The neurobiology of punishment

Ben Seymour*, Tania Singer‡ and Ray Dolan*

Abstract | Animals, in particular humans, frequently punish other individuals who behave negatively or uncooperatively towards them. In animals, this usually serves to protect the personal interests of the individual concerned, and its kin. However, humans also punish altruistically, in which the act of punishing is personally costly. The propensity to do so has been proposed to reflect the cultural acquisition of norms of behaviour, which incorporates the desire to uphold equity and fairness, and promotes cooperation. Here, we review the proximate neurobiological basis of punishment, considering the motivational processes that underlie punishing actions.

Kin selection
Evolutionary models which predict that animals should be motivated to protect their relatives, to ultimately preserve their genetic inheritance.

Direct reciprocity
So called 'tit-for-tat reciprocity', in which a beneficial act to another individual is likely to be reciprocated by that individual.

Indirect reciprocity
Sometimes to referred to as reputation formation, this involves acquiring a reputation as being beneficial towards others, and therefore being treated favourably by them.

Altruistic punishment
Punishing that involves a selfless personal cost to the punisher, which is never likely to be recovered.

*The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, 12 Queen Square, London WC1X 3BG, UK. ‡Institute of Cognitive Neuroscience, 17 Queen Square, London WC1X 3BG, UK, and the Center for Social Neuroscience and Neuroeconomics, University of Zürich, Blümlisalpstrasse 10 8006 Zürich, Switzerland. Correspondence to B.S. e-mail: bjseymour@gmail.com

Animals, including humans, shape their behaviour on the basis of experience. In social animals, the action of one individual can influence the state of another, and ecological evidence indicates that many species will engage in punishing behaviour to protect the interests of themselves and their kin[1]. In humans, however, punishment also appears to adopt an important role in promoting and preserving cooperation. From an evolutionary perspective, theoretical models have suggested that whereas cooperation in small groups ought to be sustainable by kin selection, direct reciprocity and indirect reciprocity, cooperation in large groups requires the cultural acquisition of cooperative norms of behaviour[2–7]. These norms include the motivation to altruistically punish according to common standards of fairness, and raises the question as to what, in the brain, underlies the proclivity to do so.

Drawing on a current behavioural and neurobiological understanding of motivation, we consider the neurobiological processes that underlie the motivation to punish others, addressing in particular how such behaviours may be acquired. First, we review the basic structure of motivational systems and consider principal types of action, namely innate actions (Pavlovian responses), stimulus–response actions (habits) and goal-orientated actions, and discuss how these actions are learned through experience and, importantly, through observation. This provides an appropriate basis for considering both the types of action that underlie punishing actions, and for illustrating how behaviour is adapted following receipt of punishment. We then discuss evidence concerning whether, and how, punishment mediates and promotes cooperation in social interactions by considering ecological observations in animals and by drawing on game theoretic experiments in humans. Finally, we describe recent human functional MRI (fMRI) studies involving fairness judgements and punishment in econom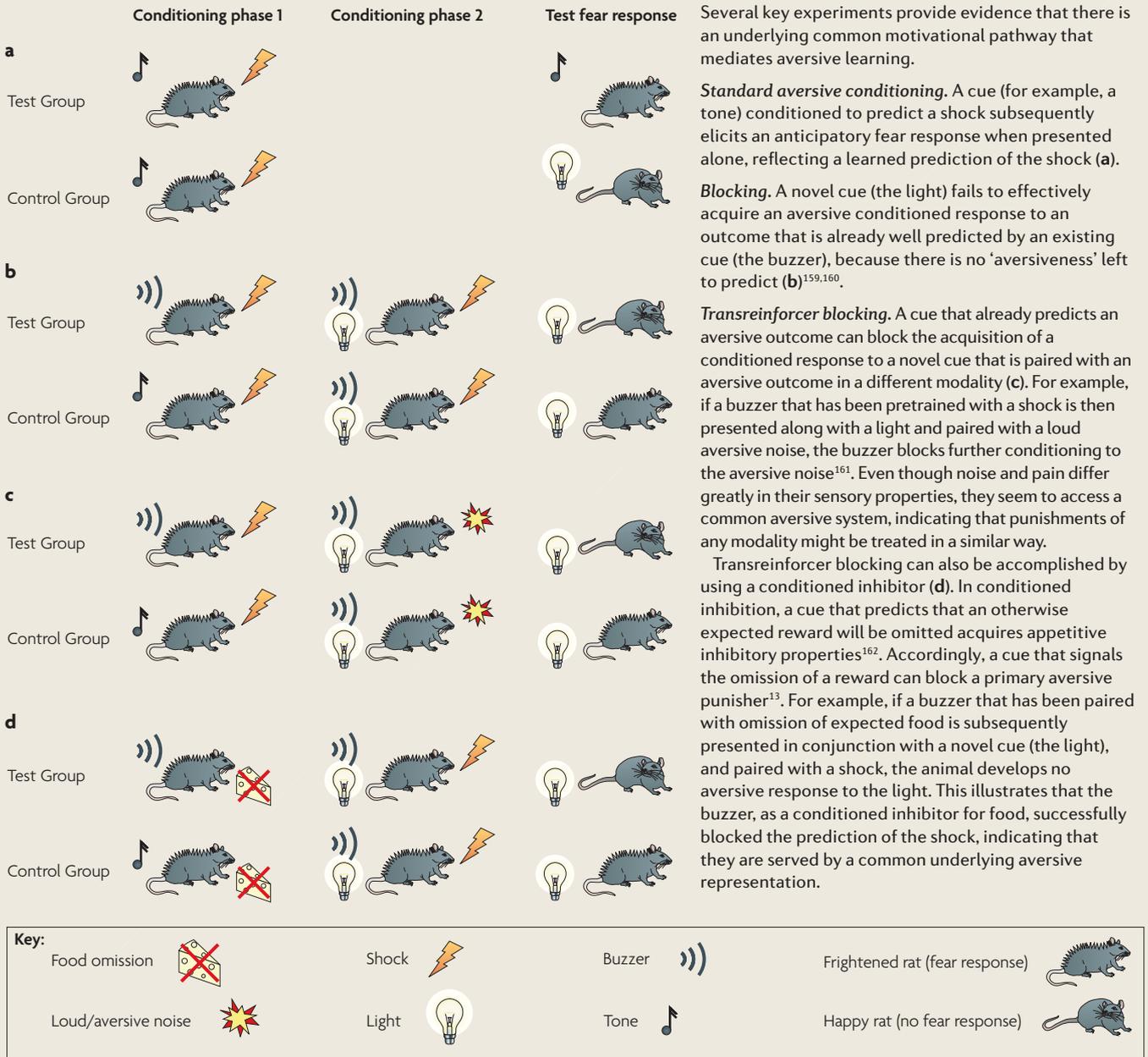ic games. Together with insights into the basic structure of motivational systems, we put forward a neurobiological model of punishment that incorporates impulsive, goal-directed and ultimately altruistic actions.

Motivation, learning and action

Our current understanding of human motivation draws strongly on an extensive animal and human literature in experimental psychology[8–11]. Motivation is characterized by action, either to increase the probability of an outcome (appetitive motivation), or to reduce it (aversive motivation). Thus, a reward can be operationally defined as an event that an animal will expend energy to bring about, whereas a punishment is an event that an animal will expend energy to reduce or avoid. Note that in this way (and typical in the experimental psychology literature), the term punishment is taken to apply to any aversive event, regardless of its aetiology. However, in social contexts (in social psychology and behavioural economics), a punishment is often assumed to refer more specifically to an aversive event administered by another individual. These semantic distinctions aside, any complete motivational account of punishment needs to consider both the propensity to administer, and the impact of receiving, a punishment.

Motivational value. The value of a reward or an aversive event can be considered in terms of an ordinal scale of preference, in which given a choice, higher-valued (or less aversive) outcomes will be selected over lower-valued ones. This concept of value is useful, as it outlines a unitary currency against which events of different modalities can be judged[12]. Furthermore, value seems to be underpinned by basic underlying appetitive and aversive systems[13] (BOX 1). The concept of value naturally extends to incorporate otherwise neutral states or cues

Box 1 | A common aversive motivational system

	Conditioning phase 1	Conditioning phase 2	Test fear response
a Test Group			
Control Group			
b Test Group			
Control Group			
c Test Group			
Control Group			
d Test Group			
Control Group			

Several key experiments provide evidence that there is an underlying common motivational pathway that mediates aversive learning.

Standard aversive conditioning. A cue (for example, a tone) conditioned to predict a shock subsequently elicits an anticipatory fear response when presented alone, reflecting a learned prediction of the shock (**a**).

Blocking. A novel cue (the light) fails to effectively acquire an aversive conditioned response to an outcome that is already well predicted by an existing cue (the buzzer), because there is no 'aversiveness' left to predict (**b**)[159,160].

Transreinforcer blocking. A cue that already predicts an aversive outcome can block the acquisition of a conditioned response to a novel cue that is paired with an aversive outcome in a different modality (**c**). For example, if a buzzer that has been pretrained with a shock is then presented along with a light and paired with a loud aversive noise, the buzzer blocks further conditioning to the aversive noise[161]. Even though noise and pain differ greatly in their sensory properties, they seem to access a common aversive system, indicating that punishments of any modality might be treated in a similar way.

Transreinforcer blocking can also be accomplished by using a conditioned inhibitor (**d**). In conditioned inhibition, a cue that predicts that an otherwise expected reward will be omitted acquires appetitive inhibitory properties[162]. Accordingly, a cue that signals the omission of a reward can block a primary aversive punisher[13]. For example, if a buzzer that has been paired with omission of expected food is subsequently presented in conjunction with a novel cue (the light), and paired with a shock, the animal develops no aversive response to the light. This illustrates that the buzzer, as a conditioned inhibitor for food, successfully blocked the prediction of the shock, indicating that they are served by a common underlying aversive representation.

Key:

Food omission		Shock	Buzzer	Frightened rat (fear response)
Loud/aversive noise		Light	Tone	Happy rat (no fear response)

that predict 'primary' rewards or aversive events. This 'state–outcome' associability is embodied within Pavlovian learning, in which a reliable, predictive pairing of a neutral stimulus or context (the conditioned stimulus) with a reward or aversive event (the unconditioned stimulus) enables an appropriate anticipatory response (the conditioned response) to be elicited when the predictive stimulus or context is encountered in the future[14,15]. For example, being bitten by a particular dog is likely to induce increased heart rate, sweating and fleeing when that dog is encountered subsequently.

The ability to predict reward (or aversive events) embodies self-evident motivational benefits as it allows an animal to deal with future interests. This predictive ability also yields a new set of possible events, namely

those associated with omission of an expected outcome. Accordingly, the omission of an expected aversive event can be rewarding (an aversive inhibitor), and the omission of reward can be aversive (an appetitive inhibitor). Importantly, this seems to be mediated by the same mechanisms as primary rewards and aversive events, respectively (BOX 1). This relationship underpins a basic architecture of motivational systems in which reward and aversive mechanisms oppose each other[13,16]. This 'Konorskian' model consists of underlying, mutually inhibitory appetitive and aversive systems, the operation of which gives rise to four basic categories of motivation — prediction of reward (hope), prediction of aversive events (fear), omission of reward (frustration) and omission of aversive events (relief)[10] (FIG. 1).

Pavlovian learning
Learning that a previously neutral environmental cue predicts a motivational salient outcome.

In the brain, the representation of appetitive and aversive value has been studied in both animals and humans, using experiments that dissociate value from basic sensory representations (for example, by altering motivational state, affective context, expected value, relative value, or counterfactual value[17–27]). From these studies, the most prominent region to emerge in the encoding of appetitive value is the orbitofrontal cortex, particularly its medial aspect, which functionally exploits reciprocal connections, especially to the amygdala and ventral striatum (nucleus accumbens). The representation of aversive value, although less well studied, has highlighted two regions in particular: the lateral orbitofrontal cortex and the anterior insula cortex[19,28–35], which are also interconnected with the amygdala and ventral striatum[36,37].

The ability to learn values leads to the question of how individuals exert control over their environment to bring about appetitive, and reduce aversive, events. Several distinct action mechanisms can be dissociated both behaviourally and neurobiologically, characterized by the way in which an action is learned and evoked (FIG. 2).

Pavlovian actions. The simplest actions are the responses directly associated with Pavlovian conditioned or unconditioned stimuli, generally regarded as 'hard-wired'. Accordingly, prediction of aversive events often produces defensive or aggressive responses that are likely to have evolved to protect the immediate welfare of the animal. Indeed, aggressive responses are often seen towards inanimate aversive cues in animal experiments, and consistent with an opponent model, these responses can even be elicited by appetitive inhibitors[38]. Pavlovian responses are often context-dependent: for example, in male rats, the prediction of a painful shock can produce freezing in a solitary animal, and aggression in the presence of another male[39].

From a neurobiological perspective, Pavlovian actions are mediated by a network of predominantly subcortical regions that coordinate the acquisition of predictive value with the execution of responses. The best understood region is the amygdala, which learns to mediate conditioned responses through connections with other brain regions such as the periaqueductal grey, hypothalamus, parabrachial nuclei, caudal pontine nuclei of the reticular formation, and ventral tegmental area[40].

Structures such as the periaqueductal grey and anterior hypothalamus mediate primitive defensive, retaliatory and offensive responses, and encode essential motor patterning mechanisms for fighting[41].

Instrumental actions. However, Pavlovian actions provide only a restricted set of options for action, and more flexible control is mediated by instrumental learning, whereby an individual learns to associate a particular action with its outcome[42]. Consequently, actions that lead to a reward are executed more frequently in the future, whereas those that lead to aversive events are executed less often. For example, discovering that pressing a lever results in food delivery will cause an animal to press that lever more often, whereas if such an action is followed by an electric shock, the animal will press the lever less often. Clearly, the latter result is often the desired endpoint when punishment is administered by another individual, and response suppression is proportional to the magnitude, certainty and imminence of an anticipated punishment[43–53]. This effect is to some extent Pavlovian: cues that were previously paired with punishment suppress instrumental responding in the absence of any instrumental contingency (conditioned suppression[54]), but adding such a contingency substantially enhances suppression[55,56].

Instrumental learning allows learning of arbitrary and potentially highly adaptive responses beyond the restrictive set that are available to Pavlovian mechanisms, and there are two distinct types of instrumental action: habits, and goal-orientated actions. Habits appear to learn the simple value of actions, by essentially collapsing the value of future outcomes into a single action-value for each choice available to the animal. Thus, although the (value of the) outcome may be directly used to reinforce or inhibit the action, the resulting habit does not encode any specific representation of that outcome (FIG. 2). In the brain, this habit-based learning system is thought to involve a dopamine dependent error-based learning mechanism, and anatomically it crucially involves the dorsolateral striatum and ventral prefrontal cortex[57–62].

Habit-based learning may be a highly effective and computationally simple way to learn and act following extensive exposure to an environment with predictable outcomes. However, it might be a less effective way to make choices given limited experience, or if the outcomes depend on more complex aspects of the action and the environment. By contrast, goal-orientated actions incorporate an internal representation of the outcome, which can be used more directly to guide actions. Experimentally, one of the hallmarks of goal-orientated action is sensitivity to outcome devaluation: if an animal learns to press a lever for food when hungry, and is subsequently fed to satiety, it presses the lever less frequently when exposed to the lever again, indicating that it appropriately represents the reduced value of the action. However, there is good behavioural evidence of a transfer of action control from goal-orientated to habit based systems over time, and on extensive training this sensitivity to outcome devaluation is reduced[63–65].

| | Outcome | |
	Excitatory	Inhibitory
Aversive	Aversive excitor (fear)	Aversive inhibitor (relief)
Appetitive	Appetitive excitor (hope)	Appetitive inhibitor (frustration)

Valence

Figure 1 | **Appetitive and aversive excitators and inhibitors.** Motivational stimuli can be excitatory or inhibitory, depending on whether they predict the occurrence or the absence, respectively, of an affective outcome or of another predictor. They can also be classified by valence, as stimuli that are associated with either appetitive or aversive outcomes or predictors. When combined, these two classifications illustrate the four basic motivational states of fear, relief, hope and frustration.
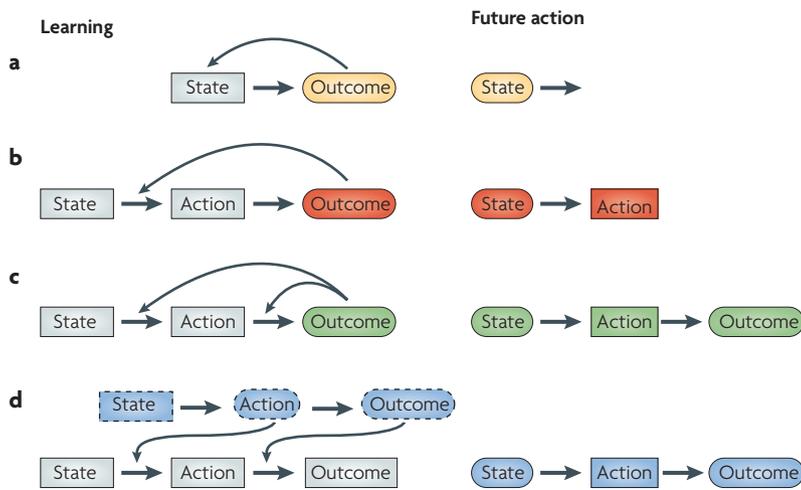
Learning   Future action

a

b

c

d

Figure 2 | **Different mechanisms of learning and action. a** | Pavlovian learning results in Pavlovian responses: repeated predictive pairing of an arbitrary state or cue with a motivationally salient outcome (a reward or aversive event) causes a conditioned and typically innate response to be emitted when the state is encountered in the future, appropriately anticipating the outcome. **b** | Instrumental habit learning causes habit formation: if an action is executed while the animal is in a certain state, and the action leads to a reward, the action is reinforced such that encountering the state in the future makes executing that action more likely. If the value of the outcome was aversive, then the action is inhibited in the future. **c** | Instrumental goal-orientated learning results in goal-directed actions: if an action from a certain state leads to a reward, then an explicit representation of the sequence is remembered, which is available to guide actions when the state is encountered in future. **d** | Learning following observation can also result in habits and goal-direction actions: Observing actions and outcomes in others may lead to reinforcement of either state-action pairs, or possibly state–action–outcome sequences, which might reinforce future actions in the observer.

In addition to simple outcome representations, goal-orientated action selection may accommodate substantial complexity, involving representation of potentially intricate sequences of actions, including those whose outcomes are governed by higher-order structures and rules. Although many animals may possess a surprisingly sophisticated ability to model the structure of their environment to guide goal-orientated behaviour[66,67], this capacity is clearly remarkably developed in humans. From a neurobiological perspective, experiments in rats have shown that goal-orientated actions are known to crucially involve the dorsomedial striatum and medial prefrontal cortex, which appear to learn actions through a dopamine-independent mechanism. In humans (and to a lesser extent in monkeys), there is evidence of a considerable capacity to form cognitive models of the world, involving recruitment of extensive regions of the prefrontal cortex[68–72].

Given this diversity of action mechanisms, how does the brain decide which mechanism to use in different circumstances? Competition between different mechanisms can be illustrated by designing experiments so that the output of different processes are in opposition[73,74]. Remarkably, apparently self-punitive behaviour can inadvertently result from instrumentally punishing an animal for emitting Pavlovian responses to that punishment[75–78]. For example, squirrel monkeys have been observed to increase pulling on a restraining leash that has been experimentally (instrumentally) arranged to

deliver painful electric shocks, with the unfortunate result that they end up getting more shocks than to the optimal response (to stay still), had they been able to learn it[78]. Similarly, if Siamese fighting fish are instrumentally punished for their characteristic (Pavlovian) aggressive fighting display, the display is reinforced rather than inhibited[77]. This illustrates that as the Pavlovian response to being punished may be (reciprocal) aggression, the institution of punishment can counterintuitively reinforce the behaviour it was intended to reduce. Accordingly, the artefacts of (Pavlovian–instrumental) competition may have important consequences for the institution and efficacy of punishment in social circumstances, and are related to a long debate about the optimum methods of punishment in modern-day institutions such as schools[79].

Pavlovian–instrumental interactions clearly cooperate in guiding avoidance learning, which in social contexts may be a key mechanism underlying ostracism — possibly one of the most effective means of punishment in many animals, including humans[80,81]. In ostracism, interactions with other individuals are reduced, which subsequently reduces the opportunity for the ostracized individual to reap the benefits of cooperation. Avoidance appears to involve the rather complex coordination of several processes: Pavlovian learning of predictors of aversive outcomes, Pavlovian and instrumental escape from these predictors, establishment of aversive inhibitory value of the avoided (safety) state, and instrumental reinforcement of the avoided action[11,82–87]. The latter instrumental component may possibly be under either, or both, habitual and goal-orientated control[88]. The neurobiological substrates reflect the operation of these multiple mechanisms, and crucially include the amygdala, orbitofrontal and ventromedial prefrontal cortex, and striatum[89–92]. Furthermore, the instrumental reinforcement of the avoidance action by the aversive inhibitor (safety state) may be dopamine-dependent[93].

Learning from observation and teaching. There is another important mechanism by which the value of actions can be learned, and which turns out to have special importance when considering social actions such as punishment. Clearly, one advantage of social living is that it allows an individual to learn about events that they never themselves experience, simply by observing others. Information gained from observation has considerable value, whether it reflects the fortunes or misfortunes of others, as it can be used to improve one's own individual future behaviour.

An important mechanism of human learning, particularly during development[94,95], observational learning is also observed in a number of other species, for example in behaviours such as determining what foods to eat and in learning to hunt[96–99]. In laboratory studies, a rat will cautiously approach a metal prod that it has previously seen administer a shock to another rat[100], and primates will acquire a fearful response to an animal (for example a snake) that they previously observed to evoke a fearful response in a conspecific[101].

REVIEWS

However, less is known about the mechanisms that underlie observational learning compared with individual learning; in particular, what precisely is being learned. This, in part, reflects difficulties performing these experiments in animals, and much of what is known comes from experimental observations in monkeys and in developing and adult humans. There is good evidence that both actions and action-goals can be learned through observation, and various models have addressed how this relates to associative and goal-orientated aspects of state–action–goal associations[102–104]. There is evidence for quite sophisticated goal-oriented imitation in human infants; for example, after observing an adult use a novel action to perform a task, infants only imitated the action when they judged it to be a rational means to achieve the goal[105]. If the goal of the action is not well understood, it appears that learning is based on reinforcing representations of the infant's own actions in an associative manner based on observed stimulus–response couplings[103]. This type of learning has been shown to influence behaviour even when in opposition to the observer's own action goals[106].

It is less clear, however, how the value of goals is learned through observation of other individuals' choices (FIG. 2). The distinction between action learning and value learning in observational contexts is important, and resembles that between motor control and instrumental learning in individual contexts. Motor control has been extensively studied with respect to imitation and skill acquisition, particularly since the discovery of mirror neurons (see below); however, value learning is less well understood. For instance, the value of many actions may be far from obvious, and for sequential actions this gives rise to the credit assignment problem[107]. In theory, one way of solving this problem is to observe the affective reaction of the executor following action execution, and use this as a surrogate value. However, this may be either poorly observed, or extinguished. This problem is no less troublesome in the aversive domain, illustrated by an 'inverse' avoidance problem: if an observer witnesses that others frequently avoid an action or state (for instance, if they were previously punished for it), long after observable conditioned responses have extinguished[108], how does he/she infer that it is aversive? In principle, it ought to be possible to either directly acquire habits (for example, through imitation), or to more explicitly model (infer) the value of goals (FIG. 2). The relative lack of evidence that addresses this appeals to the need to design the sort of experiments that have dissociated the different types of instrumental learning in individual trial-and-error contexts (such as outcome devaluation).

From a neurobiological perspective, there has been intense interest in the information accrued from observation since the discovery in monkeys that some neurons (mirror neurons) selectively fire during both observation and execution of similar actions[109,110]. These neurons are located in the inferior frontal cortex and the anterior region of the inferior parietal lobule, areas that have a crucial role in sensorimotor control. The different properties of various subtypes of mirror neurons suggest that they have a sophisticated role in action observation.

Some clearly mirror the action being observed very precisely, while others reflect primarily the goal of the action. Interestingly, recent evidence indicates that a proportion of neurons in the monkey inferior parietal lobule code specifically for the intentions of the individual being observed (such as eating or placing an object)[111], and a similar property has been observed in the human right posterior inferior frontal gyrus[112]. So, it seems plausible that learning about the motivational value of actions may access this system, which subsequently allows individuals to improve their own decisions. However, how this might be coordinated with the brain regions involved in individual motivational control is not yet understood.

One especially useful method that might help an individual to learn complex, sequential tasks is for another individual (an expert) to provide supplementary outcomes for various sub-components of the task. This comes into play during helping and teaching behaviours, which in addition to being ubiquitous in humans (where it is greatly facilitated by language), can also be observed in other species[113,114]. Meerkats, for example, teach pups how to handle potentially dangerous prey such as scorpions, using live but disarmed (by removing the sting) scorpions[115]. Thus, pups learn how to avoid outcomes (potentially fatal stings) that they never actually experience.

An important feature of learning though observation and teaching is that it represents cultural learning, allowing knowledge about observed (or possibly inferred) goals to be passed between individuals, perhaps even generations, over time[94,99]. Actions learned in this way may have the capacity to become behavioural norms, and can be adhered to in a way that incorporates some independence from actual outcomes. This may be especially important for punishment, as we discuss below, as in theory it allows the propensity to punish to be culturally acquired, and the outcome of punishment to be learned without personal transgression.

Punishment and cooperation

Motivation in social environments. One of the principal inadequacies in our current understanding of the motivational basis for punishment derives from the fact that, despite an extensive animal literature on learning about basic outcomes such as food and pain, there is much less comparable work for learning about outcomes that arise from interactions with other individuals. Clearly, learning to select optimal actions becomes increasingly difficult when the consequences of the actions of others need to be taken into account. Beyond simple retaliation or avoidance, accurately predicting the manner in which punishment might arise across the breadth of social circumstances, based on our understanding of the basic structure of motivational systems, begins to become more speculative[116].

One approach is to formalize models of individual action selection and learning, import them into multiagent contexts, and simulate the behaviours that emerge. Currently the most successful models come from the field of reinforcement learning[117,118], in which two broad classes of algorithm exist: those which learn an explicit model of the environment and search through it to deter-

Box 2 | **Economic games**

Behavioural economics and social psychology have for some time embraced the complexity that arises when single agent decision makers are embedded in multiagent contexts. Evolutionary game theory draws on hypothetical and experimental paradigms that exploit the dynamic, goal-orientated nature of social interactions. These games characteristically allow an individual to choose between purely selfish and cooperative actions.

For example, the Prisoner's Dilemma illustrates the classic tension between cooperation and competition. In its original form, two prisoners on trial must independently decide whether to stay silent about their crime, or betray the other to the prosecution. Although mutual silence results in the best cooperative outcome, the selfish choice is to betray the other[163]. Modern versions of the Prisoner's Dilemma and other games, typically involving financial payoffs, allow games to be played repeatedly to study how cooperation develops or declines over time, and can be studied in brain imaging environments[164]. They have illustrated how decision-making in social environments is underpinned by the same neurobiological substrates, and involves learning in a similar manner as non-social tasks[12,165–170].

Evolutionary models aim to explain why individuals should behave in a certain way in these games, and in the simplest case consider discrete behavioural subtypes (for example, cooperators, punishers and free-riders) interacting in simple games such as the public goods game. By specifying the costs of various actions (as well as other parameters such as immigration and emigration rates, mutation rates and so on), one can determine the overall fitness of an individual's behaviour, which determines the number of offspring it will have, each of which inherit the behavioural genotype with a certain probability. This allows the proportions of different behavioural phenotypes to be modelled or simulated over many generations, as a population grows or diminishes, to determine which are evolutionarily stable.

mine the best actions (model-based actions), and those that estimate the individual value of the actions based on eventual summed outcomes, and select according to these values (model-free actions). This distinction bears a well-supported (both behaviourally and neurobiologically) correspondence to goal-orientated, and Pavlovian and habit-like action selection, respectively[57,119–124]. In theory, individual agents can learn actions in social contexts by these simple reinforcement methods, or better still, learn the value of joint actions (that is, take the actions of others into account when learning the value of their own actions). Interestingly, simulated agents such as these will often learn and maintain stable cooperative behaviour[125–127].

However, such solipsistic methods ignore the intentionality of other agents: since the value of an action is contingent upon the action of another individual, it is governed in part by unobservable information determined by the decision policies of other agents. This may require more sophisticated goal-directed learning[71], which involves generating a model (an estimate) of other agent decision policies to guide better actions, and updating that model according to experience. In principle, this allows an agent to learn how its actions influence the actions of others. A central computational component of policy estimation may be the representation of the affective state of others, as this is a strong determinant of their response to your action. Accordingly, goal-directed instrumental decision-making towards selfish motivational goals may proceed in a way that requires modelling the behaviour and intentionally influencing the affective state of others (that is, selfish prosocial behaviour). In theory, more complex behaviours such as reputation formation can be tractably predicted, in which your model incorporates other agent's representations of your policy. However, the computational basis of these more complex goal-directed systems is currently poorly understood. So too is our understanding of how individuals learn values from observation (which can be approached from an inverse reinforcement learning

perspective[128,129]). However, important insight into more goal-directed punishment can be gained from predicting, and observing experimentally, how individuals behave in simple economic games (BOX 2).

The free-rider problem. How punishment might operate in social and reciprocal interactions is illustrated by the free-rider problem. Consider a game in which individual players invest a certain amount of their own money into a central pot (FIG. 3), which is then multiplied by a fixed amount, and the total amount subsequently divided equally amongst all players, which they add to the money they didn't invest initially. This type of game, termed a public goods game, is similar to many real-life situations, such as a business in which the earnings of each employee depend of the overall turnover of the business. The contribution of each employee increases the public good and is beneficial for everyone. More specifically, the overall benefit of the group is bigger than the individual cost of contributing, but this in turn is higher than the direct benefit for the individual. Thus, each individual has also a strong temptation not to contribute initially, that is, to free-ride (defect) on the contributions of the rest of the group because each individual also profits from the common good, even if he/she does not contribute. If everyone defects, however, cooperation breaks down and the common good is no longer realized. This problem is referred to as the first-order free-rider problem.

Punishment provides a possible solution: if contributing employees start punishing free-riders by fining them (but at personal cost), the level of cooperation increases again because free-riders want to avoid the cost of being punished[130]. If the punisher knows that he/she will interact with the free-rider again, he/she will subsequently benefit from the increased cooperation, and punishment in this case can be viewed as a (long-term) selfish form of reciprocity. However, if the punisher knows that they will not interact with the free-rider again, he/she pays the cost of punishing while others benefit from the free-rider's switch to cooperation, and thus punishing

becomes altruistic. In reality, as we discuss below, humans punish both selfishly and altruistically[131,132].

But a new problem arises: why should individuals endure the costs of punishing free-riders instead of simply cooperating and avoiding the costs of being punished by others? This is the second-order free-rider problem. One solution is to introduce higher levels of punishment, and punish those who do not punish. Another solution, that human societies maintain punishment by group selection and transmission of conformity, has also been suggested[3,133,134]. Accordingly, groups with altruistic punishers are able to enforce cooperation norms. With an increasing number of punishers, the number of defectors

Figure 3 | **Punishment in a public goods game.** Public goods games provide an experimental illustration of the utility of punishment in social economic interactions. In this example, each player receives an initial endowment of £10 (**a**), and contributes a certain proportion towards the public good (**b**), temporarily leaving each with £5. However, the red player — a free-rider — contributes nothing, and so remains with £10. The collective contribution is multiplied by a certain amount (4 times in this example), which reflects the overall economic benefit of cooperation (**c**). This amount is then equally divided amongst all players, including the free-rider, who as a result ends up with the most money: £27.5 as opposed to £22.5 (**d**). However, if another player (in blue) is allowed to fine the free-rider (£10 at a personal cost of £2 in this example), this removes the incentive to free-ride in the red player (**e**). Even though this seems irrational in the short term, the blue player may benefit from future interactions in which the red player cooperates. Thus, in the long run, short-term punishment results in long-term gain, and accordingly reflects a selfish form of reciprocity with repeated interactions. If the blue player does not interact again, however, then punishment becomes altruistic.

in these societies is minimized, as is the cost of punishment. In terms of the ultimate basis of human reciprocity and cooperation (BOX 3), group selection should favour cooperative groups, allowing punishment and cooperation to evolve. This casts the spotlight on experimental studies that probe the existence and nature of punishment in both animals and humans.

Experimental studies

Punishment in animals. Animals not infrequently behave negatively to one another. In many cases, this behaviour is driven by an immediate selfish benefit to the animal (or its kin) — for example, assertion of dominance, the establishment of mating bonds, theft, parental–offspring conflicts and retaliation[1]. In some situations, food-sharing is increased by harassment, although whether this represents cooperation is unclear[135]. For example, the sharing rate in chimpanzees and squirrel monkeys increases with increasing acts of harassment[136]. However, punishment is observed in some situations where it seems more likely to preserve or promote cooperation. For example, chimpanzees attack allies that do not support them in third party conflicts[137], and queen naked mole rats will attack workers that they judge to be lazy[138]. Cases such as these highlight behaviour that influences future, non-immediate actions of others, rather than conferring immediate self-benefit. These dispositions might represent the evolutionary precursor of more complex and ultimately altruistic punitive behaviours widely seen in humans[136].

Punishment in humans. In addition to more simple (defensive and retaliative) forms of punishments, humans also clearly use punishment to motivate others to cooperate[139]. One of the classic experimental studies looked at cooperation in a public goods game[131]. It was shown that sanctioning by means of financial penalties increased cooperation in subsequent rounds of the game, and in comparison with games in which there was no opportunity for punishment.

The existence of altruistic punishment as a proximate intentional motivation in humans is evident by demonstrations that people are willing to incur a personal cost solely to punish others whom they consider to have behaved unfairly. The simplest illustration occurs in the Ultimatum game, where a player decides whether to accept a proposed split, offered by another player, of a central pot of money. Typically, unequal (< 20%) splits are rejected, which leave both proposer and responder empty handed. This institution of costly, altruistic punishment for unfair behaviour seems to be ubiquitous across widely different societies and cultures[140,141].

Altruistic punishment robustly promotes cooperation[2,132,142]. For example, individuals allowed to choose between playing public goods games in institutions (societies) that did or did not offer the opportunity to punish and reward others initially tended towards those institutions where they could not be punished[143]. The payoffs in these groups declined as they became dominated by free-riders, and most individuals switched to play in sanctioning games where the overall level of cooperation progressively increased. Subsequent
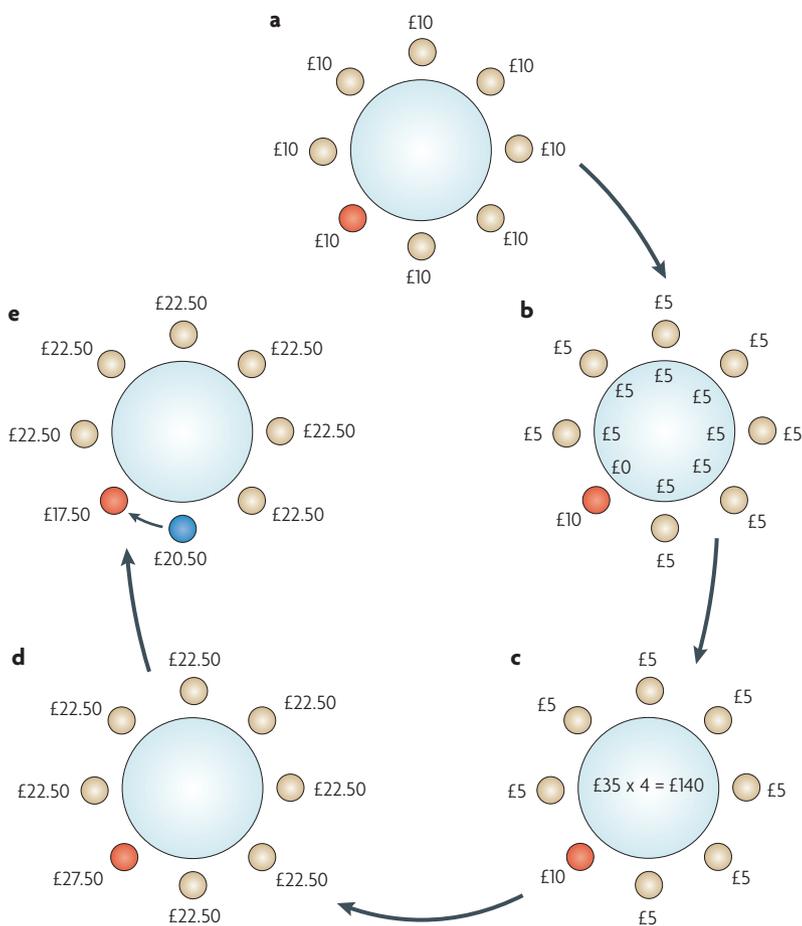
studies have indicated that cooperation may be even more robust if altruistic punishment is combined with altruistic reward, in which the cooperativeness of others is rewarded (at a personal cost)[144].

The proposed importance of cultural norms in driving behaviour predicts that individuals ought to be motivated to reward and punish those who adhere to or transgress norms towards others, even when they themselves are not involved[145]. These situations are captured by third-party punishment games, in which an observer witnesses the interactions of two other players. For example, a third-party punishment game was implemented in the context of a simultaneous Prisoner's Dilemma task[146]: an individual observed the behaviour of two players during the game, and was subsequently given the option to punish at personal cost. Players who cooperated were almost never punished, whereas almost 50% of individuals punished players who defected when their partner cooperated. When both players defected, the punishment rate decreased to 21%. This asymmetry appears to reflect the norm of conditional cooperation, which prescribes that cooperation is assumed if the other player cooperates, whereas defection is considered a more legitimate (less unfair) response in the face of defection by others. Accordingly, unilateral defection is punished more strongly than mutual defection[6]. Once a group establishes a strong reciprocating culture, interaction with other forms of (selfish) reciprocity may mean that the costs of altruistically punishing become relatively small[3,147]. In effect, the threat of punishment may become effective in maintaining cooperation.

Neuroimaging studies in humans. Recently, fMRI has been used to probe the neurobiological correlates of human cooperative behaviour in game theoretic experiments. In particular, several studies have addressed the neurobiological correlates of fairness and punishment, establishing findings which begin to shed light onto the underlying basis of punishing actions. A study of the response to fair and unfair offers in an Ultimatum game found that activity in the anterior insula correlated with the receipt of an unfair offer, and was greater when playing a human as opposed to a computerized opponent, and greater still with increasingly unfair offers[148]. Impressively, this activity predicted individuals' subsequent decisions to reject the offer, effectively (altruistically) punishing their opponent. This study also identified activity in the dorsolateral prefrontal cortex (DLPFC) in relation to fair offers, but not correlated with the degree of unfairness, indicating that it might adopt a more modulatory role. This proposition was supported by another experiment, which disrupted DLPFC activity with transcranial magnetic stimulation (TMS) during the Ultimatum game[149]. It was found that TMS applied to the right, but not left, DLPFC reduced the decisions of individuals to reject unfair offers. This behaviour was specific to human opponents, insensitive to the magnitude of the offer, and independent of subjective verbal ratings of unfairness.

These findings lead to the question of how the representation of the aversive motivational value of unfairness is linked to behavioural decisions to punish. Ultimately, the individual must choose between two outcomes: the financial value of accepting the offer, and the retributive value of punishing the opponent. We designed a task aimed at identifying the brain areas associated with retributive value by looking at the response to cues which predicted that opponents would receive painful electric shocks[150]. We compared brain activity elicited when the cues signalled that a fair or unfair opponent would receive either a high or low intensity shock, where the degree of fairness was associated with previous play in a sequential Prisoner's Dilemma game. The medial orbitofrontal cortex and nucleus accumbens were activated when cues indicated an imminent high intensity shock to unfair players, and this activity correlated with individuals' subjective feelings of anger and retribution. These findings, which were accompanied by compensatory decreases in empathic neural responses, highlight the flexible representation of retributive goals in the orbitofrontal cortex, similar to that seen for primary rewards.

While passive tasks such as this are adequate for identifying brain areas associated with retributive motivational states, they offer little insight into the question of control: that is, which brain areas are involved in learning

and executing actions to bring about punishment? One study gave individuals the opportunity to punish unfair opponents, at personal cost, in an anonymous trust game[151]. Using positron emission tomography (PET), it was found that activity in the dorsal striatum was associated with altruistic punishment acts, with greater activation associated with more severe punishments (which were tied to greater personal losses).

A neurobiological model

Taken together with an understanding of basic motivation and action selection, these findings allow one to sketch the beginnings of a neurobiological model of punishment (FIG. 4). In the simplest case, if an aversive outcome appears to be directly and predictively associated with another individual, it would seem likely to invoke a Pavlovian mechanism, centred on the
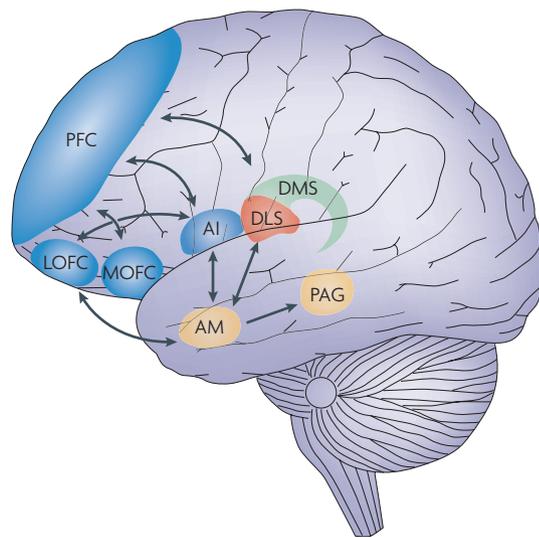

Figure 4 | **Putative neurobiological substrates of punishment.** Impulsive, predominantly Pavlovian punishment may centre on an amygdala-based circuit (depicted in yellow), in which there is associative learning between other individuals (which act as cues) and aversive outcomes. Aversive outcomes may input directly to the amygdala (for example, from brainstem nuclei associated with primitive aversive representations, such as pain[29]), or through more complex aversive representations in the anterior insula (AI) and lateral orbitofrontal cortex (LOFC). This pathway might also be important for avoidance and ostracism. Instrumental punishment may involve striatal-mediated reinforcement of actions that lead to appetitive retributive goals. This appetitive representation (depicted in blue) may involve the medial orbitofrontal cortex (MOFC), and might result from forward-planning of future interactions in broader areas of the prefrontal cortex (PFC) involved in theory or mind, agency, hidden state-estimation and working memory. Goal-directed actions may reinforce action through the dorsomedial striatum (DMS, green). Habit-based actions might reinforce action through dorsolateral striatum (DLS, red), possibly utilizing a dopamine-dependent circuit via the substantia nigra and ventral tegmental area. PAG, periaqueductal grey. Anatomical image adapted, with permission, from REF. 172 © (1996) Appleton & Lange.

Conditioned reinforcement
The process by which a Pavlovian-acquired value can reinforce instrumental action.

amygdala, that may present a relatively pre-potent or impulsive route to punishment. This pathway may direct retaliative responses towards that individual, mediated in part via aggression-related areas such as the periaqueductal grey. Furthermore, this amygdala-dependent pathway might have a central role in guiding escape and avoidance from future interactions with that individual, contributing to subsequent ostracism.

The amygdala may exploit functional connectivity with the lateral orbitofrontal cortex and anterior insula, which might be necessary for more sophisticated, context-dependent aversive representations; for example, those relating to fairness. In principle, one can import fairness-related outcomes onto the Konorskian model (FIG. 1) to specify the full range of excitatory–inhibitory fairness-related outcomes (and predictors) (FIG. 5). This would predict that the anterior insula is similarly involved in representing retributive inhibitors — that is, outcomes and predictive cues associated with the frustration of seeing a free-rider unpunished. However, at this time we know relatively little about how the brain represents observed norms of cooperative behaviour in a way that allows judgement of the fairness of others' behaviour[6,152].

Beyond these simple aversive responses, instrumental control may be dependent on an appropriate representation of the appetitive retributive value of outcomes associated with successful punishment in the medial orbitofrontal cortex. This appetitive value could reinforce punishing actions (or avoidance actions) through reciprocal connections with the dorsal striatum, in a similar manner to primary rewards. Furthermore, reinforcement may arise from complex models of future reciprocal interactions involving more widespread areas or the prefrontal cortex: this might include theory of mind areas (the anterior paracingulate cortex, the superior temporal sulci and the temporal poles) that are likely to be involved in representing the policies of others[153–155], anterior cingulate cortical subregions that are involved in representing agency[156], and more anterior prefrontal cortical areas that are involved in model-building and resolution of partial observability[71]. Ultimately, in repetitively predictable situations, such actions may become habitual responses to unfairness.

Altruistic action. The retributive value of punishment may arise from potentially sophisticated forward modelling of future interactions. However, this leads to the question of how altruistic goals are acquired — if they, by definition, ultimately result in personal cost. There are several possibilities. First, they might reflect a misassumption that future interactions are not improbable (not unreasonable in the smaller societies in human evolutionary history). Second, they could reflect the anticipated prospect that kin, possibly in subsequent generations, will interact with the individual being punished. Third, if punishment from selfish reciprocal (goal-orientated) action reliably results in eventual long-term payoffs, more proximal states following punishment may be reinforced both through habit-based learning, and through sequential Pavlovian learning[157]. The latter process allows the state immediately following punishment to acquire an
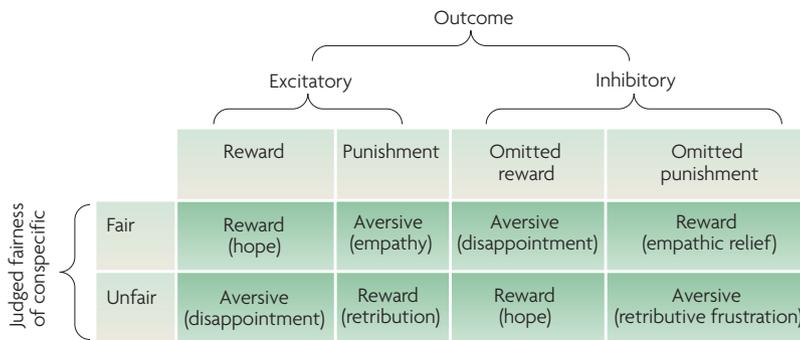
Figure 5 | **Fairness related outcome representations.** This figure extends the Dickenson and Dearing's 'Konorskian' motivational model[13] to incorporate social reinforcement made with respect to judgements of fairness. When affective outcomes are observed in conspecifics who are fair (or who are kin), the motivational value is congruent with the observer. If the individual is judged to be unfair, then the pattern of value is reversed. This illustrates the full spectrum of prosocial motives according to predicted or omitted outcomes, or their predictors.

to assume that altruistic punishment should necessarily be hard-wired as an inherited intrinsic motivational goal (that is, as an unconditioned appetitive stimulus) in the same manner as primary rewards. However, neither does it exclude the possibility. Future research might help to resolve both the role of learning and early development in the acquisition of altruistic behaviour.

Clearly, there are many potentially complex ways in which punishing behaviour, including altruistic punishment, might be acquired, and the nature of this acquisition governs the types of action by which it is mediated. Although this says nothing about why such behaviour should have evolved (that is, the ultimate basis of different forms of punishment), it illustrates (proximately) how they might be based on the operation and, importantly, the interaction of different learning systems. Furthermore, this complexity illustrates the difficulty that evolutionary models face. Since the underlying learning and decision making processes are not solely concerned with punishment behaviour, such models need to take into account the other behaviours that these systems subserve, many of which are not related to reciprocity and cooperation. This difficulty may be similarly evident in other apparently irrational punishment-related behaviour, such as self-punitive actions and reciprocal aggression. Thus, future models may need to take a more generic approach to understanding the interaction between evolution and learning[158].

appetitive value, which might then independently reinforce other actions (through conditioned reinforcement). Both these forms of control will be insensitive to the possibility that in some situations the outcome is altruistic. Fourth, it is possible that learning mechanisms that are invoked when one individual observes other individuals punishing in situations which may not necessarily be altruistic, might generalize across situations in which it is. Given that many selfish reciprocal punishing actions may stem from a long-term view of future interactions, the eventual benefits of an action are likely to be frequently obscure to a naïve observer. In other words, the appetitive value of retributive states and actions might be purely imitated or inferred through observation, since the observer does not have access to the eventual goals in the mind of the individual being observed. Thus, the motivation to punish unfair individuals may be acquired across states in a way that assumes eventual outcomes. Fifth, and in a similar manner, the value of punishment might be taught by experts to non-experts (for example, from parents to offspring, or from dominant to subordinate individuals). In this case, the appetitive value of punishment may be intricately tied in with cultural concepts of morality and justice.

Thus, the very nature of action systems, both those involved in individual and observational learning, may have an inherent tendency to generalize non-altruistic to altruistic actions. This suggests that there is no reason

Conclusions

Punishment, in its various forms, is likely to have played a key role in shaping the dynamics of social interaction in many species — humans in particular. Although many aspects of our neurobiological model are speculative, punishment is likely to involve the integration of a number of distinct representation, learning and action systems. Whatever the neural mechanism, the affirmation that punishment, including altruistic punishment, substantially promotes cooperation in human societies seems firm. Vital to furthering our knowledge will be an understanding of the behavioural and neurobiological basis of cultural and observational learning, sequential learning and model-based learning and planning in the context of other agents. This could be crucial to gaining neurobiological insight into how apparently altruistic behaviours are acquired, as well as shedding light onto more complex social aspects of punishment, such as arbitration, policing and the role of hierarchies.

1. Clutton-Brock, T. H. & Parker, G. A. Punishment in animal societies. *Nature* **373**, 209–216 (1995).
   A definitive review of the ecology and economics of punishment in animals.
2. Boyd, R. & Richerson, P. J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* **13**, 171–195 (1992).
3. Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* **100**, 3531–3535 (2003).
4. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785–791 (2003).
5. Bowles, S. & Gintis, H. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theor. Popul. Biol.* **65**, 17–28 (2004).
6. Fehr, E. & Fischbacher, U. Social norms and human cooperation. *Trends Cogn. Sci.* **8**, 185–190 (2004).
7. Gintis, H. The hitchhiker's guide to altruism: gene–culture coevolution, and the internalization of norms. *J. Theor. Biol.* **220**, 407–418 (2003).
8. Moja, E. A., Cipolla, P., Castoldi, D. & Tofanetti, O. Dose–response decrease in plasma tryptophan and in brain tryptophan and serotonin after tryptophan-free amino acid mixtures in rats. *Life Sci.* **44**, 971–976 (1989).
9. Mackintosh N. J. *Conditioning and Associative Learning.* Oxford Univ. Press, New York 1983).
   This classic text provides an essential and comprehensive introduction to animal learning theory.
10. Gray, J. A. *Problems in the Behavioural Sciences* 2nd edn Vol. 5 (Cambridge Univ. Press, Cambridge, 1991).
11. Bouton, M. E. *Learning and Behavior: A Contemporary Synthesis* (Sinauer, Sunderland, Massachusetts, 2006).
12. Montague, P. R. & Berns, G. S. Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284 (2002).
13. Dickinson, A. & Dearing M. F. Appetitive–aversive interactions and inhibitory processes in *Mechanisms of Learning and Motivation* (eds Dickinson, A. & Boakes, R. A.) 203–231 (Erlbaum, Hillsdale, New Jersey, 1979).
14. Pavlov, I. P. *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex* (Oxford Univ. Press, London, 1927).

15. Rescorla, R. A. Probability of shock in the presence and absence of CS in fear conditioning. *J. Comp. Physiol. Psychol.* **66**, 1–5 (1968).
16. Konorski, J. *Integrative Activity of the Brain: An Interdisciplinary Approach* (Chicago Univ. Press, Chicago, 1967).
17. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
18. Schultz, W. Multiple reward signals in the brain. *Nature Rev. Neurosci.* **1**, 199–207 (2000).
19. O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J. & Andrews, C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neurosci.* **4**, 95–102 (2001).
20. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
21. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nature Rev. Neurosci.* **6**, 363–375 (2005).
22. Ursu, S. & Carter, C. S. Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. *Brain Res. Cogn. Brain Res.* **23**, 51–60 (2005).
23. Nieuwenhuis, S. *et al.* Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage* **25**, 1302–1309 (2005).
24. Rolls, E. T. The orbitofrontal cortex and reward. *Cereb. Cortex* **10**, 284–294 (2000).
25. Roesch, M. R. & Olson, C. R. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* **304**, 307–310 (2004).
26. Milad, M. R. & Quirk, G. J. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* **420**, 70–74 (2002).
27. Baxter, M. G. & Murray, E. A. The amygdala and reward. *Nature Rev. Neurosci.* **3**, 563–573 (2002).
28. Calder, A. J., Lawrence, A. D. & Young, A. W. Neuropsychology of fear and loathing. *Nature Rev. Neurosci.* **2**, 352–363 (2001).
29. Craig, A. D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Rev. Neurosci.* **3**, 655–666 (2002).
30. Paulus, M. P. & Stein, M. B. An insular view of anxiety. *Biol. Psychiatry* **60**, 383–387 (2006).
31. Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C. & Jones-Gotman, M. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain* **124**, 1720–1733 (2001).
32. Seymour, B. *et al.* Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nature Neurosci.* **8**, 1234–1240 (2005).
33. Jensen, J. *et al.* Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum. Brain Mapp.* 15 Jun 2006 (doi:10.1002/hbm.20274).
34. Nitschke, J. B., Sarinopoulos, I., Mackiewicz, K. L., Schaefer, H. S. & Davidson, R. J. Functional neuroanatomy of aversion and its anticipation. *Neuroimage* **29**, 106–116 (2006).
35. Sarinopoulos, I., Dixon, G. E., Short, S. J., Davidson, R. J. & Nitschke, J. B. Brain mechanisms of expectation associated with insula and amygdala response to aversive taste: implications for placebo. *Brain Behav. Immun.* **20**, 120–132 (2006).
36. Mesulam, M. M. & Mufson, E. J. Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *J. Comp. Neurol.* **212**, 1–22 (1982).
37. Mufson, E. J., Mesulam, M. M. & Pandya, D. N. Insular interconnections with the amygdala in the rhesus monkey. *Neuroscience* **6**, 1231–1248 (1981).
38. Hutchinson, R. R., Azrin, N. H. & Hunt, G. M. Attack produced by intermittent reinforcement of a concurrent operant response. *J. Exp. Anal. Behav.* **11**, 489–495 (1968).
39. Ulrich, R. E. & Azrin, N. H. Reflexive fighting in response to aversive stimulation. *J. Exp. Anal. Behav.* **5**, 511–520 (1962).
40. Fendt, M. & Fanselow, M. S. The neuroanatomical and neurochemical basis of conditioned fear. *Neurosci. Biobehav. Rev.* **23**, 743–760 (1999).
41. Adams, D. B. Brain mechanisms of aggressive behavior: an updated review. *Neurosci. Biobehav. Rev.* **30**, 304–318 (2006).
42. Thorndike, E. L. *Animal Intelligence* (Macmillan, New York, 1911).

43. Azrin, N. H. Some effects of two intermittent schedules of immediate and non-immediate punishment. *J. Psychol.* **42**, 3–21 (1956).
44. Church, R. M., Raymond, G. A. & Beauchamp, R. D. Response suppression as a function of intensity and duration of a punishment. *J. Comp. Physiol. Psychol.* **1**, 39–44 (1967).
45. Camp, D. S., Raymond, G. A. & Church, R. M. Temporal relationship between response and punishment. *J. Exp. Psychol.* **74**, 114–123 (1967).
46. Azrin, N. H. Effects of punishment intensity during variable-interval reinforcement. *J. Exp. Anal. Behav.* **3**, 123–142 (1960).
47. Solomon, R. L., Turner, L. H. & Lessac, M. S. Some effects of delay of punishment on resistance to temptation in dogs. *J. Pers. Soc. Psychol.* **8**, 233–238 (1968).
48. Atnip, G. W. Stimulus and response reinforcer contingencies in autoshaping, operant, classical and omission training procedures in rats. *J. Exp. Anal. Behav.* **28**, 56–69 (1977).
49. Baron, A. Delayed punishment of a runway response *J. Comp. Physiol. Psychol.* **60**, 131–134 (1965).
50. Walters, G. C. & Grusec, J. E. *Punishment* (W. H. Freeman, San Francisco, 1977).
51. Solomon, R. L., Turner, L. H. & Lessac, M. S. Some effects of delay of punishment on resistance to temptation in dogs. *J. Pers. Soc. Psychol.* **8**, 233–238 (1968).
52. Azrin, N. H., Holz, W. C. & Hutchinson, R. R. Fixed-ratio escape reinforcement. *J. Exp. Anal. Behav.* **6**, 141–148 (1963).
53. Boe, E. E. & Church, R. M. Permanent effects of punishment during extinction. *J. Comp. Physiol. Psychol.* **63**, 486–492 (1967).
54. Estes, W. K. & Skinner, B. F. Some quantitative properties of anxiety. *J. Exp. Psychol.* **29**, 390–400 (1941).
55. Church, R. M. in *Punishment and Aversive Behavior* (eds Campbell, B. A. & Church, R. M.) (Appleton, New York, 1969).
56. Bolles, R. C., Holtz, R., Dunn, T. & Hill, W. Comparison of stimulus learning and response learning in a punishment situation. *Learn. Motiv.* **11**, 78–96 (1980).
57. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
    Comprehensively describes the theoretical basis of reinforcement learning models of dopamine in monkeys.
58. Wise, R. A. Dopamine, learning and motivation. *Nature Rev. Neurosci.* **5**, 483–494 (2004).
59. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
60. Yin, H. H. & Knowlton, B. J. The role of the basal ganglia in habit formation. *Nature Rev. Neurosci.* **7**, 464–476 (2006).
61. Bar-Gad, I., Morris, G. & Bergman, H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog. Neurobiol.* **71**, 439–473 (2003).
62. Niv, Y., Joel, D. & Dayan, P. A normative perspective on motivation. *Trends Cogn. Sci.* **10**, 375–381 (2006).
63. Dickinson, A. & Balleine, B. W. in *Steven's Handbook of Experimental Psychology* 3rd edn Vol. 3 (ed. Gallistel, C. R.) 497–533 (John Wiley & Sons, New York, 2002).
64. Balleine, B. W. Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol. Behav.* **86**, 717–730 (2005).
65. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neurosci.* **8**, 1704–1711 (2005).
66. Blaisdell, A. P., Sawa, K., Leising, K. J. & Waldmann, M. R. Causal reasoning in rats. *Science* **311**, 1020–1022 (2006).
67. Raby, C. R., Alexis, D. M., Dickinson, A. & Clayton, N. S. Planning for the future by western scrub-jays. *Nature* **445**, 919–921 (2007).
    This remarkable study shows that scrub-jays can plan for future goals in a manner independent from their current motivational state.
68. Duncan, J. An adaptive coding model of neural function in prefrontal cortex. *Nature Rev. Neurosci.* **2**, 820–829 (2001).
69. Koechlin, E., Ody, C. & Kouneiher, F. The architecture of cognitive control in the human prefrontal cortex. *Science* **302**, 1181–1185 (2003).

70. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
71. Yoshida, W. & Ishii, S. Resolution of uncertainty in prefrontal cortex. *Neuron* **50**, 781–789 (2006).
    Uses sophisticated behavioural modelling to show that the human prefrontal cortex is engaged to solve partially observable decision tasks, one of the central computational problems underlying multiagent games.
72. Hampton, A. N., Bossaerts, P. & O'Doherty, J. P. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* **26**, 8360–8367 (2006).
73. Hershberger, W. A. An approach through the looking-glass. *Anim. Learn. Behav.* **14**, 443–451 (1986).
    This famous experiment shows that chicks cannot learn to retreat from a food cart that moves in the same direction as them but at twice the speed.
74. Dayan, P., Niv, Y., Seymour, B. & Daw, D. The misbehavior of value and the discipline of the will. *Neural Netw.* **19**, 1153–1160 (2006).
75. Fowler, H. & Miller, N. E. Facilitation and inhibition of runway performance by hind- and forepaw shock of various intensities. *J. Comp. Physiol. Psychol.* **56**, 801–805 (1963).
76. Anson, J. E., Bender, L. & Melvin, K. B. Sources of reinforcement in establishment of self-punitive behavior. *J. Comp. Physiol. Psychol.* **67**, 376–380 (1969).
77. Melvin, K. B. & Anson, J. E. Facilitative effects of punishment on aggressive behavior in Siamese fighting fish. *Psychon. Sci.* **14**, 89–90 (1969).
78. Morse, W. H., Mead, R. N. & Kelleher, R. T. Modulation of elicited behavior by a fixed-interval schedule of electric shock presentation. *Science* **157**, 215–217 (1967).
79. Lieberman, D. A. *Learning: Behavior and Cognition* (Wadsworth, Belmont, California, 1999).
    Considers the efficacy of various forms of punishment in children, including omission of privileges and reinforcing good behaviour.
80. Cinyabuguma, M., Page, T. & Putterman, L. Cooperation under the threat of expulsion in a public goods experiment. *J. Pub. Econ.* **89**, 1421–1435 (2005).
81. Williams, K. D. Ostracism. *Annu. Rev. Psychol.* **58**, 425–452 (2007).
82. Damato, M. R., Fazzaro, J. & Etkin, M. Anticipatory responding and avoidance discrimination as factors in avoidance conditioning. *J. Exp. Psychol.* **77**, 41–47 (1968).
83. Bolles, R. C. & Grossen, N. E. Effects of an informational stimulus on acquisition of avoidance behavior in rats. *J. Comp. Physiol. Psychol.* **68**, 90–99 (1969).
84. Starr, M. D. & Mineka, S. Determinants of fear over course of avoidance-learning. *Learn. Motiv.* **8**, 332–350 (1977).
85. Crawford, M., Masterson, F. & Wilson, D. Species-specific defense reactions in escape-from-fear situations. *Anim. Learn. Behav.* **5**, 63–72 (1977).
86. Dickinson, A. *Contemporary Animal Learning Theory* (Cambridge Univ. Press, Cambridge, UK, 1980).
87. Dinsmoor, J. A. Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *J. Exp. Anal. Behav.* **75**, 311–333 (2001).
88. Hendersen, R. W. & Graham, J. Avoidance of heat by rats — effects of thermal context on rapidity of extinction. *Learn. Motiv.* **10**, 351–363 (1979).
89. Kirkby, R. J. & Kimble, D. P. Avoidance and escape behavior following striatal lesions in the rat. *Exp. Neurol.* **20**, 215–227 (1968).
90. White, I. M. & Rebec, G. V. Responses of rat striatal neurons during performance of a lever-release version of the conditioned avoidance response task. *Brain Res.* **616**, 71–82 (1993).
91. Killcross, S., Robbins, T. W. & Everitt, B. J. Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature* **388**, 377–380 (1997).
92. Kim, H., Shimojo, S. & O'Doherty, J. P. Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* **4**, 1453–1461 (2006).
93. Cook, L. & Catania, A. C. Effects of drugs on avoidance and escape behaviour. *Fed. Proc.* **23**, 818–835 (1964).
94. Bandura, A. *Social Learning Theory* (General Learning, New York, 1977).
95. Gergely, G. & Watson, J. in *Early social cognition: Understanding Others in the First Months of Life* (ed. Rochat, P.) 101–136 (Lawrence Erlbaum, Mawah, New Jersey, 1999).

96. Galef, B. G. & Whiskin, E. E. Social transmission of information about multiflavored foods. *Anim. Learn. Behav.* **20**, 56–62 (1992).

97. Stoinski, T. S., Wrate, J. L., Ure, N. & Whiten, A. Imitative learning by captive western lowland gorillas (*Gorilla gorilla gorilla*) in a simulated food-processing task. *J. Comp. Psychol.* **115**, 272–281 (2001).

98. Whiten, A. & Boesch, C. The cultures of chimpanzees. *Sci. Am.* **284**, 60–67 (2001).

99. Tomasello, M., Kruger, A. & Ratner, H. Cultural learning. *Behav. Brain Sci.* **16**, 495–552 (1993).

100. Galef, B. G. & Laland, K. N. Social learning in animals: empirical studies and theoretical models. *Bioscience* **55**, 489–499 (2005).

101. Mineka, S. & Cook, M. Mechanisms involved in the observational conditioning of fear. *J. Exp. Psychol. Gen.* **122**, 23–38 (1993).

102. Wohlschlager, A., Gattis, M. & Bekkering, H. Action generation and action perception in imitation: an instance of the ideomotor principle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**, 501–515 (2003).

103. Heyes, C. & Ray, E. What is the significance of imitation in animals? *Adv. Study Behav.* **29**, 215–245 (2000).

104. Heyes, C. Causes and consequences of imitation. *Trends Cogn. Sci.* **5**, 253–261 (2001).

105. Gergely, G., Bekkering, H. & Kiraly, I. Rational imitation in preverbal infants. *Nature* **415**, 755 (2002).

106. Heyes, C., Bird, G., Johnson, H. & Haggard, P. Experience modulates automatic imitation. *Brain Res. Cogn. Brain Res.* **22**, 233–240 (2005).

107. Bellman, R. *Dynamic Programming* (Princeton Univ. Press, Princeton, 1957).

108. Solomon, R. L., Kamin, L. J. & Wynne, L. C. Traumatic avoidance learning — the outcomes of several extinction procedures with dogs. *J. Abnorm. Soc. Psychol.* **48**, 291–302 (1953).

109. Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141 (1996).

110. Rizzolatti, G. & Craighero, L. The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).

111. Fogassi, L. *et al.* Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667 (2005).

112. Iacoboni, M. *et al.* Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* **3**, e79 (2005).

113. Tomasello, M. & Call, J. *Primate Cognition* (Oxford Univ. Press, Oxford, 1970).

114. Warneken, F. & Tomasello, M. Altruistic helping in human infants and young chimpanzees. *Science* **311**, 1301–1303 (2006).

115. Thornton, A. & McAuliffe, K. Teaching in wild meerkats. *Science* **313**, 227–229 (2006).

116. Yamagishi, T. in *Advances in Group Processes* Vol. 3 (ed. Lawler E. J.) 51–87 (Elsevier Science, Greenwich, Connecticut 1986).

117. Kaelbling, L. P., Littman, M. L. & Moore, A. W. Reinforcement learning: a survey. *J. Artif. Intell. Res.* **4**, 237–285 (1996).

118. Sutton, R. S. & Barto, A. G. *Reinforcement Learning. An Introduction* (MIT press, Cambridge, Massachusetts, 1998).

119. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).

120. O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H. & Dolan, R. J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329–337 (2003).

121. Tanaka, S. C. *et al.* Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neurosci.* **7**, 887–893 (2004).

122. Haruno, M. *et al.* A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* **24**, 1660–1665 (2004).

123. Seymour, B. *et al.* Temporal difference models describe higher-order learning in humans. *Nature* **429**, 664–667 (2004).

124. Tanaka, S. C. Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Netw.* **19**, 1233–1241 (2006).

125. Littman, M. L. in *Proceedings of the Eleventh International Conference on Machine Learning*, 157—163 (Morgan Kaufmann, San Francisco, California, 1994). (1994).

126. Hu, J. L. & Wellman, M. P. Nash Q-learning for general-sum stochastic games. *J. Mach. Learn. Res.* **4**, 1039–1069 (2004).

127. Claus, C. & Boutilier, C. The dynamics of reinforcement learning in cooperative multiagent systems. *Proc. Natl Conf. Artif. Intell.* **15**, 746–752 (1998).

128. Ng, Y. N. & Russell, S. Algorithms for inverse reinforcement learning. *Proc. Seventeenth Int. Conf. Mach. Learn.* 663–670 (2000).
    Provides a thorough analysis of the inverse reinforcement learning problem.

129. Abbeel, P. & Ng, A. Y. Apprenticeship learning via inverse reinforcement learning. *ACM Int. Conf. Proc. Series* pp1–8 (2004).

130. Yamagishi, T. & Sato, K. Motivational basis of the public goods problem. *J. Pers. Soc. Psychol.* **50**, 67–73 (1986).

131. Yamagishi, T. The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.* **51**, 110–116 (1986).
    One of the first experimental demonstrations that punishment increases cooperation.

132. Fehr, E. & Gachter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
    The first demonstration of altruistic punishment in humans.

133. Boyd, R. & Richerson, P. J. The evolution of reciprocity in sizable groups. *J. Theor. Biol.* **132**, 337–356 (1988).

134. Gintis, H. Strong reciprocity and human sociality. *J. Theor. Biol.* **206**, 169–179 (2000).

135. Stevens, J. R. & Hauser, M. D. Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* **8**, 60–65 (2004).
    Discusses some of the underlying cognitive components involved in reciprocity.

136. Stevens, J. R. The selfish nature of generosity: harassment and food sharing in primates. *Proc. Biol. Sci.* **271**, 451–456 (2004).

137. De Waal, F. B. M. *Chimpanzee Politics: Power and Sex Among Apes* (Johns Hopkins Univ. Press, Baltimore, Maryland, 1998).

138. Reeve, H. K. Queen activation of lazy workers in colonies of the eusocial naked mole-rat. *Nature* **358**, 147–149 (1992).

139. Shinada, M., Yamagishi, T. & Ohmura, Y. False friends are worse than bitter enemies: 'altruistic' punishment of in-group members. *Evol. Hum. Behav.* **25**, 379–393 (2004).

140. Henrich, J. *et al.* In search of *Homo economicus*: Behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* **91**, 73–78 (2001).

141. Henrich, J. Cooperation, punishment, and the evolution of human institutions. *Science* **312**, 60–61 (2006).

142. Fehr, E. & Gachter, S. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000).

143. Gurerk, O., Irlenbusch, B. & Rockenbach, B. The competitive advantage of sanctioning institutions. *Science* **312**, 108–111 (2006).

144. Andreoni, J., Harbaugh, W. & Vesterlund, L. The carrot or the stick: rewards, punishments, and cooperation. *Am. Econ. Rev.* **93**, 893–902 (2003).

145. Bendor, J. & Swistak, P. The evolution of norms. *Am. J. Sociol.* **106**, 1493–1545 (2001).

146. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **25**, 63–87 (2004).

147. Rockenbach, B. & Milinski, M. The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718–723 (2006).

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

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

150. Singer, T. *et al.* Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466–469 (2006).

151. de Quervain, D. J. *et al.* The neural basis of altruistic punishment. *Science* **305**, 1254–1258 (2004).

152. Moll, J., Zahn, R., Oliveira-Souza, R., Krueger, F. & Grafman, J. The neural basis of human moral cognition. *Nature Rev. Neurosci.* **6**, 799–809 (2005).

153. Brunet, E., Sarfati, Y., Hardy-Bayle, M. C. & Decety, J. A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage* **11**, 157–166 (2000).

154. Gallagher, H. L. & Frith, C. D. Functional imaging of 'theory of mind'. *Trends Cogn. Sci.* **7**, 77–83 (2003).

155. Gallagher, H. L., Jack, A. I., Roepstorff, A. & Frith, C. D. Imaging the intentional stance in a competitive game. *Neuroimage* **16**, 814–821 (2002).

156. Tomlin, D. *et al.* Agent-specific responses in the cingulate cortex during economic exchanges. *Science* **312**, 1047–1050 (2006).

157. Vlaev, I. & Chater, N. Game relativity: how context influences strategic decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 131–149 (2006).

158. Ackley, D. H. & Littman, M. L. in *Artificial Life II*, *SFI Studies in the Sciences of Complexity*, Vol. X (eds Langton, C. G., Taylor, C., Farmer, J. D. & Rasmussen, S.) 487–509 (Addison-Wesley, Reading, Massachusetts, 1991).
    Provides an excellent introduction to theory surrounding the evolution of learning systems.

159. Kamin, L. J. in *Miami Symposium on the Prediction of Behavior: Aversive Stimulation* (ed. Jones, M. R.) 9–33 (Miami Univ. Press, Miami, 1968).

160. Rescorla R. A. Variation in the effectiveness of reinforcement and non-reinforcement following prior inhibitory conditioning. *Learn. Motiv.* **2**, 113–123 (1971).

161. Bakal, C. W., Johnson, R. D. & Rescorla, R. A. The effect of change in US quality on the blocking effect. *Pavlov. J. Biol. Sci.* **9**, 97–103 (1974).

162. Bull, J. A. & Overmier, J. B. Additive and subtractive properties of excitation and inhibition. *J. Comp. Physiol. Psychol.* **66**, 511–514 (1968).

163. Camerer, C. F. *Behavioural Game Theory: Experiments in Strategic Interaction* (Princeton Univ. Press, Princeton, 2003).

164. Montague, P. R. *et al.* Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage* **16**, 1159–1164 (2002).

165. McCabe, K., Houser, D., Ryan, L., Smith, V. & Trouard, T. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* **98**, 11832–11835 (2001).

166. Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J. & Frith, C. D. Brain responses to the acquired moral status of faces. *Neuron* **41**, 653–662 (2004).

167. King-Casas, B. *et al.* Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).

168. Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* **22**, 1694–1703 (2004).

169. Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* **15**, 2539–2543 (2004).

170. Rilling, J. *et al.* A neural basis for social cooperation. *Neuron* **35**, 395–405 (2002).

171. Bowles, S. & Gintis, H. *Homo reciprocans*. *Nature* **415**, 125–128 (2002).

172. Martin, J. H. *Neuroanatomy: Text and Atlas* 2nd edn (Appleton & Lange, Stamford, Connecticut, 1996).

FURTHER INFORMATION
Wellcome Trust Centre for Neuroimaging:
http://www.fil.ion.ucl.ac.uk
Access to this links box is available online.

Biographies

Ben Seymour is a Clinical Research Fellow in Neurology at the Wellcome Trust Centre for Neuroimaging and the National Hospital for Neurology and Neurosurgery, in London, UK. His main research interests are in the neuroscience of motivation, learning and decision-making, with a particular focus on aversive learning and pain.

Ray Dolan is Director of the Wellcome Trust Centre for Neuroimaging and Kinross Professor of Neuropsychiatry, University College London, UK. He obtained his primary medical degree in Ireland and trained in psychiatry in London. His research interests span learning and decision making, emotion and motivation, and memory.

Tania Singer is an assistant professor at the Center for Social Neuroscience and Neuroeconomics, in Zurich, Switzerland, where she works on the neural basis of human social behaviour. She studied psychology and did her Ph.D. and first year of postdoctoral research at the Max Planck Institute for Human Development in Berlin, Germany, with Ulman Lindenberger and Paul Baltes, and subsequently at the Wellcome Trust Centre for Neuroimaging, University College London (UCL), UK, and the Institute of Cognitive Neuroscience, UCL, with Chris and Uta Frith, respectively. Her main research interests are in the social neurosciences with a particular focus on social emotions such as empathy, fairness and feelings of revenge.

At a glance summary

- Punishment is common across species, where it often serves to protect the personal interests of the animal concerned. In humans, it seems to promote and preserve cooperative behaviour, according to culturally acquired (moral) norms of behaviour. This extends to the existence of altruistic punishment, in which the act of punishing is personally costly to the punisher, but protects the cooperative interests of the group.
- An account of the proximate basis of punishment focuses attention on the behavioural and neurobiological basis of motivation, and an understanding how basic learning and action selection systems deal with outcomes that involve other individuals.
- Appetitive and aversive systems motivate actions through Pavlovian, habit-based and goal-directed actions, which are acquired in specific ways. Particularly important for social and altruistic actions may be learning through observation, in which naïve observers learn actions from others. This reflects the cultural acquisition of behaviour, and may be an underlying component of social norm development.
- In social decision-making situations, individuals often benefit from cooperation. However, many cooperative situations give rise to the temptation to free-ride on the cooperative behaviour of others. This can be deterred by punishing non-cooperative free-riding behaviour. This pays off in the long run if the punisher interacts with the reformed free-rider again, but is costly if they do not.
- Altruistic punishment appears to be common across diverse human cultures. It extends to third-party situations, in which passive observers punish those that they witness acting unfairly. It may be especially powerful in promoting cooperation when combined with other forms of punishment that arise from direct and indirect modes of reciprocity.
- In the brain, recent functional MRI studies have highlighted key areas involved in the recognition and representation of unfairness (anterior insula), in the establishment of retributive goals (orbitofrontal cortex), and in the execution of punishing actions (striatum).
- We outline a neurobiological model of punishment, in which an amygdala-centred pathway mediates impulsive, retaliative punishment, and a striatal system mediates instrumental punishment. This latter system may mediate goal-directed punishing actions that involve forward planning (that is, reciprocity-based punishment), as well as more habit-like punishing actions as experience becomes more extensive.
- Consideration of the learning systems that underlie punishment predict that selfish reciprocity is likely to generalize to altruistic (strong) reciprocity. Accordingly, it may not be necessary to assume that altruistic retributive goals are inherited unconditioned appetitive stimuli.
- Future research needs to focus on exploring more precisely the type actions that underlie altruistic punishment (at a proximate level), and understanding how learning and evolution interact in shaping cooperative behaviour (at an ultimate level).

Toc Blurb

Humans punish selfishly but also altruistically. Seymour, Singer and Dolan propose a neurobiological model of punishment, based on our understanding of motivational systems, observational studies that show how punishment mediates cooperation, and brain imaging data from humans playing economic games.