Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Dialogue on economic choice, learning theory, and neuronal representations

Camillo Padoa-Schioppa[1] and Geoffrey Schoenbaum[2]

In recent years, two distinct lines of work have focused on the substrates of associative learning and on the mechanisms of economic decisions. While experiments often focused on the same brain regions — most notably the orbitofrontal cortex — the two literatures have remained largely distinct. Here we engage in a dialogue with the intent to clarify the relationship between the two frameworks. We identify a potential correspondence between the concept of *outcome* defined in learning theory and that of *good* defined in neuroeconomics, and we specifically discuss the concept of *value* defined in the two frameworks. While many differences remain unresolved, a common idea is that good/outcome values are subjective, devaluation-sensitive and computed on the fly, not 'cached' or pre-computed.

Addresses
[1] Departments of Anatomy and Neurobiology, Economics and Biomedical Engineering, Washington University, St. Louis, MO 63110, United States
[2] Intramural Research Program, NIDA, 251 Bayview Drive, Baltimore, MD 21224, United States

Corresponding authors: Padoa-Schioppa, Camillo (camillo@wustl.edu) and Schoenbaum, Geoffrey (geoffrey.schoenbaum@nih.gov)

Current Opinion in Behavioral Sciences 2015, **5**:16–23

This review comes from a themed issue on **Decision making/ neuroeconomics**

Edited by **John O'Doherty** and **Colin Camerer**

http://dx.doi.org/10.1016/j.cobeha.2015.06.004

2352-1546/Published by Elsevier Ltd.

Introduction

Economic theory has had a large influence on neuroscience and has greatly informed our understanding of how neural circuits may generate decisions. However, the relationship between ideas that originated in economics and other ideas regarding how associative information is organized conceptually and represented in neural circuits is unclear. For example, the orbitofrontal cortex (OFC) has been associated with at least two classes of behaviors. One set of studies implicates the OFC in economic decisions (e.g. choices between different dishes on a restaurant menu), which reflect the values assigned to the available goods [1–4,5•]. These studies emphasize the subjective nature of value and the need to integrate across multiple dimensions. Another set of studies shows that the OFC contributes to learned behaviors dependent on the ability to represent the current value of expected outcomes (e.g. choosing not to go to the restaurant at all, if the restaurant specializes in something you just had last night) [6•,7•,8–10]. These studies emphasize the need to use inference and a model of the environment, rather than direct experience, to derive a value. These two literatures have remained largely separate, although experimental findings such as the effects of OFC lesions on behavior following outcome devaluation procedures have been interpreted in both frameworks [11•]. A potentially unifying concept is that of *value*, but the operational definitions of value used in neuroeconomics and in learning theory have been difficult to align. Here we discuss some conceptual questions with the intent to clarify the relationship between the two frameworks. The article is written in the form of a dialogue. While we refer explicitly to the OFC, many of the concepts and issues addressed here are more general and pertain also to other brain regions.

Subjective value and economic choice

GS: Lets start with general definitions. What is economic value? Is economic value adequately operationalized by current experimental approaches, which are essentially defined by choices?

CPS: Here I use the terms ''utility'', ''economic value'' and ''subjective value'' as equivalent. Subjective values measured behaviorally are always derived from choices. Economic theory tells us (roughly) the following. If an individual makes a set of choices that satisfy transitivity, her choices can be described as if they were dictated by a subjective value. Transitivity is satisfied if every time X is preferred to Y and Y is preferred to Z, X is preferred to Z. The subjective value function is defined up to a monotonic transformation [12].

In our studies, we often simplify the analysis by assuming that value functions are linear (i.e. that the subjective value of two drops of apple juice equals two times the subjective value of one drop of apple juice). If this is true, an adequate set of choices between, say, apple juice and grape juice will provide an operational measure for the subjective value of any quantity of grape juice in units of apple juice (or vice versa). Under this assumption, we and others have shown that subjective values are explicitly represented in OFC and other brain regions [11•,13–15].

Note that the linearity assumption is not conceptually critical — we could relax it, although in practice we would need many more trials for each experimental condition.

GS: Does utility or subjective value encompass everything that affects choices — costs, benefits, punishments, etc.? Can we fractionate it? For example, do neural circuits separate the impact of positive and negative valence, or that of cues and actions?

CPS: From the behavioral perspective, anything that matters to choice — cost, benefits, probabilities, delays, contingencies, etc. — can and should be integrated into subjective values [11•]. Of course, this does not mean that these quantities are not also represented independently somewhere in the brain. Whether positive and negative values are represented in the same way, and possibly by the same neurons within OFC, is not clear. More experimental work is necessary to examine this point, although the results of Morrison and Salzman [16] do suggest a unitary representation.

Different types of decisions

GS: Are all choices based on value signals at some level or can we make choices without computing a value? If so, how do we distinguish experimentally between different choice mechanisms?

CPS: This is a very important issue, and my personal view is as follows. Many behaviors can be construed as involving a decision. Indeed, the phrase ''decision making'' is used in neurobiology to describe a variety of systems, from *C. elegans* that ''decide'' when to move to a new food patch [17], to fruit flies that ''decide'' where to lay their eggs based on the local concentration of sucrose [18], to collective ''decisions'' in many species [19]. Many of these behaviors may be described using an economic optimization formalism (which, incidentally, is a very general formalism originally developed in physics). That said, I would argue that the decisions you and I make when we look at a restaurant menu or when we choose between different possible investments in our 401k are based on qualitatively different cognitive and neural mechanisms from these other kinds of decisions. So when I say or write ''economic decision'', I don't mean ''any decision that can be described using an economic formalism''; rather, I refer to a particular kind or domain of decision making.[3]

The key question is: How do we define the domain of economic decisions? My starting point is an appeal to intuition, as I just did in the examples of the restaurant and the 401k. These examples highlight two important traits of economic decisions — they involve goods that can vary on multiple dimensions, and they normally do not have an intrinsically (objectively) correct answer.[4] This allows us to proceed with experimental work and to identify candidate brain structures and neuronal mechanisms that might underlie such choices. In the long run, however, appealing to intuition or even to the behavioral criteria described above is somewhat unsatisfactory. In my view, different kinds of decisions should eventually be defined based on the underlying neural mechanisms. In other words, I hope that in a near future we will define an economic decision as a neural decision process that takes place in particular brain area(s) and according to specific neuronal mechanisms.

GS: The idea that economic decisions have no intrinsically correct answer sounds clear at first, but when I try to apply it, I find some difficulty. A friend of mine dislikes ham. This choice is subjective at first glance, however it is due to a proscription defined by her society during her upbringing. Thus the value driving the decision is actually extrinsic and thus might be seen as objective in the context of that society.

CPS: You asked whether all choices require computing values. The short answer is no, because there are many behaviors that can be described as a decision but are not necessarily value-based. Your friend's attitude toward ham seems dictated by what are sometimes called ''sacred values'', which are often associated with religious beliefs or ethnic identity. Interestingly, Berns *et al.* [20•] showed that decisions based on sacred values activate lateral prefrontal cortex (PFC), but not OFC or ventromedial PFC, which are usually associated with subjective values and economic choices. Berns' results suggest that when we make decisions based on sacred values, we don't weigh all the aspects of the situation; we simply apply a rule.

Neuronal representations of goods and values

GS: You often distinguish between good space and action space, and economic value seems to be represented in both. Why distinguish them, and how are they linked?

CPS: Let me make an important premise. Subjective values are never represented purely as such. They are always attached to something — for example a good or an action. There may be some confusion in the field on this

[3] The distinction between 'economic' decisions and other types of decisions is not one usually made in economic theory. As discussed here, this distinction ultimately rests on cognitive and neural concepts.

[4] The lack of an intrinsically correct answer is contrast to the situation found in perceptual decisions, where there is always a correct answer (or, at least, the subject believes that there is one). It may be argued that in some 'economic' decisions goods vary on a single dimension and/or one of the options is objectively dominant (e.g., the choice between $2 and $1). The critical point, however, is that economic goods can and typically do vary on multiple dimensions (including non-parametric dimensions), and that decisions are typically made in the absence of an objectively correct answer.

point, so let me elaborate. I have argued that the neuronal representation of economic values in OFC is ''abstract'' [11•]. The full phrase, however, is ''abstract from the sensorimotor contingencies of choice''. In other words, value-encoding neurons in OFC don't have a spatial response field the way neurons in many other brain areas do [21•]. So if a particular neuron in OFC encodes the value of an apple, it does so independently of the spatial location of the apple and independently of the action necessary to obtain the fruit (except for the action costs).

An action space is a representation in which values are attached to actions. For example in the lateral intraparietal area, each neuron has a response field and the entire population provides a map of all possible saccades. The activity of each cell is modulated by the value associated to the corresponding saccade [22,23]. Similar action-based representations exist in many or all premotor and motor areas [11•,24]. In contrast, a good space is a representation in which values are attached to goods, and this is how values seem to be represented in the OFC. Normally, goods exist independent of their location — for example an apple is an apple independently of its position in space. Thus good-based representations are normally action independent. However, in some experiments, goods may be defined exclusively by their spatial locations [25,26]. In such cases, abstract representations cannot be distinguished from spatial representations (for discussion see [27]).

GS: You have argued that the activity of some neurons in the OFC reflects the subjective nature of values and cannot be explained as encoding a physical property or an ingredient of the juices [3]. Yet the above seems to imply that the values are not represented as such, that they are attached to something physical. Are these two statements in contradiction?

CPS: No, they are not. The fact that a neuron is associated with a physical object does not imply that its firing rate encodes a physical property of that object. Some neurons in OFC are associated with one of the goods available for choice, for example an apple or an orange (*offer value* cells). Other neurons are associated with the good chosen by the animal, which in different trials could be an apple, or an orange, or something else (*chosen value* cells) [3,28].[5] In principle, a neuron associated to a particular good (a fruit or the chosen good) could encode the subjective value of that good or, alternatively, a physical property of the good such as the quantity of a particular ingredient. These two hypotheses can be distinguished because subjective values vary to some extent from day to day, while physical properties remain unchanged. You can think of fluctuations in subjective values as a naturally-occurring devaluation and revaluation. Our studies have shown that the activity of *offer value* cells and *chosen value* cells in the OFC reflects this variability (or vice versa) [3,29•]. Hence, the activity of these neurons reflects the subjective nature of values. (Similar results were obtained for other brain regions [23,30–32].)

To summarize, neurons in the OFC may be associated to physical goods, but they encode the subjective values of those goods.

Learning theory and neuroscience

CPS: We will discuss the relation between concepts defined in the context of economic choice and concepts defined in the context of associative learning. Let me first ask you a few general questions. What do you mean when you refer to ''reinforcement learning'' versus ''learning theory''? Is there a single, well-defined learning theory?

GS: When I refer to learning theory, I am referring to the branche of experimental psychology that is devoted to figuring out empirically the rules that govern associative learning. Despite its name, it does not refer to a single theory, but rather to the fact that the field is theoretically driven, i.e. learning models are typically instantiated in formal equations. Specific models often compete to explain the same phenomena (e.g. selective learning), but may occasionally be seen as complementary [33]. They include ''classic'' examples, such as the Rescorla–Wagner [34], Mackintosh [35], Pearce-Hall [36] and Wagner's SOP models [37], but the field has continued to produce models to integrate more recent empirical findings [38–41]. By contrast, when I refer to reinforcement learning (RL), I am referring to the field of machine learning that specializes in designing algorithms to describe how the actions of an agent — a computer construct — can maximize some reward. The former is strongly bottom up inasmuch as its goal is to explain animal learning. As such it has tried to match biological data from the beginning. The latter is more top down (normative), at least to my understanding, since originally its goal was simply to design efficient algorithms.While ideas from the two fields are remarkably convergent, there are some fundamental distinctions, which are often overlooked. For instance, temporal difference RL (TDRL) and the Rescorla–Wagner model use similar algorithms to drive learning but make different underlying assumptions. Whereas TDRL's algorithm drives changes in the ''cached'' value of the cue, Rescorla and Wagner apply their algorithm to drive changes in the associative strength between the cue and the outcome. In the Rescorla–Wagner model, the strength of this association will determine the extent to which the cue is able to activate a memory of the outcome in the future, which is not something that TDRL contemplates. This difference — between the storage of a simple value versus

[5] A third group of neurons found in OFC (*chosen juice* cells) encodes the binary choice outcome. Together, *chosen juice* cells and *chosen value* cell capture the output of the decision process.

the formation of an association that allows the activation of a memory of the reinforcer — is why TDRL is not able to explain, for instance, the effects of outcome devaluation on learned behavior (since the current value of the outcome is not part of the resultant learning for TDRL), whereas the Rescorla–Wagner model can, given certain assumptions.

CPS: Is learning theory independent of results from neuroscience, or is it informed by what neurons in the brain do or don't do?

GS: Learning theory is orthogonal to neuroscience in the sense that its goal is simply to define, empirically, the rules that govern associative learning. Some of the best practitioners have never opened up the brain or have only done so in collaboration with others. Of course, modern neuroscience would propose that rules governing learning should be related to the brain in some way. However it is an experimental question whether, to what extent, and at what level this is the case. This is the question that our lab and many others are interested in answering. Our use of learning theory is very similar to how you and others use economics, I think, inasmuch as economic concepts have been developed independently, and you are now asking whether and to what extent these concepts map onto brain systems.

Concepts of value in learning theory

CPS: I understand how value is defined in Sutton and Barto's model of reinforcement learning [42]. Are there other or more general definitions of value in learning theory?

GS: Value is tricky [43•]. Historically, the concept of value in learning theory has been intimately linked to the concept of reinforcement and reward. Thus, for instance, a stimulus has positive value to the extent that it reinforces performing an arbitrary action, like pressing a lever. However behavior cannot always be taken to reflect the current value of a reward, since it can be influenced by various kinds of associative information. Indeed this is almost certainly the rule rather than the exception.

The classic demonstration of this is the distinction between habitual and goal-directed behavior. Loosely speaking, a habit would be a behavior that does not change immediately when your desire for the outcome changes, whereas a goal-directed behavior does change, and it changes immediately and without any further learning. In each case, there is a behavior that might arguably be said to reflect value, however there are clearly different and dissociable associations underlying the behavior in the two circumstances. And we know that brain circuits respect these distinctions to some extent. So while learning theorists may use the word value, modern papers tend to qualify the use of this word, typically using it in the context of specific associations.

CPS: Thus would it be fair to say that, within learning theory, there are different types of "value"? For example, is there a "habit value" and a "goal-directed value"?

GS: Yes perhaps that is one way to think about it. If you operationalize value as being revealed by behavior, then learning theory would hold that value is not unitary but would be multidimensional, reflecting the underlying associative structure. Thus there would be a "habit value", which would reflect the strength of the stimulus–response association underlying an action, and a "goal-directed value", which would reflect the strength of the response–outcome association underlying an action [44]. Note that these are instrumental associations — more on this below.

Important in the idea that behavior (and therefore value) can be affected by different types of associations is that normal behavior will generally reflect a mixture of these influences [45]. Thus when I drive to work, I am engaging multiple types of associative structures. My decision to make a left turn at the end of my street likely reflects some "habit value" and some "goal-directed value", apportioned perhaps according to how focused I am on my morning meeting and whether I've remembered that new construction going on at the bottom of the hill that blocks traffic after 9 AM. To address this issue, learning theory has developed rigorous procedures for demonstrating what sort of underlying associations are mediating the behavior. This is the rationale for manipulations such as outcome devaluation or training manipulations designed to make behavior "habitual" [44,46,47]. It has only been through the use of such manipulations that learning theory has been able to identify and clearly dissociate neural circuits mediating behavior supported by different types of associative information [48–51].

From my perspective, the current ways of measuring economic value are not as well controlled. Specifically there is no explicit manipulation, like devaluation, designed to remove or rule out influences of prior experience or what might be considered "habit value" in the above lexicon. Notably this has not prevented the brain regions associated with economic value from aligning well with the areas we believe are important for devaluation-sensitive behavior. I think the use of relatively unique choices, made only once or at most a small handful of times likely accounts for this in many studies [4,5•,52,53]. As a result, the influence of what might be thought of as habits or policies (stimulus–response or parallel Pavlovian associations), either previously acquired or acquired in the course of the experiment, is likely limited. But in settings in which choices are presented many times, it is likely that both types of associations are formed. This is not to say that this behavior is "habitized" or driven solely by associations that do not incorporate a representation of the outcome. Indeed it is very hard to "habitize" behavior when multiple outcomes are involved. But neural correlates could reflect either type of information. In this regard, it is particularly impressive that the neural correlates of economic value that you have demonstrated in the OFC show a dependence on the relative value of the juice pairs across days [3] [see their Supplementary material]. This provides direct evidence that this activity varies with shifts in current value of the juices with minimal opportunity for attaching this value to the predictive cues. This combined with data showing that neural activity in the OFC is devaluation sensitive [7•,54] suggests that economic value correlates in the OFC are likely related to the sort of value that learning theory experiments would assign to this area [6•,8,55–58].

CPS: I see your point that the economic framework does not dissociate between different influences on choice as closely related to the above discussion about "Different types of decisions". However, the idea that you can practically dissociate between the two components of value — "goal-directed" and "habitual" — in the framework of learning theory may be illusory, I think. Take the example of you driving to work this morning. How would I actually measure from your behavior the "goal-directed value" versus the "habit value"? Presumably, I would have to run a well-controlled experiment, in which I vary one variable at the time and I measure some learning process for each of those variables. In principle that works, but in reality the value that I would measure in such experiments might not at all be the same value that your brain computed this morning when you drove to work. In this perspective, the economic approach is more agnostic — value is what I can measure from choices, right now. In the end, I think that both approaches have to "bite the bullet" and recognize that the only way to dissociate different kinds of decisions is to define them in terms of the underlying brain processes.

Goods, outcomes and values in neuroeconomics and learning theory

GS: What is the relation between economic choice and behaviors defined in learning theory?

CPS: To my understanding, what we call economic choice is closely related to what is often called goal-directed behavior. Many studies show that OFC lesions affect task performance in reward devaluation procedures [6•,8,55,57–60]. Usually, these results are interpreted saying that *the behavior elicited by the task is goal-directed*. However, in many cases the task involves a decision, and thus the computation of values, and the results can be interpreted as evidence that *values are not learned as such* (i.e. values are computed on the fly) [11•]. The two interpretations are not mutually exclusive.

GS: This makes sense, and indeed one defining feature of a "goal-directed behavior" is that it is based on the current value of the outcome. In this sense, it reflects a predicted value that is computed on the fly. However, goal-directed behavior, as it is currently defined by learning theorists such as Balleine and Dickinson [44], is instrumental or specifically based on action-outcome associations. In this view, goal-directed behavior would have to be controlled by an action-based, not a good-based, representation. In this way, goal-directed behavior seems to differ from or at least be more restricted than economic decision making.

CPS: It is true that goal-directed behavior is often discussed as based on actions. But from my perspective, the concept of goal-directed behavior could easily generalize to abstract representations that do not depend on specific actions. For example, lets say that normally you are equally prone to watch a movie or read a book. However, if you watched a movie yesterday (selective satiation) you are more likely to opt for a book today. The decision between the book and the movie would qualify as goal-directed behavior, but it is unlikely to be based on the physical action you have to undertake to turn on the TV or to walk upstairs and grab your book. That said, I take your point. Borrowing your language, it is more accurate to say that economic choice is closely related to "outcome-guided" (as distinguished from "goal-directed") behavior.

GS: What does it mean that economic value does not reflect learning?

CPS: I don't know, I never said that. What I did say and write [11•] is that economic values are not learned and stored as such; they are computed online (or on the fly) when needed. Subjective values are computed by integrating multiple determinants, of which some are external and some are internal. If I need to compute the subjective value of a green apple, I will do so by integrating information that I have learned over time (e.g., the typical taste of green apples, which is not the same as that of red apples) and also information that I cannot have possibly learned in the past (e.g. how hungry or thirsty I am right now). Learning processes are not irrelevant to the computation of value, but values are not fixed; they depend on the circumstances, and cannot be learned as such [11•].

GS: This brings the concept of economic choice closely into alignment with the idea of devaluation-sensitive behavior in learning theory. How do you view the relation between value defined in neuroeconomics and value defined in learning theory?

CPS: In principle, values defined in economic choice and values defined in learning theory have nothing to do with one another. Notably, when they perform economic choices in our experiments, animals are no longer learning anything, so it is not clear why one would invoke concepts from learning theory. At the same time, the concept of ''outcome'' as used in learning theory seems closely related to the concept of ''good'' that we often refer to. One difference is that goods in neuroeconomics are defined by multiple physical dimensions, including quantity, probability, delay, etc. [11•]. Thus one apple delivered with probability $p = 0.5$ and one apple delivered with probability $p = 1$ are different goods. In contrast, scholars of learning theory usually think of the outcome as just the apple. However, associative learning can have a hierarchical structure, for example when one first association forms an entity that enters further associations [61]. Thus one first association between an apple and a probability, or between an apple and a delay, could constitute the ''outcome'' of a further association. So the possible correspondence is between ''goods'' and the potentially-hierarchical concept of ''outcomes''.

With this understanding, your concept of signaling the predicted ''outcome value'' to guide a behavior is similar to our concept of encoding ''economic value'' at the time of a decision. In both conceptualizations, behavior is guided by the current/subjective value of the outcome or good, and it is thus sensitive to devaluation. In principle, it would seem a good idea if the brain had a single machinery that computed such values whenever needed, and then used those values to drive different behaviors in different circumstances. My working hypothesis is that neurons in the OFC are such machinery, although other brain regions such as the amygdala might contribute to this computation.

Conclusions

Experimental frameworks derived from learning theory and economics have been successfully used to examine the neural circuits mediating associative learning and decision making, respectively. While experiments often focused on the same brain regions, work conducted in the two frameworks has remained largely separated. We attempted to clarify basic concepts defined in each framework, with the goal to facilitate a comprehensive understanding of the literature. We identified a potential correspondence between the concepts of ''good'' and ''economic value'' defined in neuroeconomics and the concepts of ''outcome'' and ''outcome value'' defined in the learning literature. Common to both frameworks is the idea that these values and the behaviors they drive are sensitive to changes in the subjective desire for the good or outcome, and convergent experimental results from both traditions indicate that neurons in the OFC encode the identity of goods/outcomes and their subjective values.

Conflict of interest

Nothing declared.

Acknowledgements

The authors thank Guillermo Esber, Thomas Stalnaker, Katherine Conen and Xinying Cai for comments on the manuscript. This work was supported by funding from NIDA (GS, intramural program; CPS, R01-DA032758). The opinions expressed in this article are the authors' own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government.

References

1. Arana FS, Parkinson JA, Hinton E, Holland AJ, Owen AM, Roberts AC: **Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection**. *J Neurosci* 2003, **23**:9632-9638.

2. Roesch MR, Olson CR: **Neuronal activity in primate orbitofrontal cortex reflects the value of time**. *J Neurophysiol* 2005, **94**:2457-2471.

3. Padoa-Schioppa C, Assad JA: **Neurons in orbitofrontal cortex encode economic value**. *Nature* 2006, **441**:223-226.

4. Kable JW, Glimcher PW: **The neural correlates of subjective value during intertemporal choice**. *Nat Neurosci* 2007, **10**:1625-1633.

5. • Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW: **Ventromedial frontal lobe damage disrupts value maximization in humans**. *J Neurosci* 2011, **31**:7527-7532.
This paper shows deficits in maintaining stable preferences among baskets of goods after damage to the medial part of orbitofrontal cortex. Preference stability — transitivity — is a key feature of economic decision-making, thus this study shows a role for the orbital areas in this type of decision-making.

6. • Gallagher M, McMahan RW, Schoenbaum G: **Orbitofrontal cortex and representation of incentive value in associative learning**. *J Neurosci* 1999, **19**:6610-6614.
This paper was the first to report that changes in conditioned responding following outcome devaluation depend on orbitofrontal cortex.

7. • Gottfried JA, O'Doherty J, Dolan RJ: **Encoding predictive reward value in human amygdala and orbitofrontal cortex**. *Science* 2003, **301**:1104-1107.
This paper shows that BOLD response to predictive cues in orbitofrontal areas in humans is altered by outcome devaluation. Notably the authors employed a paradigm that closely parallels the design used in causal rodent and primate studies, thus facilitating extrapolation across species.

8. Rudebeck PH, Saunders RC, Prescott AT, Chau LS, Murray EA: **Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating**. *Nat Neurosci* 2013, **16**:1140-1145.

9. Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez G, Mirenzi A, Schoenbaum G: **Orbitofrontal cortex supports behavior and learning using inferred but not cached values**. *Science* 2012, **338**:953-956.

10. McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G: **Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning**. *J Neurosci* 2011, **31**:2700-2705.

11. • Padoa-Schioppa C: **Neurobiology of economic choice: a good-based model**. *Annu Rev Neurosci* 2011, **34**:333-359.
This paper discusses in depth many of the ideas on economic choice, goods and subjective values referred to here.

12. Kreps DM: *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press; 1990.

13. Wallis JD: **Cross-species studies of orbitofrontal cortex and value-based decision-making**. *Nat Neurosci* 2012, **15**:13-19.

14. Bartra O, McGuire JT, Kable JW: **The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value**. *Neuroimage* 2013, **76**:412-427.

15. Clithero JA, Rangel A: **Informatic parcellation of the network involved in the computation of subjective value**. *Soc Cogn Affect Neurosci* 2013.

16. Morrison SE, Salzman CD: **Representations of appetitive and aversive information in the primate orbitofrontal cortex**. *Ann N Y Acad Sci* 2011, **1239**:59-70.

17. Bendesky A, Tsunozaki M, Rockman MV, Kruglyak L, Bargmann CI: **Catecholamine receptor polymorphisms affect decision-making in *C. elegans***. *Nature* 2011, **472**:313-318.

18. Yang CH, Belawat P, Hafen E, Jan LY, Jan YN: **Drosophila egg-laying site selection as a system to study simple decision-making processes**. *Science* 2008, **319**:1679-1683.

19. Conradt L, Roper TJ: **Consensus decision making in animals**. *Trends Ecol Evol* 2005, **20**:449-456.

20. • Berns GS, Bell E, Capra CM, Prietula MJ, Moore S, Anderson B, Ginges J, Atran S: **The price of your soul: neural evidence for the non-utilitarian representation of sacred values**. *Philos Trans R Soc Lond B Biol Sci* 2012, **367**:754-762.
This paper traced the important distinction between sacred values and economic values and showed that decisions based on the former involve the lateral prefrontal cortex. Decisions based on sacred values are thus somewhat similar to rule-based decisions.

21. • Grattan LE, Glimcher PW: **Absence of spatial tuning in the orbitofrontal cortex**. *PLOS ONE* 2014, **9**:e112750.
This paper provides the most compelling evidence that the activity of neurons in the OFC is not visuo-spatial in nature. Neurons were extensively tested with protocols similar to those normally used for visual areas, but responses did not present preferred regions of space. In other words, neurons in the OFC do not have a receptive field (or a response field).

22. Platt ML, Glimcher PW: **Neural correlates of decision variables in parietal cortex**. *Nature* 1999, **400**:233-238.

23. Louie K, Glimcher PW: **Separating value from choice: delay discounting activity in the lateral intraparietal area**. *J Neurosci* 2010, **30**:5498-5507.

24. Roesch MR, Olson CR: **Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex**. *J Neurophysiol* 2005, **94**:1469-1497.

25. Abe H, Lee D: **Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex**. *Neuron* 2011, **70**:731-741.

26. Tsujimoto S, Genovesio A, Wise SP: **Monkey orbitofrontal cortex encodes response choices near feedback time**. *J Neurosci* 2009, **29**:2569-2574.

27. Padoa-Schioppa C, Cai X: **The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives**. *Ann N Y Acad Sci* 2011, **1239**:130-137.

28. Padoa-Schioppa C: **Neuronal origins of choice variability in economic decisions**. *Neuron* 2013, **80**:1322-1336.

29. • Raghuraman AP, Padoa-Schioppa C: **Integration of multiple determinants in the neuronal computation of economic values**. *J Neurosci* 2014, **34**:11583-11603.
This study generalized the concept of outcome devaluation to reflect the risk attitude. The subjective value of a certain good normally fluctuates over time depending on the internal motivation. For an uncertain good, the value will also fluctuate depending on the time-varying risk attitude. Cell activity in OFC matched the variability measured behaviorally, providing further evidence that these neurons encoded subjective values.

30. Cai X, Padoa-Schioppa C: **Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex**. *J Neurosci* 2012, **32**:3791-3808.

31. Kim S, Hwang J, Lee D: **Prefrontal coding of temporally discounted values during intertemporal choice**. *Neuron* 2008, **59**:161-172.

32. Cai X, Kim S, Lee D: **Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice**. *Neuron* 2011, **69**:170-182.

33. Pearce JM, Bouton ME: **Theories of associative learning in animals**. *Annu Rev Psychol* 2001, **52**:111-139.

34. Rescorla RA, Wagner AR: **A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement**. In *Classical Conditioning II: Current Research and Theory*. Edited by Black AH, Prokasy WF. Appleton-Century-Crofts; 1972:64-99.

35. Mackintosh NJ: **A theory of attention: variations in the associability of stimuli with reinforcement**. *Psychol Rev* 1975, **82**:276-298.

36. Pearce JM, Hall G: **A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli**. *Psychol Rev* 1980, **87**:532-552.

37. Wagner AR: **SOP: a model of automatic memory processing in animal behavior**. In *Information Processing in Animals: Memory Mechanisms*. Edited by Spear NE, Miller RR. Erlbaum; 1981:5-47.

38. Wagner AR, Brandon SE: **A componential theory of associative learning**. In *Contemporary Learning: Theory and Applications*. Edited by Mowrer RR, Klein SB. Erlbaum; 2001.

39. McLaren IPL, Mackintosh NJ: **Associative learning and elemental representation: II. Generalization and discrimination**. *Anim Learn Behav* 2002, **30**:177-200.

40. McLaren IPL, Mackintosh NJ: **An elemental model of associative learning: I. Latent inhibition and perceptual learning**. *Anim Learn Behav* 2000, **28**:211-246.

41. Esber GR, Haselgrove M: **Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning**. *Proc R Soc B* 2011, **278**:2553-2561.

42. Sutton RS, Barto AG: *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press; 1998.

43. • O'Doherty JP: **The problem with value**. *Neurosci Biobehav Rev* 2014, **43**:259-268.
A lucid critique of the literature on the neural representations of subjective value.

44. Dickinson A, Balleine BW: **Motivational control of goal-directed action**. *Anim Learn Behav* 1994, **22**:1-18.

45. Rescorla RA: **Pavlovian conditioning: it's not what you think it is**. *Am Psychol* 1988, **43**:151-160.

46. Holland PC, Straub JJ: **Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning**. *J Exp Psychol Anim Behav Process* 1979, **5**:65-78.

47. Colwill RM: **An associative analysis of instrumental learning**. *Curr Dir Psychol Sci* 1993, **2**:111-116.

48. 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.

49. Yin HH, Ostlund SB, Knowlton BJ, Balleine BW: **The role of the dorsomedial striatum in instrumental conditioning**. *Eur J Neurosci* 2005, **22**:513-523.

50. Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW: **The thalamo-striatal pathway and cholinergic control of goal-directed action: interlacing new and existing action-outcome associations in the striatum**. *Neuron* 2013, **79**:153-166.

51. 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.

52. Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW: **Neural representation of subjective value under risk and ambiguity**. *J Neurophysiol* 2010, **103**:1036-1047.

53. Plassmann H, O'Doherty J, Rangel A: **Orbitofrontal cortex encodes willingness to pay in everyday economic transactions**. *J Neurosci* 2007, **27**:9984-9988.

54. Rolls ET, Sienkiewicz ZJ, Yaxley S: **Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey**. *Eur J Neurosci* 1989, **1**:53-60.

55. Izquierdo A, Murray EA: **Combined unilateral lesions of the amygdala and orbital prefrontal cortex impair affective processing in rhesus monkeys**. *J Neurophysiol* 2004, **91**:2023-2039.

56. Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, Schoenbaum G: **Orbitofrontal cortex supports behavior and learning using inferred but not cached values**. *Science* 2012, **338**:953-956.

57. Rhodes SE, Murray EA: **Differential effects of amygdala, orbital prefrontal cortex, and prelimbic cortex lesions on goal-directed behavior in rhesus macaques**. *J Neurosci* 2013, **33**:3380-3389.

58. West EA, DesJardin JT, Gale K, Malkova L: **Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques**. *J Neurosci* 2011, **31**:15128-15135.

59. Gremel CM, Costa RM: **Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions**. *Nat Commun* 2013, **4**:2264.

60. Rudebeck PH, Murray EA: **Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior**. *J Neurosci* 2011, **31**:10569-10578.

61. Rescorla RA: **Pavlovian conditioning. It's not what you think it is**. *Am Psychol* 1988, **43**:151-160.