Values and Actions in Aversion

Peter Dayan and Ben Seymour
UCL, London.

23nd November 2007

Abstract

Losses are inextricably intertwined with gains, making it rather puzzling that the bulk of research focuses on the latter rather than the former. In this chapter, we review how the basic architecture of affective decision-making applies to losses and punishments, noting particularly the paradigmatic case of avoidance learning. One key facet of aversive processing is a collection of preparatory withdrawal responses inspired by predictors of negative outcomes; we consider how these may underpin aspects of a number of anomalies of decision making such as aggression, altruistic punishment, framing, dread, and depressive realism.

1. Introduction

It was the English scholar Jeremy Bentham who first argued that the understanding of human economic behaviour might benefit from the study of the physiological processes from which it derives (Bentham, 1823). Pertinently, he pursued an account of economic decision making that balanced the opposing motives of losses and gains, in recognition of the fact that most choices involve contemplation of comparable measures of each. Furthermore, he recognised that the immutable characteristic of the former (incarnate as his plethora of 'pains' (Bentham, 1817)) is the basic devaluing property that drives decisions to reduce or avoid them.

More prosaically, issues of loss are central to many everyday economic decisions, such as health, insurance and borrowing; further, apparent anomalies of choice such as loss aversion, framing effects and regret, all arise in aversive contexts. There is even a tight, though confusing, link between aversion and stress and psychiatric conditions such as depression. Nevertheless, partly for ethical reasons having to do with the undesirability of actually relieving human subjects of part of their own wealth in an experiment, it has been very hard to study truly aversive learning and processing in a human economic context. Fortunately, along with a number of inventive attempts along these lines, substantial data relevant to these issues have been collected in experimental psychology and behavioural neuroscience using other forms of aversive outcomes, and this chapter is underpinned by these results, along with the much more substantial understanding of reward, which is aversion's evil twin.

Through such sources, the broad outline of the architecture of decision-making is slowly emerging. There is ample evidence that a number of systems is involved in making, and learning to make, predictions about future positive (which, in the psychological literature are often called *appetitive*) and negative (*aversive*) outcomes, and in choosing actions that generally increase the former and decrease the latter (Adams and Dickinson, 1981; Daw et al., 2005; Dayan, 2008; Dickinson and Balleine, 2002). Cooperation among, and competition between, the different systems influence the responses of subjects in experiments, although the exact interactions are only beginning to become clear.

In this chapter, we first outline the components of this architecture, focusing on different systems involved in evaluating outcomes and choosing actions. Their impact in the broader field of neuroeconomics has historically been most apparent in positive cases; we therefore focus on two key asymmetries between loss-related and reward-related issues. One of these relates directly to the anomalies of choice listed above, and arises from the influence on normative, reward-maximizing and punishment-minimizing choices of innate responses to aversive predictions and outcomes. That the mere prediction of an aversive outcome can have an effect on behaviour that, paradoxically, increases the chance of attaining that outcome, is an Achilles heel of decision-making with widespread unfortunate consequences (Breland and Breland, 1961; Dayan et al., 2006).

The second asymmetry has to do with learning. In one important class of tasks, subjects are penalised for any action they take except for one, particular, choice (selected ahead of time by the experimenter). For such tasks, telling subjects that they just performed a bad action does not, in general, tell them what they

could have done instead that would have been better. By contrast, telling them that an action was not bad is much more specifically useful. The consequence of this asymmetry lies in psychological and neural aspects of the interaction between learning associated with rewards and punishments. Learning which actions to execute to avoid punishments appears to require the involvement of positive signals created through mutually opponent interactions between separate systems involved in appetitive and aversive predictions and outcomes. The positive signal arises in the light of the progression from a state in which punishment is expected, to a state in which it is not. Although this asymmetry has fewer direct consequences for existing economic tasks, for which learning often plays a somewhat restricted role, it is important in ecologically more natural settings.

We start by describing the architecture of prediction and decision-making in positive and negative contexts. We then discuss a class of so-called Pavlovian influences over choice in negative contexts; and finally consider issues to do with learning. Loss aversion itself is discussed in detail elsewhere (see the chapter by Fox and Poldrack).

2. The architecture of affective decision making.

The fields of economics, operations research, control theory and even ethology share a common theoretical framework for modelling how systems of any sort can learn about the environments they inhabit, and also can come to make decisions that maximize beneficial outcomes and minimize adverse ones (Camerer, 1995;

Mangel and Clark, 1988; Puterman, 1994; Sutton and Barto, 1998). This framework is closely associated with dynamic programming (Bertsekas, 1995), and encompasses many different algorithmic approaches for acquiring information about an unknown environment, including learning from trial and error, and using that information to specify controls. It has recently become apparent that different structures in the brain instantiate various of these approaches, in some cases in a surprisingly direct manner; producing a complex, but highly adapted and adaptive overall controller (Daw et al., 2005; Dayan, 2008; Dickinson and Balleine, 2002).

In many cases for experimental and behavioural economics, the specification of the problem includes exactly the full costs and benefits of each course of action in a stylized tableau (Camerer, 1995). However, in typical natural cases of decision making, this simplifies away at least two issues. First, feedback for a choice is usually only available after some time has elapsed, and, potentially, also additional choices (as, for instance, in a maze). This problem of delayed feedback seems to have played an important role in determining the nature of the neural controllers, with forms of prediction lying at their heart (Montague et al., 1996; Sutton and Barto, 1998). The second main difference between natural and economic decision-making is that the latter mostly involves money, which only has derived, and not intrinsic, value to the subjects. The extent to which proxies such as money, let alone more abstract outcomes such as mere points in a computer game, can entrain neural decision-making structures that are presumably evolved to handle natural rewards (‘reinforcers’) such as food, water, and intrinsic threats, is actually quite remarkable.

The essence of the solution to the problem of delayed feedback is prediction of the value of being in a particular situation (typically called a ‘state’) and/or doing a particular action at that state, in terms of the rewards and punishments that can be expected to accrue in the future. Different ways of making predictions underlie different approaches to control, leading to an overall architecture that is complicated. In particular, we have suggested that there is evidence for at least four different sorts of predictor or value system and four different sorts of controller (Dayan, 2008). However, for the present purposes, two predictors and three associated controllers are most important.

The predictors (called model-based and model-free, for reasons that we discuss below) trade off the complexity of learning for the complexity of computation. These predictors are directly associated with two of the controllers (which psychologists refer to respectively as goal-directed and habitual). The third controller (called Pavlovian) uses the model-based and model-free values but emits responses that are selected by evolution rather than learning. We argue that the Pavlovian controller plays a critical role in creating decision-theoretic anomalies (Dayan et al., 2006).

In the rest of this section, we describe these key value systems and controllers. We organize the descriptions around the simple rodent maze task shown in figure 1a (adapted from Niv et al, 2006). This has three choice points (A, B and C); and four possible outcomes (cheese, nothing, water, and carrots). When the animal is hungry, the cheese is most valuable, i.e., has the highest outcome utility, followed by the water and carrots; when thirsty, the water is most valuable. However, the cheese can be devalued, either by allowing the animal to eat it freely until it chooses to eat it no more (this is called sensory-specific satiety, since the value of the cheese is

specifically lowered), or by injecting the animal with a chemical (lithium chloride) after it eats some cheese. The latter treatment makes the animal sick, an outcome that induces a form of specific food aversion, such that, again, the cheese is no longer valuable. Figure 1b shows the utilities of each of the outcomes under the three motivational states of hunger, thirst and cheese aversion.

2.1 Model-based values; goal-directed control

One obvious way for a subject to make predictions about future punishments or rewards is to use a model of the world. This model should indicate the probability with which the subject will progress from one state to the next, perhaps dependent on what actions it takes, and what the likely outcomes are at those states, which again may depend on the actions (Sutton and Barto, 1998). Figure 1c depicts the model of the simple maze task; it is nothing more than the tree of locations in the maze, which are the states of the world, joined up according to the actions that lead between them. Not only does the model specify which outcomes arise for which actions, it should also specify the (expected, experienced) utility of those outcomes. As shown in the figure, this depends on the motivational state of the subject. The information necessary for the model can readily be acquired from direct experience in an environment, at least provided that the environment changes at most relatively slowly.

Given some systematic way of acting at each location or state (e.g., choosing to go left or right with probability 0.5), models such as that shown in figure 1c admit a conceptually very simple way of making predictions about the values of states or locations in the maze, namely searching forward in the model, accumulating

expected values all the while. Unfortunately, computing model-based values accurately when there are many different possible states and actions, places a huge burden on working memory, and also on aspects of calculation. The values can therefore only possibly be accurate in rather small environments.

Of course, it is not enough to compute the value of a random choice of action at a location; rather it is necessary to find the best action. Since the model in figure 1c actually specifies the utility consequences of the different possible actions, it can also straightforwardly be used to perform the dynamic programming step of finding the optimal action. This can, in principle, be performed either forwards or backwards in the tree.

One critical facet of this model-based method of choosing actions is that the decision utilities used to make choices, i.e., the information about the expected utilities of the actions, can depend on a calculation as to which outcomes will result, and what their expected (experienced) utility will be. Take the case that the model includes all the utilities shown in figure 1b. If the subject is trained whilst hungry, it will normally turn left at A to get the cheese. However, as soon as the cheese has been devalued through pairing with illness, the prediction of the utility of going left at A will be reduced, and the subject will turn right instead to get the carrots.

In psychological terms, since these values depend on the expected outcomes and their modelled utilities, this sort of control is considered to be goal-directed (Dickinson and Balleine, 2002) since these utilities define the animals' goals. This sort of outcome-sensitive control is a close relative of human notions of 'cognitive' control, in which individuals explicitly consider the outcome of actions, and of subsequent actions, and use some form of tree-

search to inform current actions. The brain might support different ways of doing this, for instances using propositional, linguistic structures, or, by more or less direct analogy with navigation, structures associated with spatial processing. It is closely related to the classical notion of outcome-expectancy expounded by Tolman (Tolman, 1932). Indeed, model-based prediction and control has the key characteristic of being highly flexible over the course of learning – new information about the environment can be fit into the model in exactly the right place to have an appropriate effect.

Further, one might imagine that subjects could acquire higher order information about the overall structure of the environments they experience that might generalize from one particular task to another. One example that has been highly influential in the psychological literature is that of controllability – a measure of the influence a subject might expect to have over its outcomes. There is a range of experiments into what is known as learned helplessness (Maier and Seligman, 1976) in which subjects are taught that they cannot control some particular aspect of one environment (for instance, being unable to influence a shock). They generalize this inference to other environments, failing to explore or exploit them effectively. There are various possible formalizations of controllability as Bayesian prior distributions over characteristics of the models (Huys and Dayan, 2008), but more data are necessary to pin this issue down completely.

The neural instantiation of the model and associated calculations for predictions and action choice is not completely known. However, there is evidence for the involvement of several regions of prefrontal cortex, including ventromedial prefrontal cortex (related to the prelimbic and infralimbic cortex in rats), lateral orbitofrontal cortex, and middle frontal gyrus, along with the dorsomedial

striatum (Balleine and Dickinson, 1998; Carter et al., 2006; Dayan and Balleine, 2002; Koechlin et al., 2003; Ursu and Carter, 2005; Yin et al., 2006; Yoshida and Ishii, 2006). Most of these experiments involve rewards rather than punishments, though, and the representation of model-based negative values is not wholly clear.

2.2 Model-free or cached values; habitual control

The problem with model-based prediction and control is the complex, and thus error-prone, calculations that are necessary to compute values. One way round at least some of this complexity is to collapse the total anticipated value of future state transitions or actions by storing or, to use a word taken from computer science, caching, what would be the results of this tree search (Daw et al., 2005; Sutton and Barto, 1998). In effect, a cached value provides a single simple metric, an outcome-independent neural currency, as to the overall utility of a particular state, or taking a certain action at that state.

Figure 1d shows the cached values (called Q-values) of each action at each location in the maze, assuming that the subject chooses optimally for the state of hunger. Such cached values can be used without direct reference to a model of transitions or outcomes; hence this form of prediction is often termed model-free. These values are represented by a function (the Q function) whose argument is the state (here, the location in the maze).

Of course, the cached values in figure 1d are just the same as the optimal values produced by model-based evaluation in the case of hunger. However, critically, it turns out that these values can be learned directly over the course of experience of state transitions

and utilities, without any reference to a model at all. Ways to do this, i.e., ways of implementing asynchronous, sampled, dynamic programming, are highlighted elsewhere in this volume under the guise of temporal difference methods of reinforcement learning (Barto et al., 1990; Sutton and Barto, 1981; Watkins and Dayan, 1992). Temporal difference learning works by exploiting the key property possessed by the cached values in figure 1d, namely consistency from one state to the next. For example, since no outcome is provided at state A, the value of going left at that state is just the same as the value of the state (B) consequent on going left there; the value of going right is the same as the value of the state (C) that arises for going right. The discrepancy (if any) between these successive value estimates is exactly the basis of the temporal difference learning rule. In this way, sequential estimates of values effectively transfer between adjacent states, obviating the need to wait for actual outcomes themselves.

Perhaps surprisingly, it turns out that temporal difference algorithms are not just distant abstractions over baffling neural complexities. Rather, at least in the case of positive outcomes, there is substantial evidence (also reviewed elsewhere in the volume) that the moment-by-moment (phasic) activity of cells that project the neuromodulator dopamine to the striatum matches closely the key prediction error term in temporal difference learning, providing a signal that is ideally suited for manipulating predictions appropriately (Montague et al., 1996; Nakahara et al., 2004; Satoh et al., 2003; Schultz et al., 1997). Unfortunately, the case of aversive outcomes is less well understood. fMRI studies suggest that punishments lead to prediction errors with rather similar properties to those for rewards (Jensen et al., 2006; Seymour et al., 2004), although electrophysiological evidence from animals is thinner on the ground (Belova et al., 2007).

Most importantly for model-free predictions is that the brain appears not to use the obvious representation in which rewards (and positive prediction errors) are coded by greater-than-average neural activity in a neural population, and losses (and negative prediction errors) by less-than-average neural activity in the same population. Rather, as in many other cases, it seems to use two systems that oppose each other (Dickinson and Dearing, 1979; Grossberg, 1984; Konorski, 1967; Seymour et al., 2005; Seymour et al., 2007a; Solomon and Corbit, 1974). In this arrangement, positive outcomes can inspire responses from the negative system when they are unexpectedly omitted, or when sequences of them cease. Further, stimuli which predict the absence of rewards (called appetitive inhibitors), and stimuli which predict the presence of punishments or loss (aversive excitors) are treated in a formally similar manner. For example, in terms of value representations, omission of food is intrinsically similar to painful shocks. This is demonstrable in various psychological paradigms (Dickinson and Dearing, 1979). Conversely, there is a natural similarity between appetitive excitors and aversive inhibitors.

The neural realization of the system associated with negative, model-free values that is opponent to dopamine is not completely resolved. One class of theoretical models hints at the involvement of a different neuromodulator called 5-Hydroxytryptamine (5-HT or serotonin), as a more or less direct opponent (Daw et al., 2002). However, direct evidence for this possibility is scant, there are competing theories for the role of this neuromodulator, and the fMRI studies, with their poor spatial resolution and the uncertainties about exactly what aspects of neural activity they capture in structures such as the striatum (Jensen et al., 2003; Jensen et al.,

2006; Seymour et al., 2005; Seymour et al., 2007a), leave us without a completely unified picture.

In fact, until recently, the striatum had been considered to be reward-specific in economic studies in humans. However, the findings above and others (Seymour et al., 2004, Delgado and colleagues, forthcoming), along with ample animal studies (Horvitz, 2000; Ikemoto and Panksepp, 1999; Schoenbaum and Setlow, 2003; Setlow et al., 2003; Wilson and Bowman, 2005) suggest that the striatum is involved in both appetitive and aversive processing, and indeed may be a critical point in the brain where these opposing motivational streams are integrated. Slightly clearer is the representation of the cached aversive values themselves, which evidently involves the amygdala and anterior insula cortex (Paton et al., 2006; Seymour et al., 2004).

The clear advantage of that model-free, cached, values have over model-based values is that they are represented directly, and do not need to be computed by a process of tree-based evaluation that imposes a heavy burden on working memory, and is likely to be inaccurate in even moderately complex domains. However, attending this computational benefit is statistical inefficiency over learning, and inflexibility in the face of change.

First, the drive underlying temporal difference learning is discrepancy between the predictions made at successive states. However, early in learning, the predictions at all states are wildly inaccurate, and therefore the discrepancies, and thus the temporal difference prediction error, are of little use. Thus model-free learning is statistically inefficient in the way it employs experience. To put the point another way, temporal difference learning involves bootstrapping (i.e., using one estimate to improve another one), a

procedure which is far from optimal in its use of samples from the environment.

The second problem with model-free methods is inflexibility. As we noted, cached values such as those shown in figure 1d are just numbers, divorced from the outcomes that underlie them, or the statistics of the transitions in the environment. This is why caching is computationally efficient. However, if the motivational state of the subject changes (for instance if the cheese is poisoned, as in the rightmost column of figure 1b), then the cached values will not change without further, statistically expensive, learning. By contrast, the model-based values, which are based on direct evaluation in the tree of outcomes, can change directly.

In figure 1d, the model-free values are predictions of the long-run utilities of particular actions at each location. They can thus be directly used as decision utilities, to choose between the possible actions at a location. This leads to a model-free controller, one that makes decisions without reference to a model of the environment. We pointed out above that the cached values do not change with the motivational state of the subjects without further learning, and so the model-free decisions will not change either. In psychological terms, this is exactly the characteristic of habits (Dickinson and Balleine, 2002) and so this controller is deemed habitual (compared with the goal-directed control associated with the model-based value system).

From a neural perspective, there is evidence for the involvement of the dorsolateral striatum in representing the values of taking actions at states (Yin et al., 2006), and indeed in habitual control. In the appetitive case, again, dopaminergic projections from the substantia nigra pars compacta to this region of the striatum are

believed to play a central role in learning (Montague et al., 1996; Schultz et al., 1997). The habits themselves may be represented or stored in cortico-thalamic loops (Yin and Knowlton, 2006).

The habitual controller defined above involves the competition between different actions depending on values (or other quantities depending on the values) that are the output of a function of the state (the Q function in figure 1d). An even more primitive form of habitual controller would use a function to parameterize the mapping from state to action directly, without going through the intermediate value of a range of actions (Barto et al, 1983). Psychologists would consider this to be a stimulus (i.e., state) – response (i.e., action) mapping. It is also model-free, and insensitive to motivational changes, and thus hard to distinguish behaviourally from the Q-value-dependent, model-free controller described above. There are intriguing reports of just such a controller in even more dorsolateral striatal regions (Everitt and Robbins, 2005).

The existence of multiple controllers (goal-directed and habitual) gives rise to a new choice problem, that of choosing between them. One view is that they compete for the control of behaviour according to their relative uncertainties (Daw et al., 2005). Model-based values are favoured early in the course of learning, because of their greater statistical efficiency. However, model-based values are disdained once sufficient samples have accumulated, because the computational demands of calculating them inevitably lead to extra noise.

Most work distinguishing habitual and goal-directed control has involved appetitive outcomes; we discuss some subtleties of aversive habitual control in section 4.

2.3 Pavlovian control.

Model-based and model-free controllers can, in principle, learn arbitrary actions to optimize their behaviour, at least those actions that can be expressed and explored. Indeed, these are often referred to as instrumental controllers, since their choices are learned to be instrumental for the delivery of desired outcomes. Although this flexibility is very powerful, it comes with an attendant cost of learning. Evolution appears to have endowed everything from the simplest organisms to us with powerful, pre-specified, but inflexible alternatives (Dickinson, 1980; Konorski, 1967; Mackintosh, 1983). These responses are called Pavlovian, after the famous Russian physiologist and psychologist Pavlov.

Immediately available rewards such as food or water, and immediate threats, such as pain or predators (collectively called unconditioned stimuli), elicit a range of apparently unlearnt, typically-appropriate, so-called consummatory, responses. For appetitive outcomes, these are relatively simple, although they may reflect certain specific attributes of the outcome, for instance, differing for solid and liquid outcomes. The consummatory responses associated with aversive outcomes appears to be more sophisticated than those for rewards, including increased heart rate and sweating during acute pain, fighting in the midst of a contest, and leg flexion in the face of foot-shock. The choice between the whole range of defensive and aggressive responses depends rather precisely on the nature of the outcome, the context, and particularly the effective (`defensive') distance of the threat (Blanchard and Blanchard, 1990). These responses are seemingly under the control of a brainstem structure called the periaqueductal grey (PAG),

which has a rich, topographically organised architecture (Fanselow, 1994; Fendt and Fanselow, 1999; Graeff, 2004; Mobbs et al., 2007).

However, and more relevantly for us, predictions associated with these appetitive or aversive outcomes also elicit an often somewhat different set of so-called preparatory responses. These are automatically tied to the predictions, independent of whether they are actually appropriate responses in a given circumstance. They thus provide an additional route by which the predictive mechanisms discussed in the previous subsections can generate behaviour.

Such preparatory responses are also varied. For instance, in rats, anticipation of a shock causes attempted escape if the cue underlying the anticipation is localised at a particular point in the environment (e.g., a light LED), but freezing if it is more general. Such anticipation can also lead to fighting in the presence of another male, and copulation in the presence of a female (Sachs and Barfield, 1974; Ulrich and Azrin, 1962). However, there are also preparatory responses that reflect the general positive or negative valence of the predicted outcome, and elicit non-specific responses such as approach or withdrawal. We suggest in the next section that it is these general preparatory responses, arising largely from predictions of financial gain and loss, that are associated with significant behavioural anomalies in human economic choices.

The neural realization of Pavlovian responses has been well studied (LeDoux, 2000; Maren and Quirk, 2004). As mentioned above, aversive value predictions depend critically on the amygdala (Balleine and Killcross, 2006; Cardinal et al., 2002). The amygdala is a complex and incompletely understood structure with many sub-

parts. However, it seems that one sub-area, called the central nucleus, is predominantly involved in directing non-specific preparatory responses. These include arousal and autonomic responses, and also approach/withdrawal, achieved through its extensive connections to brainstem nuclei and one part of the nucleus accumbens (called the core). Another sub-area, called the basolateral complex, is predominantly involved in much more specific responses, mediated downstream through connections to regions such as the hypothalamus and periaqueductal grey and a separate part of the nucleus accumbens (called the shell).

3. Pavlovian influences over instrumental behaviour.

The responses of the Pavlovian controller are determined by evolutionary (phylogenetic) considerations rather than (ontogenetic) aspects of the contingent development or learning of an individual. These responses directly interact with instrumental choices arising from goal-directed and habitual controllers. This interaction has been studied in a wealth of animal paradigms, and can be helpful, neutral or harmful according to circumstance. Although there has been less careful or analytical study of it in humans, we have argued that it can be interpreted as underpinning a wealth of behavioural aberrations (Dayan et al., 2006).

Crudely, predictions of future appetitive outcomes lead to engagement and approach; predictions of future aversive outcomes lead to disengagement and withdrawal. For instance, consider the phenomenon of Pavlovian-instrumental transfer (PIT) (Dickinson and Balleine, 2002; Estes, 1948; Lovibond, 1983). In this, the

speed, rate, alacrity or, more generally, vigour with which subjects perform an instrumental response for a particular positive outcome is influenced by the mere presentation of stimuli that are associated in a Pavlovian manner with either appetitive or aversive outcomes. In our terms, the stimuli signal states; the important aspect of PIT is that the predictive association of the Pavlovian stimulus occurs separately from that of the instrumental context.

Vigour is boosted the most by stimuli predicting an appetitive Pavlovian outcome that is exactly the same as the outcome of the instrumental behaviour. This so-called specific PIT depends (at least in rats) on the integrity of the basolateral amygdala and nucleus accumbens shell (Cardinal et al., 2002; Corbit et al., 2001; Corbit and Balleine, 2005). However, vigour is also boosted by stimuli predicting motivationally relevant appetitive outcomes (such as water, for a thirsty subject) that are different from the instrumental outcome. This is called general PIT, and may be seen as a general, non-selective, preparatory appetitive phenomenon. In rats, general PIT depends on the integrity of the central amygdala and nucleus accumbens core (Cardinal et al., 2002; Corbit et al., 2001; Corbit and Balleine, 2005), in keeping with the description above about the neural realization of Pavlovian conditioning.

Finally, stimuli predicting aversive Pavlovian outcomes can actually suppress appetitive instrumental responding, and lead to extraneous actions such as withdrawal. This is normally called conditioned suppression (Estes and Skinner, 1941), rather than aversive PIT, which would perhaps be the more natural term. However, it is a ubiquitous and powerful phenomenon that is in fact often used as a sensitive measure of the strengths of aversive Pavlovian predictors.

Most critically, choice, as well as vigour, is affected by these Pavlovian predictions. This is seen very clearly in a slightly complex paradigm called negative automaintenance (Williams and Williams, 1969). In one example of negative automaintenance, pigeons are shown the predictive association between the illumination of a key and the delivery of food. The Pavlovian prediction associated with the lighting of the key automatically elicits a peck response on the key, as a form of preparatory approach and engagement In fact, this part of the procedure is one of the standard forms of Pavlovian conditioning, which is called autoshaping because of the automaticity of the pecking (Brown and Jenkins, 1968). The experimenter then arranges what is called an omission schedule, so termed because on any trial in which the pigeon pecks the key when illuminated, no food will be provided. In this case, there is a battle between the Pavlovian response of pecking and the instrumental need to withhold. Pigeons cannot help themselves but peck to some degree, showing the critical, and indeed in this case, deleterious, impact of the Pavlovian prediction.

Although it has been suggested that Pavlovian responses interfere comparatively more with instrumental habits than goal-directed actions, the factorial PIT-based interactions between model-based and model-free Pavlovian predictions and model-based and model-free instrumental actions have not been systematically studied.

There appear to be fewer aversive examples of phenomena like negative auto-maintenance, which is somewhat surprising given the robustness of Pavlovian aversive responses in general. Where they can be shown to exist, they yield self-punitive behaviour. In one putative example, squirrel monkeys were punished, by way of an electric shock, for pulling on a restraining leash (Morse et al., 1967). The (instrumentally) optimal action in such a circumstance is

to stop pulling, however one Pavlovian response to shock in the time leading up to its expected delivery is to try and escape by pulling. As expected from Pavlovian misbehaviour, the monkeys did in fact pull on the leash more rather than less. A similar example is seen in Siamese fighting fish, who can be trained to swim through a hoop and perform an aggressive display. If an experimenter then tries to inhibit this display by an aversive shock, the behaviour is paradoxically augmented (Melvin and Anson, 1969). This is most likely since the aggressive display is part of the innate repertoire of defensive responses, which turns out to be extremely difficult to overcome. This is a form of self-punitive behaviour.

What then are the neuroeconomic consequences of these Pavlovian effects? After a methodological note, we briefly consider four: impulsivity, framing, depressive realism, and dread. Note that these are all complex and rich phenomena; we only focus on the subset of issues that Pavlovian control may explain. This may seem like the same sort of smorgasboard of issues to which other broad explanatory frameworks such as hyperbolic discounting have been turned; rather, we argue that it is critical to understand the breadth of phenomena associated with something as basic as Pavlovian conditioning, given its overwhelming evidentiary basis in psychology and neuroscience.

3.1 Methodology.

We must first raise a couple of methodological points about the relationship between economic and psychological paradigms. In experimental and behavioural economics, decisions are often probed in relation to options with stated parameters, that is, the magnitudes, risks and uncertainties of various options are given directly. These are likely to exert their effects mostly through

model-based predictions (and goal-directed control). By contrast, in experimental psychology, the parameters of options are typically learned through trial and error. Thus, representations of value and risk are experience-based rather than propositional, and can have an impact through model-free as well as model-based control. Of course, experience-based representations are imperative in animal experiments, and have also been highly successful in deconstructing the components of aversive (and appetitive) behaviour. However, any complete account of aversive behaviour needs to integrate both, since humans are presented with both types of situation: one shot decisions such as those regarding pensions and life insurance; and repeated decisions, such as those regarding what painkiller to take or which foods to buy.

A further difference in methodologies relates to type of aversive events used. Neuroscientists have often used pain, for instance in the form of an electric shock to hand or paw. The advantage of this is it is an immediately and relatively instantaneously consumed commodity. Furthermore, it is both potent and ecologically valid, in the sense that it is the sort of stimulus with which aversive systems evolved to deal. We should therefore say a word about the neural processing of pain itself.

Physical pain is subserved by a sophisticated system of specialised neural pathways signalling information about actual or imminent tissue damage to many areas of the spinal cord and brain (Craig, 2002; Fields, 2004; Julius and Basbaum, 2001). This results not just in the set of characteristic, involuntary, defensive responses described above, but also a perceptual representation of negative hedonic quality.

In the brain, the basic representation of aversive innate value implicates brainstem and midbrain structures, including the periaqueductal grey, parabrachial nucleus, and thalamus (Lumb, 2002). Cortical structures such as insula (particularly anterior regions), lateral orbitofrontal and mid-anterior insula cortices are more directly associated with refined aversive representations, including conscious negative hedonic experience (Craig, 2002). These correlate more closely with the subjective experience of unpleasantness, which in humans, often accompanies innate aversive outcomes. In fact, the feeling associated with loss dictates the way these systems are often described in traditional psychological accounts (Price, 1999). This can, however, be approached more formally by considering 'feeling' as a process of hedonic inference. As with many less motivationally-laden sensory systems, afferent information is rarely perfect, and a statistically informed approach is to integrate afferent input with either concomitant information from other modalities (multi-sensory integration), or prior knowledge of events (expectation) (Seymour, Daw and Dayan, forthcoming).

By contrast with these rich phenomena associated with actual threats, economists have, naturally, tended to use financial losses. Various of the other chapters capture aspects of the psychological and neural richness of money as a stimulus; for simplicity, we adopt the straightforward view of it as a conditioned reinforcer, that is, a stimulus that has undergone (extremely extensive) Pavlovian training to be associated with many different sorts of future reward. In these terms, losing money is like taking away a conditioned reinforcer; an outcome that is indeed known to be aversive.

One complicating issue is the slightly unclear relationship between the affective values of states and those associated with state

changes (Kahneman and Tversky, 2000). To take a concrete example – take the state of hunger. On one hand, this would seem to be clearly an aversive state – it poses a threat to homeostasis. On the other, the affective worth of the same morsel of food is greater when hungry than when sated, and so, for instance, the average long-run experienced utility may actually be higher (Niv et al., 2006). Is the apparently masochistic act of starving yourself actually utility maximizing in that you enjoy food in the future sufficiently more? In general, teasing apart the contribution to utility of the actual outcome and the motivational state within which it is evaluated is hard.

The answer to the masochism question is not yet quite clear. However, it does pertain to one of the other value systems that we have not yet discussed. Most economic decision-making tasks are one-shot or phasic. By comparison, many psychological paradigms for animals are on-going or continuous. For these, it often makes sense to predict and maximize the long-run average rate of rewards rather than, for instance, the more conventional long-run sum of exponentially-discounted rewards. In this case, this average rate of reward has a status as something like an opportunity cost for time. Niv and colleagues (2007) noted this, and studied a framework in which subjects were free to choose not only which actions to do, but also how fast to do them. Under the reasonable assumption that acting quickly is expensive, it turns out that the optimal speed or vigour of responding is determined by the average rate of reward. Arguing partly on the basis of the data on the control of vigour from the Pavlovian-instrumental transfer paradigms we discussed above, they suggested that the long-run, tonic, level of dopamine or dopaminergic activity should report this average reward. This is the additional value system. However, vigour is also important in cases in which signalled punishments or aversion can be avoided through

active actions. Tonic dopamine may therefore represent the sum of average rewards and avoidable punishments; bar the expectation of a long-run absence of food, hunger is exactly an example of this sort of case. Whether the tonic aversiveness of hunger is also represented by the tonic activity of another system (for instance, some subset of 5-HT cells) is not clear.

For the present, we will just consider phasic aversive outcomes, such as shocks, or immediate financial losses, together with predictions of these. Neurobiological evidence is starting to accrue that confirm that the underlying motivational processes in financial loss share strong similarities with that associated with physical pain (Delgado et al., 2005; Knutson et al., 2007; Seymour et al., 2007a). For example, Knutson and colleagues have suggested that financial amounts associated with payments in shopping transactions are correlated with activity in and around insula cortex (Knutson et al., 2007), which has also been shown to correlate with expected value of pain (Seymour et al., 2004). We have shown activation to prediction errors for financial loss in striatum, in a similar manner to those seen in studies of aversive conditioning for painful shocks (Seymour et al., 2007a). Delgado and colleagues (forthcoming) have recently shown directly the common striatal aversive processing for pain and financial loss, by engaging subjects in a task that involves both.

3.2 Impulsivity and altruistic punishment.

Impulsivity covers a broad range of phenomena. Classically, it features engagement in actions whose immediate benefits are less than those of longer term pay-offs that would accrue if the subjects could be patient (Cardinal et al., 2004). That is, subjects exhibit

temporal short-sightedness. Impulsivity is best described in the appetitive domain, but similar notions may apply in aversive domains too. In the appetitive case, we have argued that the effect of a Pavlovian approach response associated with a proximally available beneficial outcome can be to boost early, and thus impulsive, responding at the expense of what would be favoured by goal-directed or habitual instrumental systems (Dayan et al., 2006). Treating this form of impulsivity in Pavlovian terms amounts to a subtly different explanation of the behaviour from accounts appealing to (or data fitting with) hyperbolic discounting or indeed ideas about differences between (model-based) rational and (model-free or perhaps neuromodulator-based) emotional cognition, which conventionally ignore the normative intent of model-free control.

In the aversive case, one example of apparent impulsiveness is altruistic punishment, in which subjects punish others (typically free-riders who fail to cooperate in various forms of group interactions, but nevertheless take advantage of the group effort) at a pure cost to themselves (i.e., with negative immediate benefit), without any prospect of a direct return on this investment of effort or risk (i.e., with no long term pay-off at all). Although the nature of the actions which subserve altruistic punishment remain unclear (Seymour et al., 2007b), there is good evidence that humans readily engage in such actions (Fehr and Gachter, 2002; Yamagishi, 1986), (see chapter by Fehr in this volume).

Certainly, some aspects of apparent altruism can be explained by reputation formation (a form of indirect reciprocity) and tit-for-tat (a form of direct reciprocity). These can be captured by model-based and even model-free instrumental mechanisms. The argument that altruistic punishment is partly a Pavlovian anomaly is

that (a) punishment is a form of aggression, whose innate roots we explored above, and that (b) in highly social species such as humans, there is an evolutionary imperative to prevent exploitation by free-riders that is satisfied by making non-cooperation expensive. First, innate aggression is evidently a potentially life saving mechanism of defence in the face of predators, and in within-species contests, can be important for protecting food, territory and mating partners (Clutton-Brock and Parker, 1995). Second, in humans, and possibly some other primate species, aggressive responses can also serve to promote cooperation, since they provide a negative incentive for members of a group to exploit each other, and protect various forms of reciprocity (Boyd and Richerson, 1992; De Waal, 1998; Stevens, 2004). Thus innate responses to perceived unfairness may have evolved on the basis of punishment in these sorts of non-altruistic circumstances, such as in groups or societies of small enough size such that individuals (and certainly their kin) would be likely to interact repeatedly with offenders, rendering the punishment non-altruistic (ie. 'selfish'). However, once established as an innate response, punishing non-cooperators could have become blind to its proximal consequences for the individual (like other Pavlovian responses), thus appearing impulsive.

There is also the alternative possibility that altruistic punishment arises from the structural inefficiency of instrumental control associated with habits, rather than the interference of Pavlovian imperatives over instrumental ones. Crudely, the idea is that choosing precisely whom to punish in a circumstance requires the detailed calculations of the consequences of punishment and likelihood of future interactions that only the goal-directed system could entertain. However, the habit system can engage in instrumental punishment in reciprocal cases and may therefore gain

control over all similar such conditions, as discussed above. Its inability to calculate in detail the consequences of its output can then lead it to punish 'inappropriately' in altruistic situations. This type of 'error' resembles that seen in devaluation experiments, when habitually trained animals fail to reduce responding to outcomes that have been separately paired with punishment.

3.3 Framing effects.

Framing effects are a rather well-studied peculiarity of human (and non-human; Santos and Chen, this volume) choice in which the decision between options is influenced by subtle features of the way in which those options are presented. Typically, the language used to describe an option is manipulated in a valance related manner, whilst the expected value remains unchanged.

The so-called 'Asian disease dilemma' is a popular example. In this, subjects are asked to choose between two options relating to the management plan of an epidemic, one of which contains risk, and the other not (Tversky and Kahneman, 1981). The risky option is fixed, such as 'Option A has 2/3 chance of curing all 600 affected people', but the non-risky option is presented in either a positive or negative frame, as either 'With Option B, 200 people will be *saved* ' or 'With Option B, 400 people will *die*'. Subjects tend to choose the risky option when the sure option is presented in terms of people dying, and the sure option when presented in terms of the numbers who will be saved.

Similarly, De Martino et al (2006) conducted a study involving loss/gain framing of non-risky, alongside risky, financial options, matched for expected value. Subjects showed a risk preference

reversal from risk aversion to risk seeking when the choice was switched to a loss frame. This change in behaviour was positively correlated with amygdala activity.

Given the role of the amygdala in Pavlovian-instrumental transfer, and thus the untoward influence of predictions on instrumental actions (Corbit and Balleine, 2005), results such as this are consistent with a Pavlovian component to framing. That is, an option that is presented as involving sure deaths will automatically engage a Pavlovian aversive withdrawal response decreasing its propensity to be chosen, that is absent for the option involving sure survival. The latter might generate an appetitive approach response instead. As we have seen above, model-based evaluation mechanisms, which could compute the equality between the options, are not the only source of predictions; model-free mechanisms, which lack such computational power, also exert their influence, in this case in just the direction shown. Indeed one can look at the classic trolley moral dilemmas (Thomson, 1986) in a similar light. Even if subjects didn't have any choice, but just had to execute an action to register a single option, we would predict that the same Pavlovian effect would make their reaction times slower, an effect seen in other experiments (Shidara et al., 2005; Sugase-Miyamoto and Richmond, 2005).

3.4 Depressive realism.

In comparisons between healthy volunteers and patients with depression, a (not completely uncontroversial) finding is that the volunteers are unduly optimistic about the appetitive value of, and the degree of control they exert over, artificial, experimentally-created environments. By contrast, the depressed subjects make

more accurate assessments, and so are more realistic. This phenomenon is called depressive realism (Abramson et al., 1979). Further, by comparison with control subjects, depressed patients ruminate on negative outcomes.

It has been suggested that Pavlovian withdrawal associated with predictions of negative outcomes is an important route to the over-optimism of the volunteers, and that one of the underlying neural malfunctions associated with depression is associated with a weakening of this withdrawal, thereby leading to more accurate, but more pessimistic, evaluations (Huys and Dayan, 2008). Consider a healthy subject entertaining chains of thought about the future. Any chain of thought leading towards a negative outcome engenders a Pavlovian withdrawal response, which may lead to its being terminated or (in the jargon of tree-based search) pruned. Thus if healthy subjects contemplate the future, they will tend to favour samples with more positive outcomes, and will therefore be more optimistic. Given the possibility that this form of Pavlovian withdrawal is mediated by 5-HT, as the putative aversive opponent to dopamine (Daw et al., 2002), and the pharmacological suggestion that depressed patients have low effective 5-HT levels (Graeff et al., 1996), it is conceivable that this withdrawal mechanism is impaired in the depressed subjects. This would, of course, lead to the basic phenomenon of depressive realism. Indeed, boosting 5-HT, which is the ultimate effect of the standard treatment for depression, namely selective 5-HT reuptake inhibitors, helps restore the original optimism.

Altered levels of 5-HT are also associated with other phenomena such as impulsivity (Cardinal, 2006; Chamberlain and Sahakian, 2007) which have been argued to have Pavlovian roots.

3.5 Dread.

In an aversive domain, many subjects show an additional sort of impulsivity in the form of dread (Berns et al., 2006). They prefer a larger electric shock that comes sooner to a weaker shock that comes later, reportedly because of the misery of aversive anticipation (Caplin and Leahy, 2001; Loewenstein, 1987; Loewenstein, 2006). Indeed, during the anticipation phase in the study by Berns and colleagues, brain regions commonly associated with physical pain are activated, as if the anticipation was indeed actually miserable. Subjects also exhibit related behaviours such as not collecting free information if it is likely to provide bad news. These phenomena can be decision-theoretically rebadged by appealing to a psychologically rich utility model (Caplin and Leahy, 2001). The question for us is the psychological context of these utilities.

Three Pavlovian issues appear to be important. First, the activation of the primary pain system is consistent with a Pavlovian phenomenon called stimulus substitution, in which predictors of particular outcomes are treated in many respects like those outcomes themselves. Although the neural foundations of this are not clear, let alone its evolutionary rationale, it is an effect that is widely described, particularly in appetitive circumstances. For instance, the way that a pigeon treats a key which has a Pavlovian association with an appetitive outcome depends directly on whether it is food or water that is predicted. The pecks that result are recognisably associated with the specific outcome itself. The activation of the primary pain areas may arise through model-based stimulus substitution. If this then leads to an effective overcounting

of the temporally distant shock, it can make the subject prefer the immediate one.

The other two Pavlovian effects are related to those discussed in the context of depressive realism. Not seeking information that is likely to be aversive is exactly akin to not exploring, or actually pruning, paths of thought that are likely to lead to negative outcomes. For dread itself, we can speculate as to the effects of the guaranteed prospect of a substantially delayed, future aversive outcome whose occurrence cannot be accurately predicted because of the inaccuracy in timing intervals (Gibbon et al., 1997). This has both model-based and model-free consequences for the Pavlovian mechanism that creates optimism through pruning. From a model-based perspective, it creates a prior expectation of environments that are relatively unpleasant because they contain unpredictable aversive outcomes. Such environments are in general associated with larger average aversive values and so lead to Pavlovian avoidance (Huys and Dayan, 2008). From a model-free perspective, the persistent expectation of an aversive outcome might set a baseline level for the Pavlovian mechanism that prunes negative lines of thought. Since this baseline would be substantially more negative than usual, it would permit substantially more negative paths than normal to be explored, and therefore lead to net aversion.

4. Aversively motivated behaviour.

We have so far used the analysis of the architecture of choice to highlight how Pavlovian predictions of aversive outcomes can lead to aberrant influences over instrumental choices in a wide variety of

circumstances. However, there is an important instrumental component to aversive behaviour too. Despite the apparent lack of current neuroeconomic interest in the topic, we will discuss avoidance, which is perhaps the most important such paradigm.

In an avoidance experiment, animals (or humans) learn actions that reliably lead to their avoiding incurring losses or pains. Typically, an animal receives a warning stimulus (such as a tone or light), that precedes delivery of an aversive stimulus, such as prolonged electrification of the floor of one compartment of the experimental apparatus. At first, the individual responds only during the aversive stimulus, for instance escaping the shock by jumping into a neighbouring compartment. Conventionally, the warning stimulus will be extinguished following this escape response. After several presentations, the escape response is executed more quickly, and eventually, the individual learns to jump when observing the warning stimulus (again with the effect of turning off this stimulus), thus completely avoiding the shock.

Consideration of the problems that must be solved in avoidance hints that such behaviour may not be straightforward. For instance, how are successful avoidance actions reinforced, if by definition they lead to no outcome? (How) does a subject ever realise that the threat is gone, if it is never sampled?

Mowrer famously suggested that learning to avoid involves two processes: predicting the threat, and learning to escape from the predictor (Mowrer, 1947). These processes, proposed respectively to be under Pavlovian and instrumental control, comprise two-factor theory, which in one form or another has survived well over the past decades. Although there are many unanswered questions

about precisely how the different action systems are orchestrated in different avoidance situations, some key facts are well supported.

In particular, Pavlovian mechanisms play a critical (and multifarious) role in avoidance, and indeed Pavlovian responses to the warning stimulus alone are often capable of implementing successful avoidance. For example, jumping out of an electrified chamber, blinking in anticipation of an eye-puff, leg flexion to an electric foot plate can all completely remove an aversive stimulus, without any need for an instrumental component. That they do pays tribute to their evolutionary provenance, and led some to question the involvement of instrumental responses at all (Mackintosh, 1983) for review). The latter is implied by the experimenter-controlled arbitrariness of the required avoidance actions (although more arbitrary ones are slower to learn (Biederman et al., 1964; Ferrari et al., 1973; Hineline, 1977; Riess, 1971)).

Further, there is good evidence that the safety state that arises from successful avoidance acts as a Pavlovian aversive inhibitor (Candido et al., 2004; Dinsmoor, 2001; Morris, 1975; Rescorla, 1969; Weisman and Litner, 1969a), that is a state that predicts the absence of otherwise expected punishment. Importantly, as mentioned above, the values of aversive inhibitors at least partly share a common representation with those of appetitive excitators, ie. predictors of rewards, as is demonstrated by their ability to affect subsequent learning in appetitive domains (a phenomenon knows as transreinforcer blocking). That the safety state plays an important role in control is suggested by the fact that avoidance responses continue long after the Pavlovian aversive responses to the discriminative stimulus have extinguished, as they will of course do if avoidance is successful (Weisman and Litner, 1969b).

This places in the spotlight the role of the value attached to the warning stimulus (Bersh and Lambert, 1975; Biederman, 1968; De Villiers, 1974; Kamin et al., 1963; Mineka and Gino, 1980; Overmier et al., 1971a; Overmier et al., 1971b; Starr and Mineka, 1977). On one hand it has the power to initiate Pavlovian preparatory responses. It is also known to be able to suppress appetitive instrumental behaviour, in a similar fashion to conditioned suppression by an aversive Pavlovian predictor. On the other hand, it has the instrumental power to initiate an appropriate avoidance response.

The dissociation of components in avoidance is supported by neural data. For instance, selective lesions of the central amygdala selectively impair conditioned suppression (aversive PIT) (Killcross et al., 1997). Further, neuroleptics, which are dopamine antagonists interfere with learning avoidance responses, but not acquisition of instrumental escape responses (Cook and Catania, 1964). This effect is of particular interest, since it suggests that it may only be the dopaminergically-reported appetitive outcome of reaching the safety state that can control instrumental learning of the avoidance response, as if the reduction of the aversive prediction itself is insufficient. This would be a very strange asymmetry between appetitive and aversive systems, and merits closer investigation.

In human studies, in support of the role of appetitive pathways, dorsal striatum and ventromedial prefrontal cortex display reward-signed activities during avoidance (Kim et al., 2006; Pessiglione et al., 2006). Furthermore, they do so in a manner predicted by reinforcement learning models.

There are known to be model-based components to avoidance learning. As we discussed in section 2, one signature of this is the

immediate sensitivity of actions to changes in the state of the subject that change the values of outcomes. An example of this outcome-sensitivity is an experiment that manipulated body temperature. Henderson and Graham (1979) trained rats to avoid a heat source when rats were themselves hot. They then made the animals cold before testing them, and found that avoidance was attenuated, provided the rats had had the opportunity to experience the heat source in their new, cold state, thereby learning that it was rewarding. Selective lesions that dissociate goal-directed and habit-based components of the avoidance action are, however, currently lacking.

Sampling biases also pose a particular problem for aversive learning, since subjects will be unwilling to try options with aversive consequences in order to hone their behaviour (Denrell and March, 2001). In fact, the sloth of extinction in avoidance is an example of this -- if successful avoidance becomes reliably executed, how will the organism know if the threat has disappeared (termed the 'hot stove effect' in economics). This contrasts with the appetitive case in which extinction is immediately frustrating. Pavlovian withdrawal will also severely hinder learning actions that lead to small, immediate, losses, but large, delayed, gains.

Of course, unnecessary avoidance is only economically problematic if there is some non-negligible cost to performing the action, or if unbeknownst to the organism, the action now leads to rewards. The problem of correctly navigating this issue is an example of the famous exploration-exploitation dilemma, which is raised elsewhere in the volume. Briefly, the battle is between exploiting existing knowledge, namely the lack of punishment that evidently ensues from performing the avoidance action, and exploring the possibility that the environment has changed such that the punishment is no

longer present. The optimal solution to this dilemma is radically computationally intractable, since it depends on calculations associated with the uncertainties of unknown change. One conventional approximate approach is to behave non-deterministically, thus constantly sampling apparently lower-valued options stochastically. Another, sometimes more proficient alternative is specifically to target actions whose consequences are more uncertain, as in uncertainty 'bonus' schemes. The effect of these, in either appetitive or aversive domains, is to make subjects less risk- (and indeed ambiguity-)averse.

In sum, there is a substantial, subtle, literature on learned avoidance showing a range of intricate effects. Presently, little of this has had an impact in neuroeconomic paradigms, but it is a ripe area for exploration.

5. Conclusions

Aversion is not merely reward viewed through a looking glass. As we have reviewed here, aversion poses its own range of critical representational and learning phenomena, and exerts an important influence over a wealth of ecologic and economic tasks. We focused on just a few of these – the substantial Pavlovian effects over experimental-, behavioural- and neuro-economic constructs, and the intricate complexities of avoidance learning; but there are also many other central issues that are being actively explored. From an economic perspective, much flows from the basic finding that mere monetary losses act in a very wide range of ways like real pains, thus allowing direct generalization from (and indeed to) an extensive psychological and neural literature.

Opponency has been a central concept in this chapter, as indeed it has been over a wealth of psychological investigations. Unfortunately, although it is relatively uncontroversial that one of the opponents is dopamine, the identity, nature and even exact functional role of the other is much less clear. We and others have argued in favour of the involvement of 5-HT, however, this is not yet totally accepted. Further, whether 5-HT, or the opponent, reports all punishments, or, for instance, only those punishments that are uncontrollable, or something else, is not yet evident.

Aversion is critical, pervasive, and interesting. Most relevantly, it is in clear need of the theoretical sophistication of neuroeconomic methods and analyses that are evidently on offer.

Figure caption. Model-based and model-free actions in a simplified maze task. 1a shows a simple maze with three states (S1, S2 and S3) from which the animal has to make left-right decisions, with the terminal states yielding outcomes of cheese, nothing, water or carrots. **1b** shows the values of these outcomes under three different motivational states: hunger, thirst, and cheese devaluation. This latter state results from cheese ingestion with vomiting (artificially induced by Lithium Chloride injection in most experiments). **1c** specifies a tree-based model of the state-action environment, which can be used to guide decisions at each state by a model-based controller. **1d** specifies the cached values available to a model-free, habitual controller. Immediately after cheese devaluation, these values do not change, in contrast to the model-based controller). It is only after direct experience with the devalued cheese that the value associated with Left (S2), and subsequently Left (S1), is reduced. Figure adapted from Niv, Joel & Dayan (2006).

Acknowledgements

We are very grateful for discussions and ideas to our collaborators in these studies: Richard Bentall, Y-Lan Boureau, Nathaniel Daw, Ray Dolan, Quentin Huys, Michael Moutoussis, Yael Niv, and John O'Doherty, and for comments on an earlier version of this chapter to Paul Glimcher and Antonio Rangel. Funding was from the Gatsby Charitable Foundation.

Reference List

1. Abramson,L.Y., Metalsky,G.I., and Alloy,L.B. Judgment of contingency in depressed and nondepressed students: sadder but wiser? J.Exp.Psychol.Gen 108[4], 441-485. 1979.

2. Adams,C.D. and Dickinson,A. (1981). Instrumental Responding Following Reinforcer Devaluation. Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology *33*, 109-121.

3. Balleine,B.W. and Dickinson,A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology *37*, 407-419.

4. Balleine,B.W. and Killcross,S. (2006). Parallel incentive processing: an integrated view of amygdala function. Trends in Neurosciences *29*, 272-279.

5. Barto, AG, Sutton RS, and Anderson CW (1983). Neuronlike elements that can solve difficult learning problems. IEEE Transactions on Systems, Man, and Cybernetics 13, 834-846..

6. Barto,A.G., Sutton,R.S., and Watkins,C.J.C.H (1990). Learning and sequential decision making. In M Gabriel & J Moor, eds. *Learning and Computational Neuroscience: Foundations of Adaptive Networks.* Cambridge, MA: MIT press. 539-602.

7. Belova,M.A., Paton,J.J., Morrison,S.E., and Salzman,C.D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. Neuron *55*, 970-984.

8. Bentham,J. (1817). A Table of the springs of Action, Showing the Several Species of Pleasures and Pains, of which Man's Nature is Susceptible (London, R. & A. Taylor).

9. Bentham,J. (1823). An Introduction to the Principles of Morals and Legislation (London, T.Payne, 1789).

10. Berns,G.S., Chappelow,J., Cekic,M., Zink,C.F., Pagnoni,G., and Martin-Skurski,M.E. (2006). Neurobiological substrates of dread. Science *312*, 754-758.

11. Bersh,P.J. and Lambert,J.V. (1975). Discriminative Control of Free-Operant Avoidance Despite Exposure to Shock During Stimulus Correlated with Nonreinforcement. Journal of the Experimental Analysis of Behavior *23*, 111-120.

12. Bertsekas,D.P. (1995). Dynamic Programming and Optimal Control, Athena Scientific.

13. Biederman,G. (1968). Discriminated Avoidance Conditioning - Cs Function During Avoidance Acquisition and Maintenance. Psychonomic Science *10*, 23-&.

14. Biederman,G., D'Amato,M.R., and Keller,D. (1964). Facilitation of discriminated avoidance learning by dissociation of CS and manipulandum. Psychonom.Sci. 1, 229-230.

15. Blanchard,R.J. and Blanchard,D.C. (1990). Fear and Defence, P. F. Brain, R. J. Blanchard, S. Parmigiani, Eds. (Harwood Academic, London, 1990), pp. 89-108.

16. Boyd,R. and Richerson,P.J. (1992). Punishment Allows the Evolution of Cooperation (Or Anything Else) in Sizable Groups. Ethology and Sociobiology *13*, 171-195.

17. Breland,K. and Breland,M. (1961). The misbehavior of organisms. AmericanPsychologist 16, 681-684..

18. Brown,P.L. and Jenkins,H.M. (1968). Auto-shaping of the pigeon's key-peck. J.Exp.Anal.Behav. *11*, 1-8.

19. Camerer,C. (1995). Individual decision making, in John H. Kagel, and Alvin E. Roth, ed.: The Handbook of Experimental Economics ~Princeton University Press, Princeton NJ.

20. Candido,A., Gonzalez,F., and de Brugada,I. (2004). Safety signals from avoidance learning but not from yoked classical conditioning training pass both summation and retardation tests for inhibition. Behavioural Processes *66*, 153-160.

21. Caplin,A. and Leahy,J. (2001). Psychological expected utility theory and anticipatory feelings. Quarterly Journal of Economics *116*, 55-79.

22. Cardinal,R.N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. Neural Netw. *19*, 1277-1301.

23. Cardinal,R.N., Parkinson,J.A., Hall,J., and Everitt,B.J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci.Biobehav.Rev. *26*, 321-352.

24. Cardinal,R.N., Winstanley,C.A., Robbins,T.W., and Everitt,B.J. (2004). Limbic corticostriatal systems and delayed reinforcement. Adolescent Brain Development: Vulnerabilities and Opportunities *1021*, 33-50.

25. Carter,R.M., O'Doherty,J.P., Seymour,B., Koch,C., and Dolan,R.J. (2006). Contingency awareness in human aversive

conditioning involves the middle frontal gyrus. Neuroimage. *29*, 1007-1012.

26. Chamberlain,S.R. and Sahakian,B.J. (2007). The neuropsychiatry of impulsivity. Curr.Opin.Psychiatry *20*, 255-261.

27. Clutton-Brock,T.H. and Parker,G.A. (1995). Punishment in animal societies. Nature *373*, 209-216.

28. Cook. L. and Catania,A.C. (1964). Effects of drugs on avoidance and escape behaviour. Fed.Proc. *23*, 818-835.

29. 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.

30. Corbit,L.H., Muir,J.L., and Balleine,B.W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. J.Neurosci. *21*, 3251-3260.

31. Craig,A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. Nat.Rev.Neurosci. *3*, 655-666.

32. Daw,N.D., Kakade,S., and Dayan,P. (2002). Opponent interactions between serotonin and dopamine. Neural Netw. *15*, 603-616.

33. Daw,N.D., Niv,Y., and Dayan,P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat.Neurosci. *8*, 1704-1711.

34. Dayan,P. (2008). The role of value systems in decision making. In: Better Than Conscious? Implications for Performance and Institutional Analysis, ed. C. Engel and W. Singer. Strungmann Forum Report. Cambridge, MA: MIT Press.

35. Dayan,P. and Balleine,B.W. (2002). Reward, motivation, and reinforcement learning. Neuron *36*, 285-298.

36. Dayan,P., Niv,Y., Seymour,B., and Daw,D. (2006). The misbehavior of value and the discipline of the will. Neural Netw.

37. De Villiers,P.A. (1974).The law of effect and avoidance: a quantitative relationship between response rate and shock-frequency reduction. J Exp Anal Behav 21, 223-235.

38. De Waal,F.B.M. (1998). Chimpanzee politics: Power and sex among apes. (Baltimore, MD: Johns Hopkins University Press).

39. Delgado,M., Labouliere,C., and Phelps,E. (2005). Fear of losing money? Aversive conditioning with secondary reinforcers. Journal of Cognitive Neuroscience 196.

40. Denrell,J. and March,J.G. (2001). Adaptation as information restriction: The hot stove effect. Organization Science *12*, 523-538.

41. Dickinson,A. (1980). Contemporary animal learning theory. Cambridge, England: Cambridge University Press. 1980.

42. Dickinson,A. and Balleine,B.W. (2002). The role of learning in motivation.In: Gallistel, C.R., Editor, , 2002.Learning, Motivation and Emotion, Volume 3 of Steven's Handbook of

Experimental Psychology (Third Edition ed.), John Wiley & Sons, New York in press.

43. Dickinson,A. and Dearing,M.F. (1979). Appetitive-aversive interactions and inhibitory processes. *In Mechanisms of Learning and Motivation.* Dickinson A and Boakes RA eds. Erlbaum, Hillsdale, NJ. 203-231.

44. Dinsmoor,J.A. (2001). Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. J.Exp.Anal.Behav. *75*, 311-333.

45. Estes,W.K. (1948). Discriminative Conditioning .2. Effects of A Pavlovian Conditioned Stimulus Upon A Subsequently Established Operant Response. Journal of Experimental Psychology *38*, 173-177.

46. Estes,W.K. and Skinner,B.F. (1941). Some quantitative properties of anxiety. Journal of Experimental Psychology 29, 390-400.

47. Everitt,B.J. and Robbins,T.W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat.Neurosci. *8*, 1481-1489.

48. Fanselow,M.S. (1994). Neural Organization of the Defensive Behavior System Responsible for Fear. Psychonomic Bulletin & Review *1*, 429-438.

49. Fehr,E. and Gachter,S. (2002). Altruistic punishment in humans. Nature *415*, 137-140.

50. Fendt,M. and Fanselow,M.S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. Neurosci.Biobehav.Rev. *23*, 743-760.

51. Ferrari,E.A., Todorov,J.C., and Graeff,F.G. (1973). Nondiscriminated avoidance of shock by pigeons pecking a key. J.Exp.Anal.Behav 19, 211-218.

52. Fields,H. (2004). State-dependent opioid control of pain. Nat.Rev.Neurosci. *5*, 565-575.

53. Gibbon,J., Malapani,C., Dale,C.L., and Gallistel,C.R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. Current Opinion in Neurobiology *7*, 170-184.

54. Graeff,F.G. (2004). Serotonin, the periaqueductal gray and panic. Neuroscience and Biobehavioral Reviews *28*, 239-259.

55. Graeff,F.G., Guimaraes,F.S., DeAndrade,T.G.C.S., and Deakin,J.F.W. (1996). Role of 5-HT in stress, anxiety, and depression. Pharmacology Biochemistry and Behavior *54*, 129-141.

56. Grossberg,S. (1984). Some normal and abnormal behavioral syndromes due to transmitter gating of opponent processes. Biol.Psychiatry *19*, 1075-1118.

57. Henderson,R.W. and Graham,J. (1979). Avoidance of Heat by Rats - Effects of Thermal Context on Rapidity of Extinction. Learning and Motivation *10*, 351-363.

58. Hineline,P.N. (1977). Negative reinforcement and avoidance. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 364-414). Englewood Cliffs, NJ: Prentice Hall.

59. Horvitz,J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience *96*, 651-656.

60. Huys,Q. and Dayan,P. (2008). A Bayesian formulation of behavioral control. submitted .

61. Ikemoto,S. and Panksepp,J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. Brain Res.Brain Res.Rev. *31*, 6-41.

62. Jensen,J., McIntosh,A.R., Crawley,A.P., Mikulis,D.J., Remington,G., and Kapur,S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. Neuron *40*, 1251-1257.

63. Jensen,J., Smith,A.J., Willeit,M., Crawley,A.P., Mikulis,D.J., Vitcu,I., and Kapur,S. (2006). Separate brain regions code for salience vs. valence during reward prediction in humans. Hum.Brain Mapp.

64. Julius,D. and Basbaum,A.I. (2001). Molecular mechanisms of nociception. Nature *413*, 203-210.

65. Kahneman,D. and Tversky,A. (2000). *Choices, Values, and Frames* (Cambridge University Press, 2000).

66. Kamin,L.J., Black,A.H., and Brimer,C.J. (1963). Conditioned Suppression As A Monitor of Fear of Cs in Course of Avoidance Training. Journal of Comparative and Physiological Psychology *56*, 497.

67. Killcross,S., Robbins,T.W., and Everitt,B.J. (1997). Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. Nature *388*, 377-380.

68. Kim,H., Shimojo,S., and O'Doherty,J.P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. Plos Biology *4*, 1453-1461.

69. Knutson,B., Rick,S., Wirnmer,G.E., Prelec,D., and Loewenstein,G. (2007). Neural predictors of purchases. Neuron *53*, 147-156.

70. Koechlin,E., Ody,C., and Kouneiher,F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science *302*, 1181-1185.

71. Konorski,J. (1967). Integrative Activity of the Brain: An Interdisciplinary Approach (Chicago: University of Chicago Press).

72. LeDoux,J.E. (2000). Emotion circuits in the brain. Annu.Rev.Neurosci. *23*, 155-184.

73. Loewenstein,G. (1987). Anticipation and the Valuation of Delayed Consumption. Economic Journal *97*, 666-684.

74. Loewenstein,G. (2006). The pleasures and pains of information. Science *312*, 704-706.

75. Lovibond,P.F. (1983). Facilitation of Instrumental Behavior by A Pavlovian Appetitive Conditioned-Stimulus. Journal of Experimental Psychology-Animal Behavior Processes *9*, 225-247.

76. Lumb,B.M. (2002). Inescapable and escapable pain is represented in distinct hypothalamic-midbrain circuits: specific roles for A delta- and C-nociceptors. Experimental Physiology *87*, 281-286.

77. Mackintosh,N.J. (1983). Conditioning and associative learning. New York: Oxford University Press.

78. Maier,S.F. and Seligman,M.E.P. (1976). Learned Helplessness - Theory and Evidence. Journal of Experimental Psychology-General *105*, 3-46.

79. Mangel,M. and Clark,C.W. (1988). Dynamic Modelling in Behavioral Ecology. Princeton, New Jersey: Princeton University Press.

80. Maren,S. and Quirk,G.J. (2004). Neuronal signalling of fear memory. Nat.Rev.Neurosci. *5*, 844-852.

81. Melvin,K.B. and Anson,J.E. (1969). Facilitative Effects of Punishment on Aggressive Behavior in Siamese Fighting Fish. Psychonomic Science *14*, 89-&.

82. Mineka,S. and Gino,A. (1980). Dissociation Between Conditioned Emotional Response and Extended Avoidance Performance. Learning and Motivation *11*, 476-502.

83. Mobbs,D., Petrovic,P., Marchant,J.L., Hassabis,D., Weiskopf,N., Seymour,B., Dolan,R.J., and Frith,C.D. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. Science *317*, 1079-1083.

84. Montague,P.R., Dayan,P., and Sejnowski,T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J.Neurosci. *16*, 1936-1947.

85. Morris,R.G.M. (1975). Preconditioning of Reinforcing Properties to An Exteroceptive Feedback Stimulus. Learning and Motivation *6*, 289-298.

86. Morse,W.H., Mead,R.N., and Kelleher,R.T. (1967). Modulation of Elicited Behavior by A Fixed-Interval Schedule of Electric Shock Presentation. Science *157*, 215-&.

87. Mowrer,O.H. (1947). On the dual nature of learning: A re-interpretation of" conditioning" and" problem-solving. Harvard Educational Review 17, 102-148.

88. Nakahara,H., Itoh,H., Kawagoe,R., Takikawa,Y., and Hikosaka,O. (2004). Dopamine neurons can represent context-dependent prediction error. Neuron *41*, 269-280.

89. Niv,Y., Daw,N.D., Joel,D., and Dayan,P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology *191*, 507-520.

90. Niv,Y., Joel,D., and Dayan,P. (2006). A normative perspective on motivation. Trends Cogn Sci. *10*, 375-381.

91. Overmier,J.B., Bull,J.A., and Trapold,M.A. (1971b). Discriminative Cue Properties of Different Fears and Their Role in Response Selection in Dogs. Journal of Comparative and Physiological Psychology *76*, 478-&.

92. Overmier,J.B., Bull,J.A., and Trapold,M.A. (1971a). Discriminative Cue Properties of Different Fears and Their Role in Response Selection in Dogs. Journal of Comparative and Physiological Psychology *76*, 478-&.

93. Paton,J.J., Belova,M.A., Morrison,S.E., and Salzman,C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature *439*, 865-870.

94. Pessiglione,M., Seymour,B., Flandin,G., Dolan,R.J., and Frith,C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature *442*, 1042-1045.

95. Price,D.D. (1999). Psychological Mechanisms of Pain and Analgesia. IASP press, Seattle USA. 1999.

96. Puterman,M. (1994). Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, Inc. New York, NY, USA.

97. Rescorla,R.A. (1969). Establishment of A Positive Reinforcer Through Contrast with Shock. Journal of Comparative and Physiological Psychology *67*, 260-&.

98. Riess,D. (1971). Shuttleboxes, Skinner boxes, and Sidman aoidance in rats: acquistion and terminal performance as a function of response topography. Psychonom.Sci. 25, 283-286.

99. Sachs,B.D. and Barfield,R.J. (1974). Copulatory behavior of male rats given intermittent electric shocks: theoretical implications. J.Comp Physiol Psychol. *86*, 607-615.

100. Satoh,T., Nakai,S., Sato,T., and Kimura,M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. J.Neurosci. *23*, 9913-9923.

101. Schoenbaum,G. and Setlow,B. (2003). Lesions of Nucleus Accumbens Disrupt Learning about Aversive Outcomes. J.Neurosci. *23*, 9833-9841.

102. Schultz,W., Dayan,P., and Montague,P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593-1599.

103. Setlow,B., Schoenbaum,G., and Gallagher,M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning. Neuron *38*, 625-636.

104. Seymour,B., Daw,N., Dayan,P., Singer,T., and Dolan,R. (2007a). Differential encoding of losses and gains in the human striatum. J.Neurosci. *27*, 4826-4831.

105. Seymour,B., O'Doherty,J.P., Dayan,P., Koltzenburg,M., Jones,A.K., Dolan,R.J., Friston,K.J., and Frackowiak,R.S. (2004). Temporal difference models describe higher-order learning in humans. Nature *429*, 664-667.

106. Seymour,B., O'Doherty,J.P., Koltzenburg,M., Wiech,K., Frackowiak,R., Friston,K., and Dolan,R. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. Nat.Neurosci. *8*, 1234-1240.

107. Seymour,B., Singer,T., and Dolan,R. (2007b). The neurobiology of punishment. Nature Reviews Neuroscience *8*, 300-311.

108. Shidara,M., Mizuhiki,T., and Richmond,B.J. (2005). Neuronal firing in anterior cingulate neurons changes modes across trials in single states of multitrial reward schedules. Exp.Brain Res. *163*, 242-245.

109. Solomon,R.L. and Corbit,J.D. (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. Psychol.Rev. *81*, 119-145.

110. Starr,M.D. and Mineka,S. (1977). Determinants of Fear Over Course of Avoidance-Learning. Learning and Motivation *8*, 332-350.

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

112. Sugase-Miyamoto,Y. and Richmond,B.J. (2005). Neuronal signals in the monkey basolateral amygdala during reward schedules. J.Neurosci. *25*, 11071-11083.

113. Sutton,R.S. and Barto,A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. Psychol.Rev. *88*, 135-170.

114. Sutton,R.S. and Barto,A.G.(1998). Reinforcement Learning. An introduction. MIT press (Cambridge MA). 1998.

115. Thomson,J.J. (1986). Rights, Restitution and Risk (Harvard Univ. Press, Cambridge), pp. 94-116. 1986.

116. Tolman,E.C. (1932). Purposive Behavior in Animals and Men (Century, New York, 1932).

117. Tversky,A. and Kahneman,D. (1981). The framing of decisions and the psychology of choice. Science *211*, 453-458.

118. Ulrich,R.E. and Azrin,N.H. (1962). Reflexive fighting in response to aversive stimulation. J.Exp.Anal.Behav. *5*, 511-520.

119. Ursu,S. and Carter,C.S. (2005). Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. Brain Res.Cogn Brain Res. *23*, 51-60.

120. Watkins,C.J.C.H. and Dayan,P. (1992). Q-Learning. Machine Learning *8*, 279-292.

121. Weisman,R.G. and Litner,J.S. (1969a). Positive Conditioned Reinforcement of Sidman Avoidance Behavior in Rats. Journal of Comparative and Physiological Psychology *68*, 597-603.

122. Weisman,R.G. and Litner,J.S. (1969b). The course of Pavlovian excitation and inhibition of fear in rats. J Comp Physiol Psychol *69*, 667-672.

123. Williams,D.R. and Williams,H. (1969). Auto-Maintenance in Pigeon - Sustained Pecking Despite Contingent Non-Reinforcement. Journal of the Experimental Analysis of Behavior *12*, 511-&.

124. Wilson,D.I. and Bowman,E.M. (2005). Rat nucleus accumbens neurons predominantly respond to the outcome-related properties of conditioned stimuli rather than their behavioral-switching properties. J.Neurophysiol. *94*, 49-61.

125. Yamagishi,T. (1986). The Provision of A Sanctioning System As A Public Good. Journal of Personality and Social Psychology *51*, 110-116.

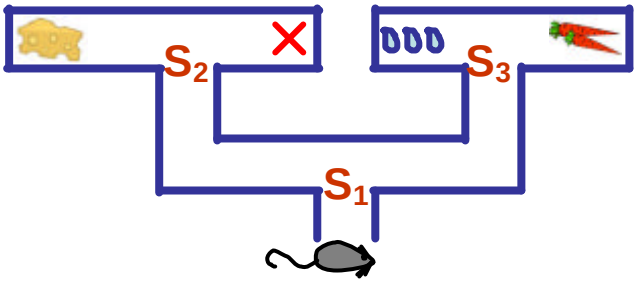
126. Yin,H.H. and Knowlton,B.J. (2006). The role of the basal ganglia in habit formation. Nature Reviews Neuroscience *7*, 464-476.

127. Yin,H.H., Knowlton,B.J., and Balleine,B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav.Brain Res. *166*, 189-196.

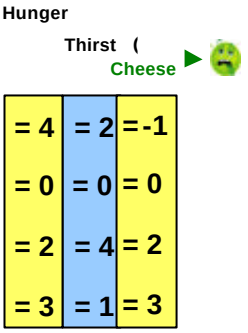
128. Yoshida,W. and Ishii,S. (2006). Resolution of uncertainty in prefrontal cortex. Neuron *50*, 781-789.

Figure 1:

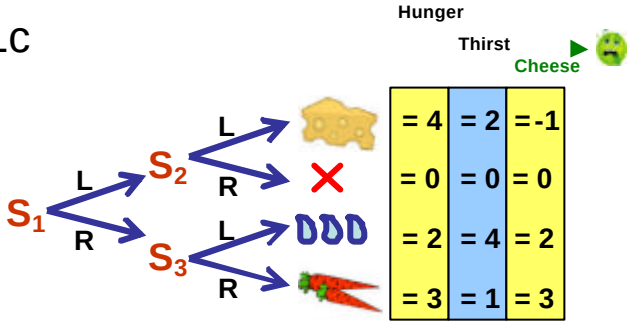
1a



1b



1c



1d