



Published in final edited form as:

Neurosci Biobehav Rev. 2014 June ; 43: 259–268. doi:10.1016/j.neubiorev.2014.03.027.

The problem with value

John P. O'Doherty

California Institute of Technology

Abstract

Neural correlates of value have been extensively reported in a diverse set of brain regions. However, in many cases it is difficult to determine whether a particular neural response pattern corresponds to a value-signal per se as opposed to an array of alternative non-value related processes, such as outcome-identity coding, informational coding, encoding of autonomic and skeletomotor consequences, alongside previously described “salience” or “attentional” effects. Here, I review a number of experimental manipulations that can be used to test for value, and I identify the challenges in ascertaining whether a particular neural response is or is not a value signal. Finally, I emphasize that some non-value related signals may be especially informative as a means of providing insight into the nature of the decision-making related computations that are being implemented in a particular brain region.

Keywords

Reward; neuroeconomics; decision-making; learning

Introduction

Interest in the neurobiological substrates of value-learning and value-based decision-making has surged in the past decade, following the emergence of nascent disciplines such as the fields of neuroeconomics and decision-neuroscience (Camerer, 2008; Fehr and Camerer, 2007; Glimcher and Rustichini, 2004; Levy et al., 2010; Montague and Berns, 2002; Sanfey et al., 2006). The prevailing assumption in these domains is that the brain encodes a representation of the expected value or utility of stimuli and/or of actions, and that in decision-making situations, those representations are used to guide choice such that actions are taken to maximize future expected rewards. Consistent with this proposed framework, experiments in humans using neuroimaging methods, and in animals using neurophysiological recordings, have uncovered evidence for value-related neuronal activity in a wide array of neural structures during learning and decision-making tasks. These findings suggest that a diverse network of brain regions participate in the encoding of value, and have led to proposals that some of these structures participate directly in the decision-process whether over goods (or stimuli) or over actions linked to selection of those goods.

However, ascertaining whether a neuronal response truly corresponds to a value or subjective utility signal is a rather challenging endeavor. Here I outline some of the problems in inferring that a particular neuronal response pattern encodes a value signal per se or else a number of other non-value related processes. A point that has frequently been

made before is that reward-related responses may be confounded with attentional mechanisms, sometimes also referred to as “attention”, “motivation” or “salience” (Horvitz, 2000; Leathers and Olson, 2012; Maunsell, 2004; Roesch and Olson, 2004). I will consider this possibility here, but also identify other less often highlighted but equally problematical confounding signals to a valuation account. These include differential encoding of sensory information about an outcome, informational signaling of an outcome, and representation of behavioral responses. I then consider viable steps to determining whether a particular neuronal response truly corresponds to a value signal. Finally, I argue that even if signals hitherto presumed to correspond to value turn out to represent something else, such signals should not be ignored but instead properly categorized as they might still play an important and perhaps even critical role in the processes of learning, value computation and choice.

Summary of different types of putative value signals

Before embarking on consideration of the type of signals that may confound value, it is worth briefly first considering how value can be defined, and then summarizing the different types of value signals that have been reported in the brain.

What is value?

There are multiple approaches to the definition of value. Here I discuss a variety of approaches:

One approach is to define value as some function which pertains to the relative attractiveness of a particular good at the time of choice, and which is maximized as a result of the decision process (Rangel et al., 2008).

A related approach, is to adopt the notion of utility as used in economics and to apply this to the definition of value in neuroscience. In economics, “utility” is a function that describes a set of preferences an individual has over a set of goods. If the individual prefers good A over good B, then by definition good A will have a higher utility than good B. Translating this to neuroscience, we might expect that some neural process encoding utility would show an ordered relationship in its responses (such as for example by changes in average firing rates), to the stimuli presented to the animal, such that the neuronal responses are greater for a good that is more preferred by the animal compared to a good that is less preferred. Here, utility/value is inferred directly from behavior, and neural representations of this function are assumed to reflect behavioral preferences (Dean, 2013).

Another strongly related approach to the preference approach in economics arises from the behavioral neuroscience literature, and that is to define value in terms of the motivating properties of a stimulus for instrumental action (Rolls, 2007). The degree to which an animal is prepared to work (i.e. perform some kind of effortful action) to obtain a stimulus, relates to the degree to which the animal finds that stimulus rewarding.

Yet another approach is to define value is in terms of the subjective “pleasure” that is engendered by a particular stimulus. The experienced utility of a good, is the pleasure or happiness that arises from its consumption (Bentham, 2007; Kahneman et al., 1997) The

challenge in using such a definition in experimental work is that it is very difficult to access such subjective “pleasure” states, particularly in animals, although some have argued for the existence of clear behavioral proxies for such subjective evaluations even in animals (Berridge, 1996). In humans, one can simply ask participants to verbally or otherwise rate their subjective pleasure. However this does require adopting the assumption that human participants have reliable insight into the nature of their affective disposition.

Although these definitions of value differ, by and large we would expect them to make similar predictions about when a particular stimulus would be deemed valuable or not, as well as to make similar predictions about the expected pattern of neural responses tracking value. However, under some situations, the definitions may lead to divergent predictions. Most notably, under situations where behavior becomes habitized, an action can be selected resulting in attainment of a good that is not actually preferred by the animal (Dickinson, 1985; see below for further discussion of habits). Under those conditions, both the revealed preference and work motivation operationalizations of value would yield an inference that this particular good has high value for the animal, because that is what is reflected in the action-selection behavior. However, once the good is actually attained by the animal as a result of the habitual action(s), the animal would not actually consume the good. Some have also proposed that the motivation to work for a good or “wanting” can be neurally and sometimes behaviorally dissociable from its subsequent evaluation (or “liking”) (Berridge, 1996). Thus, it is clear that how one defines value, has non-trivial implications for how one interprets value-signals in the brain.

Different types of value signals

Now I will consider the existence of different types of value signal as they have been described in the brain:

Outcome value codes

Valuation responses that are presumed to occur in response to the receipt of an outcome, have been referred to as an experienced value or outcome value (O'Doherty, 2004). Many studies have reported neuronal responses to the receipt of rewarding or aversive stimuli in the literature, in both human neuroimaging studies and in animal neurophysiological recordings. As I will consider later, probably only a subset of these can unambiguously be ascertained to correspond to experienced value per se.

Predictive value codes

We now consider value codes that are elicited on the basis of predictions about the value of future outcomes. I will call these collectively predictive value codes. This type of signal can be further subdivided into a number of unique forms of valuation code:

One form of predictive value-signal are Pavlovian values— these correspond to a representation of the value of an expected outcome signaled by a discriminative stimulus. Such signals have been widely reported in the orbitofrontal cortex and amygdala as well as the ventral striatum in both rats, monkeys and humans (Gottfried et al., 2002; Paton et al., 2006; Schoenbaum et al., 1998).

Another class of predictive value-signal corresponds to what are variously described as “offer values”, “stimulus values”, or “goal-values” – these correspond to the expected value of a prospective outcome or goal as this goal is being evaluated at the point of choice, typically under situations where other prospective outcomes or goals are also available. Such signals have been reported in the monkey central orbitofrontal cortex, as well as in human ventromedial prefrontal cortex (Padoa-Schioppa and Assad, 2006; Plassmann et al., 2007).

Yet more value signals arise when an animal or human must render a choice over different actions in order to obtain a goal outcome. Collectively these signals can be called “action-values”. Action-value signals have been reported in a number of brain regions including the striatum, lateral intra-parietal cortex and supplementary motor cortices (Lau and Glimcher, 2007; Platt and Glimcher, 1999; Samejima et al., 2005; Sohn and Lee, 2007; Sugrue et al., 2004; Wunderlich et al., 2009).

Finally, post-decision value signals have been reported corresponding to the value of the option that is ultimately chosen in a decision-task, particularly in medial and central orbitofrontal cortex (Hampton et al., 2006; Padoa-Schioppa and Assad, 2006; Wunderlich et al., 2009, 2010).

I will now consider competing explanations for different value signals.

Outcome identity coding vs outcome valuation

Any outcome whether a rewarding, aversive or affectively neutral event, has perceptual properties: attributes that distinguish it from other stimuli in the world. Thus, any difference found in neural activity in response to different outcomes might reflect these sensory properties as opposed to the underlying value of those outcomes. This problem is particularly stark under situations where outcomes differ in their sensory modalities such as for example by comparing responses to juice reward vs a painful cutaneous stimulation as a means of determining rewarding vs aversive responding outcome values. However, the problem is not overcome even when using reward stimuli in the same sensory modality by e.g. comparing a sweet vs a salty taste, and is not unique to appetitive vs aversive comparisons, as it is equally evident even when comparing responses to two different rewarding stimuli (such as a more vs less preferred reward).

One way to attempt to circumvent this difficulty would be to use the same stimulus (such as a particular juice reward), and instead manipulate the intensity or magnitude of the outcome provided. However, once again, any variation in outcome magnitude or intensity will result in a change not only in outcome utility, but also in the sensory properties, such that those properties will be either more “concentrated”, or more plentiful at the time of perception. Thus, variation in any sensory quality of the stimulus will typically involve a conflation between those stimulus properties and the ensuing valuation changes. Compounding this issue is the fact that sensory (non-value related) encoding of outcome representations have been reported even in brain regions traditionally associated in value processing such as the orbitofrontal cortex (Grabenhorst and Rolls, 2011; Klein-Flugge et al., 2013; McDannald et al., 2014; McNamee et al., 2013).

Predictions of outcome identity vs predictive value codes

Leaving aside the difficulty in discriminating outcome identity from outcome value, lets turn instead to predictive coding of value. Lets first consider Pavlovian values. Pavlovian value signals depend on associations being formed between a conditioned stimulus or cue and an outcome value, thus when the cue is presented, a representation is elicited of the value of the outcome. One approach to assess value representations elicited by Pavlovian cues, is to pair one cue with the subsequent delivery of a rewarding outcome, and another cue with the subsequent delivery of either an aversive, neutral outcome or even the absence of an outcome. Neuronal responses exhibiting differences between responses elicited to the cue paired with the rewarding outcome vs the cue paired with the aversive outcome might be taken as evidence for predictive value coding. However, as with outcome codes, it is very difficult to rule out the possibility that any cue showing differential responses based on its association with a particular outcome is being driven not by differences in value for that outcome but instead by differences in the outcome's sensory properties. In other words, instead of indexing a stimulus→outcome value code, one might be identifying responses engendered by a stimulus → outcome identity association (i.e. a stimulus→stimulus code). Note that this is the not the same issue as attributing neuronal activity to the perceptual identity of the CS itself. Effects of CS-identity can be disambiguated from outcome signals by simply altering the contingencies between pairs of CSs and respective outcomes so that for example a CS initially paired with a rewarding outcome now predicts an aversive outcome and a CS initially paired with an aversive outcome now predicts a rewarding outcome. However, such a manipulation cannot rule out effects arising from the sensory properties of the outcome itself that could mimic a value response. Manipulating outcome contingency (by altering outcome probability) or outcome magnitude can also not exclude an outcome identity account, as the former would not only alter expected value but also potentially alter the associative strength of the link between the cue stimulus and the outcome identity, while the latter would simply alter the intensity of the sensory features of the outcome representation (Fig. 1). Outcome identity effects can also be manifested for signals observed in more complex learning and choice situations in which predictive information about a potential outcome are being represented, including goal-values, action-values and chosen-values.

An informational signaling role for outcomes and cues

In order to solve a learning or decision problem, an animal or human not only needs to know about the expected subjective value of different possible actions or goals, but also needs to have the ability to infer which state of the world (or of the decision problem in particular) they are in. This is particularly so under situations where decision-problems have a hidden structure. This is perhaps best illustrated by an example. In a probabilistic instrumental reversal-learning task as used in a number of studies (Cools et al., 2002; Hampshire et al., 2012; O'Doherty et al., 2003; O'Doherty et al., 2001), there are two available options, one of which a subject can select on a given trial. One of these options if chosen typically yields a monetary reward with a high probability and a loss outcome otherwise, while the other option if chosen typically yields a monetary reward with a low probability, otherwise with a high probability it yields a monetary loss. After a period of time according to some arbitrary

and often probabilistic rule, the contingencies reverse so that the previously highly rewarded stimulus now yields losses, while the previously highly loss predicting stimulus now yields gain with high probability. The hidden structure in this task is the knowledge that the contingencies on the two options are fully anti-correlated, so that when the expected value of one is high the other is low and vice-versa, as well as the knowledge that periodically the contingencies will reverse. Once a reversal is known to occur then the EVs on the two options can simply be switched by the subject, without the need to re-learn the expected values. In this context then, an outcome (such as winning or losing money) not only yields experienced value in and of itself, but is also informative as to what state of the decision-problem the subject is in. More precisely, if an outcome is a monetary loss, then this may provide some evidence in support of the possibility that a reversal has occurred (because monetary losses are more likely to occur if the formerly good option is now the bad option), whereas if an outcome is a monetary gain, this may provide evidence in support of the possibility that the correct stimulus is being selected and a reversal has not occurred. It is easy to see therefore that outcomes in this task are serving two very distinct purposes: one is simply related to experienced value, the other is providing evidence about the underlying state of the task, and in particular where the participant is in the task state-space. Thus if one is comparing neural responses to e.g. a monetary gain vs a monetary loss on such a paradigm, then differential activity observed to gains vs losses could reflect either experienced value, or else informational signaling about the probability with which a reversal has occurred (see also O'Doherty, 2007).

Going beyond the reversal learning example, informational signaling effects can manifest on any task in which there are multiple states and in which different rules may operate to drive reinforcement as a function of the configuration of the state-space. Neural activity observed in response to either a cue or an outcome during performance of any such task may potentially correspond not just to expected value but also to informational signaling.

Response-related coding elicited by predictive cues and/or outcomes

Another important feature of outcomes that is likely to be uncontroversial to most, yet one that is rarely considered in the course of interpreting neural data, is that in addition to generating a subjective experienced value, a highly valued outcome will also yield an array of unconditioned reflexes, both skeletomotor and autonomic. The precise patterning of these will depend on the specific outcome involved. For instance, a food outcome will generate consummatory activity (increased salivation, licking), along with increased physiological arousal including increased heart rate, along with insulin release etc, while a different type of reward such a kiss by your partner may yield a different pattern of skeletomotor and physiological responding. Consequently, it is challenging to separate out neural activity related to valuation per se, from activity related to the effects of valuation on motor and physiological responses. Even more stark differences in responses are going to be evident when comparing appetitive vs aversive outcomes, where one type of outcome promotes consummation and/or approach while the other promotes expulsion and/or avoidance.

Of course we have known since Pavlov that skeletomotor and behavioral responses also come to be elicited by cues as a function of associative learning (Pavlov, 1927). Thus, when

interpreting responses to Pavlovian predictive cues, or indeed any other predictive value signal such as option-values, chosen-values or action-values, it is entirely feasible that differentiable neuronal activity observed correspond not to a valuation representation per se, but instead to neural activity related to the generation of, or representation of a conditioned response elicited by a cue yielding a valuable outcome.

“Salience”, attention and valuation

A long appreciated confound of value in the literature is the fact that valuable stimuli are likely to be “salient” to an animal in the sense that such an item can draw attentional resources, and result in enhanced perceptual and/or cognitive processing of that item (Horvitz, 2000; Maunsell, 2004; Zink et al., 2006).

As with value itself, the construct of salience is often not very precisely defined. One type of salience-type mechanism that has received careful and specific definitional treatment is the notion that changes in uncertainty in the predictiveness of a cue during associative learning can be used to modulate the rate of learning involving that cue (see (Behrens et al., 2007; Mackintosh, 1975; Payzan-LeNestour and Bossaerts, 2011; Pearce and Hall, 1980; Roesch et al., 2012). However, the construct of salience is often used to refer to other effects such as changes in the emotional intensity of a response to a stimulus, or in terms of the motivational properties of a stimulus for action, that would not be captured by such a specific definition (Berridge, 2012; Gray et al., 2007).

How does one separate out a salience account from valuation? One approach that has been used is to compare neural responses to appetitive outcomes or to cues that predict appetitive outcomes to responses elicited by aversive outcomes or predictors of aversive outcomes. If a neuron or fMRI BOLD response scales differently for appetitive and aversive stimuli, such as for example responding only to appetitive and not aversive, or increasing to appetitive or decreasing to aversive stimuli, then the claim could be made that this is a value signal. On the other hand, if the neuronal response and/or BOLD activation increases as a function of the presentation of **both** appetitive and aversive stimuli (but not to affectively neutral stimuli), then the argument could be made that this is an attentional and/or arousal signal. In fMRI studies, a bivalent signal along these lines could occur even if the underlying neuronal population encodes faithfully encodes a value signal so long as the underlying neuronal population within that region contains distinct but spatially inter-mixed sub-populations of neurons, some of which positively encode value, and others which negatively encode value. Average activity at the level of fMRI might therefore suggest an arousal code, whereas in fact these neurons might implement a faithful value code.

However, for single or multi-unit neurophysiological recording studies it is possible to ascertain the response properties of individual neurons and determine whether they respond uniquely to appetitive or aversive stimuli or both. Studies performed along these lines have claimed evidence for value signals in some areas (e.g. orbitofrontal cortex), while other brain regions such as supplementary motor cortex and more recently intraparietal sulcus have been suggested to encode salience/attentional signals (Leathers and Olson, 2012; Roesch and Olson, 2004). However, there is another important caveat to this interpretation

that needs to be born in mind. Specifically, the type of task used to assess responses to aversive predictors and aversive outcomes is crucial. One possibility is to use an instrumental avoidance paradigm in which the animal is performing an action in order to *avoid* obtaining the aversive outcome. However, the “relief” that follows from avoidance of an aversive outcome can act as a reward in a similar manner to the way in which missing out on an expected rewarding outcome can be actually aversive (Solomon and Corbit, 1974). If an animal therefore succeeds in responding to avoid aversive outcomes, then that animal may come to predict in that situation that there is a high prospect of avoiding that outcome and hence obtaining the reward that follows from the relief of successful avoidance. Neural signals encoding expected value in response to discriminative stimuli during the avoidance paradigm could therefore represent the expected reward that would follow from successfully avoiding the aversive outcome. Thus, neurons found to respond to both predictors of reward and predictors of aversive outcomes in an instrumental context could simply be representing expected reward and not expected punishment. Therefore, an instrumental avoidance paradigm is problematical as a means of discriminating valence from salience/attentional accounts. The other paradigm that could be deployed to test for these different types of response profiles is a Pavlovian one, in which the animal is not required to make any type of response but instead a particular cue is followed reliably by an aversive outcome while another cue is followed reliably by an appetitive outcome. In these circumstances, finding overlapping scaling neural activity in response to both types of cues would be more convincing evidence for a salience code, because there is no possibility to ascribe to the aversive cue, activity related to encoding of the positive hedonic consequences arising from avoiding the now inescapable aversive outcome.

In the case of the theoretically more constrained notion of salience as reflecting predictiveness around cue-uncertainty, value can arguably be more definitively clearly separated out from such cue uncertainty signals and indeed a number of studies have accomplished precisely this (Behrens et al., 2007; Payzan-LeNestour et al., 2013; Roesch et al., 2010).

Reinforcer Devaluation/Revaluation

Another approach to measuring neural responses to valuation is to measure activity to a particular outcome, or a cue, or action associated with a given outcome, before and after inducing a change in the experienced utility of that outcome through a procedure called reinforcer devaluation. This involves feeding the subject to satiety on a particular outcome, thereby inducing a change in the value of that outcome, or alternatively separately pairing the outcome with an aversive event such as illness (Rolls et al., 1981). The advantages of this procedure are that any changes in activation measured in response to the stimulus following the devaluation procedure, can be assumed to be related to a change in the reward-value of the associated outcome, as opposed to the sensory features of the outcome simply because the sensory features of the outcome remain constant from pre to post devaluation. Perhaps the biggest obstacle to reinforcer devaluation manipulations as a practical means for studying predictive values or decision-values in most experimental contexts but especially in single-unit neurophysiology, is that testing must be done in extinction (that is by presenting the cue or action but without presenting the outcome), in order that activity observed reflects

the retrieval of the outcome value from the previously learned association as opposed to merely reflecting re-learning of an association between a given cue or action and the outcome in its now current devalued state. However, if an outcome is presented in extinction, then humans and animals will very quickly stop responding to the outcome in an instrumental context or stop exhibiting conditioned responses in a Pavlovian context. As a consequence, there may be only a very small number of trials prior to extinction reaching asymptote during which it is possible to measure the effects of the devaluation procedure on cues or actions associated with a given outcome. This small number of trials is anathema to the trial averaging needed for extracting meaningful signals with either neurophysiological or neuroimaging methods, although in some cases it has proved possible to obtain significant effects in spite of those limitations (Gottfried et al., 2003; Valentin et al., 2007). One possible work-around to the trial averaging problem is to perform multiple sequential devaluation protocols, as a means of building up sufficient numbers of measurements. However, such a procedure also engenders the complication that if the devaluation procedure comes to be expected (i.e. incorporated into meta-knowledge about the task procedure), then ensuing neural activity may not simply reflect the value of the expected outcome but rather task-related cognitive signals as discussed previously. Furthermore, it could be argued that reinforcer devaluation does not completely rule out stimulus-confounds, as habituation of the sensory features of the stimulus may occur during the devaluation process, particularly if the means of devaluing of the outcome is through feeding to satiation. Evidence mitigating against this possibility is that the subjective ratings of the sensory intensity of a devalued outcome during selective satiation procedures in humans typically does not change from pre to post devaluation (Rolls et al., 1981). However, evidence from subjective ratings of this sort may not suffice to rule out the possibility that sensory habituation could occur somewhere in the sensory pathways following a devaluation episode.

Revealed preferences

Yet another approach to establishing neural correlates of value is to make use of the choice behavior or “revealed” preferences of an individual and use that information to derive an underlying subjective utility for certain goods or decision options which can in turn be related to neural activity elicited by those goods. A good example of this approach is the work by Tremblay and Schultz (1999), who presented monkeys with blocks of trials in which in a given block two out of three different types of juice reward were presented to monkey while activity was recorded from a region of central orbitofrontal cortex. Each monkey had a clear preference ordering over the three juices, and it was found that activity of some neurons in OFC scaled according to those preferences. In particular, some neurons responded strongly to the most preferred juice out of the pair of juices that was presented in a given block. Importantly, the main factor driving the activation of those neurons was which juice was preferred by the monkey out of a given pair within a block, such that if the most and middle preferred juice were presented, the neurons responded strongly to the most preferred juice and only weakly to the middle preferred juice, but if the middle and the least preferred juice were presented, the neurons now responded strongly to the middle preferred juice. Thus, this class of orbitofrontal neuron appeared to encode the relative preference the

monkey has for a given juice outcome within a block. Such a response profile is unlikely to be accounted for in terms of the sensory features of the outcomes because the same middle preferred outcome with the same sensory features is presented in both types of trial blocks, yet the neuron responds differently depending on the relative preference the monkey has for that item within a block.

Another example of this approach is a study by Padoa-Schioppa and Assad (2006) who repeatedly presented choices involving pairs of juice rewards in differing quantities to monkeys. For a given pair of items they then plotted the choice data as a function of different offer quantities for the two goods, and fit a choice sigmoid function to that data as a means of determining the relative subjective value the monkey has for each good. This subjective choice function was also fit to neural data measured in the orbitofrontal cortex, in order to find populations of neurons corresponding to the subjective value for each option (offer values), as well as for the option that is chosen on a given trial (chosen values). The fitting of a subjective preference function derived from behavior to neural data can potentially overcome issues of stimulus-identity confounds, provided that changes in the stimulus properties do not strongly correlate with changes in subjective preference. In the Padoa-Schioppa and Assad example, in which the monkey is given choices between different quantities of two different goods, there is potential for a confound to emerge between the magnitude of the expected outcome which will be associated with varying stimulus properties and the subjective value of that outcome. To overcome this, Padoa-Schioppa and Assad were able to make use of the fact that the same food items were presented on multiple sessions whereby the monkey's relative preferences changed from session to session (in essence implicitly taking into account devaluation and revaluation of the food items on the part of the monkey for the food items), thereby enabling simple effects of stimulus properties to be taken into account.

A similar approach has been adopted in a number of studies examining inter-temporal choice, in which an individual is given a choice between a smaller reward delivered sooner and a larger reward delivered later. By varying the amount of rewards on offer at the two points, and the length of time before the later reward is delivered, and presenting choices between these sooner and later options, it is possible to derive a subjective function describing an individual's subjective preferences for rewards at different time delays. Typically such a function shows a fall off in the subjective value of the reward the longer the interval between the distal reward and the soonest available reward. This typical profile can be fit using hyperbolic (or exponential) function(s), enabling inference to be made about the subjective inter-temporal preferences for a given individual. Those same functions can then be fit to neural data (fMRI or neurophysiology), in order to ascertain areas correlating with subjective utility for different choice options. For example, using this approach Kable and Glimcher (2007), presented human subjects with repeated choices between a fixed immediate monetary reward and a time and magnitude varying distal monetary reward, finding activity in a number of areas including medial prefrontal cortex correlating with the subjective value of the chosen option derived from the subjective preference function. Stimulus confounds are unlikely in the case of inter-temporal studies involving abstractly presented decision options and monetary rewards simply because it is implausible that there

are stimulus features which correlate in an orderly fashion with the magnitude of the abstract reinforcer.

A related approach that has been used in human studies is to obtain a proxy behavioral measure of subjective decision-utility, by assessing individual's "willingness to pay" (WTP) for a particular item (Chib et al., 2009; Plassmann et al., 2007). In these types of paradigms, many different types of items such as food items or trinkets are presented to participants, and WTP is collected for each item and then correlated against the fMRI data in order to isolate regions correlating with subjective decision-utility. In the case of the WTP approach, simple stimulus confounds are perhaps not likely given that the items can vary considerably along a number of stimulus dimensions yet still have similar WTPs (e.g. a food item vs a t-shirt). Nevertheless, within some domains such as food-items, it is likely that some elemental stimulus features of a good such as its sweetness, saltiness, caloric content etc. are going to correlate to some extent (albeit perhaps not in a simple linear fashion) with the overall value assigned to that good.

These points notwithstanding, while the revealed preference approach offers a lot of potential for ruling out stimulus-feature accounts of neural value correlates, the approach still faces challenges when it comes to separating out "value" per se from consequences of value such as skeletomotor and autonomic responses.

Relevance of non-value related outcome representations

Non-value related features of outcomes are important in their own right for understanding the computations underpinning learning and decision-making and the contribution of specific brain areas to these computations:

Outcome identity in goal-directed and Pavlovian control

For goal-directed instrumental control, representing outcome identity is a crucial interim step in enabling the associated incentive value of an outcome to be retrieved. According to associative theories of goal-directed learning, associations between states, actions and outcome identity are used to facilitate retrieval of an associated outcome value (Balleine and Dickinson, 1998). Computational theories of goal-directed control such as model-based reinforcement-learning (Daw et al., 2005), also require outcome-identity to be represented within the cognitive model of the world, before the value of that outcome can be accessed and an action-value can be computed. Thus, neural signals of outcome identity elicited by cues or actions if found in the brain during performance of a goal-directed task, may be a critical component signal for goal-directed action. Characterizing where and how such signals are represented is therefore essential for unraveling how goal-directed behavior is implemented at the neural level. At least some forms of Pavlovian conditioning involve associations between stimuli and outcome-identity (stimulus-stimulus associations) that in turn retrieve outcome values in a manner parallel to that which occurs in the goal-directed instrumental system. This type of Pavlovian association is devaluation sensitive (Colwill and Motzkin, 1994).

Somatic consequences of outcomes

Once outcome-identities are elicited, it is an open question how outcome-values are then subsequently retrieved. In the phenomenon of “incentive-learning” (Balleine and Dickinson, 1998), it has been shown that for the purposes of goal-directed control, rats are not able to construct the value of an outcome in a particular motivational state without first having to experience the outcome in that state. For example, if the rat learns to perform an action to obtain a novel food when hungry, if the rat is subsequently sated so that the food should no longer be valuable, the rat doesn't adjust its behavior in the sated state unless it first experiences the novel food in that sated state. The rat may therefore need to link being in a sated context with the somatic consequences of experiencing a food in that state in order to compute a valuation. This process, which resembles that postulated in the somatic-marker hypothesis (Bechara et al., 1994), could suggest that the mechanism by which outcome-values are used to guide action-selection depends on ultimately retrieving somatic-states associated with those outcomes contingent on being in a particular motivational state. The implication of this notion is that, understanding how the brain encodes somatic consequences elicited by outcomes in a given motivational state is also a key element in understanding how it is that values for goal-directed actions are computed, as is establishing how and via which learning-mechanisms does outcome-identity become linked to such somatic representations.

For Pavlovian conditioning, visceral and autonomic signals elicited to cues predicting particular outcomes (or UCSs) are key conditioned responses – present during both appetitive and aversive Pavlovian conditioning.

Skeletomotor responses

In Pavlovian conditioning a series of reflexive skeletomotor responses can also come to be elicited by the conditioned stimulus alongside visceral and autonomic responses. Perhaps the most well characterized are the responses of approaching and/or orienting to an appetitive stimulus, or avoiding and/or orienting away from an aversive stimulus (Brown and Jenkins, 1968; Jarvik and Kopp, 1967). Another example are consummatory responses elicited in anticipation of the onset of the outcome, such as chewing or sucking behavior in anticipation of a food or liquid reward respectively. A critical distinction has been made in the conditioning literature between outcome-general conditioned responses and outcome-specific conditioned responses (Balleine and Killcross, 2006; Konorski, 1948). Approach and avoidance are good examples of outcome-general conditioned responses because they can be elicited very generally in response to cues associated with many different outcomes, so long as those outcomes are appetitive and aversive respectively. On the other hand, consummatory responses such as chewing or sucking may be very specific to the type of outcome with which a cue is associated (Jenkins and Moore, 1973). Evidence is emerging that outcome-specific and outcome-general coding may be a very important distinction at the neural level too, with specific structures such as the basolateral nucleus of the amygdala and the shell of the nucleus accumbens implicated in the former, while the centromedial nucleus and core of the accumbens is implicated in the latter (Balleine and Killcross, 2006). As a consequence, characterizing the extent to which neural activity in a given area in

response to cues that predict outcomes represent or are associated with the elicitation of conditioned responses is vital for understanding the type of conditioning a given brain region is involved in mediating and for addressing the type of outcome representation a given set of neurons are involved in implementing.

Multiple routes to behavior: the role of historical values and non-value related behavioral response systems

It is also important to consider that behavior can be controlled in an adaptive manner using mechanisms that eschew value representations based on the current value of an associated outcome entirely. This can happen in both instrumental and Pavlovian learning situations. In instrumental conditioning, a distinction has been made between goal-directed actions which are outcome sensitive and habitual actions which are insensitive to outcome value, in that habitual responding will persist on an action that leads to a previously valued outcome even if that outcome is no longer considered valuable to the organism (Dickinson, 1985). It has been suggested that habitual actions are acquired via the formation of stimulus-response associations without involving any explicit associative link to the outcome produced by that response (Balleine and Dickinson, 1998). Importantly, habits are shaped by reinforcement, that is, according to Thorndike's law of effect, they are strengthened under situations where responses lead to rewarding outcomes and weakened when they lead to non-rewarding outcomes (Thorndike, 1898). Thus, the extent to which a habit develops will be determined by the extent to which a given response led to a rewarding outcome in the past. However, because habits are not sensitive to the current incentive value of the outcome i.e. because they are devaluation insensitive, they can be thought of as reflecting historical or cached value, but not reflecting current outcome value.

Behavioral control divorced from current outcome value can also occur in Pavlovian conditioning. Analogous to stimulus-response habits in the instrumental domain, Pavlovian cues can become associated directly with conditioned reflexes (i.e. CS→CR), without necessitating any intervening representation of the associated outcome (UCS) or the value of such an outcome (Everitt et al., 2003). The behavioral expression of such an association would be devaluation insensitive. Although prevailing behavioral evidence appears to suggest the dominance of devaluation sensitivity in the expression of Pavlovian conditioned responses (Colwill and Motzkin, 1994; Holland and Straub, 1979), such tests have been applied to only some classes of conditioned responses, and thus the presence of a direct associative route between conditioned stimuli and conditioned reflexes remains an open possibility. The possible presence of different learning mechanisms spanning instrumental and Pavlovian behavioral control, underlines the importance of discriminating outcome-sensitive value-signals from other types of relevant signals when measuring neural activity.

Dopamine, value and prediction error

Another value-related signal reported in the brain concerns the phasic activity of dopamine neurons that have been found to resemble a prediction error signal from formal computational models (Hollerman and Schultz, 1998; Mirenowicz and Schultz, 1994; Morris et al., 2006; Roesch et al., 2007; Schultz, 1998). In the context of this discussion, one

question that could arise is whether or not dopamine neurons are reflecting a value-related response (i.e. the derivative of value with respect to time), or whether these neurons are instead encoding some other non-value-related feature such as salience/arousal or stimulus-identity coding. The relative homogeneity of reward-selective dopamine neurons (Hollerman and Schultz, 1998) as showing an increasing response to unpredicted reward outcomes, and a decreasing response to omitted reward outcomes would exclude a simple salience/arousal account for those neurons (although such an interpretation has been proposed for a subset of dopamine neurons in the posterior lateral substantia nigra, see Matsumoto and Hikosaka, 2009). For reward-selective dopamine neurons, the homogeneity in those responses also suggests that it is unlikely that such neurons encode prediction errors about detailed sensory features of a stimulus, that for example could be used to learn different stimulus-stimulus associations unrelated to value. Consistent with this, (Lak et al., 2014), reported that dopamine neurons encode prediction errors that correlate better with subjective value over a range of different goods (different juices and solid food), as opposed to sensory features of those outcomes.

As yet unknown is whether or not the prediction side of the computation being used to generate prediction errors (i.e. where $PE = \text{outcome value} - \text{predicted value}$), is a signal that indexes predictions about current outcome values, or whether instead those predictions are “model-free”, which though representing historically the degree of reward associated with that stimulus or response, might not index the current value of the outcome to the animal. The way to address this question is to test whether these neurons decrease their activity in response to a cue that is associated with an outcome that was previously highly rewarding but is now devalued (Balleine et al., 2008). Such a manipulation should help to elucidate whether or not dopamine neurons are involved in facilitating learning of goal-directed or “model-based” value signals that index the current incentive value of an associated outcome, or whether these neurons contribute instead to the learning of “model-free” or historical value signals such as habits (Daw et al., 2005). It is important to note that even if the dopamine neurons are receiving an input from historical value predictions as opposed to predictions about current outcome value, re-exposure of the animal to the actual outcome itself following devaluation, would nevertheless still induce a gradual updating of the value of an associated cue or action via dopamine-mediated learning. This is because the now changed outcome value would still be predicted to result in a change in the prediction error code following re-exposure to the outcome, thereby facilitating incremental convergence of historical value signals to the current outcome value.

Causal manipulations of neuronal systems and circuits as a means of testing for value

So far we have focused on evidence about value-representations that can be garnered from correlative measures of brain function such as neurophysiological recordings or neuroimaging. This leaves open the extent to which methods that can elucidate causal relationships between neuronal activity and behavior can discriminate value from value-related computations. There is a very large literature describing the use of experimental lesion approaches in rodents and non-human primates as a means of ascertaining the causal

role of brain regions such as the orbitofrontal cortex, amygdala, ventral striatum and elsewhere in value-related behavior, as well as complementary approaches examining the effects of circumscribed brain lesions in human patients on value-related behavior. For example, lesions of the orbitofrontal cortex and amygdala have been found to render animals insensitive to outcome value following reinforcer devaluation based on discriminative stimuli (Baxter et al., 2000; Pickens et al., 2003), and effects have also been reported following lesions in this area on other aspects of value-related behavior (Noonan et al., 2010). Lesions of rodent prelimbic cortex, and dorsomedial striatum abolish the selection of instrumental actions based on current outcome value (Balleine and Dickinson, 1998; Yin et al., 2004; Yin et al., 2005). Relatedly, human patients with lesions of the ventromedial prefrontal cortex have been found to be impaired on a variety of value-related decision making and learning tasks (Bechara et al., 1994; Fellows and Farah, 2005; Hornak et al., 2004; Rolls et al., 1994). The recent emergence of molecular technologies for specific optical (Williams and Deisseroth, 2013) and pharmacological stimulation (Shapiro et al., 2012), most typically in rodents, but also potentially in monkeys, has opened up a whole new avenue for fine-grained spatially and temporally specific causal manipulations in brain circuits. Can these approaches provide insight into the question of whether or not a given neuronal signal is involved in encoding value per se or some other variables?

Let's consider a few scenarios. First of all, imagine a situation where a lesion in a given area X, does cause impairment in the performance of behavior on a value-related task. Can we assume that this implies the area is involved in encoding value? It is entirely possible that a lesion which impacts on or abolishes signals that are precursors for generating a value signal, such as stimulus identity representations, could also result in an impairment in value-related behavior. It does not follow that this region is necessarily involved in encoding the value signal per se, but rather that this area could be providing an input into a value-signal computation being performed further downstream. Similarly, even if we imagine the situation where a specific set of neurons are optogenetically, electrically or otherwise stimulated, which in turn exert an effect on value-related behavior, this also does not necessarily imply the neurons are necessarily computing value per se. So long as the identified neurons are encoding some variable relevant to the subsequent computation of value, or to the manner in which a value signal is subsequently transformed to guide behavior, then a manipulation of those neurons might impact on value-related behavior, without necessitating a direct role for those neurons in encoding value.

Conversely, we could consider the situation where a given region or set of neurons have been causally manipulated and there is no reported effect on value-related behavior. Does the absence of an effect allow us to exclude the possibility that these neurons are involved in value computations? Due to the possibility of redundancy in neural systems, and because of the hypothesized existence of multiple control and learning systems for guiding behavior, which often may yield similar behavioral responses (e.g. goals, habits and Pavlovian control) (Balleine et al., 2008; Balleine and Dickinson, 1998), it is entirely possible that the absence of an effect in a given stimulation or lesion protocol on behavior might occur due to the fact that one of the systems left intact after the lesion continues to drive behavior in a similar manner to that which occurred prior to the lesion. For example, a lesion to a region known to impact on goal-directed control, might still leave the habitual and Pavlovian

systems in place which would continue to drive adaptive value-related behavior in many circumstances. Of course, with a behavioral protocol that is sufficiently sensitive to the distinct contributions from different learning systems, it may be possible to disambiguate effects of lesions to one or other system (e.g. Yin et al. 2004; 2005). However, in general, in causal brain manipulations as with any other experimental approach, one needs to be cautious about interpreting the absence of a significant effect as evidence in support of the null hypothesis.

It is important to emphasize that this discussion is no way arguing against causal manipulations as an essential complement to correlative measures. The argument being made here is only that one cannot necessarily pin-point the precise computation being implemented by a given neural process (such as in encoding value representations per se) solely from a causal manipulation.

How in practice is it possible to distinguish value from its consequences?

Out of several methods discussed here for assessing value, the most successful approaches are those that attempt to explicitly divorce non-value related outcome-features from subjective value by experimentally manipulating the value of the outcome itself while leaving sensory features constant through devaluation/revaluation of the outcome, or in a revealed preference approach, particularly under situations where preferences to the same objects change over time, or where preferences to different types of goods are measured without obvious stimulus-features that correlate with those preferences.

However, such approaches still face challenges from the fact that concomitant changes in visceral, autonomic and skeletomotor activity will follow as a function of changes in value. Neural responses to these signals can potentially be misattributed as value. One essential step to resolving this challenge would be to obtain detailed physiological measurements of autonomic and skeletomotor responses during an experiment (such as by monitoring pupil dilation, skin conductance, facial electromyography, body movement etc.) as this could enable such effects to be taken into account as potential drivers of neural activity. By pitting such correlates against a value signal during statistical analysis it may be possible to differentiate value per-se from its sequelae. One natural approach to this question is to deploy a similar approach used to attempt to disambiguate attentional effects from value, which is to compare and contrast neural, skeletomotor and autonomic responses to both appetitive and aversive predicting cues, actions or outcomes. Some autonomic patterns are common across appetitive and aversive situations, thus enabling value signals which respond differentially to appetitive and aversive situations to be dissociated from such autonomic responses (Cacioppo et al., 2000; Lang et al., 1993). Neurons encoding such non-specific signals can be dissociated from value signals because value signals will show differential activity to aversive vs appetitive predictions. However, other responses such as faciomotor reactions, consummatory reactions, and approach or avoidance behaviors are known to be much more specific to particular appetitive vs aversive situations, and may therefore be more difficult to fully differentiate from value per se (Kreibig, 2010; Lang et al., 1993).

Another viable approach is to make use of the fact that such autonomic and skeletomotor reactions are more likely to be driven by the average expected reward or punishment in a given situation, as opposed to being elicited in response to the value of specific available possible actions or goals prior to choice, particularly when multiple actions and/or multiple goals are available. For instance, if a given neuronal population is found to encode the value for a given action but not the value of other actions, or a set of neurons is found to encode the value of a specific available goal but not other goals, such neural activity can be feasibly de-correlated from the overall expected reward or punishment in a given scenario, especially if the action or goal in question ends up not being chosen by the animal on a given trial. In this way, it should be possible to disambiguate value signals for particular actions and/or goals from skeletomotor and autonomic reactions elicited in anticipation of the outcome as a consequence of the choice situation.

Concluding remarks: What are the implications for decision neuroscience/neuroeconomics?

We have considered a variety of experimental protocols used in the literature for testing whether or not activity in a particular neuron, set of neurons, or BOLD signal relates to encoding of value signals in the brain. As we have seen, it is surprisingly challenging to definitively determine whether a given measured neural signal corresponds to a value response per se or a myriad of other possible signals, including the sensory features of an associated outcome or responses elicited by the outcome or cues predicting that outcome whether physiological, cognitive or skeletomotor. We have also considered the possibility that some classes of behavioral controllers (in both Pavlovian and instrumental domains) 18 may not even require access to predictions about current outcome value in order to produce adaptive responses, such as in the case of stimulus-response habits and Pavlovian cue-reflex associations.

Overall, the main argument being made here is the importance of distinguishing between different possible accounts for neural response patterns found in experimental manipulations of value. By doing so, it will be possible to not only gain a better understanding of how and where value-signals are represented in the brain, but also to better characterize how such signals are constructed from non-value precursors, as well as to begin to establish the way in which value signals are ultimately transformed in order to influence behavior.

Acknowledgments

I would like to thank Wolfgang Pauli for helpful comments on the manuscript, and Antonio Rangel for helpful discussions. The preparation of this manuscript was supported by grants to JOD from NIDA, (DA033077-01 supported by Oppnet), and the NIMH Conte Center for the neurobiology of Social Decision Making.

References

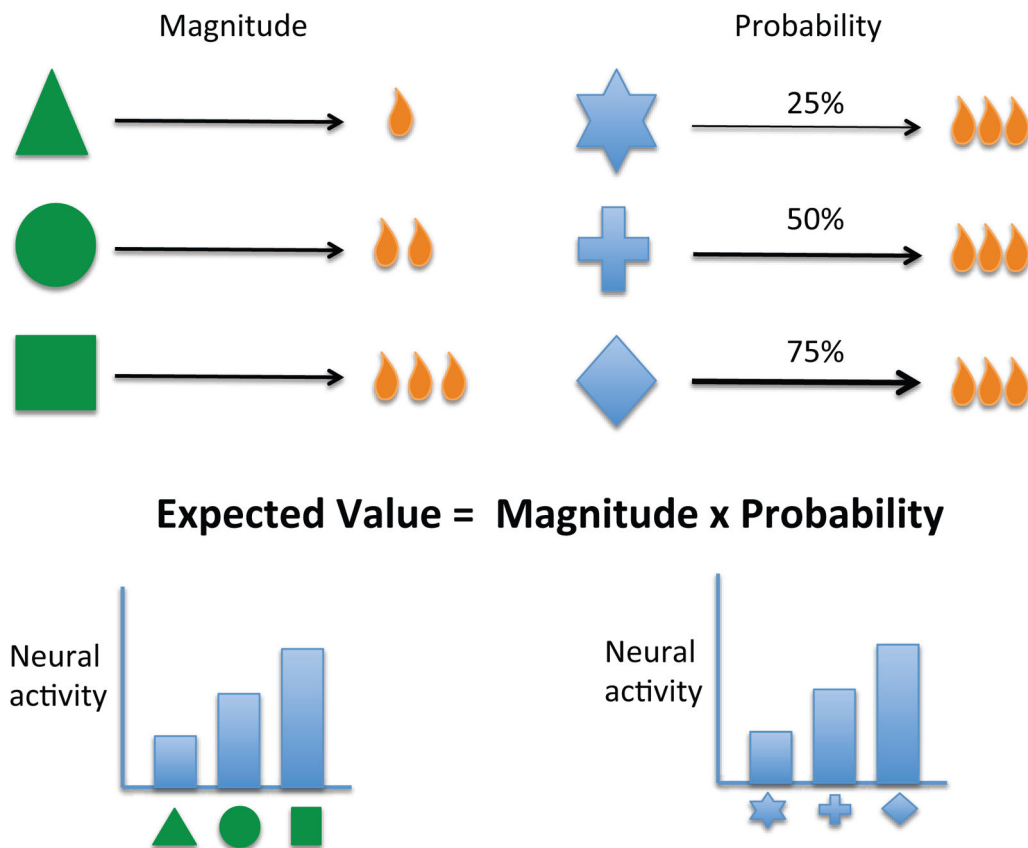
- Balleine, BW.; Daw, ND.; O'Doherty, JP. Multiple forms of value learning and the function of dopamine. In: Glimcher, PW.; Camerer, C.; Fehr, E.; Poldrack, RA., editors. *Neuroeconomics: decision making and the brain*. Elsevier; New York: 2008. p. 367-385.
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998; 37:407-419. [PubMed: 9704982]

- Balleine BW, Killcross S. Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci.* 2006; 29:272–279. [PubMed: 16545468]
- Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci.* 2000; 20:4311–4319. [PubMed: 10818166]
- Bechara A, Damasio AR, Damasio H, Anderson SW. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition.* 1994; 50:7–15. [PubMed: 8039375]
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nature neuroscience.* 2007; 10:1214–1221.
- Bentham, J. *An Introduction to the Principles of Morals and Legislation* (Dover Philosophical Classics). Dover Publications; New York: 2007.
- Berridge KC. Food reward: brain substrates of wanting and liking. *Neuroscience and biobehavioral reviews.* 1996; 20:1–25. [PubMed: 8622814]
- Berridge KC. From prediction error to incentive salience: mesolimbic computation of reward motivation. *The European journal of neuroscience.* 2012; 35:1124–1143. [PubMed: 22487042]
- Brown PL, Jenkins HM. Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior.* 1968; 11:1–8. [PubMed: 5636851]
- Cacioppo JT, Berntson GG, Larsen JT, Poehlmann KM, Ito TA. The psychophysiology of emotion. *Handbook of emotions.* 2000; 2:173–191.
- Camerer CF. Neuroeconomics: opening the gray box. *Neuron.* 2008; 60:416–419. [PubMed: 18995815]
- Chib VS, Rangel A, Shimojo S, O'Doherty JP. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci.* 2009; 29:12315–12320. [PubMed: 19793990]
- Colwill RM, Motzkin DK. Encoding of the unconditioned stimulus in Pavlovian conditioning. *Animal Learning & Behavior.* 1994; 22:384–394.
- Cools R, Clark L, Owen AM, Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci.* 2002; 22:4563–4567. [PubMed: 12040063]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience.* 2005; 8:1704–1711.
- Dean, M. Chapter 7: What Can Neuroeconomics Tell Us About Economics (and Vice Versa)?. In: Zentall, TR.; Crowley, PH., editors. *Comparative Decision-Making.* Oxford University Press; New York: 2013.
- Dickinson A. Actions and habits: the development of a behavioural autonomy. *Philos Trans R Soc Lond B Biol Sci.* 1985; 308:67–78.
- Everitt BJ, Cardinal RN, Parkinson JA, Robbins TW. Appetitive behavior: impact of amygdala-dependent mechanisms of emotional learning. *Ann N Y Acad Sci.* 2003; 985:233–250. [PubMed: 12724162]
- Fehr E, Camerer CF. Social neuroeconomics: the neural circuitry of social preferences. *Trends in cognitive sciences.* 2007; 11:419–427. [PubMed: 17913566]
- Fellows LK, Farah MJ. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb Cortex.* 2005; 15:58–63. [PubMed: 15217900]
- Glimcher PW, Rustichini A. Neuroeconomics: the consilience of brain and decision. *Science.* 2004; 306:447–452. [PubMed: 15486291]
- Gottfried JA, O'Doherty J, Dolan RJ. Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J Neurosci.* 2002; 22:10829–10837. [PubMed: 12486176]
- Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science (New York, NY).* 2003; 301:1104–1107.
- Grabenhorst F, Rolls ET. Value, pleasure and choice in the ventral prefrontal cortex. *Trends in cognitive sciences.* 2011; 15:56–67. [PubMed: 21216655]

- Gray MA, Harrison NA, Wiens S, Critchley HD. Modulation of emotional appraisal by false physiological feedback during fMRI. *PloS one*. 2007; 2:e546. [PubMed: 17579718]
- Hampshire A, Chaudhry AM, Owen AM, Roberts AC. Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *Neuroimage*. 2012; 59:4102–4112. [PubMed: 22075266]
- Hampton AN, Bossaerts P, O'Doherty JP. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci*. 2006; 26:8360–8367. [PubMed: 16899731]
- Holland PC, Straub JJ. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1979; 5:65–78. [PubMed: 528879]
- Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature neuroscience*. 1998; 1:304–309.
- Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J Cogn Neurosci*. 2004; 16:463–478. [PubMed: 15072681]
- Horvitz JC. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*. 2000; 96:651–656. [PubMed: 10727783]
- Jarvik ME, Kopp R. An improved one-trial passive avoidance learning situation. *Psychological Reports*. 1967; 21:221–224. [PubMed: 6078370]
- Jenkins HM, Moore BR. The form of the auto-shaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior*. 1973; 20:163–181.
- Kable JW, Glimcher PW. The neural correlates of subjective value during intertemporal choice. *Nature neuroscience*. 2007; 10:1625–1633.
- Kahneman D, Wakker PP, Sarin R. Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*. 1997; 112:375–406.
- Klein-Flugge MC, Barron HC, Brodersen KH, Dolan RJ, Behrens TE. Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2013; 33:3202–3211. [PubMed: 23407973]
- Konorski, J. Conditioned reflexes and neuron organization. 1948.
- Kreibig SD. Autonomic nervous system activity in emotion: A review. *Biological psychology*. 2010; 84:394–421. [PubMed: 20371374]
- Lak, A.; Stauffer, WR.; Schultz, W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America*; 2014.
- Lang PJ, Greenwald MK, Bradley MM, Hamm AO. Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology*. 1993; 30:261–273. [PubMed: 8497555]
- Lau B, Glimcher PW. Action and outcome encoding in the primate caudate nucleus. *J Neurosci*. 2007; 27:14502–14514. [PubMed: 18160658]
- Leathers ML, Olson CR. In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science*. 2012; 338:132–135. [PubMed: 23042897]
- Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW. Neural representation of subjective value under risk and ambiguity. *Journal of neurophysiology*. 2010; 103:1036–1047. [PubMed: 20032238]
- Mackintosh NJ. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*. 1975:82.
- Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459:837–841. [PubMed: 19448610]
- Maunsell JH. Neuronal representations of cognitive state: reward or attention? *Trends in cognitive sciences*. 2004; 8:261–265. [PubMed: 15165551]

- McDannald MA, Jones JL, Takahashi YK, Schoenbaum G. Learning theory: A driving force in understanding orbitofrontal function. *Neurobiology of learning and memory*. 2014; 108C:22–27. [PubMed: 23770491]
- McNamee D, Rangel A, O'Doherty JP. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature neuroscience*. 2013; 16:479–485.
- Mirenowicz J, Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol*. 1994; 72:1024–1027. [PubMed: 7983508]
- Montague PR, Berns GS. Neural economics and the biological substrates of valuation. *Neuron*. 2002; 36:265–284. [PubMed: 12383781]
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. Midbrain dopamine neurons encode decisions for future action. *Nature neuroscience*. 2006; 9:1057–1063.
- Noonan MP, Walton ME, Behrens TE, Sallet J, Buckley MJ, Rushworth MF. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A*. 2010; 107:20547–20552. [PubMed: 21059901]
- O'Doherty J, Critchley H, Deichmann R, Dolan RJ. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J Neurosci*. 2003; 23:7931–7939. [PubMed: 12944524]
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature neuroscience*. 2001; 4:95–102.
- O'Doherty JP. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol*. 2004; 14:769–776. [PubMed: 15582382]
- O'Doherty JP. Lights, camembert, action! The role of human orbitofrontal cortex in encoding stimuli, rewards, and choices. *Annals of the New York Academy of Sciences*. 2007; 1121:254–272. [PubMed: 17872386]
- Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature*. 2006; 441:223–226. [PubMed: 16633341]
- Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*. 2006; 439:865–870. [PubMed: 16482160]
- Pavlov, IP. *Conditioned reflexes*. Oxford University Press; Oxford: 1927.
- Payzan-LeNestour E, Bossaerts P. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS computational biology*. 2011; 7:e1001048. [PubMed: 21283774]
- Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP. The neural representation of unexpected uncertainty during value-based decision making. *Neuron*. 2013; 79:191–201. [PubMed: 23849203]
- Pearce JM, Hall G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*. 1980; 87:532–552. [PubMed: 7443916]
- Pickens CL, Saddoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci*. 2003; 23:11078–11084. [PubMed: 14657165]
- Plassmann H, O'Doherty J, Rangel A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci*. 2007; 27:9984–9988. [PubMed: 17855612]
- Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. *Nature*. 1999; 400:233–238. [PubMed: 10421364]
- Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. *Nature reviews Neuroscience*. 2008; 9:545–556.
- Roesch MR, Calu DJ, Esber GR, Schoenbaum G. All that glitters ... dissociating attention and outcome expectancy from prediction errors signals. *Journal of neurophysiology*. 2010; 104:587–595. [PubMed: 20554849]
- Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature neuroscience*. 2007; 10:1615–1624.
- Roesch MR, Esber GR, Li J, Daw ND, Schoenbaum G. Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *Eur J Neurosci*. 2012; 35:1190–1200. [PubMed: 22487047]

- Roesch MR, Olson CR. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science*. 2004; 304:307–310. [PubMed: 15073380]
- Rolls BJ, Rolls ET, Rowe EA, Sweeney K. Sensory specific satiety in man. *Physiol Behav*. 1981; 27:137–142. [PubMed: 7267792]
- Rolls, ET. *Emotion Explained*. Oxford University Press; Oxford: 2007.
- Rolls ET, Hornak J, Wade D, McGrath J. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J Neurol Neurosurg Psychiatry*. 1994; 57:1518–1524. [PubMed: 7798983]
- Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science (New York, N Y)*. 2005; 310:1337–1340.
- Sanfey AG, Loewenstein G, McClure SM, Cohen JD. Neuroeconomics: cross-currents in research on decision-making. *Trends in cognitive sciences*. 2006; 10:108–116. [PubMed: 16469524]
- Schoenbaum G, Chiba AA, Gallagher M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature neuroscience*. 1998; 1:155–159.
- Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol*. 1998; 80:1–27. [PubMed: 9658025]
- Shapiro MG, Frazier SJ, Lester HA. Unparalleled control of neural activity using orthogonal pharmacogenetics. *ACS chemical neuroscience*. 2012; 3:619–629. [PubMed: 22896806]
- Sohn JW, Lee D. Order-dependent modulation of directional signals in the supplementary and presupplementary motor areas. *J Neurosci*. 2007; 27:13655–13666. [PubMed: 18077677]
- Solomon RL, Corbit JD. An Opponent-Process Theory of Motivation: I. Temporal Dynamics of Affect. *Psychological Review*. 1974; 81:119–145. [PubMed: 4817611]
- Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science (New York, N Y)*. 2004; 304:1782–1787.
- Thorndike, EL. *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*. Macmillan; New York: 1898.
- Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. *Nature*. 1999; 398:704–708. [PubMed: 10227292]
- Valentin VV, Dickinson A, O'Doherty JP. Determining the neural substrates of goal-directed learning in the human brain. *J Neurosci*. 2007; 27:4019–4026. [PubMed: 17428979]
- Williams SC, Deisseroth K. Optogenetics. *Proc Natl Acad Sci U S A*. 2013; 110:16287. [PubMed: 24106299]
- Wunderlich K, Rangel A, O'Doherty JP. Neural computations underlying action-based decision making in the human brain. *Proc Natl Acad Sci U S A*. 2009; 106:17199–17204. [PubMed: 19805082]
- Wunderlich K, Rangel A, O'Doherty JP. Economic choices can be made using only stimulus values. *Proc Natl Acad Sci U S A*. 2010; 107:15005–15010. [PubMed: 20696924]
- Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci*. 2004; 19:181–189. [PubMed: 14750976]
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci*. 2005; 22:513–523. [PubMed: 16045504]
- Zink CF, Pagnoni G, Chappelow J, Martin-Skurski M, Berns GS. Human striatal activation reflects degree of stimulus saliency. *Neuroimage*. 2006; 29:977–983. [PubMed: 16153860]

**Figure 1.**

Schematic of a simple experiment aimed at detecting neural responses to expectation of magnitude and probability of a juice outcome. Even if a cue is retrieving sensory features of the outcome (e.g. its sweetness, odor, texture or some combination thereof) but not its value, putative neural responses to the cue would still scale with both magnitude (the intensity of the sensory experience) and probability (the strength of the stimulus-stimulus association formed). Thus, distinguishing neural signals encoding cue-outcome associations that are entirely sensory based (i.e. cue → outcome sensory features) from cue-outcome associations that retrieve underlying values (cue → value(outcome)), is challenging using this type of manipulation.