, except that in the prisoner’s dilemma game both players simultaneously choose whether to cooperate with the other or not, without knowing their partner’s choice. The payoff for both players depends on the interaction of their choices, such that a player’s individual earnings are highest when he or she defects and the other cooperates, with the cooperating player receiving only a small amount. Mutual cooperation, in contrast, is rewarding for both players, although individual earnings are lower than during unilateral defection. Game theory would predict that in the prisoner’s dilemma game both players should decide to defect, but again, this appears to be not in accordance with the experimental data, as behavioral experiments typically find mutual cooperation occurring about half of the time. Given that standard game theoretic models are often inaccurate at describing individual decision making, how can neuroeconomic studies shed light on why human behavior appears at odds with classical economic predictions?

One potentially interesting contribution may come from how the brain perceives and computes the intentions of others, suggesting that we value not only the monetary outcome of a social interaction, but are also concerned with the message our partner conveys by his or her behavior [13,14]. Several neuroimaging studies have shown a consistent network of areas to be involved during reciprocal exchange in both the trust game and prisoner’s dilemma game [11,15,16]. For example, one of the earliest neuroeconomic studies on the trust game found that activity in the paracingulate cortex was higher when subjects played the game with another human being as opposed to a computer opponent [16] (Fig. 1a). Increased activation in this same region was also found in a following MRI study using the trust game [15]. In this study, trust building was examined by having several pairs of strangers play a nonanonymous multiround trust game, while alternating their roles as trustor and trustee. Both players were each in a separate MRI scanner, with both brains scanned simultaneously during the game. Using within- and between-brain analyses, the results showed that the paracingulate cortex was particularly active during the first phase of the multiround game for players that reciprocated their partner’s decision to trust. In later stages of the game, the activity in this region decreased for this group of reciprocators. This result suggested that the paracingulate cortex is important during the early trust building phases of a relationship and that it plays less of a role during the maintenance stage of trust.

Other areas that have been consistently associated with reciprocal exchange are the posterior superior temporal sulcus and the temporal-parietal junction (Fig. 1b). Together with the paracingulate cortex, these regions have been strongly implicated in the cognitive capacity of perspective taking, an ability often termed as theory of mind [17,18]. The concept of theory of mind is defined as the understanding that others have of their own intentions and individual perspective of the world, and that these may differ from your own. The finding that regions related to theory of mind are involved in reciprocal exchange games suggests that players in the trust game and prisoner’s dilemma game attempt to infer the strategy of their opponent and subsequently act on the belief they have about their partner’s intentions. Therefore, if we perceive that a partner treats
Figure 1. Brain regions involved in decision making: (a) PCC, paracingulate cortex; VMPFC, ventromedial prefrontal cortex; \(x = 46\); (b) TPJ, temporal–parietal junction; PSTS, posterior superior temporal sulcus; \(x = 7\).

us badly when they have the opportunity to do otherwise, we may punish that person. However, if we believe that our partner had no choice in their behavior, we may decide to be more lenient [19]. Furthermore, these regions also seem to play a role in people’s ability to determine whether the actions of others are meaningful. Many neuroimaging studies have found that game interactions with human partners produce stronger activity in theory of mind regions than identical interactions with purported computer partners [16,20,21].

Interestingly, areas important for theory of mind may also be involved in the decision to behave selfishly or altruistically. A study on altruism found that activity in the posterior superior temporal cortex correlated with individual’s tendency to engage in helpful behavior [22]. Of course, in addition to the cortical network related to theory of mind, other areas may be relevant for reasoning about the intentions of others. For example, the cingulate cortex seems to represent information about which player in a trust game is responsible for a specific outcome (themselves or their partner), as its activity correlates with the agent-responsibility (“me” or “not me”) of an action [23]. Future research is expected to increase our understanding on the neural mechanism of theory of mind even further by investigating what additional areas may be involved in the network as well. Furthermore, our knowledge of the specific functions of these regions is still limited. This situation will improve as more studies in this domain will be conducted.

Social Reward

In addition to the involvement of brain structures related to theory of mind, research on the trust game and prisoner’s dilemma game have indicated another area that is critical to social decision making: the striatum [7,10,11,24] (Fig. 2a). The striatum is located in the center of the brain and is a major input station of midbrain dopamine neurons that play a primary role in the processing of rewards. Single-cell recordings in nonhuman primates have demonstrated that dopamine neurons in the striatum track reward prediction errors, with unexpected rewards increasing the firing rate of dopamine cells and the omission of expected rewards leading to a decrease in the firing rate [25,26]. This reinforcement–learning mechanism is thought to be important for the learning of reward values of stimuli in the environment, thereby allowing for the improvement of choices over time.

In parallel to these primate studies, the striatum in humans is known to be involved in reward processing. Using neuroimaging and PET, researchers have discovered that this area responds not only to primary rewards such as liquids, foods, and sexual stimuli [27–30] but also to money [31,32] or more abstract rewards like reputation or
status [33,34]. Importantly, reward-related activity has also been associated with social decision-making behavior. Several neuroimaging studies have demonstrated that this area responds to the decision of one’s partner to either reciprocate or not reciprocate cooperation [7,10,11,24]. For example, in the prisoner’s dilemma game, the striatum is activated when cooperation is reciprocated and deactivated if not reciprocated [10,11].

In addition, the striatum appears to update the trustworthiness of a partner, as activity in this area has been associated with increased cooperation or reciprocity in subsequent rounds of a prisoner’s dilemma [10] or trust game [7]. A further study using the trust game paradigm showed that reward prediction errors in the striatum are significantly dependent on prior experiences, such as the moral reputation of players [24]. Providing investors with a fictional profile of the moral reputation of their partners (either morally good, bad, or neutral) before the start of a trust game resulted in reduced striatal activity when people interacted with morally good or bad partners, but not with neutral partners. Presumably this modulation of striatal activity was absent when players interacted with neutral partners because the behavior of the neutral partners was more unpredictable, as no prior information was given on their reputation. This finding demonstrates that prior beliefs may influence the processing of social feedback information, thereby guiding future cooperation.

Reward-related activity in social decision making has been associated not only with positive, mutual cooperation. A study using PET showed that signals in the striatum were also enhanced when players punished free riders (nonreciprocators that benefit from others in a trust game), suggesting that players may derive satisfaction from punishing defectors [35]. In addition, studies have reported activity in reward networks when participants engaged in charitable donations at a personal cost to themselves [36,37]. For example, a recent study [38] provided participants with an endowed amount of money and then observed them using fMRI while they engaged in a variety of donation situations. The researchers found that reward areas were activated when participants observed a donation to a charity and, importantly, that this activation was enhanced when the donation was made voluntarily.

These neuroeconomic results are important for game theoretic approaches, as they show that regions underlying social decision making overlap with classical reward circuitry, thereby suggesting that cooperative and reciprocal behavior may be experienced directly as rewarding by players, independent of any monetary reward. This provides valuable evidence that social motivations are
indeed important when engaging in social decision making, and that we need to consider rewards other than money when constructing models of behavior in these contexts. Furthermore, the results on charitable donations are intriguing as they have relevant implications for decision making at a societal level. The fact that mandatory transfers to a charity elicit reward-related activity suggests that even taxation may produce satisfaction in certain situations.

The Role of Emotions

In addition to research on strategizing and social reward, another fruitful direction within neuroeconomics is the growing insight that emotions play a crucial role in social decision making. This idea is in contrast to classical models of economic decision making, which assume that decisions are the results of a rational and deliberate evaluation of the available choice options. However, although these models have provided a strong foundation for the development of central decision-making theories, more recent research showed that this rationality assumption is often violated [39, 40].

Early work in this domain showed that patients with prefrontal brain damage and associated emotional deficits performed worse on a gambling task than healthy participants [41]. The area that was damaged in these patients was the ventromedial prefrontal cortex, a region that later was found to be essential for the activation and successful integration of emotion-related memories in the anticipation of future consequences of actions [42] (Fig. 1a). Building from this pioneering research, researchers have continued to discover that emotions play a vital role in determining decisions.

A standard game that is frequently used to illustrate that emotions have an effect on decision making is the ultimatum game. In this well-studied task, two players are required to split a sum of money that is provided by the experimenter. One of the players is deemed the proposer, and the other the responder. The proposer is instructed to specify the division of the money between the two players, and is allowed to make any split he or she wants. The responder can then accept or reject the offer; accepting means the money is split as proposed and rejecting means that neither player receives anything. Since game theory predicts people are motivated purely by self-interest, the standard game theoretic solution for the responder is to accept any amount that is offered. The proposer should anticipate this and therefore offer the responder the smallest amount possible. However, a considerable amount of behavioral research showed that this prediction is at odds with observed behavior in the laboratory and the field: proposers generally propose an equal split and responders reject offers that are 20% of the total amount or less [43]. This very robust finding indicates that people’s behavior in the ultimatum game appears not to be motivated solely by financial self-interest.

Neuroeconomic studies have attempted to specify the processes underlying the reactions of responders to unfair offers in the ultimatum game. An early study by Sanfey and colleagues [44] found that activity in the anterior insula correlated with the degree of unfairness of an offer (Fig. 2a). In addition, this region was more active for offers made by a human partner as opposed to that by a computer partner. In terms of the responder’s decision, activation in this area predicted the player’s decision to either accept or reject the offer, with a higher level of activity for rejections than for acceptances. The activation of the insula in the ultimatum game is particularly interesting in light of its suggested association with negative emotional states. Previous studies showed that activation in this area is consistently seen in studies to pain and disgust, and to autonomic arousal [45, 46]. On the basis of these findings, anterior insula activity in the ultimatum game can be interpreted as a reflection of the responder’s negative emotional state to an unfair offer.

Another brain region that was engaged during unfair offers in the ultimatum game was the right dorsolateral prefrontal cortex, a region that, in contrast to the anterior insula, usually has been associated with cognitive processes, such as the implementation of goal-directed behavior (Fig. 2b). Activation in this area was greater than activation in the insula when responders accepted unfair
offers, and less when they rejected these offers. To investigate the function of the dorsolateral prefrontal cortex in the decision to accept or reject an unfair offer, activity in this brain region has been disrupted using TMS [47,48]. After receiving TMS at the right dorsolateral prefrontal cortex, the acceptance rate of unfair offers increased, suggesting that the dorsolateral prefrontal cortex plays a causal role in the implementation of fairness-related behavior.

The finding of separable neural correlates for decisions based on emotion (rejections of an unfair offer) and deliberation (acceptance of an unfair offer) suggests that there is no unitary system at the neural level that is in control of decision making. Indeed, several neuroimaging studies showed that emotional processes seem to reliably engage a different set of brain structures than cognitive processes [49,50]. This distinction provides some support for dual-process theories, which hypothesize that decision making is the result of an interaction between an automatic, affective system and a more controlled, deliberative system. However, explaining decision making in terms of the interaction between different subsystems violates traditional economic theory, as standard economic models typically describe human behavior as being governed by a unitary process. Future economic models of decision making could usefully take into account the neuroeconomic evidence for the differential processing of emotional and deliberative information.

THE FUTURE

Many of our most important decisions are made in a social context. In the preceding sections we presented an overview of how neuroeconomics has contributed thus far to the understanding of social decision-making behavior. These findings have demonstrated that the brain has the ability to rapidly recognize and act on the intentions of others, that key regions involved in reward-based learning also underlie choices in social settings, and that emotions play a crucial role in social decision-making situations. Despite these advances, the current state of knowledge regarding the mechanisms underlying social decision making is still rather limited. However, owing to the explosion of research in this field, and to continual refinement and advances in neuroscientific methodology, this area of research offers much promise in the years to come.

One potentially fruitful avenue of research is the investigation of how parameters of decision processes are represented in the brain using a formal modeling approach [51,52]. By connecting the parameters of formal (game theoretic) models with neural activity, a more precise characterization of the mechanisms underlying decision making can be described. In turn, by examining whether correlates of game theory parameters are present or absent in the brain, this approach may provide additional constraints on models that aim to predict social decision-making behavior.

Future studies are also likely to benefit from advancements in neuroscientific methods. For instance, fMRI measurement techniques are being improved continuously. The traditional, and still most common, approach to investigate brain function is the localization of function to specific isolated brain areas. However, a more recent way of examining neural function is to look at the structural and functional connectivity between brain areas [53]. Structural connectivity methods, such as diffusion tensor imaging, aim at inferring with the anatomical connectivity between brain regions by tracking bundles of white matter fibers linking cortical areas. Functional methods like Granger causality mapping or dynamic causal modeling investigate the effective relationship between brain areas by modeling the direct or indirect influence of one brain area over another (possibly more spatially remote) brain area. Currently, these brain connectivity methods are becoming increasingly popular and neuroeconomic data generated by this alternative approach can yield useful insights into the neural networks that are relevant for decision making.

A final interesting direction for future research is to investigate the role of hormonal and genetic factors in decision making. For instance, oxytocin, a peptide that functions both as a neurotransmitter and as a hormone,
appears to modulate a broad profile of human social behaviors. In a trust game, intranasal administration of this peptide increased the amount of money transferred by the investor [8]. A related neuroimaging study showed that this change in trusting behavior was associated with a modulation of activity in the neural systems mediating both fear and reward processing [54]. In addition, higher levels of oxytocin have been associated with more generous offers toward strangers in an ultimatum game [55,56]. Another hormone potentially of importance for social decision making is testosterone. A recent study showed that men high in testosterone were more likely to reject relatively low offers in an ultimatum game [57]. Subsequent research also examined whether testosterone levels affect the decision of the proposer. However, as is often the case in early studies contrasting findings have been reported, as one study found that artificially raised testosterone levels decreased the generosity of proposers [58], whereas another study found opposite results [59]. Researchers have also begun to examine whether genetic variation within people can explain individual differences in social decision-making behavior. Results from a twin study of the heritability of ultimatum game behavior showed that more than 40% of the variation of responder’s rejection rate was explained by genetic effects. Furthermore, studies of oxytocin and vasopressin—a peptide with a structure very similar to that of oxytocin—have shown that the genetic polymorphisms for oxytocin and vasopressin receptors are associated with the allocation of funds in economic games that measure altruistic behavior [60–62].

Clearly, the future looks promising for a neuroeconomic approach of social decision making. Exploration of new research domains and the development of more advanced techniques offer fruitful opportunities for the study that combines the mathematical models of game theory with techniques of modern neuroscience. This approach will benefit the predictive accuracy of economic models by constraining their parameters based on neural substrates as well as further our knowledge of how the brain functions.

REFERENCES

14. Rudoy JD, Paller KA. Who can you trust? Behavioral and neural differences between


60. Israel S, Lerer E, Shalev I, et al. Molecular genetic studies of the arginine vasopressin 1a receptor (AVPR1a) and the oxytocin receptor (OXTR) in human behaviour; from autism to altruism with some notes in between. Prog Brain Res 2008;170:435–449.
