Goal-directed, habitual and Pavlovian prosocial behavior

*Filip Gęsiarz and Molly J. Crockett **

Department of Experimental Psychology, University of Oxford, Oxford, UK

Although prosocial behaviors have been widely studied across disciplines, the mechanisms underlying them are not fully understood. Evidence from psychology, biology and economics suggests that prosocial behaviors can be driven by a variety of seemingly opposing factors: altruism or egoism, intuition or deliberation, inborn instincts or learned dispositions, and utility derived from actions or their outcomes. Here we propose a framework inspired by research on reinforcement learning and decision making that links these processes and explains characteristics of prosocial behaviors in different contexts. More specifically, we suggest that prosocial behaviors inherit features of up to three decision-making systems employed to choose between self- and other-regarding acts: a goal-directed system that selects actions based on their predicted consequences, a habitual system that selects actions based on their reinforcement history, and a Pavlovian system that emits reflexive responses based on evolutionarily prescribed priors. This framework, initially described in the field of cognitive neuroscience and machine learning, provides insight into the potential neural circuits and computations shaping prosocial behaviors. Furthermore, it identifies specific conditions in which each of these three systems should dominate and promote other- or self- regarding behavior.

Keywords: model-based, model-free, Pavlovian, reinforcement learning, dictator game, prosocial behavior, altruism, warm-glow

The existence of prosocial behaviors—actions that increase the welfare of others, often at cost to oneself—remains an enduring scientific puzzle. At a first glance such behaviors are inconsistent with the axiom of rational self-interest in neo-classical economics, the law of natural selection in evolutionary biology and the law of effect in behavioral psychology. Nevertheless, prosocial behaviors are widespread across cultures and also found in the animal kingdom (Waal, 1997; Henrich et al., 2001; Engel, 2011). One persisting set of questions concerns the extent to which such behaviors are guided by an "altruistic" motivation to improve the welfare of others. For decades, scientists have debated whether altruistic motivation even exists, and if so, whether it is "rational" in the sense of satisfying real preferences, or rather is a by-product of our evolutionary history. We suggest that to answer both of these questions it is necessary to examine different motivations, and the prosocial behaviors they give rise to, in terms of their underlying cognitive and neural mechanisms.

Here we will show that many theories about the causes of prosocial behaviors can be organized and integrated under a reinforcement learning and decision-making (RLDM) framework, initially developed in the field of cognitive neuroscience and machine learning (Sutton and Barto, 1998; Daw et al., 2005; Dayan, 2008; Dolan and Dayan, 2013). We will argue that this scheme not only streamlines the seemingly heterogeneous landscape of motivations driving prosocial behaviors, but also provides insight into the mechanisms governing them. In a broader context, this proposition also complements recent suggestions that an RLDM framework can help explain patterns of

moral judgments (Crockett, 2013; Cushman, 2013) and elucidate computations underlying social cognition (Dunne and O'Doherty, 2013).
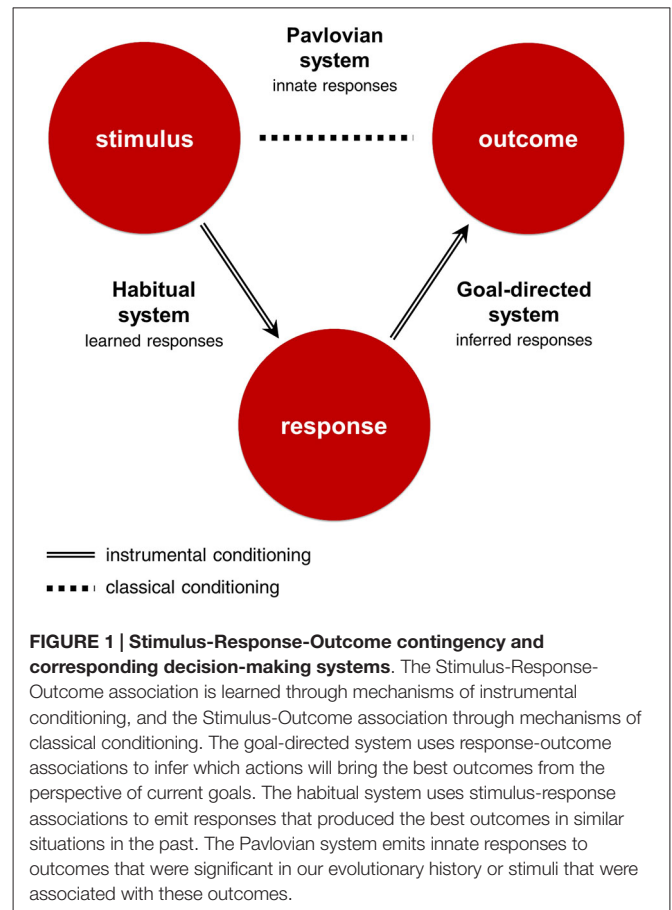
As prosocial behaviors can be expressed in many ways and describing them all is beyond the scope of this paper, we will focus here on sharing, consoling, helping and cooperating. To tackle the problem more formally, we will attempt, where possible, to use examples from game theory—most notably the *Dictator Game*, in which a participant receives a certain endowment and must decide whether to transfer some portion of it to another participant (Forsythe et al., 1994). We will start our considerations with a brief outline of the RLDM framework and its underlying computations. Subsequently, we will consider how three decision systems described by it, either in isolation or through interacting with one another, can give rise to different characteristics of prosocial behavior.

The RLDM Framework

The RLDM framework addresses the problem of how artificial agents should make choices and learn from interactions with the environment to achieve some goal (Sutton and Barto, 1998). It was built on the Markov decision processes framework, according to which every decision-making problem can be decomposed into four elements: the agent's situation (state), which defines currently available outcomes; the agent's choices (actions), which define currently available behaviors; the agent's goal (reward function), which defines how rewarding given outcomes are, and finally the model of the environment (transition function), which defines how given choices lead to certain situations (Sutton and Barto, 1998; van Otterlo and Wiering, 2012). This formalization has been used in three classes of algorithms aiming to optimize decision-making: *model-based planning*, which infers the best decisions from knowledge of the environment; *model-free learning*, which learns the best decisions from the outcomes of past actions; and *a priori programming*, which defines the best decisions for each situation beforehand, for example on the basis of performed simulations (Sutton and Barto, 1998; van Otterlo and Wiering, 2012).

Reinforcement learning algorithms, in principle, can be employed in any domain, and each of them, given enough information, can prescribe an optimal policy for a wide range of problems. One could speculate that such universal tools would be advantageous for any organism struggling for survival and therefore their emergence should be promoted by evolution. Indeed, a large body of evidence suggests that similar algorithms are present in the mammalian brain and are embedded in the *goal-directed*, *habitual*, and *Pavlovian* decision-making systems (Daw et al., 2005; Dayan, 2008; Rangel et al., 2008; Balleine and O'Doherty, 2010; Dolan and Dayan, 2013).

All three RLDM systems learn about some part of the stimulus-response-outcome contingency, and use this knowledge to make decisions (**Figure 1**; **Table 1**). The *goal-directed system* uses response-outcome associations to infer which responses will bring the best outcomes from the perspective of current goals. It can be characterized as deliberate, dominating at the beginning of learning, dependent on working

FIGURE 1 | Stimulus-Response-Outcome contingency and corresponding decision-making systems. The Stimulus-Response-Outcome association is learned through mechanisms of instrumental conditioning, and the Stimulus-Outcome association through mechanisms of classical conditioning. The goal-directed system uses response-outcome associations to infer which actions will bring the best outcomes from the perspective of current goals. The habitual system uses stimulus-response associations to emit responses that produced the best outcomes in similar situations in the past. The Pavlovian system emits innate responses to outcomes that were significant in our evolutionary history or stimuli that were associated with these outcomes.

memory and sensitive to sudden changes in motivational states. The *habitual system* uses stimulus-response associations to emit responses that produced the best outcomes in similar situations in the past. It dominates in later stages of learning, is independent from working memory and insensitive to sudden changes in motivational states. These two systems are called 'instrumental' as they use associations learned through actions. In contrast, the *Pavlovian system* emits reflexive responses to outcomes that were significant in our evolutionary history or stimuli that were associated with these outcomes through the mechanisms of classical conditioning. For example, pavlovian system can emit approach reaction to stimuli associated with food and withdrawal reaction to stimuli associated with pain. Importantly, these responses might be highly sophisticated and sensitive to contextual cues, as in the case of a flight reaction to distal threat and a fight reaction to proximal threat (McNaughton and Corr, 2004). Pavlovian responses, unlike those of the instrumental systems, are inborn, inflexible and preprogrammed by evolution. As such, this system is unable to update its responses when they produce undesirable outcomes. Instead, Pavlovian responses are beholden to the evolutionary context in which they evolved. As a result, Pavlovian responses are efficient solutions to a range of situations that were important in our phylogeny, but may sometimes produce counterproductive behaviors when the current environment demands a more tailored response.

TABLE 1 | Properties of three decision-making systems.

Goal-directed system	Habitual system	Pavlovian system
• Employs model-based planning algorithms	• Employs model-free learning algorithms	• Employs a priori programmed solutions
• Deliberate	• Automatic/Learned	• Automatic/Inborn
• Dominating at the beginning of learning	• Dominating in late stages of learning	• Can dominate at all stages of learning
• Dependent on working-memory	• Independent from working-memory	• Independent from working-memory
• Sensitive to sudden changes in motivational states	• Insensitive to sudden changes in motivational states	• Sensitive to sudden changes in motivational states
• Sensitive to consequences of actions	• Insensitive to consequences of actions	• Insensitive to consequences of actions

The RLDM framework shares many similarities with dual-process accounts of judgment and decision making, in which one system is usually described as emotional, intuitive, domain-specific and automatic, and a second system as cognitive, reflective, domain-general and controlled (Stanovich and West, 1998; Evans, 2008). However, neither of these systems can be directly mapped to the RLDM framework because of a few important differences. First of all, the RLDM systems do not distinguish between "emotion" and "cognition"; rather, all of the RLDM systems rely on emotions, in the sense of processing the affective valence of events. Furthermore, the RLDM systems use well-specified algorithms that do not have an equivalent in dual-process frameworks. Finally, the RLDM framework emphasizes a distinction between inferred, learned and inborn responses—one that is often overlooked by other frameworks. Therefore, it can be concluded that, despite some overlap, the RLDM framework is distinct from traditional dual-process accounts in psychology. In the following sections, we will describe the computational properties and neural substrates of the goal-directed, habitual and Pavlovian systems, as well as procedures used to differentiate between them.

The Goal-Directed System

Model-based planning algorithms select the best decision on the basis of available information—extracted, for instance, from task instructions (Daw, 2012). The *tree-search* algorithm is one of the main examples of this approach. It utilizes a model of the environment to simulate the outcomes of each possible sequence of actions and then evaluates the cumulative value of them in the light of current goals (Daw et al., 2005; Daw, 2012). By considering each possible scenario, this approach ensures making an optimal decision. However, it has some limitations. The first problem is that the agent might not have enough information about the environment to foresee the consequences of each action. Computer scientists deal with this issue by adding a component to the above algorithm that infers the unknown contingencies (Littman, unpublished doctoral dissertation). The second problem is intractability—the more potential sequences of actions and the more complex relationships between them, the more probable it is that the agent will not have enough time and computational power to evaluate all possible outcomes. To prevent this, model-based algorithms use heuristics to narrow down the extent of considered scenarios (Daw, 2012). Other approaches propose that model-based planning, rather than investigating the consequences of each action, could also start with the desirable end state and try to infer, for example through

a procedure known as Bayesian model inversion, the actions that could lead to this state (Botvinick and Toussaint, 2012; Solway and Botvinick, 2012).

Employing tree-search and Bayesian model inversion algorithms to behavioral tasks shows that these algorithms share many characteristics of the goal-directed system, including sensitivity to changing circumstances and having an advantage at the beginning of learning over other RLDM systems (Daw et al., 2005; Keramati et al., 2011; Solway and Botvinick, 2012). Perhaps the most surprising common feature of model-based algorithms and the goal-directed system is the slow pace of operation. Serial processing of standard computer processors greatly limits how many sequences of actions can be evaluated in a unit of time. In the brain, which mostly relies on parallel processing (Alexander and Crutcher, 1990), this problem should be much less pronounced. However, taxing participants' working memory with a demanding task impairs functioning of the goal-directed system (Otto et al., 2013), and suggests that, at least in part, the goal-directed system also employs serial processing (Zylberberg et al., 2011).

Although many brain regions underlie the goal-directed system, the dorsolateral prefrontal cortex (DLPFC) stands out as one of its main neural substrates. First, fMRI studies show that the DLPFC is engaged in tasks involving cognitive processes related to model-based computations, such as: forward planning (Kaller et al., 2011; Wunderlich et al., 2012), organizing working memory content (Owen et al., 2005) and updating a model of the environment (Gläscher et al., 2010). Second, single-unit recordings in monkeys' DLPFC show that neurons in this region encode all variables crucial for performing tree-search and Bayesian model inversion algorithms—namely the potential outcomes, actions and goals (Abe and Lee, 2011; Genovesio et al., 2014). Third, modeling of DLPFC activity suggests that it shows some characteristics of serial processing (Yildiz and Beste, 2014). Finally, disrupting DLPFC function using TMS impairs participants' performance in tasks requiring model-based computations (Smittenaar et al., 2013).

The orbitofrontal cortex and anterior caudate nucleus have also been identified as important components of the goal-directed system (Balleine and O'Doherty, 2010; Gläscher et al., 2010). However, recent evidence suggests that these regions might integrate information from all three decision-making systems (Daw et al., 2011; Liljeholm and O'Doherty, 2012; Wunderlich et al., 2012; Lee et al., 2014). For these reasons in further sections we will concentrate on DLPFC and treat its activation as being consistent with an

involvement of the goal-directed system, although we note this assumption should be treated with caution as it represents reverse inference—i.e., the logical fallacy of inferring the involvement of a particular cognitive function from brain region activation, when this region is not engaged exclusively by this cognitive function (Poldrack, 2006; but see: Hutzler, 2014).

The Habitual System

Model-free learning algorithms ignore the model of the environment and instead integrate the history of consequences of a given action into a *cached action value* (Sutton and Barto, 1998; Dayan, 2008). Although there are many different procedures describing this process, here we will focus on the family of actor-critic models that have been inspired by neuroscience (Joel et al., 2002). In essence, in these models the *actor* automatically choses the action with the highest expected value. The *critic*, in turn, evaluates the outcomes of this action and "teaches" the actor how to update the expected value—by adding to the previous expectation a fraction of a *prediction error* (the difference between the actual and expected value). As this algorithm relies on a single cached value, refined incrementally, it is much more computationally efficient than its model-based alternative.

Efficiency of the model-free algorithms comes at a cost: as model-free algorithms require extensive experience to optimize their policies, they are outcompeted by model-based algorithms when rapid changes in the environment invalidate what has been learned so far (Carmel and Markovitch, 1998). This property is related to the insensitivity of the habitual system to sudden changes in motivational states and the gradual transition from goal-directed to habitual control with experience. Both of these features are well illustrated by the example of the devaluation procedure (Gottfried et al., 2003). In this procedure rats are first trained to make an action (such as pressing a lever) to obtain a rewarding outcome, e.g., sweetened water. At some point, the value of water is artificially diminished (i.e., devalued) by pairing it with nausea-inducing chemicals, which makes the previously desired outcome aversive. If the devaluation procedure is carried out early in training, when the habit of pressing the lever is still weak, rats will not perform the action that delivers the now-devalued sweetened water in this situation. But if the devaluation procedure is employed after extensive training, rats will keep pressing the lever in this situation, even though they are no longer interested in the outcome of this action.

Neuroscientific evidence in general supports the actor-critic model as a plausible computational approximation of the habitual system, although details of how it is actually implemented in the brain are still under debate (Dayan and Balleine, 2002; Joel et al., 2002). First, the division between the actor and the critic is mimicked by the dissociation between the action-related dorsal and reward-related ventral parts of the striatum (O'Doherty et al., 2004; FitzGerald et al., 2012). Furthermore, responses of neurons in both of these regions resemble prediction errors (Schultz, 1998; Joel et al., 2002; Stalnaker et al., 2012). Finally, parallel processing in

the striatum (Alexander and Crutcher, 1990; Yildiz and Beste, 2014), its dense connections with sensorimotor cortex (Ashby et al., 2010) and increasing involvement of its dorsal part with training (Tricomi et al., 2009) explains the fast responses of the habitual system, in comparison to its goal-directed counterpart.

The Pavlovian System

Instead of letting an algorithm infer or learn the best policy, one can simply program *a priori* the best action for any given situation and execute it automatically whenever this situation is encountered (van Otterlo and Wiering, 2012). This could be done either on the basis of the programmer's knowledge or algorithmically, for example using the Monte Carlo method, which identifies the best responses by simulating random action sequences in a given environment and averages the value of outcomes for each response in a given situation (Sutton and Barto, 1998). The main shortcomings of this strategy are its specificity and inflexibility. As the variety of situations in the real world is potentially infinite, it is unfeasible to pre-program appropriate responses to all of them, and therefore one has to focus on some subset of events. One solution to this problem is to generalize rules defining when the given action should be executed. However, such generalizations increase the risk of encountering exceptions to the rule, where the triggered action is inappropriate in the given context.

The Pavlovian system bears many similarities to this strategy, as it reflexively executes unconditional approach and withdrawal responses to classes of stimuli that were important in our evolutionary history. Characteristics of cues triggering these unconditional responses were probably determined by relative costs of omissions vs. false alarms in our phylogeny (Schmajuk, 1987; Parker and Smith, 1990). For example, it might have been more adaptive to overgeneralize features triggering reactions to potential threats, as in the case of a startle response induced by suddenness, because omissions could result in death, whereas false alarms merely cost energy. Classical conditioning can be thought of as a mechanism that helps the organism to generalize inborn Pavlovian responses to situations not consistently paired with unconditional stimuli in our evolutionary history, but nevertheless predicting their occurrence in the current environment.

Inflexibility of Pavlovian responses can have maladaptive consequences in certain contexts—classically illustrated in the negative auto-maintenance procedure (Williams and Williams, 1969). In the first phase of this procedure, food is reliably paired with a conditioned stimulus, until the animal starts to approach not only the food, but also the conditioned stimulus. In the second phase, food is delivered only if the animal refrains from approaching the conditioned stimulus. Although in this context approach behaviors bring negative consequences, animals will repeat them for thousands of trials without obtaining any reward (Killeen, 2003). This procedure not only demonstrates the rigidity of the Pavlovian system, but also its strength when it is pitted against the two other systems (Dayan et al., 2006).

The Pavlovian system has many similarities with the habitual system, and in many cases their influences might be hard to distinguish, as both systems involve responses that are automatic, independent of working memory and insensitive to the predicted consequences of actions (**Table 1**). To prove that a response has a Pavlovian rather than a habitual character one has to show that it is inborn, instead of learned. Furthermore, Pavlovian responses are sensitive to the current motivational state of the organism, in contrast to habitual responses (Dayan and Berridge, 2014). Specifically, conditioned stimuli associated with a particular outcome will trigger automatic approach reactions only if the animal is currently in a state in which the associated outcome is rewarding. For example, rats are typically attracted to a lever associated with sucrose solution delivery and repulsed by a lever associated with a highly saline solution delivery (Robinson and Berridge, 2013). However, rats that are injected with a drug mimicking the state of salt deprivation start to express Pavlovian approach responses towards the lever previously associated with saline solution, such as sniffing, grasping and nibbling. These reactions occur even though approaching the lever does not deliver any outcome and therefore has no instrumental value or even has a negative action value because it was previously associated with an aversive outcome.

Importantly, the Pavlovian system can invigorate or inhibit the responses of the instrumental systems—a phenomon known as Pavlovian-to-instrumental transfer (PIT; Talmi et al., 2008; Lewis et al., 2013). Specifically, the presence of appetitive stimuli has been shown in many experiments to invigorate instrumental approach reactions and inhibit instrumental withdrawal reactions (Talmi et al., 2008; Corbit and Balleine, 2011; Huys et al., 2011; Guitart-Masip et al., 2014). For example, Huys et al. (2011) have shown that visual cues previously associated with monetary rewards speeded movement towards the target stimulus, and slowed movement away from the target stimulus. In contrast, visual cues previously associated with monetary losses have been shown to inhibit instrumental approach reactions and invigorate instrumental withdrawal reactions (Huys et al., 2011; Lewis et al., 2013). The precise mechanisms underlying PIT are still not well understood. It has been proposed that PIT could modulate instrumental approach and withdrawal reactions either through increasing the expectation of a specific outcome or increasing positive and negative arousal (Corbit and Balleine, 2005, 2011).

At the neural level, the most important substrates of the Pavlovian system are the amygdala, which is crucial for acquiring associations between conditioned and unconditioned stimuli (Savage and Ramos, 2009), and the ventral striatum, which takes part in processing the value of primary rewards and punishments, as well as the value of conditioned stimuli (Liljeholm and O'Doherty, 2012). Both of these structures also play a crucial role in PIT (Corbit and Balleine, 2005, 2011; Talmi et al., 2008; Lewis et al., 2013). At the level of neurotransmitters, Pavlovian approach reactions have been predominantly associated with dopamine and Pavlovian inhibition with serotonin (Boureau and Dayan, 2011; Crockett et al., 2012; Guitart-Masip et al., 2014).

An RLDM Framework for Prosocial Behavior

Having characterized the three RLDM systems in more detail, it is important to ask why the RLDM framework is suitable for describing and explaining prosocial behaviors. It could be argued that choice between other- and self-regarding acts is just an ordinary decision-making problem for the brain, and therefore it should be resolved by general-purpose decision-making systems. In this scenario, processes underlying prosocial behaviors would face the same challenges as any other decision and in consequence inherit the exact characteristics of whichever system is primarily responsible for them.

An alternative perspective suggests that, due to the importance of social interactions for human survival, selective pressures could have produced dedicated brain circuits responsible for other-regarding acts, such that they could be motivated by unique processes extending beyond reinforcement learning mechanisms (Field, 2004). We do not exclude this possibility; however we argue that a strong separation between decision-making systems and circuits responsible for prosocial behaviors is unlikely in light of the substantial overlap between social and economic decisions on the neural and behavioral level (Ruff and Fehr, 2014). Following the debate about *common currency* in neuroeconomics—according to which the brain makes choices using a single scale that represents the values of options irrespective of the social or non-social nature of stimuli (Levy and Glimcher, 2012; Ruff and Fehr, 2014)—we suggest instead that brain circuits specialized for prosocial behaviors, if such circuits exist, could either be embedded within the general-purpose RLDM systems or constitute an input and output for them.

In the following sections, we will review evidence showing that many instances of other-regarding acts resemble goal-directed, habitual or Pavlovian decisions. Furthermore, we will suggest in what contexts each of these systems should promote or suppress prosocial behaviors from the perspective of reinforcement learning. Future work will need to address to what extent this framework is sufficient to explain the broad array of observed patterns of prosocial behavior and to what extent it needs to be supplemented by other mechanisms.

Goal-Directed Prosocial Behavior

A desire to achieve some goal, through the means of other-regarding acts, is perhaps the most straightforward motivation driving prosocial behaviors. Evolutionary biologists and neoclassical economists proposed that the superordinate goal of all behaviors is to propagate one's own genes and maximize one's own utility, respectively (Hamilton, 1964; Hollander, 1977). Consequently, according to these perspectives, all other-regarding acts are ultimately selfish. Alternative accounts proposed that some people might have genuine preferences for others' welfare or act in accordance with moral principles (Batson, 1994; Fehr and Fischbacher, 2002). In this section we will review how self-interest can motivate prosocial behavior and show that to appreciate the benefits of other-regarding acts, people must simulate the short- and long-term consequences

of their behavior on the basis of knowledge about the environment—an ability constituting a hallmark of the goal-directed system, requiring model-based computations and likely implemented by the DLPFC. Furthermore, we will suggest that the same mechanisms are employed in the pursuit of non-egoistic goals.

The first mechanism through which self-interest can motivate prosocial behavior is direct reciprocity, where helping someone increases the likelihood that they will return the favor (Trivers, 1971). Direct reciprocity has been mostly studied using the repeated prisoner's dilemma, in which two players have to decide whether to cooperate or defect (Rapoport, 1965). If both cooperate, each gets a moderate reward; if both defect, each gets only a small reward. If one defects while the other cooperates, the defector gets a large reward while the cooperator gets nothing.

If the game is played only once, from the perspective of an individual it is always better to defect, because this either exploits the other's cooperativeness or avoids exploitation of the individual. If the game is repeated, however, in the long run mutual cooperation maximizes the outcomes of both players. Therefore, each player has to establish when cooperative moves have a chance of being reciprocated and adjust their strategy accordingly. The most successful strategies ("tit-for-tat") always start with cooperative move and copy responses of the opponent from the previous encounter thereafter (Axelrod and Hamilton, 1981). Furthermore optimal strategy should be also sensitive to the probability of future interactions and switch from the above "tit for tat" behavior to "always defect" when this probability is low (Rand and Nowak, 2013).

Direct reciprocity is common in humans but surprisingly rare in other animals (Clutton-Brock, 2009). One reason for this might be that it requires sophisticated cognitive abilities absent in simpler organisms (Stevens and Hauser, 2004). A well-developed goal-directed system might be one such ability. In the repeated prisoner's dilemma an agent has to resolve a conflict between smaller rewards now, resulting from defection, and cumulatively larger rewards later, resulting from long-term cooperation—a task reportedly hard for animals (Green et al., 1995). The goal-directed system has the capacity to promote optimal strategies for the current situation as it is able to evaluate the cumulative value of outcomes of different action sequences and override automatic responses. Therefore it can choose a tit-for-tat strategy when the probability of future interactions is high, but switch to defection when it is low—a pattern often observed in behavioral experiments (Bó, 2005; Rand and Nowak, 2013). Consistent with the involvement of the goal-directed system in direct reciprocity, holding a belief that one's interaction partner will reciprocate in an iterated prisoner's dilemma, relative to lacking insight into the partner's strategy, is associated with greater activity in the DLPFC (Sakaiya et al., 2013). The same brain region was shown to be engaged in a prisoner's dilemma by prosocial individuals when they decided to defect, as well as in antisocial individuals when they decided to cooperate, suggesting that it might be involved in goal-directed adjustments of dominant behaviors (Rilling et al., 2007).

Another mechanism through which self-interest could motivate prosocial behavior is indirect reciprocity—that is,

gaining personal benefits from having a good reputation (Nowak and Sigmund, 2005). Laboratory experiments show that being publicly generous pays back, as third parties tend to reward those who are kind to others (Wedekind and Braithwaite, 2002; Servátka, 2010). Behaving in line with social norms also improves one's public image (Andreoni and Bernheim, 2009; Bereczkei et al., 2010) and being altruistic increases one's sexual attractiveness (Farrelly et al., 2007; Barclay, 2010). Perhaps the strongest evidence that people are in fact driven by such motivations comes from the studies that eliminate the opportunity to improve one's reputation by making all prosocial acts anonymous, which greatly decreases the willingness to share an endowment (Bereczkei et al., 2010; Franzen and Pointner, 2012; but see: Barmettler et al., 2012). Importantly, prosocial behaviors are performed more vigorously in public only if they signal to the audience intrinsic prosocial motivations; this vigor is diminished if the person could appear to be acting prosocially to obtain external rewards (Ariely et al., 2009). Differential prosocial behavior between public and private conditions can already be observed in 5-year-olds (Engelmann et al., 2012; Leimgruber et al., 2012). Moreover, this effect is sensitive to the features of the observer: 5-year-olds share more resources when the person looking can potentially reward them for good deeds, in comparison to the situation when they cannot, suggesting that this behavior is, at least in part, deliberate and strategic (Engelmann et al., 2013). Such reputation management probably depends on the development of theory of mind, understood as an ability to attribute mental states to others, as it enables individuals to judge how their actions will be evaluated by others. Consistent with this, chimpanzees and children with autism, both characterized by an underdeveloped theory of mind, do not seem to be concerned about their own reputation (Izuma et al., 2011; Engelmann et al., 2012). On the other hand, studies investigating influence of individual differences in theory of mind on prosocial behaviors found mixed results (Edele et al., 2013; Artinger et al., 2014).

How are concerns about one's reputation incorporated into prosocial decisions? We speculate that the goal-directed system treats others' minds as a part of the environment and simulates their contents in order to determine the consequences of one's actions for one's own reputation. In line with this idea, some studies suggest that in economic games the engagement of theory of mind is related to activity in the DLPFC, among other areas (Yoshida et al., 2010), and involves computations similar to tree-search (Yoshida et al., 2008) and Bayesian model inversion algorithms (Baker et al., 2009; Moutoussis et al., 2014). Consistently, disruption of DLPFC by TMS impairs the accuracy of theory of mind (Costa et al., 2008) and diminishes concerns about one's reputation (Knoch et al., 2009).

Avoiding punishments for violating social norms is another factor motivating people to behave prosocially. This problem has been studied using the ultimatum game, in which a proposer decides how to divide a sum of money (e.g., $10) between themselves and another participant, similarly to the dictator game (Güth et al., 1982). However, unlike in the dictator game, the recipient can reject the offer and then both of the participants get nothing. Recipients often reject offers perceived

to be unfair—a behavior interpreted as costly punishment of a fairness norm violation (Oosterbeek et al., 2004; Henrich et al., 2006) or, in the case of some individuals, as spite (Brañas-Garza et al., 2014). Perhaps anticipating this, proposers usually share money to a higher degree than in the standard dictator game. On the other hand, reducing the negative consequences of offer rejection for the proposer proportionally decreases their offers (Handgraaf et al., 2008). Together these findings suggest that there is a strategic component to proposers' prosocial behavior.

Children playing the role of proposer in the ultimatum game dramatically increase their offers as soon as they are able to pass the false-belief task, indicating the development of a theory of mind (Sally and Hill, 2006; Castelli et al., 2010; Takagishi et al., 2010). However, it is not clear if this developmental milestone increases a preference for fairness or the ability to strategically adjust fair behavior to benefit oneself. To extract the purely strategic component of prosocial behavior, researchers have compared offers in the dictator game and the ultimatum game made by the same individual. The difference in offers between these two games rises substantially from early to middle childhood (Sally and Hill, 2006) and is associated with maturation of the DLPFC (Steinbeis et al., 2012), suggesting involvement of the goal-directed system. Consistent with this, prosocial behavior in the ultimatum game, relative to the dictator game, is associated with stronger activation in the right DLPFC, lateral OFC and caudate nucleus (Spitzer et al., 2007). Furthermore, stimulation of the right DLPFC by tDCS not only increases donations in the ultimatum game, but also decreases donations in the dictator game (Ruff et al., 2013), suggesting that the right DLPFC plays a causal role in calculations aiming to maximize personal benefits.

If the norm of fairness was the only factor regulating how much people share in ultimatum games, people should never give more than half of their initial endowment. In reality, however, some subjects do (Chang et al., 2011), which implies the existence of additional motives. It can be argued that compliance with others' expectations, rather than compliance with cultural norms, might be a more sensitive strategy for building a good reputation and avoiding punishments. Consistent with this, in a study by Chang et al. (2011) participants playing the ultimatum game tried to make offers meeting the expectations of the other person, rather than splitting the endowment evenly. Moreover, this behavior was related to activation in the DLPFC, among other regions.

So far we have reviewed studies suggesting that prosocial actions are motivated at least in part by strategic self-interest and likely fall within the purview of a goal-directed RLDM system. Nevertheless, there is also evidence that even in the absence of personal incentives to behave prosocially, some people are still willing to help others (Batson et al., 1999; Franzen and Pointner, 2012). As the goal-directed system enables the pursuit of any goal, one potential explanation for these selfless behaviors is that some people are simply motivated to act in accordance with moral principles.

Several different types of moral values inform human social behavior and there is an ongoing debate about which ones can be considered universal (Haidt, 2007). In the context of

sharing, three values seem to be particularly important: equality, meritocracy and effectiveness (Charness and Rabin, 2002; Fong, 2007; Konow, 2010). People seem to incorporate these values into decisions to share resources, giving more money to the less fortunate, those who deserve it and those for whom the transfers are more effective, respectively (Brañas-Garza, 2006; Dawes et al., 2007; Hsu et al., 2008; Almås et al., 2010). Moreover, some people reject offers favoring themselves over the other person (Blake and McAuliffe, 2011), are more willing to donate money to charities than to students (Konow, 2010) and are willing to pay money to ensure the implementation of the most effective charity option (Null, 2011). Although these studies do not exclude an involvement of egoistic motivations, they clearly show that people are concerned about the consequences of their actions for other people from the perspective of moral principles.

Habitual Prosocial Behavior

Previous work combining the RLDM framework with game theory has demonstrated that simple model-free algorithms, which gradually increase the probability of successful actions and decrease the probability of unsuccessful actions, better describe human behavior than *a priori* programmed optimal strategies in a variety of two-player non-cooperative economic games (Erev and Roth, 1998; Sarin and Vahid, 2001). However, without making any additional assumptions, these same model-free algorithms predict a decrease in cooperation over time in a repeated prisoner's dilemma, in sharp contrast to observed human behavior, which is characterized by an increasing tendency to cooperate over time (Erev and Roth, 2001). Computer simulations suggest that model-free algorithms are able to learn to cooperate in a variety of cooperative games under the assumption that outcomes of cooperation are satisfactory for both partners of interaction, and are guaranteed to do so if in addition cooperation is more satisfactory than actions maximizing one's own payoffs at the cost of the other player (Sarin, 1999; Macy and Flache, 2002). What mechanism could ensure that cooperation is satisfactory for both players and more satisfactory than the maximizing option? Social norms of reciprocity and fairness, creating additional utility from acting according to these norms, could be one possibility (Fehr and Schmidt, 1999). Alternatively, but not exclusively, the goal-directed system could interact with the habitual system and reinforce prosocial actions which fulfill some goals, for example actions that boost one's reputation or are in line with some moral values. Consistently, such actions are often found to be associated with increased activity in ventral and dorsal striatum (Hsu et al., 2008; Izuma et al., 2010; Tricomi et al., 2010). According to the RLDM framework, frequent rewarding of a given action should lead to a gradual transition from goal-directed to habitual control of that action (Daw et al., 2005). Consequently, with extensive experience, certain actions can become automated and valued in and of themselves, irrespective of their consequences. In the following, we will show that many reported observations of prosocial behaviors suggest that these behaviors have features of habits and intrinsically valued actions.

The notion of prosocial habits is similar to the social heuristics hypothesis, according to which other-regarding acts in one-shot anonymous games stem from intuitive processes, shaped by successful strategies in social interactions and internalization of cultural norms (Rand et al., 2014). In line with both accounts, playing a repeated prisoner's dilemma, in which the payoff structure promoted cooperation, was shown to increase other-regarding behavior in a subsequent battery of one-shot economic games, in comparison to a condition where the payoff structure promoted defection (Peysakhovich and Rand, 2013). It is important to note, however, that interpreting this result as evidence for habit acquisition requires making a few assumptions, as in the classic RLDM literature habits are usually tied to a specific situation and action, rather than a general behavioral tendency expressed across different contexts. Although the generalization of actions across situations has been observed in the case of motor habits (Krakauer et al., 2006; Hilario et al., 2012), it is very limited in scope. Therefore, future studies need to clarify if prosocial habits can spill over into novel situations and generalize to similar actions to a much greater extent than motor plans, or if these findings can be explained by other phenomena.

The possibility that some prosocial actions might be habitual and chosen without regard for their consequences fits many findings in behavioral economics. According to public goods theory, if rational individuals were interested in achieving some desirable state of the social environment, then government spending on that cause should diminish their willingness to financially support it, an assumption known as the "crowding out" hypothesis (Steinberg, 1991; Andreoni, 1993). Indeed, some experiments demonstrated this effect by forcefully taking money from a participant's endowment and, in a transparent manner, transferring it to a given cause (Eckel et al., 2005). However, other experimental and field studies, using slightly different procedures, found incomplete crowding out (Andreoni, 1993; Ribar and Wilhelm, 2002). Similarly, satisfying the norm of fairness in a dictator games by providing an equal endowment to the dictator and the recipient does not completely diminish dictators' willingness to share the endowment (Konow, 2010; Korenok et al., 2013). Furthermore, although people are eager to donate money to charities, most are unwilling to spent money to learn which actions support charities efficiently (Null, 2011). These studies suggest that the aid itself is not the main purpose of these acts.

In line with this, some people are willing to donate money to charity even if they know that their actions are completely ineffective. For example, Crumpler and Grossman (2008) introduced a variation of the dictator game in which any action of the dictator was counterbalanced by the experimenter, such that a donation of $10 to charity would result in the experimenter donating nothing, whereas a donation of $0 would result in the experimenter donating $10. Nevertheless, subjects still gave away about 20% of their endowment in this situation, although there was no reason to do so if they were solely motivated by concern about the welfare of the charity. This effect was still present, albeit smaller, when researchers controlled for the concern about the welfare of the experimenter (Tonin and Vlassopoulos, 2014). It is unlikely that the above findings were driven by the expectations of the experimenter as the procedure was double-blind and also because other studies examining dictator giving, performed without the presence of the experimenter, have reported similar proportions of giving (Barmettler et al., 2012).

The above results have been explained in the terms of a "warm glow"—i.e., utility derived from the act of giving itself, regardless of its outcomes (Andreoni, 1990, 1993). Support for a quite literal interpretation of the warm glow hypothesis comes from the studies showing that giving money away to others produces positive feelings (Konow, 2010; Aknin et al., 2013). Although both 7-year-olds and 9-year-olds share their resources in a fair way, only the latter group feels better after doing this—suggesting that warm glow might require some experience to develop (Kogut, 2012). However, there is also some evidence for warm glow in children as young as two years (Aknin et al., 2012). Before this age children have already begun to engage in spontaneous helping behavior (Brownell and Carriger, 1990; Zahn-Waxler et al., 1992; Warneken and Tomasello, 2007; Liszkowski et al., 2008; Brownell et al., 2009; Bischof-Köhler, 2012) and are able to use model-free representations to guide behavior (Klossek et al., 2008), so it is plausible that warm glow effects could rely on the habitual system, but their developmental trajectory remains to be established. Results from experiments investigating warm glow bear striking similarity to those usually seen in the devaluation procedure—that is, persistence in performing an action because of its intrinsic value, despite diminished value of its outcome. Therefore one can speculate that they both in fact describe the same process. Consistently, on the neural level both habitual actions and warm glow-giving engage the ventral and dorsal striatum (Harbaugh et al., 2007).

Another feature shared by habits and some forms of prosocial behavior is automaticity, characterized by effortlessness and rapidness. One of most popular procedures used to study automaticity is working memory load, which is thought to impair the functioning of the goal-directed system and increase reliance on the habitual system (Otto et al., 2013). Schulz et al. (2014) used this manipulation in a series of mini-dictator games, in which participants had to make binary choices between arbitrarily defined equal and unequal divisions of money. They found that working memory load increased the proportion of fair choices. Importantly, this increase was present for all decisions, irrespective of the level of unfairness of the alternative, in sharp contrast to the control condition, where participants' decisions were highly dependent on the degree of advantageous inequity—an effect that mirrors the insensitivity of the habitual system to the consequences of one's actions. Consistently, other studies have found that working memory load also decreased the strategic tendency to defect near the end of the repeated prisoner's dilemma, reflecting blindness of the habitual system to future (Duffy and Smith, 2014).

Conceptualizing some prosocial behaviors as habitual actions can also potentially explain variation in prosocial behavior. Individual differences in prosocial orientation might stem to some extent from varying levels of the automatization of other-regarding acts, emerging due to different personal experiences. In support for this claim, working memory load

enhances prosocial behavior only for individuals with a prosocial orientation measured by questionnaires, but has an opposite effect on individuals with a proself orientation (Cornelissen et al., 2011)—consistent with the notion that prosocial behaviors of the first group might be more habitual and the second group more goal-directed. Studies using an ego-depletion procedure found similar results for proself oriented individuals, and slightly less consistent results for prosocial individuals (Balliet and Joireman, 2010; Halali et al., 2013).

Complementing these findings, some studies have shown that prosocial decisions are faster (Rand et al., 2012; Lotito et al., 2013; Rand et al., 2013) and are increased under time pressure, especially for subjects who do not have experience with anonymous one-shot economic games promoting self-interested acts (Cappelletti et al., 2011; Rand et al., 2012, 2014). However, it is important to note that some studies failed to replicate these results (Tinghög et al., 2013; Verkoeijen and Bouwmeester, 2014), and prosocial decisions concerning harm to others are slower than selfish decisions (Crockett et al., 2014).

We have shown many parallels between habits and prosocial acts on the behavioral level. Do habits and prosocial acts also engage similar neurocomputational mechanisms? Support for the applicability of the critic component of the actor-critic model in the context of prosocial behaviors comes from studies showing that outcomes of social interactions in economic games (Rilling et al., 2004; Zhu et al., 2012), violations of social norms (Klucharev et al., 2009; Xiang et al., 2013), signs of social approval (Jones et al., 2011) and even consequences of transferring money to charity (Kuss et al., 2013) are encoded in the form of reward prediction errors in the ventral striatum, in line with the possibility that this signal is used to update the expected value of other-regarding acts. Much less is known about the dorsal striatum and its role as an actor in this context, although some studies indeed found that activity of this brain part can be predictive of some other-regarding acts (Rilling et al., 2004; de Quervain et al., 2004; King-Casas et al., 2005; Harbaugh et al., 2007).

Separate lines of research focused on the influence of intuition, warm glow and habitual control in promoting prosocial behavior seem to converge in showing that other-regarding acts can be reinforced by experience, automated and have an intrinsic value. Future work will need to assess to what extent these disparate findings are actually characterizing the same process, vs. unique phenomena—an endeavor that can be facilitated by well-defined computational and neural characteristics of the habitual system. An important caveat is that automaticity and independence of responses from working memory are also features of the Pavlovian system, and therefore many of the above findings could be also attributed to Pavlovian control. To resolve this issue, future experiments will need to carefully control for current motivational states and experience with the given type of social interactions, as the habitual system should be insensitive to the former but sensitive to the latter. In the next section, we will discuss the potential contribution of a reflexive Pavlovian system that both complements and competes with the goal-directed and habitual systems for control of prosocial behavior.

Pavlovian Prosocial Behavior

Recent advances in developmental psychology have shown that infants are probably closer to Rousseau's noble savages than Locke's moral blank slates, as they are armed from birth with mechanisms allowing them to evaluate moral acts and favor, in many cases, good over evil (Bloom, 2013). However, beyond judging other's behavior, are infants also predisposed to behave prosocially? In this section we will review evidence suggesting that some other-regarding acts might be inborn and triggered by evolutionary old mechanisms embedded in the Pavlovian system.

First we consider the possibility that some prosocial tendencies expressed early in development might have a flavor of innate Pavlovian reflexes. There is ample evidence showing that children around the age of 15 months start to engage in sharing, cooperating and consoling (Brownell and Carriger, 1990; Zahn-Waxler et al., 1992; Warneken and Tomasello, 2007; Brownell et al., 2009; Bischof-Köhler, 2012). Helping can be observed even earlier, at the age of 12 months (Liszkowski et al., 2008). These behaviors could be driven by a goal-directed system and a desire to increase others' welfare. However, children before the age of 24 months do not seem to choose actions based on the predicted value of their outcomes (Klossek et al., 2008; Kenward et al., 2009), suggesting they are unlikely to engage in prosocial behaviors due to valuing their consequences. Alternatively, early social experiences and interactions with parents could reinforce prosocial behaviors and promote formation of prosocial habits. However, parental encouragement does not increase helping (Warneken and Tomasello, 2013) and external rewards can even hinder it in 20-month-old infants (Warneken and Tomasello, 2008). The last possibility is that prosocial behaviors are driven by some inborn factors. In line with this, researchers have observed similar developmental patterns of sharing and cooperating in early childhood across different cultures (House et al., 2013), as well as examples of helping and consolation in different species, including apes (Warneken and Tomasello, 2006; Romero et al., 2010), rats (Bartal et al., 2011) and birds (Seed et al., 2007).

What innate mechanism could potentially drive prosocial behaviors? Affective empathy constitutes a likely candidate (de Waal, 2008; Bischof-Köhler, 2012). It develops on the basis of emotional contagion—i.e., the automatic matching between one's own emotional state and the state of the perceived other (Preston and de Waal, 2002). Notably, emotional contagion is present from birth and also found in other mammals (Dondi et al., 1999; Langford et al., 2006; Nakashima et al., 2015). When children develop a self-other distinction around the age of 15 months, they also start to be aware that shared feelings originate from the state of the other person and are able to volitionally attend to it or not—an ability that constitutes an essence of affective empathy (Preston and de Waal, 2002; Bischof-Köhler, 2012). From the age of 18 months children are also able to infer the emotional states of others not only from emotional expressions but also from situational contexts (Vaish et al., 2009), implying that from early on we possess sophisticated capabilities of affective perspective taking.

Affective empathy has been associated with other-regarding acts in many studies. First, the occurrence of various prosocial behaviors correlates with the development of a self-other

distinction and, in consequence, with the development of affective empathy in children (Brownell and Carriger, 1990; Zahn-Waxler et al., 1992; Bischof-Köhler, 2012). Second, self-reported measures of affective empathy correlate with various prosocial behaviors in adults (Eisenberg and Miller, 1987). Third, observing another's suffering is a potent motivator of other-regarding acts: rats pull levers to terminate the distress of other rats (Bartal et al., 2011), monkeys refuse to pull a lever delivering food if it also delivers electric shocks to another monkey (Masserman et al., 1964), and humans are willing to swap places with a suffering person receiving shocks (Batson et al., 1987). Finally, impairment of affective empathy may play a causal role in antisocial behavior in psychopathy (Blair, 2005; Shamay-Tsoory et al., 2010).

Why would affective empathy promote prosocial behaviors? Feeling empathy towards a suffering individual is a source of a negative arousal and therefore other-regarding acts could be potentially driven by an instrumental motivation to eliminate it—either by bringing relief to someone or escaping from the source of distress (Cialdini et al., 1987). Consequently, habitual system should reinforce the action that lead to removal of aversive stimulus. However, humans exposed to other's suffering are more willing to help, even when they can avoid the whole situation in an easy and costless way—which suggests that, at least in case of humans, empathy might trigger an approach rather than a withdrawal reaction (Batson et al., 1987; Stocks et al., 2009). This approach reaction is consistent with the empathy-altruism hypothesis, according to which feeling empathic concern for someone in need can evoke a genuine preference for the other's well-being—a claim which has received solid support throughout the years (Batson et al., 1987; Batson, 2011). Importantly, dependence of this reaction on the state of empathic concern is inconsistent with the involvement of habitual system, as this system is insensitive to motivational states.

We speculate that the mechanism described by the empathy-altruism hypothesis has a Pavlovian character. More specifically, we propose that cues signaling harm or need, such as sad faces, may trigger an automatic urge to help, but only if a person is in the appropriate motivational state, that is, feels empathic concern for the other person. In line with this, 10 month old infants do not withdraw from victims of aggression, but instead show a preference for them, in comparison to both neutral objects and aggressors—an effect that might be interpreted as a rudimentary and perhaps inborn form of concern about the other's well-being (Kanakogi et al., 2013).

Further support for the notion of a Pavlovian urge to help triggered by empathic concern comes from experiments demonstrating that inducing empathic concern can eclipse other goals and lead to maladaptive behaviors—just as in the case of the negative auto-maintenance procedure. For example, people make unfair decisions in favor of a person for whom they feel empathy, even when some other person is in greater need (Batson et al., 1995a); they unconditionally cooperate with an empathized target in the prisoner's dilemma, even when the target has already defected (Batson and Moran, 1999; Batson and Ahmad, 2001); and they allocate more money to an empathized target, even at the cost of lower payouts for the whole group and damaging their own reputation (Batson et al., 1995b, 1999). Moreover, empathy induction in one context does not increase willingness to help the empathized target in other contexts—excluding the possibility that this procedure leads to generalized concern about other's well-being (Dovidio et al., 1990). These examples demonstrate a Pavlovian-like inflexibility and specificity of the empathy-induced other-regarding reaction.

If this reaction is indeed Pavlovian, what could be its evolutionary origin? One proposition is that empathic concern stems from an over-generalization of the parental care instinct (Preston and de Waal, 2002; de Waal, 2008; Batson, 2010). Caretaking in mammals has a very strong reflexive component—as illustrated by a study in which both male and female virgin rats, with paralyzed voluntary muscle control, showed nursing behavior when exposed to unfamiliar pups (Stern, 1991). Furthermore, many animals have been observed to adopt unrelated orphans, suggesting that in some cases childlike features might be sufficient to evoke a parental care reflex and altruistic behaviors (Boesch et al., 2010). In line with this, it has been found that people are more likely to help and care for others possessing childlike facial and vocal characteristics, irrespective whether they are children or adults (Keating et al., 2003; Lishner et al., 2008; Glocker et al., 2009). It is possible that, due to some environmental pressures, a Pavlovian system evolved in humans to trigger caretaking reactions to a wider range of stimuli than only to infants. What makes this claim plausible is that humans show signs of alloparenting and cooperative breeding—that is, taking care of children that are not their own and are often genetically unrelated (Burkart et al., 2014). Crucially, cooperative breeding requires increased responsiveness and an attentional bias towards signals of need. These requirements may have predisposed us to feel empathic concern in a broad array of situations (Burkart and van Schaik, 2010). Consistently, across 15 species of primates, the extent of engagement in cooperative breeding is one of the best predictors of other-regarding preferences in social interactions with strangers (Burkart et al., 2014).

In addition to guiding prosocial behaviors directly through inborn reflexes, the Pavlovian system may modulate habitual or goal-directed other-regarding tendencies through Pavlovian-to-instrumental transfer (PIT). We assume that prosocial behaviors have an approach character, and as such can be invigorated by presence of appetitive cues and inhibited by aversive cues. Some preliminary evidence in support of this claim comes from studies measuring reaction times of prosocial decisions. In general, it has been found that other-regarding acts are faster than self-regarding acts in the context of rewards—an effect that was interpreted as evidence for the automaticity of such responses (Rand et al., 2012, 2013; Lotito et al., 2013). However, a recent study has shown that altruistic individuals make slower decisions when they decide for others in the context of punishments (Crockett et al., 2014), suggesting that the difference in reaction times between rewarding and punishing contexts might stem from Pavlovian invigoration and inhibition of instrumental approach reactions.

Aside from modulating the vigor of responding, could PIT also increase the tendency to act prosocially? We suggest that indeed prosocial dispositions could be enhanced by a range of Pavlovian cues triggering approach reactions towards people, either through evoking positive arousal or increasing expectation of positive outcomes—effects which could be also interpreted as changes in mood and inferences about outcomes of social interaction. Happy expressions and direct eye-gaze could be examples of such Pavlovian cues: two-day-old newborns look longer at happy faces, in comparison to fearful and neutral ones (Farroni et al., 2007), and also at faces making direct eye-contact with them, in comparison to the ones with averted gaze (Farroni et al., 2002). These same cues also increase prosocial behaviors later in life: smiling faces increase helping and cooperating in one-shot social interactions (Scharlemann et al., 2001; Guéguen and De Gail, 2003; Reed et al., 2012; Mussel et al., 2013); and pictures of eyes increase prosocial behaviors in anonymous dictator games and charitable donations in field experiments (Haley and Fessler, 2005; Rigdon et al., 2009; Powell et al., 2012; but see: Fehr and Schneider, 2010). As most of these studies focused on happy expressions and compared them to neutral expressions, future work will need to address the question if also other signs of experiencing emotions can work as Pavlovian cues.

Cues of familiarity and similarity might also enhance prosocial tendencies through PIT, as they also trigger reflexive approach reactions: newborns and infants prefer familiar faces (Barrile et al., 1999; Kelly et al., 2005) and 10 month olds prefer individuals with similar tastes to themselves (Mahajan and Wynn, 2012; Hamlin et al., 2013). Attraction towards familiar and similar others probably evolved as a heuristic for identifying and favoring kin—a highly beneficial ability from the perspective of spreading copies of one's genes (Hamilton, 1964; Lieberman et al., 2007). However, these cues also enhance prosocial behaviors in many other situations. For example, seeing a picture or knowing a surname of the recipient in the dictator game increases willingness to share the endowment (Bohnet and Frey, 1999; Burnham, 2003; Charness and Gneezy, 2008); and membership in the same group (Ahmed, 2007; Halevy et al., 2012) or having similar facial features with another person (DeBruine, 2002; Krupp et al., 2008) promotes other-regarding acts in various economic games.

It could be argued, that aggression and urge to punish somebody are an approach reactions and therefore, according to the above account, should also be enhanced by appetitive cues. However, aggression and punishment can have a dual character: either prosocial, as in the case of punishments in the ultimatum game for violating social norms, or antisocial, as in the case of spite. We speculate that prosocial or antisocial nature of these actions provides a higher order context for the Pavlovian system. Consequently, we predict that appetitive cues will invigorate prosocial punishment and will inhibit antisocial punishment. As none of the studies so far has directly tested this hypothesis, future work will need to fill in this gap.

Other findings can also be re-interpreted through the lens of classical conditioning and PIT effects. Earlier we discussed the study by Peysakhovich and Rand (2013), in which repeated play of a prisoner's dilemma in settings promoting defection increased

a general tendency to act in a self-interested manner in other economic games. Involvement of the habitual system in the above findings might be questionable in light of the low generalizability of habits across contexts in other experiments using non-social stimuli (Krakauer et al., 2006; Hilario et al., 2012). An alternative explanation proposes that participants associated defecting anonymous players with a negative feeling through classical conditioning. This negative association could then stain subsequent interactions with other anonymous players in other games through PIT. To test this directly, future work will need to measure physiological reactions to anonymous players, while participants gradually acquire negative association. Supporting this idea, cooperation and defection in the prisoner's dilemma has been shown to increase and decrease, respectively, the likeability of other player's faces, as well as modulate amygdala responses to these faces in subsequent task (Singer et al., 2004).

In this section we have shown that many theories about the causes of prosocial behaviors can be re-interpreted in terms of Pavlovian reflexes and the mechanism of PIT. According to this view, the Pavlovian system can compete with other RLDM systems for behavioral control and trigger automatic prosocial behaviors in response to perceiving signals of need and feeling empathic concern for others. Alternatively, the Pavlovian system could interact with other RLDM systems by enhancing the likelihood and vigor of prosocial acts in the presence of stimuli evoking approach reactions towards other people.

Discussion

In this review we summarized evidence showing how the RLDM framework can integrate diverse findings describing what motivates prosocial behaviors. We suggested that the goal-directed system, given sufficient time and cognitive resources, weighs the costs of prosocial behaviors against their benefits, and chooses the action that best serves one's goals, whether they be to merely maintain a good reputation or to genuinely enhance the welfare of another. We also suggested that to appreciate some of the benefits of other-regarding acts, such as the possibility of reciprocity, agents must have a well-developed theory of mind and an ability to foresee the cumulative value of future actions—both of which seem to involve model-based computations.

Furthermore, we reviewed findings demonstrating that the habitual system encodes the consequences of social interactions in the form of prediction errors and uses these signals to update the expected value of actions. Repetition of prosocial acts, resulting in positive outcomes, gradually increases their expected value and can lead to the formation of prosocial habits, which are performed without regard to their consequences. We speculated that the expected value of actions on a subjective level might be experienced as a 'warm glow' (Andreoni, 1990), linking our proposition to the behavioral economics literature. We also suggested that the notion of prosocial habits shares many features of the social heuristics hypothesis (Rand et al., 2014), implying that the habitual system could be a possible neurocognitive mechanism explaining the expression of social heuristics.

Finally, we have posited that the Pavlovian system, in response to another's distress cues, evokes an automatic approach response towards stimuli enhancing another's well-being—even if that response brings negative consequences. We have also proposed that presence of appetitive and aversive stimuli can increase or decrease the vigor of prosocial reactions through the mechanism of Pavlovian-to-instrumental transfer. Pavlovian-to-instrumental transfer could also be responsible for the enhancing effects of familiarity, similarity, happy expressions and pictures of eyes on prosocial acts.

In addition to organizing a diverse set of findings on patterns of prosocial behavior, the RLDM framework also provides insight into possible sources of individual differences, developmental changes and interspecies variability in prosocial tendencies. Furthermore, by connecting behavioral economics, psychology, cognitive neuroscience, evolutionary biology and machine learning, this scheme opens new avenues of research at the boundaries of these disciplines.

However, explaining prosocial behavior within the RLDM framework is far from complete. There is an ongoing debate concerning the basic neural circuitry of the goal-directed, habitual and Pavlovian system, and researchers have only recently begun to uncover how these systems cooperate and compete with one another (Dolan and Dayan, 2013; Lee et al., 2014). Meanwhile, there is still relatively little work elucidating the neural substrates of prosocial behaviors, and almost none of this research has attempted to explain prosocial behaviors explicitly in terms of RLDM mechanisms. Future work will need to especially focus on the instances of prosocial behaviors which could be under the control of more than one system and utilize paradigms used in classical RLDM literature to disentangle influence of each of the three systems.

Here we mainly focused on the role of the DLPFC and striatum in motivating prosocial behaviors—the former being a crucial hub of model-based computations used for goal-directed behavior, and the latter responsible for the formation of habits and approach reactions towards stimuli. It is important to note that many of the studies cited in this review also reported the involvement of other neural circuits associated with the RLDM framework, such as the orbitofrontal cortex and the amygdala. However, their precise functional role in other-regarding decisions is less clear than the role of the DLPFC and striatum. Furthermore, it is known that brain regions involved in affective processing and social cognition, such as the anterior insula, anterior cingulate cortex, medial prefrontal cortex and temporo-parietal junction, also play a vital role in prosocial behaviors—although traditionally they are not considered to be a part of the RLDM neural circuitry (Singer et al., 2006; Hare et al., 2010; Morishima et al., 2012; Waytz et al., 2012; Smith et al., 2014). Following other authors, we suggest that information encoded by these regions serves as an input for the three decision-making systems used to predict the consequences of one's actions in social situations and compute the values of different states of the world (Phelps et al., 2014; Ruff and Fehr, 2014).

Under the assumption that the three decision-making systems described here indeed govern prosocial behaviors, it is possible

to generate a number of specific predictions that have yet to be tested. First, according to the RLDM framework, the goal-directed system might use heuristics to narrow down the range of considered scenarios, such as discarding action sequences which produce immediate and substantial negative outcomes—a process described as Pavlovian 'pruning' (Huys et al., 2012). Such a process could be responsible for selfish decisions in situations involving immediate personal costs, despite much greater potential social benefits. Therefore it is speculated that costly prosocial behaviors could be enhanced by situating the personal costs later in the action sequence. Second, irrelevant cues evoking approach and withdrawal reactions towards another person could potentially invigorate or inhibit prosocial tendencies towards this person, through the mechanism of PIT (Bray et al., 2008)—a feature that could be used, for example, in fund-raising advertisements. Third, the formation of habits requires in the training phase that the learner experience an instrumental contingency between responses and outcomes; in other words, the learner has to feel that given actions are associated with some positive value (Keramati et al., 2011). Actions that simultaneously bring counterbalanced appetitive and aversive consequences will have a net value close to zero and therefore will be immune to habitization. From this one could predict that costless other-regarding acts will be particularly prone to becoming habits, while prosocial behaviors requiring difficult trade-offs will probably stay under the control of the goal-directed system. Finally, it is well established in the RLDM literature that random schedules of reinforcements, that is reinforcements delivered at unpredictable intervals, lead to rapid habit formation (Derusso et al., 2010). Therefore one could speculate that uncertainty embedded in social interactions is particularly well-suited to automatize prosocial behaviors, as other-regarding acts are not always and not immediately rewarded. Knowledge of which conditions are the most effective in creating habits could potentially be used in designing interventions to promote prosocial tendencies.

We began this review by referring to a question about the true motivation behind prosocial behaviors. Perhaps not surprisingly, dissecting the mechanisms shaping other-regarding acts reveals a blend of altruistic and egoistic motives at their source. What is important, however, is that reinforcement learning mechanisms are able to transform egoistic motivations into prosocial behaviors, as in the case of prosocial habits formed on the basis of repetition of egoistically motivated other-regarding acts, and altruistic motivations into antisocial behaviors, as in the case of empathic concern for one person eclipsing the well-being of people in greater need. These insights, among others described here, demonstrate the potential for this line of research to help improve society.

Acknowledgments

References

Abe, H., and Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* 70, 731–741. doi: 10.1016/j.neuron.2011.03.026

Ahmed, A. M. (2007). Group identity, social distance and intergroup bias. *J. Econ. Psychol.* 28, 324–337. doi: 10.1016/j.joep.2007.01.007

Aknin, L. B., Barrington-Leigh, C. P., Dunn, E. W., Helliwell, J. F., Burns, J., Biswas-Diener, R., et al. (2013). Prosocial spending and well-being: cross-cultural evidence for a psychological universal. *J. Pers. Soc. Psychol.* 104, 635–652. doi: 10.1037/a0031578

Aknin, L. B., Hamlin, J. K., and Dunn, E. W. (2012). Giving leads to happiness in young children. *PLoS One* 7:e39211. doi: 10.1371/journal.pone.0039211

Alexander, G. E., and Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13, 266–271. doi: 10.1016/0166-2236(90)90107-l

Almås, I., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science* 328, 1176–1178. doi: 10.1126/science.1187300

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *Econ. J.* 100, 464–477. doi: 10.2307/2234133

Andreoni, J. (1993). An experimental test of the public-goods crowding-out hypothesis. *Am. Econ. Rev.* 83, 1317–1327.

Andreoni, J., and Bernheim, B. D. (2009). Social image and the 50–50 norm: a theoretical and experimental analysis of audience effects. *Econometrica* 77, 1607–1636. doi: 10.3982/ecta7384

Ariely, D., Bracha, A., and Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *Am. Econ. Rev.* 99, 544–555. doi: 10.1257/aer.99.1.544

Artinger, F., Exadaktylos, F., Koppel, H., and Sääksvuori, L. (2014). In others' shoes: do individual differences in empathy and theory of mind shape social preferences? *PLoS One* 9:e92844. doi: 10.1371/journal.pone.0092844

Ashby, F. G., Turner, B. O., and Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci.* 14, 208–215. doi: 10.1016/j.tics.2010.02.001

Axelrod, R., and Hamilton, W. D. (1981). The evolution of cooperation. *Science* 211, 1390–1396. doi: 10.1126/science.7466396

Baker, C. L., Saxe, R., and Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition* 113, 329–349. doi: 10.1016/j.cognition.2009.07.005

Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69. doi: 10.1038/npp.2009.131

Balliet, D., and Joireman, J. (2010). Ego depletion reduces proselfs' concern with the well-being of others. *Group Process. Intergroup Relat.* 13, 227–239. doi: 10.1177/1368430209353634

Barclay, P. (2010). Altruism as a courtship display: some effects of third-party generosity on audience perceptions. *Br. J. Psychol.* 101, 123–135. doi: 10.1348/000712609x435733

Barmettler, F., Fehr, E., and Zehnder, C. (2012). Big experimenter is watching you! Anonymity and prosocial behaviour in the laboratory. *Games Econ. Behav.* 75, 17–34. doi: 10.1016/j.geb.2011.09.003

Barrile, M., Armstrong, E. S., and Bower, T. G. R. (1999). Novelty and frequency as determinants of newborn preference. *Dev. Sci.* 2, 47–52. doi: 10.1111/1467-7687.00053

Bartal, I. B.-A., Decety, J., and Mason, P. (2011). Empathy and pro-social behaviour in rats. *Science* 334, 1427–1430. doi: 10.1126/science.1210789

Batson, C. D. (1994). Why act for the public good? Four answers. *Pers. Soc. Psychol. Bull.* 20, 603–610. doi: 10.1177/0146167294205016

Batson, C. D. (2010). The naked emperor: seeking a more plausible genetic basis for psychological altruism. *Econ. Philos.* 26, 149–164. doi: 10.1017/s0266267110000179

Batson, C. D. (2011). *Altruism in Humans*. New York, NY: Oxford University Press.

Batson, C. D., and Ahmad, N. (2001). Empathy-induced altruism in a prisoner's dilemma II: what if the target of empathy has defected? *Eur. J. Soc. Psychol.* 31, 25–36. doi: 10.1002/ejsp.26

Batson, C. D., Ahmad, N., Yin, J., Bedell, S. J., Johnson, J. W., and Templin, C. M. (1999). Two threats to the common good: self-interested egoism

and empathy-induced altruism. *Pers. Soc. Psychol. Bull.* 25, 3–16. doi: 10.1177/0146167299025001001

Batson, C. D., Batson, J. G., Todd, R. M., Brummett, B. H., Shaw, L. L., and Aldeguer, C. M. (1995b). Empathy and the collective good: caring for one of the others in a social dilemma. *J. Pers. Soc. Psychol.* 68, 619–631. doi: 10.1037/0022-3514.68.4.619

Batson, C. D., Fultz, J., and Schoenrade, P. A. (1987). Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. *J. Pers.* 55, 19–39. doi: 10.1111/j.1467-6494.1987.tb00426.x

Batson, C. D., Klein, T. R., Highberger, L., and Shaw, L. L. (1995a). Immorality from empathy-induced altruism: when compassion and justice conflict. *J. Pers. Soc. Psychol.* 68, 1042–1054. doi: 10.1037/0022-3514.68.6.1042

Batson, C. D., and Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *Eur. J. Soc. Psychol.* 29, 909–924. doi: 10.1002/(SICI)1099-0992(199911)29:7<909::AID-EJSP965>3.0.CO;2-L

Bereczkei, T., Birkas, B., and Kerekes, Z. (2010). Altruism towards strangers in need: costly signaling in an industrial society. *Evol. Hum. Behav.* 31, 95–103. doi: 10.1016/j.evolhumbehav.2009.07.004

Bischof-Köhler, D. (2012). Empathy and self-recognition in phylogenetic and ontogenetic perspective. *Emot. Rev.* 4, 40–48. doi: 10.1177/1754073911421377

Blair, R. J. R. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718. doi: 10.1016/j.concog.2005.06.004

Blake, P. R., and McAuliffe, K. (2011). "I had so much it didn't seem fair": eight-year-olds reject two forms of inequity. *Cognition* 120, 215–224. doi: 10.1016/j.cognition.2011.04.006

Bloom, P. (2013). *Just Babies: The Origins of Good and Evil.* New York, NY: Random House LLC.

Bó, P. D. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *Am. Econ. Rev.* 95, 1591–1604. doi: 10.1257/000282805775014434

Boesch, C., Bolé, C., Eckhardt, N., and Boesch, H. (2010). Altruism in forest chimpanzees: the case of adoption. *PLoS One* 5:e8901. doi: 10.1371/journal.pone.0008901

Bohnet, I., and Frey, B. S. (1999). The sound of silence in prisoner's dilemma and dictator games. *J. Econ. Behav. Organ.* 38, 43–57. doi: 10.1016/S0167-2681(98)00121-8

Botvinick, M., and Toussaint, M. (2012). Planning as inference. *Trends Cogn. Sci.* 16, 485–488. doi: 10.1016/j.tics.2012.08.006

Boureau, Y.-L., and Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36, 74–97. doi: 10.1038/npp.2010.151

Brañas-Garza, P. (2006). Poverty in dictator games: awakening solidarity. *J. Econ. Behav. Organ.* 60, 306–320. doi: 10.1016/j.jebo.2004.10.005

Brañas-Garza, P., Espín, A. M., Exadaktylos, F., and Herrmann, B. (2014). Fair and unfair punishers coexist in the Ultimatum Game. *Sci. Rep.* 4:6025. doi: 10.1038/srep06025

Bray, S., Rangel, A., Shimojo, S., Balleine, B., and O'Doherty, J. P. (2008). The neural mechanisms underlying the influence of pavlovian cues on human decision making. *J. Neurosci.* 28, 5861–5866. doi: 10.1523/jneurosci.0897-08.2008

Brownell, C. A., and Carriger, M. S. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Dev.* 61, 1164–1174. doi: 10.2307/1130884

Brownell, C. A., Svetlova, M., and Nichols, S. (2009). To share or not to share: when do toddlers respond to another's needs? *Infancy* 14, 117–130. doi: 10.1080/15250000802569868

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., et al. (2014). The evolutionary origin of human hyper-cooperation. *Nat. Commun.* 5:4747. doi: 10.1038/ncomms5747

Burkart, J. M., and van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates? *Anim. Cogn.* 13, 1–19. doi: 10.1007/s10071-009-0263-7

Burnham, T. C. (2003). Engineering altruism: a theoretical and experimental investigation of anonymity and gift giving. *J. Econ. Behav. Organ.* 50, 133–144. doi: 10.1016/s0167-2681(02)00044-6

Cappelletti, D., Güth, W., and Ploner, M. (2011). Being of two minds: ultimatum offers under cognitive constraints. *J. Econ. Psychol.* 32, 940–950. doi: 10.1016/j.joep.2011.08.001

Carmel, D., and Markovitch, S. (1998). Model-based learning of interaction strategies in multi-agent systems. *J. Exp. Theor. Artif. Intell.* 10, 309–332. doi: 10.1080/095281398146789

Castelli, I., Massaro, D., Sanfey, A. G., and Marchetti, A. (2010). Fairness and intentionality in children's decision-making. *Int. Rev. Econ.* 57, 269–288. doi: 10.1007/s12232-010-0101-x

Chang, L. J., Smith, A., Dufwenberg, M., and Sanfey, A. G. (2011). Triangulating the neural, psychological and economic bases of guilt aversion. *Neuron* 70, 560–572. doi: 10.1016/j.neuron.2011.02.056

Charness, G., and Gneezy, U. (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *J. Econ. Behav. Organ.* 68, 29–35. doi: 10.1016/j.jebo.2008.03.001

Charness, G., and Rabin, M. (2002). Understanding social preferences with simple tests. *Q. J. Econ.* 117, 817–869. doi: 10.1162/003355302760193904

Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., and Beaman, A. L. (1987). Empathy-based helping: is it selflessly or selfishly motivated? *J. Pers. Soc. Psychol.* 52, 749–758. doi: 10.1037/0022-3514.52.4.749

Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature* 462, 51–57. doi: 10.1038/nature08366

Corbit, L. H., and Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J. Neurosci.* 25, 962–970. doi: 10.1523/jneurosci.4507-04.2005

Corbit, L. H., and Balleine, B. W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J. Neurosci.* 31, 11786–11794. doi: 10.1523/jneurosci.2711-11.2011

Cornelissen, G., Dewitte, S., and Warlop, L. (2011). Are social value orientations expressed automatically? Decision Making in the Dictator Game. *Pers. Soc. Psychol. Bull.* 37, 1080–1090. doi: 10.1177/0146167211405996

Costa, A., Torriero, S., Oliveri, M., and Caltagirone, C. (2008). Prefrontal and temporo-parietal involvement in taking others Perspective: TMS evidence. *Behav. Neurol.* 19, 71–74. doi: 10.1155/2008/694632

Crockett, M. J. (2013). Models of morality. *Trends Cogn. Sci.* 17, 363–366. doi: 10.1016/j.tics.2013.06.005

Crockett, M. J., Clark, L., Apergis-Schoute, A. M., Morein-Zamir, S., and Robbins, T. W. (2012). Serotonin modulates the effects of Pavlovian aversive predictions on response vigor. *Neuropsychopharmacology* 37, 2244–2252. doi: 10.1038/npp.2012.75

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., and Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proc. Natl. Acad. Sci. U S A* 111, 17320–17325. doi: 10.1073/pnas.1408988111

Crumpler, H., and Grossman, P. J. (2008). An experimental test of warm glow giving. *J. Public Econ.* 92, 1011–1021. doi: 10.1016/j.jpubeco.2007.12.014

Cushman, F. (2013). Action, outcome and value: a dual-system framework for morality. *Pers. Soc. Psychol. Rev.* 17, 273–292. doi: 10.1177/1088868313495594

Daw, N. D. (2012). "Model-based reinforcement learning as cognitive search: neurocomputational theories," in *Cognitive Search: Evolution, Algorithms and the Brain*, eds P. M. Todd, T. T. Hills, and T. W. Robbins (Cambridge, MA: MIT Press), 195–208.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215. doi: 10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioural control. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., and Smirnov, O. (2007). Egalitarian motives in humans. *Nature* 446, 794–796. doi: 10.1038/nature05651

Dayan, P. (2008). "The role of value systems in decision making," in *Better Than Conscious? Implications for Performance and Institutional Analysis*, eds C. Engel and W. Singer (Cambridge, MA: MIT press), 51–70.

Dayan, P., and Balleine, B. W. (2002). Reward, motivation and reinforcement learning. *Neuron* 36, 285–298. doi: 10.1016/s0896-6273(02)00963-7

Dayan, P., and Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: revaluation, revision and revelation. *Cogn. Affect. Behav. Neurosci.* 14, 473–492. doi: 10.3758/s13415-014-0277-8

Dayan, P., Niv, Y., Seymour, B., and Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Netw.* 19, 1153–1160. doi: 10.1016/j.neunet.2006.03.002

DeBruine, L. M. (2002). Facial resemblance enhances trust. *Proc. Biol. Sci.* 269, 1307–1312. doi: 10.1098/rspb.2002.2034

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., et al. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258. doi: 10.1126/science.1100735

Derusso, A. L., Fan, D., Gupta, J., Shelest, O., Costa, R. M., and Yin, H. H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Front. Integr. Neurosci.* 4:17. doi: 10.3389/fnint.2010.00017

de Waal, F. B. M. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59, 279–300. doi: 10.1146/annurev.psych.59.103006.093685

Dolan, R. J., and Dayan, P. (2013). Goals and habits in the brain. *Neuron* 80, 312–325. doi: 10.1016/j.neuron.2013.09.007

Dondi, M., Simion, F., and Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant? *Dev. Psychol.* 35, 418–426. doi: 10.1037/0012-1649.35.2.418

Dovidio, J. F., Allen, J. L., and Schroeder, D. A. (1990). Specificity of empathy-induced helping: evidence for altruistic motivation. *J. Pers. Soc. Psychol.* 59, 249–260. doi: 10.1037/0022-3514.59.2.249

Duffy, S., and Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *J. Behav. Exp. Econ.* 51, 47–56. doi: 10.1016/j.socec.2014.01.006

Dunne, S., and O'Doherty, J. P. (2013). Insights from the application of computational neuroimaging to social neuroscience. *Curr. Opin. Neurobiol.* 23, 387–392. doi: 10.1016/j.conb.2013.02.007

Eckel, C. C., Grossman, P. J., and Johnston, R. M. (2005). An experimental test of the crowding out hypothesis. *J. Public Econ.* 89, 1543–1560. doi: 10.1016/j.jpubeco.2004.05.012

Edele, A., Dziobek, I., and Keller, M. (2013). Explaining altruistic sharing in the dictator game: the role of affective empathy, cognitive empathy and justice sensitivity. *Learn. Individ. Differ.* 24, 96–102. doi: 10.1016/j.lindif.2012.12.020

Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91–119. doi: 10.1037/0033-2909.101.1.91

Engel, C. (2011). Dictator games: a meta study. *Exp. Econ.* 14, 583–610. doi: 10.1007/s10683-011-9283-7

Engelmann, J. M., Herrmann, E., and Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PloS One* 7:e48433. doi: 10.1371/journal.pone.0048433

Engelmann, J. M., Over, H., Herrmann, E., and Tomasello, M. (2013). Young children care more about their reputation with ingroup members and potential reciprocators. *Dev. Sci.* 16, 952–958. doi: 10.1111/desc.12086

Erev, I., and Roth, A. E. (1998). Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* 88, 848–881.

Erev, I., and Roth, A. E. (2001). "Simple reinforcement learning models and reciprocation in the prisoner's dilemma game," in *Bounded Rationality: The Adaptive Toolbox*, eds G. Gigerenzer and R. Selten (Cambridge, MA: MIT Press), 215–231.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annu. Rev. Psychol.* 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629

Farrelly, D., Lazarus, J., and Roberts, G. (2007). Altruists attract. *Evol. Psychol.* 5, 313–329.

Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U S A* 99, 9602–9605. doi: 10.1073/pnas.152159999

Farroni, T., Menon, E., Rigato, S., and Johnson, M. H. (2007). The perception of facial expressions in newborns. *Eur. J. Dev. Psychol.* 4, 2–13. doi: 10.1080/17405620601046832

Fehr, E., and Fischbacher, U. (2002). Why social preferences matter - the impact of non-selfish motives on competition, cooperation and incentives. *Econ. J.* 112, C1–C33. doi: 10.1111/1468-0297.00027

Fehr, E., and Schmidt, K. M. (1999). A Theory of Fairness, Competition and Cooperation. *Q. J. Econ.* 114, 817–868. doi: 10.1162/003355399556151

Fehr, E., and Schneider, F. (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity? *Proc. Biol. Sci.* 277, 1315–1323. doi: 10.1098/rspb.2009.1900

Field, A. J. (2004). *Altruistically Inclined? The Behavioral Sciences, Evolutionary Theory and the Origins of Reciprocity.* Ann Arbor, MI: University of Michigan Press.

FitzGerald, T. H. B., Friston, K. J., and Dolan, R. J. (2012). Action-specific value signals in reward-related regions of the human brain. *J. Neurosci.* 32, 16417–16423a. doi: 10.1523/jneurosci.3254-12.2012

Fong, C. M. (2007). Evidence from an experiment on charity to welfare recipients: reciprocity, altruism and the empathic responsiveness hypothesis. *Econ. J.* 117, 1008–1024. doi: 10.1111/j.1468-0297.2007.02076.x

Forsythe, R., Horowitz, J. L., Savin, N. E., and Sefton, M. (1994). Fairness in Simple Bargaining Experiments. *Games Econ. Behav.* 6, 347–369. doi: 10.1006/game.1994.1021

Franzen, A., and Pointner, S. (2012). Anonymity in the dictator game revisited. *J. Econ. Behav. Organ.* 81, 74–81. doi: 10.1016/j.jebo.2011.09.005

Genovesio, A., Tsujimoto, S., Navarra, G., Falcone, R., and Wise, S. P. (2014). Autonomous encoding of irrelevant goals and outcomes by prefrontal cortex neurons. *J. Neurosci.* 34, 1970–1978. doi: 10.1523/jneurosci.3228-13.2014

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States vs. rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595. doi: 10.1016/j.neuron.2010.04.016

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Gur, R. C., and Sachser, N. (2009). Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults. *Ethology* 115, 257–263. doi: 10.1111/j.1439-0310.2008.01603.x

Gottfried, J. A., O'Doherty, J., and Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107. doi: 10.1126/science.1087919

Green, L., Price, P. C., and Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: control by immediate consequences. *J. Exp. Anal. Behav.* 64, 1–17. doi: 10.1901/jeab.1995.64-1

Guéguen, N., and De Gail, M. (2003). The effect of smiling on helping behaviour: smiling and good samaritan behaviour. *Commun. Rep.* 16, 133–140. doi: 10.1080/08934210309384496

Guitart-Masip, M., Duzel, E., Dolan, R., and Dayan, P. (2014). Action vs. valence in decision making. *Trends Cogn. Sci.* 18, 194–202. doi: 10.1016/j.tics.2014.01.003

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388. doi: 10.1016/0167-2681(82)90011-7

Haidt, J. (2007). The new synthesis in moral psychology. *Science* 316, 998–1002. doi: 10.1126/science.1137651

Halali, E., Bereby-Meyer, Y., and Ockenfels, A. (2013). Is it all about the self? The effect of self-control depletion on ultimatum game proposers. *Front. Hum. Neurosci.* 7:240. doi: 10.3389/fnhum.2013.00240

Halevy, N., Weisel, O., and Bornstein, G. (2012). 'In-group love' and 'Out-group hate' in repeated interaction between groups. *J. Behav. Decis. Mak.* 25, 188–195. doi: 10.1002/bdm.726

Haley, K. J., and Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.* 26, 245–256. doi: 10.1016/j.evolhumbehav.2005.01.002

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *J. Theor. Biol.* 7, 1–16. doi: 10.1016/0022-5193(64)90038-4

Hamlin, J. K., Mahajan, N., Liberman, Z., and Wynn, K. (2013). Not like me = bad infants prefer those who harm dissimilar others. *Psychol. Sci.* 24, 589–594. doi: 10.1177/0956797612457785

Handgraaf, M. J. J., Van Dijk, E., Vermunt, R. C., Wilke, H. A. M., and De Dreu, C. K. W. (2008). Less power or powerless? egocentric empathy gaps and the irony of having little vs. no power in social decision making. *J. Pers. Soc. Psychol.* 95, 1136–1149. doi: 10.1037/0022-3514.95.5.1136

Harbaugh, W. T., Mayr, U., and Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622–1625. doi: 10.1126/science.1140738

Hare, T. A., Camerer, C. F., Knoepfle, D. T., and Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* 30, 583–590. doi: 10.1523/jneurosci.4089-09.2010

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2001). In search of homo economicus: behavioural experiments in 15 small-scale societies. *Am. Econ. Rev.* 91, 73–78. doi: 10.1257/aer.91.2.73

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al. (2006). Costly punishment across human societies. *Science* 312, 1767–1770. doi: 10.1126/science.1127333

Hilario, M., Holloway, T., Jin, X., and Costa, R. M. (2012). Different dorsal striatum circuits mediate action discrimination and action generalization. *Eur. J. Neurosci.* 35, 1105–1114. doi: 10.1111/j.1460-9568.2012.08073.x

Hollander, S. (1977). Adam Smith and the self-interest axiom. *J. Law Econ.* 20, 133–152. doi: 10.1086/466895

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., et al. (2013). Ontogeny of prosocial behaviour across diverse societies. *Proc. Natl. Acad. Sci. U S A* 110, 14586–14591. doi: 10.1073/pnas.1221217110

Hsu, M., Anen, C., and Quartz, S. R. (2008). The right and the good: distributive justice and neural encoding of equity and efficiency. *Science* 320, 1092–1095. doi: 10.1126/science.1153651

Hutzler, F. (2014). Reverse inference is not a fallacy per se: cognitive processes can be inferred from functional imaging data. *Neuroimage* 84, 1061–1069. doi: 10.1016/j.neuroimage.2012.12.075

Huys, Q. J., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., et al. (2011). Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Comput. Biol.* 7:e1002028. doi: 10.1371/journal.pcbi.1002028

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., and Roiser, J. P. (2012). Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput. Biol.* 8:e1002410. doi: 10.1371/journal.pcbi.1002410

Izuma, K., Matsumoto, K., Camerer, C. F., and Adolphs, R. (2011). Insensitivity to social reputation in autism. *Proc. Natl. Acad. Sci. U S A* 108, 17302–17307. doi: 10.1073/pnas.1107038108

Izuma, K., Saito, D. N., and Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *J. Cogn. Neurosci.* 22, 621–631. doi: 10.1162/jocn.2009.21228

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw.* 15, 535–547. doi: 10.1016/s0893-6080(02)00047-3

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., et al. (2011). Behavioural and neural properties of social reinforcement learning. *J. Neurosci.* 31, 13039–13045. doi: 10.1523/jneurosci.2972-11.2011

Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., and Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cereb. Cortex* 21, 307–317. doi: 10.1093/cercor/bhq096

Kanakogi, Y., Okumura, Y., Inoue, Y., Kitazaki, M., and Itakura, S. (2013). Rudimentary sympathy in preverbal infants: preference for others in distress. *PLoS One* 8:e65292. doi: 10.1371/journal.pone.0065292

Keating, C. F., Randall, D. W., Kendrick, T., and Gutshall, K. A. (2003). Do babyfaced adults receive more help? The (cross-cultural) case of the lost resume. *J. Nonverbal. Behav.* 27, 89–109. doi: 10.1023/A:1023962425692

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., et al. (2005). Three-month-olds, but not newborns, prefer own-race faces. *Dev. Sci.* 8, F31–F36. doi: 10.1111/j.1467-7687.2005.0434a.x

Kenward, B., Folke, S., Holmberg, J., Johansson, A., and Gredebäck, G. (2009). Goal directedness and decision making in infants. *Dev. Psychol.* 45, 809–819. doi: 10.1037/a0014076

Keramati, M., Dezfouli, A., and Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7:e1002055. doi: 10.1371/journal.pcbi.1002055

Killeen, P. R. (2003). Complex dynamic processes in sign tracking with an omission contingency (negative automaintenance). *J. Exp. Psychol. Anim. Behav. Process.* 29, 49–61. doi: 10.1037//0097-7403.29.1.49

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83. doi: 10.1126/science.1108062

Klossek, U. M. H., Russell, J., and Dickinson, A. (2008). The control of instrumental action following outcome devaluation in young children aged

between 1 and 4 years. *J. Exp. Psychol. Gen.* 137, 39–51. doi: 10.1037/0096-3445. 137.1.39

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., and Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron* 61, 140–151. doi: 10.1016/j.neuron.2008.11.027

Knoch, D., Schneider, F., Schunk, D., Hohmann, M., and Fehr, E. (2009). Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proc. Natl. Acad. Sci. U S A* 106, 20895–20899. doi: 10.1073/pnas. 0911619106

Kogut, T. (2012). Knowing what I should, doing what I want: from selfishness to inequity aversion in young children's sharing behaviour. *J. Econ. Psychol.* 33, 226–236. doi: 10.1016/j.joep.2011.10.003

Konow, J. (2010). Mixed feelings: theories of and evidence on giving. *J. Public Econ.* 94, 279–297. doi: 10.1016/j.jpubeco.2009.11.008

Korenok, O., Millner, E. L., and Razzolini, L. (2013). Impure altruism in dictators' giving. *J. Public Econ.* 97, 1–8. doi: 10.1016/j.jpubeco.2012.08.006

Krakauer, J. W., Mazzoni, P., Ghazizadeh, A., Ravindran, R., and Shadmehr, R. (2006). Generalization of motor learning depends on the history of prior action. *PLoS Biol.* 4:e316. doi: 10.1371/journal.pbio.0040316

Krupp, D. B., Debruine, L. M., and Barclay, P. (2008). A cue of kinship promotes cooperation for the public good. *Evol. Human Behav.* 29, 49–55. doi: 10.1016/j. evolhumbehav.2007.08.002

Kuss, K., Falk, A., Trautner, P., Elger, C. E., Weber, B., and Fliessbach, K. (2013). A reward prediction error for charitable donations reveals outcome orientation of donators. *Soc. Cogn. Affect. Neurosci.* 8, 216–223. doi: 10.1093/scan/ nsr088

Langford, D. J., Crager, S. E., Shehzad, Z., Smith, S. B., Sotocinal, S. G., Levenstadt, J. S., et al. (2006). Social modulation of pain as evidence for empathy in mice. *Science* 312, 1967–1970. doi: 10.1126/science.1128322

Lee, S. W., Shimojo, S., and O'Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81, 687–699. doi: 10.1016/j.neuron.2013.11.028

Leimgruber, K. L., Shaw, A., Santos, L. R., and Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS One* 7:e48292. doi: 10.1371/journal.pone.0048292

Levy, D. J., and Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* 22, 1027–1038. doi: 10.1016/j.conb. 2012.06.001

Lewis, A. H., Niznikiewicz, M. A., Delamater, A. R., and Delgado, M. R. (2013). Avoidance-based human Pavlovian-to-instrumental transfer. *Eur. J. Neurosci.* 38, 3740–3748. doi: 10.1111/ejn.12377

Lieberman, D., Tooby, J., and Cosmides, L. (2007). The architecture of human kin detection. *Nature* 445, 727–731. doi: 10.1038/nature05510

Liljeholm, M., and O'Doherty, J. P. (2012). Contributions of the striatum to learning, motivation and performance: an associative account. *Trends Cogn. Sci.* 16, 467–475. doi: 10.1016/j.tics.2012.07.007

Lishner, D. A., Oceja, L. V., Stocks, E. L., and Zaspel, K. (2008). The effect of infant-like characteristics on empathic concern for adults in need. *Motiv. Emot.* 32, 270–277. doi: 10.1007/s11031-008-9101-5

Liszkowski, U., Carpenter, M., and Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition* 108, 732–739. doi: 10.1016/j.cognition.2008.06.013

Lotito, G., Migheli, M., and Ortona, G. (2013). Is cooperation instinctive? Evidence from the response times in a public goods game. *J. Bioecon.* 15, 123–133. doi: 10. 1007/s10818-012-9141-5

Macy, M. W., and Flache, A. (2002). Learning dynamics in social dilemmas. *Proc. Natl. Acad. Sci. U S A* 99(Suppl. 3), 7229–7236. doi: 10.1073/pnas.092080099

Mahajan, N., and Wynn, K. (2012). Origins of "us" vs. "them": prelinguistic infants prefer similar others. *Cognition* 124, 227–233. doi: 10.1016/j.cognition. 2012.05.003

Masserman, J. H., Wechkin, S., and Terris, W. (1964). "Altruistic" behaviour in rhesus monkeys. *Am. J. Psychiatry* 121, 584–585. doi: 10.1176/ajp.121.6.584

McNaughton, N., and Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neurosci. Biobehav. Rev.* 28, 285–305. doi: 10.1016/j.neubiorev.2004.03.005

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., and Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron* 75, 73–79. doi: 10.1016/j.neuron. 2012.05.021

Moutoussis, M., Trujillo-Barreto, N. J. P., El-Deredy, W., Dolan, R., and Friston, K. (2014). A formal model of interpersonal inference. *Front. Hum. Neurosci.* 8:160. doi: 10.3389/fnhum.2014.00160

Mussel, P., Göritz, A. S., and Hewig, J. (2013). The value of a smile: facial expression affects ultimatum-game responses. *Judgm. Decis. Mak.* 8, 381–385.

Nakashima, S. F., Ukezono, M., Nishida, H., Sudo, R., and Takano, Y. (2015). Receiving of emotional signal of pain from conspecifics in laboratory rats. *R. Soc. Open Sci.* 2:140381. doi: 10.1098/rsos.140381

Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* 437, 1291–1298. doi: 10.1038/nature04131

Null, C. (2011). Warm glow, information and inefficient charitable giving. *J. Public Econ.* 95, 455–465. doi: 10.1016/j.jpubeco.2010.06.018

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454. doi: 10.1126/science.1094285

Oosterbeek, H., Sloof, R., and van de Kuilen, G. (2004). Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Exp. Econ.* 7, 171–188. doi: 10.1023/b:exec.0000026978.14316.74

Otto, A. R., Gershman, S. J., Markman, A. B., and Daw, N. D. (2013). The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol. Sci.* 24, 751–761. doi: 10.1177/09567976124 63080

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm. 20131

Parker, G. A., and Smith, J. M. (1990). Optimality theory in evolutionary biology. *Nature* 348, 27–33. doi: 10.1038/348027a0

Peysakhovich, A., and Rand, D. G. (2013). Habits of virtue: creating norms of cooperation and defection in the laboratory. Available online at: http://ssrn.com/abstract=2294242

Phelps, E. A., Lempert, K. M., and Sokol-Hessner, P. (2014). Emotion and decision making: multiple modulatory neural circuits. *Annu. Rev. Neurosci.* 37, 263–287. doi: 10.1146/annurev-neuro-071013-014119

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63. doi: 10.1016/j.tics.2005.12.004

Powell, K. L., Roberts, G., and Nettle, D. (2012). Eye images increase charitable donations: evidence from an opportunistic field experiment in a supermarket. *Ethology* 118, 1096–1101. doi: 10.1111/eth.12011

Preston, S. D., and de Waal, F. B. M. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20; discussion 20–71. doi: 10. 1017/s0140525x02000018

Rand, D. G., Greene, J. D., and Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature* 489, 427–430. doi: 10.1038/nature11467

Rand, D. G., Greene, J. D., and Nowak, M. A. (2013). Rand et al. reply. *Nature* 498, E2–E3. doi: 10.1038/nature12195

Rand, D. G., and Nowak, M. A. (2013). Human cooperation. *Trends Cogn. Sci.* 17, 413–425. doi: 10.1016/j.tics.2013.06.003

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., et al. (2014). Social heuristics shape intuitive cooperation. *Nat. Commun.* 5:3677. doi: 10.1038/ncomms4677

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. doi: 10.1038/nrn2357

Rapoport, A. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation.* (Vol. 165). Ann Arbor, MI: University of Michigan Press.

Reed, L. I., Zeglen, K. N., and Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evol. Human Behav.* 33, 200–209. doi: 10.1016/j.evolhumbehav.2011. 09.003

Ribar, D. C., and Wilhelm, M. O. (2002). Altruistic and joy-of-giving motivations in charitable behaviour. *J. Polit. Econ.* 110, 425–457. doi: 10.1086/338750

Rigdon, M., Ishii, K., Watabe, M., and Kitayama, S. (2009). Minimal social cues in the dictator game. *J. Econ. Psychol.* 30, 358–367. doi: 10.1016/j.joep.2009. 02.002

Rilling, J. K., Glenn, A. L., Jairam, M. R., Pagnoni, G., Goldsmith, D. R., Elfenbein, H. A., et al. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* 61, 1260–1271. doi: 10.1016/j. biopsych.2006.07.021

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* 15, 2539–2543. doi: 10.1097/00001756-200411150-00022

Robinson, M. J. F., and Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational "wanting". *Curr. Biol.* 23, 282–289. doi: 10.1016/j.cub.2013.01.016

Romero, T., Castellanos, M. A., and de Waal, F. B. M. (2010). Consolation as possible expression of sympathetic concern among chimpanzees. *Proc. Natl. Acad. Sci. U S A* 107, 12110–12115. doi: 10.1073/pnas.1006991107

Ruff, C. C., and Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* 15, 549–562. doi: 10.1038/nrn3776

Ruff, C. C., Ugazio, G., and Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science* 342, 482–484. doi: 10.1126/science.1241399

Sakaiya, S., Shiraito, Y., Kato, J., Ide, H., Okada, K., Takano, K., et al. (2013). Neural correlate of human reciprocity in social interactions. *Front. Neurosci.* 7:239. doi: 10.3389/fnins.2013.00239

Sally, D., and Hill, E. (2006). The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27, 73–97. doi: 10.1016/j.joep.2005.06.015

Sarin, R. (1999). Simple play in the Prisoner's Dilemma. *J. Econ. Behav. Organ.* 40, 105–113. doi: 10.1016/s0167-2681(99)00044-x

Sarin, R., and Vahid, F. (2001). Predicting how people play games: a simple dynamic model of choice. *Games Econ. Behav.* 34, 104–122. doi: 10.1006/game.1999.0783

Savage, L. M., and Ramos, R. L. (2009). Reward expectation alters learning and memory: the impact of the amygdala on appetitive-driven behaviours. *Behav. Brain Res.* 198, 1–12. doi: 10.1016/j.bbr.2008.10.028

Scharlemann, J. P. W., Eckel, C. C., Kacelnik, A., and Wilson, R. K. (2001). The value of a smile: game theory with a human face. *J. Econ. Psychol.* 22, 617–640. doi: 10.1016/s0167-4870(01)00059-9

Schmajuk, N. A. (1987). Classical conditioning, signal detection and evolution. *Behav. Processes* 14, 277–289. doi: 10.1016/0376-6357(87)90074-x

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

Schulz, J. F., Fischbacher, U., Thöni, C., and Utikal, V. (2014). Affect and fairness: dictator games under cognitive load. *J. Econ. Psychol.* 41, 77–87. doi: 10.1016/j.joep.2012.08.007

Seed, A. M., Clayton, N. S., and Emery, N. J. (2007). Postconflict third-party affiliation in rooks, Corvus frugilegus. *Curr. Biol.* 17, 152–158. doi: 10.1016/j.cub.2006.11.025

Servátka, M. (2010). Does generosity generate generosity? An experimental study of reputation effects in a dictator game. *J. Socio Econ.* 39, 11–17. doi: 10.1016/j.socec.2009.08.006

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., and Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex* 46, 668–677. doi: 10.1016/j.cortex.2009.04.008

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662. doi: 10.1016/s0896-6273(04)00014-5

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469. doi: 10.1038/nature04271

Smith, D. V., Clithero, J. A., Boltuck, S. E., and Huettel, S. A. (2014). Functional connectivity with ventromedial prefrontal cortex reflects subjective value for social rewards. *Soc. Cogn. Affect. Neurosci.* 9, 2017–2025. doi: 10.1093/scan/nsu005

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., and Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron* 80, 914–919. doi: 10.1016/j.neuron.2013.08.009

Solway, A., and Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychol. Rev.* 119, 120–154. doi: 10.1037/a0026435

Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., and Fehr, E. (2007). The neural signature of social norm compliance. *Neuron* 56, 185–196. doi: 10.1016/j.neuron.2007.09.011

Stalnaker, T. A., Calhoon, G. G., Ogawa, M., Roesch, M. R., and Schoenbaum, G. (2012). Reward prediction error signaling in posterior dorsomedial striatum is action specific. *J. Neurosci.* 32, 10296–10305. doi: 10.1523/jneurosci.0832-12.2012

Stanovich, K. E., and West, R. F. (1998). Individual differences in rational thought. *J. Exp. Psychol. Gen.* 127, 161–188.

Steinbeis, N., Bernhardt, B. C., and Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron* 73, 1040–1051. doi: 10.1016/j.neuron.2011.12.027

Steinberg, R. (1991). Does government spending crowd out donations? *Ann. Public Coop. Econ.* 62, 591–612. doi: 10.1111/j.1467-8292.1991.tb01369.x

Stern, J. M. (1991). Nursing posture is elicited rapidly in maternally naive, haloperidol-treated female and male rats in response to ventral trunk stimulation from active pups. *Horm. Behav.* 25, 504–517. doi: 10.1016/0018-506x(91)90017-c

Stevens, J. R., and Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* 8, 60–65. doi: 10.1016/j.tics.2003.12.003

Stocks, E. L., Lishner, D. A., and Decker, S. K. (2009). Altruism or psychological escape: why does empathy promote prosocial behaviour? *Eur. J. Soc. Psychol.* 39, 649–665. doi: 10.1002/ejsp.561

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., and Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *J. Exp. Child Psychol.* 105, 130–137. doi: 10.1016/j.jecp.2009.09.005

Talmi, D., Seymour, B., Dayan, P., and Dolan, R. J. (2008). Human pavlovian-instrumental transfer. *J. Neurosci.* 28, 360–368. doi: 10.1523/JNEUROSCI.4028-07.2008

Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., et al. (2013). Intuition and cooperation reconsidered. *Nature* 498, E1–E2; discussion E2–E3. doi: 10.1038/nature12194

Tonin, M., and Vlassopoulos, M. (2014). An experimental investigation of intrinsic motivations for giving. *Theory Decis.* 76, 47–67. doi: 10.1007/s11238-013-9360-9

Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232. doi: 10.1111/j.1460-9568.2009.06796.x

Tricomi, E., Rangel, A., Camerer, C. F., and O'Doherty, J. P. (2010). Neural evidence for inequality-averse social preferences. *Nature* 463, 1089–1091. doi: 10.1038/nature08785

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57. doi: 10.1086/406755

Vaish, A., Carpenter, M., and Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Dev. Psychol.* 45, 534–543. doi: 10.1037/a0014322

van Otterlo, M., and Wiering, M. (2012). "Reinforcement learning and markov decision processes," in *Reinforcement Learning*, eds M. Wiering and M. van Otterlo (Berlin, Heidelberg: Springer), 3–42.

Verkoeijen, P. P. J. L., and Bouwmeester, S. (2014). Does intuition cause cooperation? *PLoS One* 9:e96654. doi: 10.1371/journal.pone.0096654

Waal, F. (1997). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals.* Cambridge, Mass.: Harvard University Press.

Warneken, F., and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science* 311, 1301–1303. doi: 10.1126/science.1121448

Warneken, F., and Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy* 11, 271–294. doi: 10.1111/j.1532-7078.2007.tb00227.x

Warneken, F., and Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Dev. Psychol.* 44, 1785–1788. doi: 10.1037/a0013860

Warneken, F., and Tomasello, M. (2013). Parental presence and encouragement do not influence helping in young children. *Infancy* 18, 345–368. doi: 10.1111/j.1532-7078.2012.00120.x

Waytz, A., Zaki, J., and Mitchell, J. P. (2012). Response of dorsomedial prefrontal cortex predicts altruistic behaviour. *J. Neurosci.* 32, 7646–7650. doi: 10.1523/JNEUROSCI.6193-11.2012

Wedekind, C., and Braithwaite, V. A. (2002). The long-term benefits of human generosity in indirect reciprocity. *Curr. Biol.* 12, 1012–1015. doi: 10. 1016/s0960-9822(02)00890-4

Williams, D. R., and Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J. Exp. Anal. Behav.* 12, 511–520. doi: 10.1901/jeab.1969.12-511

Wunderlich, K., Dayan, P., and Dolan, R. J. (2012). Mapping value based planning and extensively trained choice in the human brain. *Nat. Neurosci.* 15, 786–791. doi: 10.1038/nn.3068

Xiang, T., Lohrenz, T., and Montague, P. R. (2013). Computational substrates of norms and their violations during social exchange. *J. Neurosci.* 33, 1099–1108. doi: 10.1523/JNEUROSCI.1642-12.2013

Yildiz, A., and Beste, C. (2014). Parallel and serial processing in dual-tasking differentially involves mechanisms in the striatum and the lateral prefrontal cortex. *Brain Struct. Funct.* doi: 10.1007/s00429-014-0847-0 [Epub ahead of print].

Yoshida, W., Dolan, R. J., and Friston, K. J. (2008). Game theory of mind. *PLoS Comput. Biol.* 4:e1000254. doi: 10.1371/journal.pcbi.1000254

Yoshida, W., Seymour, B., Friston, K. J., and Dolan, R. J. (2010). Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* 30, 10744–10751. doi: 10.1523/JNEUROSCI.5895-09.2010

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., and Chapman, M. (1992). Development of concern for others. *Dev. Psychol.* 28, 126–136. doi: 10. 1037/0012-1649.28.1.126

Zhu, L., Mathewson, K. E., and Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proc. Natl. Acad. Sci. U S A* 109, 1419–1424. doi: 10.1073/pnas.1116783109

Zylberberg, A., Dehaene, S., Roelfsema, P. R., and Sigman, M. (2011). The human turing machine: a neural framework for mental programs. *Trends Cogn. Sci.* 15, 293–300. doi: 10.1016/j.tics.2011.05.007