c00032

# Choice: Towards a Standard Back-pocket Model

*Paul W. Glimcher*

s0010

## INTRODUCTION

p0010    The goal of neuroeconomics is an algorithmic description of the human mechanism for choice. How far have we proceeded towards this goal? This volume reveals just how much information has been gathered. The studies presented here have leveraged existing scholarship to describe the mechanisms by which the values of actions are learned, how and where these values are encoded, how these valuations govern our actions, and how neural measurements can be used to constrain social scientific models of human behavior. With this information in hand, can we define the gross features of the human choice system in a way that will be of use to economists, psychologists, and neuroscientists? Or, to put it more precisely, can we use the lens of economic theory and experiment to better understand the neurobiological and psychological data in a way that will benefit all three disciplines? My suspicion is that many of the scholars contributing to this volume would say that the answer to this question is yes, and that the back-pocket models that most of these scholars use to guide their research are remarkably similar.

p0020    With that in mind, this chapter seeks a fairly precise definition of a standard back-pocket model of human

decision-making that incorporates the bulk of what we know today. Of course such a model, if taken seriously by policy-makers, could be dangerous. Many of the details of such a framework, even if all of those details were supported by contemporary data, would be both controversial and wrong. But a framework that organizes the information we have, and serves as a target for future challenges, may well maximize the forward movement of our discipline. Such a scaffold might make clearer what we do know and what we do not know; where we have made progress, and where critical avenues remain unexplored. In that spirit, and with the certain knowledge that the details of the following framework are wrong, what follows is a fairly formal presentation of a "standard back-pocket model" for choice.

p0030    My goal in presenting this model is to explicitly link neurobiological, psychological, and economic studies of choice so that we can examine the implications of this structure for all three of our parent disciplines. Simply relating a chooser's options to her choices or simply specifying patterns of brain connectivity and activation would run counter to the goals and spirit of neuroeconomics: What is called for is a hybrid approach that rigorously mixes the strategies and traditions of our fields so as to explicitly maximize the number of constraints these parent disciplines can impose on our understanding of choice. In interpreting this approach, scholars trained in only one of the parent disciplines may be initially troubled. For classical neurobiologists, this approach may seem to include an overly formal definition of conceptual objects to no particular end. For economists, the emphasis may seem overly algorithmic and unnecessarily focused on cardinality. My own feeling, however, is that the interaction of these constraints is what makes neuroeconomics powerful. As I hope will become clear in this chapter, the explicit ties to economics will allow neurobiologists to rule out whole classes of theories that have heretofore seemed reasonable. The explicit ties to neurobiological data will reveal that only a tiny space in the vast landscape of economic theory can be viewed as compatible with the human neuroarchitecture. In any case, I ask the reader's forbearance in this regard. I mean the presentation to be *neuroeconomic*. I hope the presentation will reveal two things: (1) that only a very specific set of economic theories is supported by the available data, and (2) that several very specific pieces of neurobiological and psychological data are required to complete the theoretical constraints on the architecture of human choice.

p0040    What follows, then, is a presentation in five parts. The first section provides a quick overview of the

basic mechanism for which I intend to argue. This section is not meant to be a defensible piece of evidence-based reasoning, but rather an opportunity to sketch out the shape of the coming argument. The second section provides a formal definition of the mathematical and empirical objects used in the rest of the presentation. For some this may seem superfluous, and for others it may seem *ad hoc* or overly restrictive. I hope it will become clear as the exposition develops that we require these particular objects to link existing economic theory to empirical psychological and neuroscientific data. The third section will provide a detailed description of the evidence for a generalized neural mechanism of valuation – a detailed description of what we do and do not know about this process, and a description of its surprisingly unitary and linear nature. The fourth section provides an overview of the choice mechanism itself; those circuits that take as their inputs the outputs of the valuation system and give as their output a plan of action – a choice. The chapter concludes by highlighting both the strengths and weaknesses of this standard back-pocket model.

## THE BASIC TWO-STAGE MODEL

Growing evidence suggests that the basic mechanism for producing choices in primates of all kinds (a group which necessarily includes humans) involves a two-stage mechanism. The first of these stages is concerned with the valuation of all goods and actions; the second is concerned with choosing amongst the goods or actions presented in a given choice set. At a very basic level, one can think of the valuation mechanism as being associated with learning and representing the values of objects and actions. The choice mechanism can be viewed performing a transformation that takes as an input the values of the options under current consideration – the choice set – and stochastically returns a high-valued option used to guide physical action. Of course, the details of these mechanisms are subtle. Some features of the valuations we infer from behavior (what an economist would call the *preference function*) seem to be attributable to mechanical processes embedded in the choice mechanism itself. For example, the model suggests that hyperbolic temporal discounting may arise from a mixture of exponential temporal discounting within the valuation system and a set of divisive computations embedded in the choice mechanism (see, for example, Glimcher *et al.*, 2007). The mapping between

the physical valuation mechanisms of the brain and psychological notions of valuation will occasionally be complicated. In a similar way, the anatomical boundaries between the choice and valuation mechanisms may not be entirely discrete. Our mixture of theoretical and empirical approaches will make it clear that neural activity in the *choice structures* both should and does influence activity in the valuation structures. For example, some neurons in key valuation areas like the striatum carry signals that encode choice. However, at a global level, it now seems extremely likely that the architecture is organized around this basic two-stage framework.

p0060    The neurobiological evidence for a two-stage process, which will be reviewed in greater detail below, arises from several key observations which are summarized only very briefly here. Perhaps the first explicit evidence for this segregation came from the work of Platt and Gimcher (1999). In the first half of that study, the authors recorded from neurons in the posterior parietal cortex while thirsty monkeys viewed two visual targets. In a typical trial, or round, the two targets might be associated with different magnitudes of reward, and after an initial delay the animal was informed which one of the two targets would yield that reward on this trial. From a choice-related point of view, each round was a decision between a response that yielded no reward and a response that yielded a small positively-valued reward. What Platt and Glimcher found, however, was that during the early part of the trial (before the zero-valued target was identified for that trial) these neurons produced firing rates almost linearly proportional to the average value of the rewards that had previously been earned for selecting that target. These firing rates cardinally encoded, in action potentials (or spikes) per second, the average value of the targets, but in a way that *did not* influence choice on that trial. It was "as if" the mean expected utility of the action *look at the right target* was linearly encoded by neuronal firing rates independent of choice. This is a point that will be developed in greater detail below, but the point I want to make here is that at the time this was seen as a major limitation of the study. In retrospect it provides some of the first compelling evidence that valuation and choice are dissociable.

p0070    At the same time, a huge number of studies from many different sub-areas of neuroscience began to suggest that broad swaths of the striatum and the frontal cortex both learn and represent the values of goods and actions even when learning is passive (Figure 32.1). Delgado *et al*. (2000) and Knutson *et al*. (2000), for example, found that when humans passively viewed events that resulted in unpredictable gains or losses, the level of neural activation in several striatal and fronto-cortical areas was linearly correlated with the magnitudes of these gains and losses. This constituted preliminary evidence that these value-encoding structures operate in the absence of choice[1]. Subsequent studies confirm this initial observation. During both choice and non-choice tasks, when humans face risky or certain gains, when they face delayed or immediate gains, the activation of discrete frontal and striatal nuclei is almost always near-linearly related to (subjective) measures of value. The linearity of this relationship may be surprising to economists, but this is not theory, it is simply an empirical (and surprisingly common neuroscientific) observation. These existing studies have made it clear that the neural systems for valuation are both neurochemically and anatomically localized.

p0080    The critical first step towards this realization was the identification of reinforcement learning mechanisms in the forebrain, and it is an understanding of these learning mechanisms that has paved the way towards a broader understanding of valuation. In the early 1990s, Wolfram Schultz and his colleagues (see, for example, Romo and Schulz, 1990; Schultz and Romo, 1990; Schultz, *et al*., 1993; also Chapter 21 in this volume) demonstrated that midbrain dopaminergic neurons encode a *reward-prediction error*. These are highly specialized and anatomically localized neurons that broadcast a signal throughout the striatum and the frontal cortex. Montague and colleagues (1997; see also Chapter 22 in this volume) provided the next step when they recognized that this class of signal could be used to construct a mechanism that learns, through trial-and-error, the values of actions or objects. What followed was 10 years of work which established the existence of at least three inter-related subsystems in these brain areas that employ distinct mechanisms for learning and representing value, and which interact to produce the valuations that guide choice (as summarized in Chapters 12 and 24). This provided a set of landmark findings summarized by the many chapters in Part 3 of this volume.

p0090    In a similar way, studies of the movement control systems of the brain revealed both the need for and existence of a discrete choice mechanism. Although it may not be obvious to non-neuroscientists, the process of producing a movement at the biomechanical level is extremely complicated. Checking one of two boxes on a retirement fund contract, or signaling

---

[1] Of course, this analysis presumes that experiencing rewards and anticipating them for the purposes of decision share a common neural substrate. That this is true is now largely beyond dispute at the neurobiological level – a point that is developed later in the chapter.
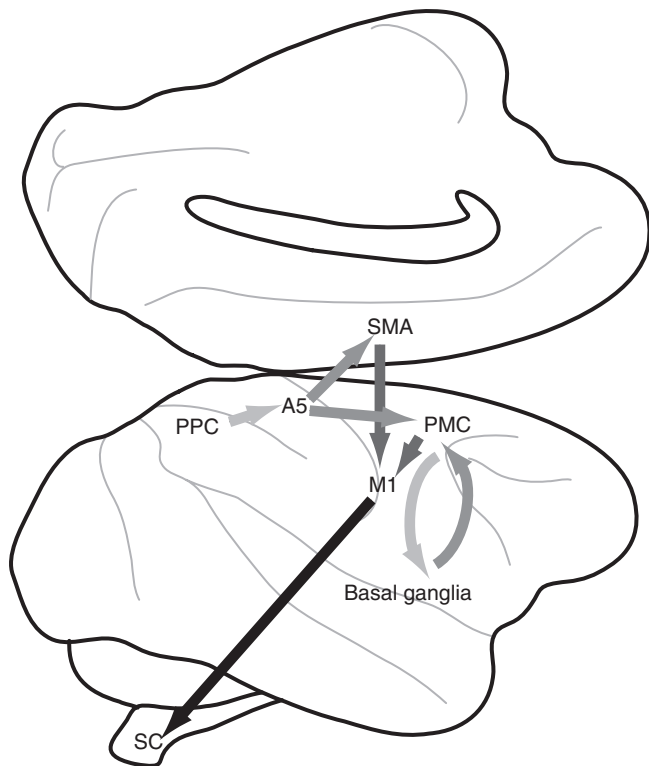
FIGURE 32.1    A highly simplified view of the skeleto-motor output system of the monkey brain. The key feature communicated by the figure is that behavioral outputs must, in principle, converge on a final common pathway for movement control. PPC, posterior parietal cortex; A5, Broadmann's Area 5; SMA, supplementary motor area; PMC, premotor cortex; M1, motor cortex.

a choice with a rightward eye movement, requires quite precise coordination of what may be literally dozens of independent muscles. Each of these movements, though, is a unitary object that must be selected and planned before it is executed. The hypothesis that option values influence muscles directly is easily falsified. When a subject moves her pencil towards a checkbox on a page, the tip of the pencil is moved by over 30 muscles and joints with more than 7 degrees of freedom. Still, the tip of that pencil traces a straight line from start to checkbox with a Gaussian velocity profile that peaks about halfway through the movement. Reaching for that checkbox reflects a movement richly planned before it is executed. The most introductory neuroscience textbook reveals this fact. Of course, we know where much of the neural hardware that plans and regulates movements is located. Areas like the motor cortex, the premotor cortex, the supplementary motor area, and Broadmann's area 5 all coordinate the generation of goal directed movements of the arms. A similar (and better understood) system coordinates movements of the eyes. What this means is that value signals must be turned into action

control signals somewhere within the nervous system, presumably at or above the level of motor cortex. It is this process that neurobiologists refer to as a *choice mechanism*. What I point out here is that it is critical to keep in mind that choice must be accomplished before movements are designed. The movement control systems reflect a *final common path* before which choice must, under normal circumstances, be complete.

Our current evidence suggests that the choice system involves large portions of the parietal cortex, amongst other areas (as summarized in Part 5 of this volume). These parietal areas receive both direct and indirect projections from the valuation areas, and project directly to the movement control areas. One issue that remains unclear, however, is how much the frontal cortex and basal ganglia participate directly in the choice process working in concert with these parietal areas. We now know that specific neurons in the orbitofrontal cortex (as reviewed here in Chapter 29) and the dorsal striatum (Samejima, *et al*., 2005; Lau and Glimcher, 2008) of the monkey also represent goods and actions that have been chosen before these choices are executed, but whether these neurons participate directly in choice is not known at this time.

This then, is a minimal working outline of the primate choice system: a valuation system that learns through repeated sampling of the environment and stores the values of actions and/or goods; a choice system that uses these values to select (from amongst a current choice set) a single option; and a motor control system that executes the physical responses dictated by the choice. Of course future experiments will enrich this description – for example, it may well be the case that perceptual systems influence the valuation systems in ways that we are just beginning to understand – but these seem to be the fundamental components of the primate architecture for choice as we understand it today.

Before beginning to examine the valuation and choice systems in detail, however, it is critical that we link these components to economic theory. This will provide both important constraints on how these systems operate and a common language for thinking about these problems. In what follows, definitions for conceptual objects explicitly linked to economic theory are presented. While these definitions may initially seem opaque to neurobiologists, neurobiological readers are urged to take them seriously. The role of these objects is to serve, in essence, as mapping rules that connect existing theoretical tools to the empirical measurements of neuroscience. At an earlier point in the history of neuroeconomics, it may have been valuable to speak in broadly metaphorical terms when saying things like "this neuronal firing rate is like a utility

signal[2]". But as this field transitions to formal tests of explicit and powerful theories, this kind of metaphorical relationship between theory and data becomes more and more untenable. If we are to leverage the precise and highly testable (if rarely tested) theories of economics and psychology, then we must be able to specify clearly *how* we would test those theories with neurobiological data – not just in a general way, but in a specific computational sense. The kinds of objects described below, I argue here, are what is required.

p0130
To many economists, a feature that will stand out about these objects is that they will seem unnecessarily restrictive. Linear relationships will be postulated which do not seem necessary. From a logical point of view, I agree. These, however, are the objects for which our empirical data calls. To many neurobiologists, the objects I have selected for definition may seem arbitrary. In the sections below, I hope to convince you that the empirical data argue that these are also called for by our current data.

## DEFINING OBJECTS

### Expected Utility Theory

p0220
As several chapters in this volume make clear, the axiomatic approach in general and expected utility theory in particular have both good and bad features (for an overview of its advantages, see Chapter 3; for an overview of its weaknesses see Chapter 11). Formally, the theory of expected utility (von Neumann and Morgenstern, 1944) rests on four axioms (or three, in Savage's 1954 formulation). For our purposes, I want to stress why these axioms are not some set of strange and arbitrary assumptions about how people *must* behave, which is an interpretation often given to them by critics. The axioms are a statement not about people (or the brain) in any sense; the axioms are a precise definition of a theory. It is reasonable to dislike any theory, but it is important to stress that, counter to what many lay people believe, this is a very minimalistic theory – much, much less restrictive and much more intuitive than, for example, temporal difference theories of learning. If what neurobiologists studying decision-making want is a simple theory of how people value things, then it is important for them to realize that economists already have several such theories, and that the implications of these theories have been very well explored. When a class of behavior obeys the axioms of a given economic theory, then we already

p0230

know quite a bit about the valuation systems that can, in principle, underlie that behavior. It is for this reason that neurobiologists need to link their measurements to economic frameworks. This is the only way that neurobiologists can rigorously exploit what economists have already learned about valuation.

p0150
To make this clear, consider expected utility theory, which will serve as my initial focus in the presentation that follows. Expected utility theory proposes that choosers should (1) show complete and transitive preferences, and (2) obey a choice separability constraint in a way that seems quite reasonable (amongst other things). Saying that a chooser obeys the axiom of complete and transitive preferences is simply saying formally that she could not be induced to:

1. pay *1 cent and an apple* for an orange, then     o0010
2. pay *1 cent and that orange* for a pear, **and**     o0020
3. pay *1 cent and that pear* for the original apple.     o0030

p0190
By the same token, saying that a chooser obeys the separability axiom (a variant of the more widely known independece axiom) is simply the assertion that she cannot:     s0030

1. prefer an apple to an orange, **and**     o0040
2. prefer 1 cent and an orange to 1 cent and an apple.     s0050/o0050

p0140
This is the reason that these axioms were included in expected utility theory. What is interesting and powerful about the theory, though, is that any chooser who obeys these rules (and the other axioms of the theory) behaves exactly as if she had a stable monotonically increasing utility function and as if her choice behavior was aimed at maximizing her net utility according to that function. Saying someone behaves according to these sensible rules is mathematically equivalent to saying that it looks as if she is trying to maximize some specific utility function. That is an important insight into valuation that neurobiologists cannot afford to ignore[3].

So what are these inferred utility functions like? A subject who behaves according to the axioms behaves "as if" she is maximizing some utility function, but how heavily does even an infinitely large dataset of choices constrain our understanding of this function? To understand the answer to this question, it is

---

[2] And in fairness, I have probably been more guilty of this kind of metaphor than anyone else – a looseness in language that I now regret.

[3] Let me stress here a point that may not be entirely obvious: "utility" really *is* "choice" when these axioms are obeyed. Utility is not "a feeling," or "happiness," or "a hedonic impulse." Utility is a common scale for valuation which gives rise to choice when choice obeys these axioms. If choice obeys these axioms, it is just *as if* a utility function gave rise to these choices. Or, put the other way, if you had a measured function having these properties, and it could perfectly predict choice then it *would be* a utility function and the choices would of necessity obey the axioms. Period.

necessary to understand three important features of utility that have not always been well enough appreciated by neurobiologists.

s0050 **1.   Utility is ordinal**

p0240      Observations of choice constrain the shape, but not the scaling, of the utility function. We might be able to show that a set of observed choices and an assumption that the chooser being studied obey the axioms of the theory are consistent with the idea (to take a simple example) that the utility of money to that chooser could be described as:

$$Utility = Dollars^{0.6}$$

p0250      With this equation, we can predict whether that chooser prefers a sure gain of $10 over a 50% chance of $22, and if the assumptions are correct for this chooser then our prediction is also correct. What is interesting to note, though, is that all of our predictions would also be correct if we had written her utility function as:

$$Utility = 50 \times Dollars^{0.6}$$

or

$$Utility = 1000 + Dollars^{0.6}$$

s0060

p0270      This means that there are multiple equivalent representations of this subject's utility function. We can predict choices by using any of these equivalent utility representations, but the one we employ in a given set of calculations is arbitrary. It is for this reason that economists refer to utility as an *ordinal scale* rather than as a discrete cardinal scale. To make the importance of this insight clear, consider a chooser who prefers apples over oranges and oranges over pears. If we assume the axioms of expected utility theory for this person, we can say that this behavior ranks the utility of the three objects. We can even arbitrarily assign apples a utility of 1 and pears a utility of 0 (for this chooser). Next, we could use lotteries (which of the following do you prefer: *a 50% chance of an apple and a 50% change of a pear* **or** *an orange*?) to place oranges on

s0070      this same scale, for example at a utility of 0.3. But consider what happens when we suddenly introduce kiwi

p0280      fruits to the chooser and it turns out that she prefers kiwis to apples. Then the entire scale must be regenerated. This *is* what is meant when an economist writes the word utility and it is no problem mathematically, but it points up an important feature. Utility functions are not cardinal sets of numbers that have definite values that can be added and subtracted. They are ordinally arranged relations between choice objects, and

this places important mathematical limits on what you can (and cannot) do with utility functions as objects. The theory of value imposed by expected utility does not include predictions about the cardinal relations of utilities, *by design*. For a neurobiologist, this imposes a particularly burdensome constraint. It implies that is meaningless to say that a neuron's firing rate *is the representation of* utility because such a statement lies outside the domain of expected utility theory. Neurons yield to us a fully cardinal measurement when we observe their firing rates. Firing rates are numbers that can be added and subtracted in a way that utilities cannot. If a neuron had a firing rate that revealed the desirability of an apple irrespective of the other objects placed before that chooser (as in the study of Padoa-Schioppa and Assad, 2008), that neuron could *not* be said to encode the utility of apples, because one feature of this powerful theory is that the utility of apples is not a unique number[4]. Of course that firing rate could be linearly proportional to utility. If we increase the number of apples presented to the subject until the firing rate doubles we might be able to conclude that utility has doubled, but utility and firing rate would remain distinct in this important way.

s    **2.   The Axioms of Utility Theory are not Always Consistent with Choice**

p0260      Humans do, on occasion, both prefer an apple to an orange and also prefer 1 cent and an orange to 1 cent and an apple. This is (speaking a bit imprecisely) what the Allais (1954) paradox shows (see Chapters 1 and 11 for more on this paradox). Of course, this has implications for economics, but it also has a huge implication for neuroscientists. If a neuron had a firing rate that was always linearly proportional to utility, then the firing rate of that neuron *could not be used to always predict real human choice*. A neuron with a firing rate proportional to utility would – by definition – obey all of the axioms of expected utility. It could not generate the Allais (1953) paradox, because that is the nature of what is meant by *utility*.

s    **3.   Utility Implies Agent Welfare/Agent Well-being**

      One of the most important functions of economics is to tell us whether a change in policy or government will make people better off. Economists have often argued that expected utility theory helps them make

---

[4]It is important to point out that this is not a limitation only of expected utility theory; it is a feature of almost all economic theories of value.

this determination. If, for example, all citizens were to obey the axioms of expected utility theory, then governments would have an easy time keeping them happy or, more formally, maximizing their welfare. Since a chooser who obeys expected utility always acts to maximize her own utility, then we can maximize her welfare by allowing her the freedom to choose whatever she desires. This is a common (though not necessarily ubiquitous) approach to figuring out how to design policies that maximize the well-being (or technically the *welfare*) of individuals.

It is important that neuroeconomists be aware of this issue, because if a neurobiologist was to argue that the firing rate of a neuron was linearly proportional to utility he might be heard as saying that maximizing the firing rates of those neurons in citizens would be maximizing their welfare, even if the firing rate of that neuron *could not be used to predict real human choice*. I think that neuroscientists need to very carefully avoid making such a claim for the foreseeable future. In almost all of the neuroeconomics studied to date, we have tried to link activity in the nervous system to choice behavior. This volume shows how much we know about the neural circuits that give rise to choice. By contrast, we know very little today about the neural circuits that give rise to an individual's sense of well-being. If we did understand those neural circuits, then we might be able to make some claims related to welfare issues in economics. In any case, we do not have such expertise at this time, and I want to take care to emphasize that it is the neural mechanism of choice and not the neural mechanism for experiencing well-being that this standard back-pocket model attempts to describe.

## Defining Subjective Value (SV)

The task before us is to ask, how can we relate neuronal firing rates, or measurements of the BOLD signal, to the valuations of actions and objects that we believe guide behavior? One way to proceed is to try to relate these activation patterns to expected utility. Under many conditions expected utility theory does predict choice, and that seems to be an observation that we do not want to overlook. On the other hand, one of the reasons that we want to develop an algorithmic model of decision-making is that we have every reason to believe that such an algorithmic model would predict choice behavior even when expected utility theory cannot. So how do we gain access to the theoretical power of expected utility theory without becoming burdened with its failures, and in a way that respects the two-stage model for valuation and choice that is developing

today? One has to note here that for many economists this is a critical point – and one about which there has been much confusion. To resolve this confusion, I suggest the following definition:

**Subjective value**: Subjective values, at least for the purposes of this initial argument, are real numbers ranging from 0 to 1000. They take as their natural units action potentials per second. Subjective values have the following properties:

1. Subjective values are equal to (or better yet defined as) the mean firing rates of specific populations of neurons, the identification of which follows. For this reason, subjective values are linearly proportional to the BOLD[5] signal as measured in these same populations.
2. Subjective values predict choice stochastically. More formally, I define them as the sum of true subjective value and a noise term (see below). This means that subjective value theory will be most closely allied with random utility-type models from economics.
3. When expected utilities predict choice behavior, subjective values are linearly proportional to those expected utilities.
4. Subjective values are *always* consistent with choice, though stochastically, even when choice is not consistent with expected utility theory.
5. Subjective values have a unique *reference-dependent* anchoring point called the baseline firing rate. All subjective values are encoded cardinally in firing rates relative to this baseline. This means that subjective value theory will be most closely allied with reference dependent forms of utility-type models from economics.

Of course I recognize that some of these properties will have to be relaxed, but probably not in important ways. The BOLD signal and mean firing rates, for example, are not exactly linear in their relation, but these five statements capture the central features of subjective value around which our definition of the choice architecture will be organized.

First and foremost, the definition I suggest here allows us to be clear about why expected utility theory will be enormously valuable to the neuroeconomic enterprise. Expected utility theory provides a compact definition that, under at least some circumstances, describes patterns of choices. *Where that is true*, measurements of utilities tell us unambiguously what SV *must* look like to within a linear transformation. (This, of course, assumes that we can find a mean neuronal firing rate that looks like a candidate for encoding SV, but we turn to that in the next section. For now, we simply seek clear definitions of our goals.) Second, this definition says that if we could actually measure SV, we would be able to use those measurements to predict behavioral violations of expected utility theory, like the Allais paradox, as well as human

---

[5]The Blood Oxygen Level Dependent Signal of functional Magnetic Resonance Imaging (fMRI).

choice behavior well captured by other approaches like prospect theory. Third, SV must be subject-specific. This follows, of course, from its relation to the utilities of expected utility theory.

SV is defined in units of average spikes per second as an object that predicts the choices of individuals. Finally, I want to be clear that measurements of SV *do not* have clear welfare implications. Because SV does not (at the very least) obey the independence axiom globally (since human choice does not obey this axiom), maximizing SV will not yield a maximization of something like a complete and transitive preference function. Further, and probably more importantly, SV predicts choice. Because we are modeling at an algorithmic level, this does not mean, *ex ante*, that SV is related to a chooser's sense of well-being. That may be mediated by other neural systems. Welfare maximization and SV maximization should not be equated.

To summarize, I define here the concept of subjective value which is meant to be a fully cardinal object with several important restrictions. At least initially, it cannot take negative values (an important point to which we will return). It has both a finite range and finite (and large) variance. The importance of this point will be immediately clear to neurobiologists. For economists, it means that errors and stochasticity in choice are unavoidable features of the architecture. This suggests properties related both to random utility models and to stochastic errors inchoice. The importance of this point will be immediately clear to economists. For neurobiologists, it means that whenever choice behavior obeys the axioms of random utility models we know a tremendous amount about how a final common valuation system ought to be behaving. Following this line of reasoning, then, my hypothesis is that SV is encoded directly in the valuation mechanisms of the human brain and that existing economic theory tells us much about how this representation must behave. As we make measurements to prove this, we will be able to place additional important restrictions on SV.

Finally, I need to make it clear that what I am suggesting is that one central goal of neuroeconomics should be to develop a complete theory of SV. As that theory is enriched, it will continue to refine our understanding of which economic theories are better than others at predicting SV. Random utility models, for example, will be shown below to be better predictors of SV than traditional utility models. Reference-dependent utility models will also be shown to be better predictors of SV than traditional consumption utility models. Whether traditional economists will care that empirical constraints on SV can be used to identify some utility-based models as closer fits to the human choice architecture will, of course, be a matter of taste.

## Relative Subjective Value (RSV)

We define the relative subjective value of a single option *j* as:

$$RSV_j = \frac{SV_j}{\sum SV_i + c}$$

where *i* is the set of all options in a choice set (including *j*) and c is an empirically measurable normalization constant of the type first described in the cerebral cortex by Heeger (1992). Our current evidence suggests that choices are actually made between options by comparing RSVs after corruption by noise. The evidence for this arises from work in parietal cortex, which is summarized in Chapters 4, 28, 29, and 31.

## Obtained Subjective Value (ExperSV)

ObtainedSV is a pattern of neuronal firing, much like SV, which encodes the subjective value of current states of the world[6]. The neural location of ExperSV is not known, though the activity of dopamine neurons provides overwhelming evidence that it is present as one of the midbrain inputs to those neurons. For reasons that will be described below, Obtained Subjective Value actually serves as one source of the utility-like properties of SV.

## Reward Prediction Error (RPE)

RPE is defined here as in learning studies and as summarized in Chapter 22. It is:

$$RPE = \alpha(SV_{forecast} - ExperSV)$$

## Stochastic Terms

The existing neural data suggest two sources of stochasticity that influence choice: one at the level of the valuation system and one at the level of the choice system. In economic terms, the first can be viewed

---

[6] Let me stress again that ObtainedSV is not necessarily a welfare measurement.

as roughly corresponding to random utility distributions[7] (McFadden, 1974; Gul and Pesendorfer, 2008) and the second as corresponding to the Trembling Hand[8] (Selten, 1975) notion of stochastic behavior. There is compelling evidence for both such sources.

p0450    *Subjective value variability* is a random term drawn from distribution assumed to be Gaussian and added to mean *subjective value* to yield SV. It is always present. The variance of this term may or may not be adjustable.

p0460    *Cortical noise before choice* is a final noise source added (as a stochastic time series) to RSV before choice occurs. The source of this term is noise intrinsic to cortical neurons, which requires that it be Poisson in distribution at the mechanistic point of addition (see, for example, Tolhurst *et al.*, 1983). Neuronal pooling that occurs during the choice process, and adjustability of the inter-neuronal correlation term, may be used to reduce this variance (Krug *et al.*, 2004). For more on the theoretical implications of this, see Glimcher (2005).

### Valuation Mechanisms and Subjective Value

p0510    Formally (and of course too simplistically), subjective value (and terms that inherit properties from subjective value) can be seen as a neuronal sum of the form:

$$SV_j = \frac{\sum_i \omega_i x_{ij}}{\sum_i \omega_i}$$

p0520    where the term *i* indexes each of the neurons in the brain, $x_i$ is the firing rate of the *i*th neuron, and $\omega_i$ is a weight ranging from 0 to 1 describing the additive contribution of that neuron to the SV of object or action *j*. This object places into the language of economics the standard neurobiological insight that a weighted sum of neurons in topographic maps encodes behaviorally relevant variables (for more details on this neurobiological issue, see Lee *et al.*, 1988.) The subjective value of a particular object in

the external world, *j*, is thus simply represented as the average weighted firing rate of a subpopulation of neurons that encodes the subjective value of that object[9]. In a topographically mapped action-encoding region like the superior colliculus, this is equivalent to saying that activity in a restricted region of the map encodes value for a particular action. I should note, however, that this definition specifically excludes distributed non-linear encoding schemes[10]

p0480    For an empirical neurophysiologist or functional magnetic resonance imager looking for $SV_j$ in the brain, two questions then become paramount:

1. Is there a firing rate pattern (or a BOLD activation in the case of fMRI) we can identify in the brain that is linearly correlated with the utility of actions or objects (when utility predicts choice)?
2. What is the most compact population of neurons (both in number of neurons and in anatomical extent of the population) that can maintain this linear correlation with $SV_j$ (i.e. the smallest population of neurons for which $\omega_i$ is not equal to zero).

The data we have available today suggest that two brain areas seem likely to contain all the neurons we require to extract SV for any object: the ventral striatum and the medial prefrontal cortex.

## THE BASIC STRUCTURE OF THE VALUATION SYSTEM

If one accepts that (1) mammals evolved to learn the values of different states of the world both when their actions influenced those states and when they did not, and (2) that choices (the selection amongst available options) must be complete in the nervous system before actions can be planned and executed, then one can hypothesize that valuation must be at least partially autonomous of choice and that the process of choice must be complete before action is produced. These hypotheses seem to have been validated by a wealth of empirical research in neuroscience conducted over the past couple of decades. Indeed, amongst neurobiologists there is essentially universal agreement that a group of neural systems for valuation has been identified.

---

[7] Note that random utility theories form a class of economic models in which it is assumed that the subjective valuations of options vary stochastically, but that choosers always select the current best option from that stochastically varying set. In these models, it is the perception of value itself that is hypothesized to vary. Perhaps surprisingly, these models place some very interesting constraints on the relationship between value representations and choice.

[8] In contrast, trembling hand models propose that stochasticity in choice arises from errors during the choice process which lead to the selection of suboptimal elements from the choice set. These models place other interesting constraints on choice. One interesting signature of models of this type is a dependency of errors on choice set construction.

---

[9] This of necessity excludes non-linear interactions like those encountered in a game theoretic specification of SV. While it is not necessarily my intent to exclude these other kinds of interactions, the available data suggest that SV actually is linear with these firing rates. In any case, this definition could be relaxed.

[10] A constraint that could, at a later date, also be relaxed.

p0530    Almost certainly, the critical historical event that pointed towards a common neural system for valuation was the study of learning and dopamine, a topic reviewed in detail in Part 3 of this volume. Understanding why the study of dopamine led to postulates about valuation, however, requires an examination of the history of dopamine. In the 1920s, the German physiologist Otto Loewi established that neurons communicated with each other through a chemical mechanism we now call neurotransmission. His groundbreaking work established the existence, however, of only one neurotransmitter through which it was believed all neurons communicated. The existence of multiple neurotransmissive systems was revealed in 1964, when Dahlström and Fuxe (1964) visualized adrenaline-, noradrenaline-, and dopamine-containing neurons. These measurements revealed a set of anatomically and neurochemically discrete brain systems, a structural feature of the nervous system that had not been previously identified. Of particular interest, for our purposes, was the discovery of two to three groups of dopamine containing cell bodies that projected from the midbrain (from two areas called the *substantia nigra pars compacta*, *SNc*, and the *ventral tegmental area*, *VTA*) to the basal ganglia and the frontal cortex (Figure 32.2). By the 1970s and 1980s it had become clear that many drugs of abuse acted through this system, suggesting a role for dopamine in hedonic experience.

## Learning Subjective Values

    The critical breakthrough that allowed modern studies of valuation to crystallize around the midbrain dopaminergic pathways, however, was the work of Schultz and colleagues (1993). These authors measured the spiking activity of single dopamine neurons while monkeys passively received rewards during a classical conditioning task (see Chapter 22 for more details). They found that *unconditioned* rewards produced a strong response in these neurons, while conditioned rewards did not. This was an important finding, because it revealed that the activity of dopamine neurons *could not* simply code hedonic experience. This led Montague and colleagues (1997) to propose that dopamine neurons encoded the difference between expected and obtained rewards; the *reward-prediction error* of learning theory. What followed were a host of papers that established that dopamine firing rates could be described as:

$$DA(spks/s) = \alpha(ExpectedR - ExperiencedR)$$

where DA is the instantaneous firing rate of VTA and SNc dopamine neurons, *ExpectedR* is the magnitude
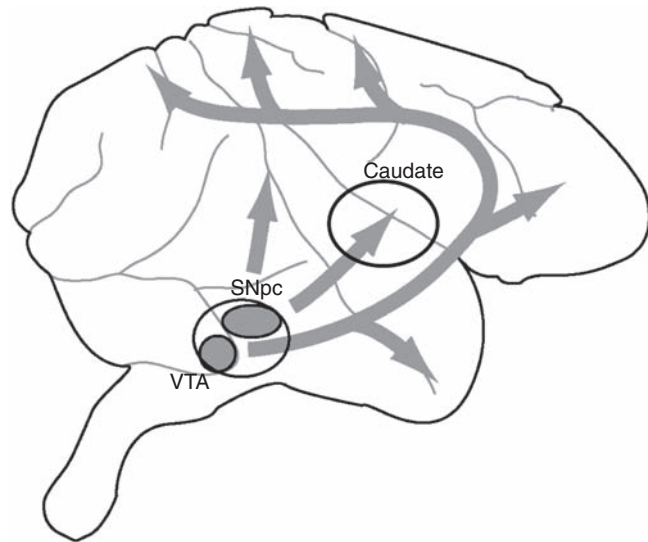
FIGURE 32.2 The principal dopaminergic pathways of the midbrain. SNpc, substantia nigra pars compacta; VTA, ventral tegmental area.

of the reward expected by the subject at this time, *ExperiencedR* is the magnitude of the reward being experienced by the subject at this time, and $\alpha$ is a scaling parameter that controls the subject's learning rate as described below and in Chapter 22. If ExperiencedR is reward value in an arbitrary currency and some other computational element simply recomputes after every expected or unexpected reward

$$Q_{j,t} = Q_{j,t-1} + DA$$

where $j$ indexes actions or goods and $t$ indexes time, then $Q$ represents a current estimate of the expected value of action or good $j$. For economists, I note this is a recursive form of the Bellman equation computing a reversed discount function in which $\alpha$ describes a *forgetting rate*.

    A very interesting feature of this system, which is, however, often overlooked, is the units in which *ExperiencedR* encodes the magnitudes of rewards being received. Consider a primary reward like water. If *ExperiencedR* encodes water in milliliters, then $Q_j$ converges towards a representation of expected value in milliliters. If, however, *ExperiencedR* encodes something like the utility of water, or more formally it encodes *ObtainedSV*, then the system converges not towards a representation of expected value but towards a representation of SV. This point is critical because neural systems of valuation must estimate real-world values through the sensory responses of neural transducers, and any upstream transformation

of information by these transducers is propagated through the system – thus, if volume encoding afferents from the tongue, like all sensory afferents that have ever been studied, encode a Stevens Power Law (Stevens, 1970) compressed representation of volume, then the set of $Q_j$s computed from dopaminergic activity would encode (or inherit) a power function representation of magnitude. It seems almost certain that this has to be the case, given that every sensory system ever studied employs a power law for compression (see, for example, Stevens, 1970). This leads to speculation that the encoding mechanism for primary rewards serves as at least one (if not the) source for the curvature of the utility function amongst primary rewards. It is also a possibility that this may offer some insight into the sources of reference dependence in human choice behavior.

p0560　For these reasons, it is tempting to speculate that dopamine neurons should receive as an input *ObtainedSV* from which their target neurons learn SV directly, thus accounting for the source of curvature in the utility functions for primary rewards with the neural hardware for sensory encoding. Of course this places some interesting constraints on how the DA neurons should work. They cannot, for example, code RSV (relative subjective value) with regard to a finite choice set, because if they did then SV could not be computed from them. The pre-existing body of economic theory makes this clear. If they did code RSV, the stored SV of a good or action would be dependent on the choice set within which it was learned. As a result, the choice mechanism would be unable to obey the axioms of complete and transitive preference which both humans and animals often obey.

p0570　This is an observation, however, that may seem at first blush to contradict data in the literature (Tobler *et al.*, 2005), and this apparent contradiction is important because it highlights the power of economic theory in neuroscience. These data suggest that the RPE signal measured in the dopamine neurons is variance dependent. As the variance of the reward stream increases the magnitude of the dopamine firing rate for a given *ExpectedR − ExperiencedR* goes down. Theory tells us, however, of the importance of complete and transitive preferences and what they imply for valuation mechanisms, and seems to suggest that *ExpectedR − ExperiencedR* cannot scale with the variance of the choice set and still preserve transitivity in the stored representation of value. The resolution of this apparent paradox is that it must be the learning rate, and not *ExpectedR − ExperiencedR* itself, which scales with variance. In other words, we can say that if subjects are transitive behaviorally, we can reject the hypothesis that *ExpectedR − ExperiencedR* scales with

variance[11]. Our ability to make this statement comes from the explicit linkage of theory and measurement. Of course, this also raises the possibility that violations of complete and transitive preference – when these violations do occur – may reflect features of this measurable set of computations.

p0580　In any case, the dopamine neurons broadcast this signal throughout the frontal cortex and basal ganglia – also suggesting a role for all of these areas in learning. The observation that the firing rates of dopamine neurons encode the difference between expected and obtained reward is critical, because it reveals that the inputs to the dopamine neurons include a signal both of both the value of the reward that was received and the value of the reward that was expected – clear evidence that a valuation signal of some kind, an object we have defined as *ObtainedSV*, must exist[12].

p0590　Though most of the work described above was conducted in animals, there is clear evidence that these dopaminergic neurons behave in the same manner in humans as they do in all other living mammals. Like other mammals, humans find dopaminergic drugs reinforcing. Like other mammals, humans have these same dopaminergic pathways. Like other mammals, dopaminergic drugs can be shown to bind to receptors in the terminal fields of these neurons. But the best evidence for the notion that a circumscribed learning-based valuation system associated with dopamine occurs in humans comes from fMRI studies of humans engaged in learning about rewards. In 2002, two groups (O'Doherty *et al.*, 2002; Pagnoni *et al.*, 2002) demonstrated simultaneously that activity in the dopaminergic terminal fields of the striatum and the frontal cortex during both gustatory and

---

[11] In fact, this scaling of the learning rate with variance is a feature of efficient learning systems – a fact well described in any mathematical treatment of the Kalman filter.

[12] To be more precise, there is now some compelling evidence that dopamine firing rates encode only positively valued, or near positively valued, reward-prediction errors (Bayer and Glimcher, 2005). There have been hints of this in the literature for some time (Hollerman and Schultz, 1998). The suggestion here is that positive and negative RPEs may be encoded separately in the nervous system (Daw *et al.*, 2002). The idea of splitting the RPE term into negative and positive elements should be naturally attractive to behavioral economists. We have known since the work of Kahneman and Tversky (see, for example, Kahneman and Tversky, 1979; also Chapter 11 of this volume) that human choosers are more sensitive to losses than to gains – a feature known as loss aversion. If positive and negative RPEs are coded by different systems and those systems map positive and negative values of *ExpectedR − ExperiencedR* to firing rate with different gain terms, then the ratio of these two independent gain terms could well account for some features of loss aversion.

monetary reward tasks behaved exactly as predicted. This indicated, basically beyond a reasonable doubt, that there existed a valuation-learning system in the striatum and frontal cortex of humans. So to summarize, this leads me to suggest that dopamine neurons lead to the direct computation of SV under some conditions:

$$SV_{jt} = SV_{j(t-1)} + \alpha(SV_{j(t-1)} - ExperSV)$$

where $SV_{jt}$ is the subjective value of object or good $j$, which is learned from repeated experience, as estimated at time $t$. Note that, as mentioned above, $ExperSV$ for primary rewards is a compressive function of $ExperiencedR$, as is really required by what we know of sensory encoding systems. This means that risk aversion, at least over primary rewards, is the product of Weber-type encoding mechanisms in our sensory apparatus.

What remains, then, is to understand where and how SV is mechanistically computed and stored. Two lines of evidence contribute to our understanding of these issues: neuronal recording studies in animals

p0630

and fMRI studies in humans. The recording studies in animals have now established that the basal ganglia contain essentially all of the computational elements required for the execution of reinforcement learning (or, more precisely, *temporal difference* learning) algorithms. There are, for example, neurons within the basal ganglia that encode the magnitude of reward that an animal expects to receive for producing a particular behavioral action, neurons that encode the

p0640

actions that have just been executed, and neurons with firing rates dependent on the current state of the environment, amongst other things. These neurons are located in the striatum and project out of the basal ganglia largely through the ventrolateral nucleus of the thalamus, which in turn projects back to the frontal cortex. Single unit recording studies in the frontal cortex have also demonstrated the existence of neurons that encode values, but this time the values of goods, not of actions (see Chapter 29). fMRI studies in humans tell a similar story (see Chapters 23 and 24 for more details), suggesting that frontal and basal ganglia circuits form the core of the human mechanism for RPE-based value learning.

There is, however, evidence for other learning mechanisms in these same structures which interact with this well studied RPE-style learning mechanism. The details of these other learning systems are still being worked out, but what is known to date is described in Chapters 12 and 24 of this volume. In essence, these studies suggest that a set of mechanisms,

most if not all interacting with dopamine, provide tools for learning and representing value in the frontal cortex and the basal ganglia.

For a neuroeconomist, then, these studies constitute overwhelming evidence that a value system exists and can be functionally localized. Where, then, is the final point of convergence at which SVs are passed to the choice system? Put more formally, in a preceding section I argued that subjective value can be seen as a neuronal sum of the form

p0620

$$SV_j = \frac{\sum_i \omega_i x_{ij}}{\sum_i \omega_i}$$

where the term $i$ indexes each of the neurons in the brain, $x_i$ is the firing rate of the $i$th neuron, and $\omega_i$ is a weight ranging from 0 to 1 describing the additive contribution of that neuron to the SV of object or action $j$. The question we need to answer is whether there is an anatomically discrete neuronal population that can supply all the non-zero values for $\omega$ required by the choice system.

p0600

One way to begin to answer this question is to look at the existing fMRI data and ask: is there a small number of areas that are actively correlated with SV under essentially all reward and choice conditions that have ever been studied? Perhaps surprisingly, the answer to this question seems to be yes: the ventral striatum and the medial prefrontal cortex show up in dozens of studies under essentially all choice conditions as coding something like SV.

Activity in the ventral striatum has been shown to be correlated with both rewards and punishments (Delgado *et al.*, 2000), the magnitude of cumulative rewards (Elliot *et al.*, 2000), the anticipation of reward (Knutson, 2000; Knutson *et al.*, 2003), the expectation of monetary reward (Breiter *et al.*, 2001), the expectation of primary rewards (O'Doherty *et al.*, 2002), the receipt of monetary rewards (Elliott *et al.*, 2003), monetary expected values (Knutson *et al.*, 2005), behavioral preference rankings amongst rewards (O'Doherty *et al.*, 2006), potential gain magnitude and loss magnitude as scaled by subject-specific levels of loss aversion (Tom *et al.*, 2007), and discounted reward value at delays ranging from minutes to 6 months (Kable and Glimcher, 2007). Single unit recording studies of the

p0610

dorsal striata of monkeys, both in the caudate (Lau and Glimcher, 2006) and in the putamen (Samejima, *et al.*, 2005), tell a similar story. Neurons in these areas have been identified which code action values. All of these data suggest that whenever rewards are received or preferences are expressed, activity in the ventral

striatum encodes the magnitudes of those rewards or preferences[13].

A similar story seems to hold in the medial prefrontal cortex. Activity in this area has been shown to be correlated with monetary reward magnitude (Knutson *et al*., 2000, 2003), preference ordering amongst primary rewards (McClure *et al*., 2004a), the expected value of a lottery (Knutson, *et al*., 2005), the subject-specific valuation of gains and losses (Tom *et al*., 2007), the subject-specific discounted reward value (Kable and Glimcher, 2007), and willingness to pay (Plassman *et al*., 2007). Activity in this area appears to be correlated with valuation under all of these conditions.

This leads me to propose that mean activity in the medial prefrontal cortex and the ventral striatum encodes SV. Different neuronal subpopulations in these areas must encode different options, and so these areas must employ a complex topographic encoding scheme which segregates the populations that encode the SVs of different actions or goods. The details of this encoding scheme, which probably lies beneath the resolution of fMRI, are only just beginning to be understood, and the encoding schemes employed by the two areas are almost certainly different. The medial prefrontal cortex, because of its closer relationship to goods-related encoding areas like the orbitofrontal cortex, may well encode SV in terms of goods, while the VS may employ an action-based encoding scheme. But in any case, I propose that these two areas serve as the final common representation of SV for use by the choice mechanism.

To be quite precise, I propose that the mean activity in subpopulations of the medial prefrontal cortex and the ventral striatum encodes SV when options are under consideration for choice or the objects of current learning. It is this activity which, I argue, both guides choice and encodes the reward prediction (SV) that is used in learning (probably as resident in the ventral striatum). This information, I suggest, is stored throughout a much larger network of areas spanning the frontal cortex and the basal ganglia in the synaptic strengths connecting neurons in these areas, the strengths of these synapses being set by the well understood biophysical mechanisms of dopamine-dependent long-term potentiation and long-term depression (and perhaps based on the actions of other plasticity generating neurotransmitters like serotonin). When instantaneous subjective value is represented, it reflects, I propose, the sum of activity passing through these synapses located in areas including the inferior frontal sulcus, the insula, the amygdala, the posterior cingulate, the superior temporal sulcus, the caudate, the putamen, and the dorsolateral prefrontal cortex, and impinging on the ventral striatum and the medial prefrontal cortex.

What we know about the biophysics that would be required to instantiate this process have two important implications that need to be mentioned. First, recall that all neurons have a limited dynamic range and a significant finite level of stochasticity in firing rate. This means that instantaneous SV is necessarily drawn at each instant from an underlying distribution. This therefore requires that the notion of SV be closely related to random utility models (and not traditional von Neumann-Morgenstern utility) from economics. Second, it needs to be noted that all neurons have a "baseline" firing rate, and neurons in these areas are no exception. Recently, Tom and colleagues (2007) have shown that activation in these areas continuously represents gains and losses on a common scale with an inflection point at a zero-gain point in those experiments. This suggests that baseline spike rate in these populations is the unique representation of the reference point for SV's reference-dependent encoding of value. Of course this conclusion, even if correct, does not constitute a theory of the reference point; it simply identifies an empirical technique for direct measurement of the reference point.

## CHOICE

Unlike valuation, which has been extensively studied in both humans and other animals, choice has been the subject of study principally in awake-behaving monkeys in neuroscience. This may reflect the fact that the temporal dynamics of choice make it difficult to study with fMRI. In any case, an understanding of choice requires an understanding of existing work in non-human primates.

Initial studies of choice in monkeys evolved almost simultaneously from studies of sensory-perceptual systems (see, for example, Newsome *et al*., 1989) and movement control studies (e.g. Glimcher and Sparks, 1992). The most important of these studies examined how monkeys used noisy visual-sensory signals to identify one of two orienting eye movements, or saccades, as reinforced. They did this by leveraging an extensive pre-existing literature on the structure of the visual and eye-movement systems to search for the decision-making circuits which connected them in

---

[13] However, one possibility that needs to be ruled out is that activity in the ventral striatum encodes only reward-prediction errors and not SV *per se*. Available single unit data rule this out in all other areas of the striatum, but the definitive study has not yet been conducted in the ventral striatum.

these tasks (details of this line of study can be found in Chapters 4, 28, 29, and 31 of this volume). Subsequent work has generalized many, but not all, of these findings to arm-movement control systems and to studies of humans.

p0710    We have to begin, therefore, with a review of the basic structure of the saccadic control system (Figure 32.3). The lateral intraparietal area (LIP) in the posterior parietal cortex is one of the critical elements in this system, and it consists of a roughly topographic map both of objects in the visual world and the eye movements that would required to align gaze with those objects (for a review, see Glimcher, 2003). Thus, a particular location on the map (or more precisely the neurons on the map at that location) activates when a visual stimulus appears 10° to the right of fixation, *and* that region might become particularly active milliseconds before an eye movement which shifts gaze 10° to the right. This area, in turn, projects both to the frontal eye-fields and the midbrain superior colliculus, two additional topographic maps that are broadly similar in function. The frontal eye-fields project, as well, to the superior colliculus directly. A final note is that many of these areas are reciprocally connected (for a review of this anatomy see Platt *et al.*, 2003), a fact which is probably important for understanding choice.

p0730    Finally, the colliculus is connected to brainstem circuits that actually govern eye movements in real time. The connection between these brainstem systems and the colliculus are mediated by a class of collicular neurons called "burst" neurons. Burst neurons have the interesting biophysical property that they can fire action potentials in either of two states: a continuous low-frequency state in which many different firing rates are observed, and a burst state characterized by a fixed and extremely high firing rate.

It is widely assumed that actual generation of a movement involves driving the collicular burst neurons above a specific firing-rate threshold, after which a burst occurs that is self-perpetuating and persists until the movement is complete. Inhibitory interconnections in the collicular map seem to preclude burst-like activity occurring at more than one location at a time, suggesting that the collicular architecture allows only a single movement to be executed

p0740    at a time. Studies in area LIP, the frontal eye-fields, and the superior colliculus all indicate that low-frequency firing in all three is related to the probability that a movement will be executed by the animal. To be more specific, if a particular movement is likely to yield a reward, then activity in all three maps at the locations associated with that movement is elevated. Of these three maps, the one that has been most studied with regard to decision is LIP. In LIP, it has been
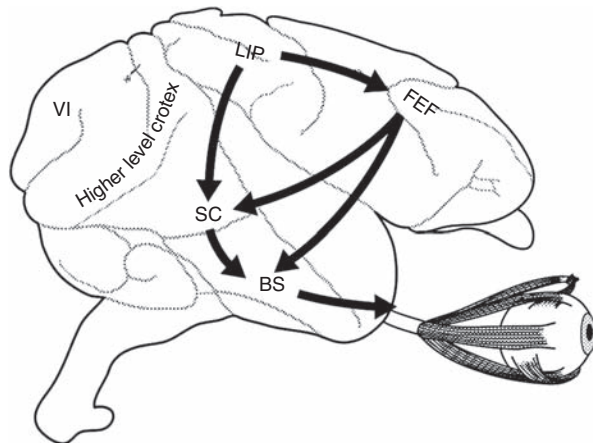


FIGURE 32.3    The saccadic control system of the rhesus monkey in which most studies of the choice mechanism have been studied. V1, primary visual cortex; LIP, lateral intraparietal area; FEF, frontal eye fields; SC, superior colliculus; BS, brainstem eye movement control circuits.    f0030

shown that if the magnitude of a reward or the likelihood of a reward is systematically manipulated, then firing rates in these areas are a roughly linear function of those variables under many conditions. To be yet more precise, current data suggest that activity in this map encodes relative subjective value (RSV) of the type defined above.

Together, these data suggest the following model for eye-movement generation. At any moment in time, neurons in LIP represent the instantaneous RSV of each movement in the saccadic repertoire. Movements that have non-zero values are thus each represented by local activity on the map that is linearly proportion to RSV (see, for example, Dorris and Glimcher, 2004). I hypothesize that the representation of SV localized in the medial prefrontal cortex and the ventral striatum serve as the initial source of this signal. Previous studies have noted that other (and perhaps all) cortical areas perform a divisive normalization on their input data (Heeger, 1992; Schwartz and Simoncelli, 2001). It has now been observed that (at least to a first approximation) this also occurs in area LIP (Dorris and Glimcher, 2004; Sugrue *et al.*, 2004; Louie *et al.*, 2007), and the result is likely a shift from SV to RSV in the posterior parietal cortex.    p0720

RSV, it should be noted, would serve to map SV into the limited dynamic range of the LIP neurons. LIP neurons are limited in number, fire over a roughly 100-Hz dynamic range, and have (errors that are drawn from a) Poisson-like distribution. This means that the representation of RSV, rather than SV, in this structure may solve an important problem. The shift to RSV guarantees a distribution of the SVs of the current choice-set over the limited dynamic range of these

neurons. Unfortunately, the finite dynamic range and noise associated with these neurons may also impose a constraint. As the choice set becomes larger, noise may swamp the signal, leading to profound inefficiencies when selecting amongst large numbers of possible movements. One is tempted to speculate that this may, in fact, be a neural account for the choice-set size effects which have recently been examined in human choosers (see, for example, Iyengar and Lepper, 2000). It may also be that processes like choice set editing are tools used by frontal areas to winnow the size of the choice set operated on in parietal areas.

p0750    It should also be noted that the Poisson variance of these neurons may serve a useful function by allowing for stochasticity in behavior under conditions in which behaviors like mixed strategy equilibria arise. It was Nash who noted that mixed strategy equilibria arise when the expected utilities of the strategies being mixed are equivalent. In a similar way, when the RSVs of two options are equivalent it might be expected that the stochastic nature of these neurons yields mixed-strategy behavior. If these neurons are always stochastic in their behavior – a hypothesis that has been largely documented in monkeys – patterns of activity in LIP may be related to economic notions of the trembling hand (Selten, 1975).

In summary, then, the available data suggest that at all three of these areas – LIP, FEF, and SC – carry signals encoding RSV, and that movements occur when activity associated with one of the positively-valued options drives its associated collicular neurons into their burst mode. A tremendous amount of work
p0800    (again summarized in Chapters 4, 28, 29, and 31) has examined this process of movement-triggering under conditions in which animals are instructed to make movements as quickly as possible. Less is known about how movement selection is triggered in non-reaction time settings. One important possibility is that an input to one or more of these areas alters the inhibitory interactions within the map, forcing convergence to a single action.

The basic model proposed for selecting eye movements is thus that signals encoding SV project to these areas, probably through LIP, which nor-
p0810    malizes those signals to represent RSV which is further contaminated by local noise, the degree of which across the entire population may be regulated by adjustable inter-neuronal correlations (Glimcher, 2005). These signals propagate recursively through these networks while reflecting SV inputs that may be entering the maps at many locations. An external signal then permits, or forces, convergence of the network to a single choice which occurs when the collicular neurons are driven above their burst threshold.

p0780    Preliminary evidence for this hypothesis has been gathered by Louie and Glimcher (2006), who have shown that early in a trial the neurons of LIP represent the RSV of discounted gains associated with specific saccades, and that it is only later in the trial that these same neurons come to encode information about the actual choice made by the animal. This seems to suggest that the basic model is sound, at least for tasks of this type.

p0790    Two questions, however, immediately arise: how does this system achieve choice amongst more abstract objects that do not have specific movements associated with them, and does this model generalize to humans and non-eye movement conditions? A limited amount of data exists which suggests that this general class of system does operate under conditions in which choices are made between more abstract objects. Gold and Shadlen, for example, demonstrated that when animals must choose between red and green targets that constantly interchange locations, activity in the superior colliculus reflects the instantaneous mapping between color and value even if this changes from trial to trial (Gold and Shadlen, 2000; see also Horwitz and Newsome, 2001; Sugrue et al., 2004). This clearly indicates that the saccadic choice circuit has access to instantaneous mapping information relating abstract properties to actions. It cannot tell us however, how
p0760    choice is accomplished (or if it can be accomplished) in the absence of any mapping to motor circuitry of any kind.

We do, however, have some interesting hints that these choice circuits are interconnected with important valuation areas in the frontal cortex and basal ganglia. Padoa-Schioppa and Assad (2006), for example, have demonstrated the existence of neurons in the orbitofrontal cortex that encode an animal's choice before the movement expressing that choice is executed. In a similar way, Lau and Glimcher (2006) have observed choice neurons in the dorsal striatum. At the very least, this suggests that the choice circuit can send information about decisions frontally,
p0770    but it may also indicate that these areas participate directly in the convergence process by which choice is accomplished.

The question of whether these circuits that have been so well studied in monkeys can be generalized to other classes of movements and other species is one about which we have much less information. We do know that adjacent to area LIP are areas specialized for arm, hand, and face movements. Standard theories suggest that a group of areas lining the intraparietal sulcus serve as movement-control interfaces for all of the body, although there are problems still being resolved with those hypotheses (cf. Levy et al., 2007).

But it does seem clear that the general theories of movement control advanced for the oculomotor system do have analogues in the skeletomuscular system. Further, injuries to any of these systems, in either humans or monkeys, leads to permanent deficits not in the musculature but in the ability to produce movements. Finally, a small number of fMRI studies have shown value-related signals in the posterior parietal cortex, although these signals are almost always of weaker magnitude than in more frontal areas. This, of course, raises the possibility that the weaker fMRI signal reflects the temporal dynamics of choice observed in the Louie and Glimcher (2006) study. Because subjective value is only represented until a decision is made, in these areas the magnitude of the SV signal, integrated over an entire trial, may be much less than in areas located more frontally where SV is represented throughout a trial.

## ALTERNATIVES TO THE TWO-STAGE MODEL

### Choice Probabilities

Some early models of the primate choice system proposed that when choosing between two actions a choice probability was computed directly from the identity of the option pair, rather than by comparing something like the utilities of the two options under consideration. This choice probability was then proposed to stochastically direct action. When these models were introduced, some argued that they could serve as an alternative to preference-based models.

Two factors argue against models of this type. The first is axiomatic. Consider an agent who has been asked repeatedly to choose between chocolate and apples. Then she is asked to choose repeatedly between apples and crackers. We can, of course, represent the behavior of the agent with two choice probabilities. If we begin, however, with the assumption that the chooser represents *only* choice probabilities, then we must necessarily remain agnostic about what the agent will select if we offer her a choice between chocolate and crackers. If, on the other hand, we hold a belief that knowing her choices under these first two conditions reveals her likely choice under the third condition, then we are basically assuming complete and transitive preferences that invoke a utility-like representation. In other words, we invoke a system which behaves "as if" abstract valuations, subjective values, are represented.

The second factor arguing against this possibility is empirical. We now have compelling neurobiological evidence that subjective values of some kind are represented in the brains of monkeys (Dorris and Glimcher, 2004; Sugrue *et al.*, 2004, Louie and Glimcher, 2006; Padoa-Schioppa and Assad, 2008). In those experiments and others like them it has been demonstrated that the subjective values of individual options, and not choice probabilities, are represented by neuronal firing rates.

For these two reasons we can consider choice-probability-only based systems as empirically and theoretically falsified. Of course, it may well be that groups of neurons (or the local ensemble connections of those neurons) do explicitly represent choice probabilities. Some evidence suggests that this may be the case in posterior parietal cortex, but we now have sufficient evidence to conclude that the representation of subjective values – or something much like them – occurs within the central nervous systems of primates.

### Multiple Selves

The principle alternatives to the standard back-pocket model presented here are the multiple-self models that employ summation of some kind. These models typically propose the existence of two largely independent decision-making systems; one associated with so called "limbic" areas of the brain and the other with so called "rational" areas of the brain. While tremendously interesting from an economic point of view, these models are, for the most part, at variance with the majority of the existing corpus of neurobiological data . However, it is still germane to ask whether the existing evidence supports a two-agent model of decision-making of the type proposed by Laibson and colleagues (see, for example, Laibson, 1997; McClure *et al.*, 2004b). In that model, it is argued that the basal ganglia and medial prefrontal cortex form an emotional decision-making module which interacts (additively) with a second system organized around posterior parietal cortex and the dorsolateral prefrontal cortex, which form a rational decision-making module. Anatomical considerations that weigh against this hypothesis aside, we must ask whether or not there is compelling evidence that the division of brain areas into emotional and rational subgroups as can be supported by the available data. My answer is no. In monkeys, it has now been conclusively shown that activity in the posterior parietal cortex predicts preferences under all conditions that have been studied – for immediate rewards and for delayed rewards (Janssen and Shadlen, 2005;

Louie and Glimcher, 2006), for large rewards and for small rewards (Platt and Glimcher, 1999; Dorris and Glimcher, 2004), and for high-probability and low-probability rewards (Shadlen and Newsome, 1996; Platt and Glimcher, 1999). The data from animals seem to be unambiguous here – LIP activity predicts choices for both rational and emotional decision-making. To take another example, let us turn to the basal ganglia. This is an area that a number of neuroeconomists have argued is associated with emotional decision-making, but there is almost no evidence for this claim. Diseases of the basal ganglia are only very weakly associated with emotional dysfunction. The many dopaminergic forms of learning described in Part 3 of this volume, although largely mediated by the basal ganglia, do not seem to capture any clear notion of emotionality. A similar case can be made for studies of the medial prefrontal cortex. As noted above, there is evidence that this structure encodes monetary and primary rewards, preference, expected values, and gains and losses, and at least one study reports that it encodes long-delayed monetary gains. Indeed, even loss-aversion seems to be encoded in the unitary activity of this structure. Together, these data paint a picture of a structure globally involved in valuation – not a structure driven exclusively by immediacy, fear, or emotionality.

p0900

In summary, then, our available evidence seems to suggest that existing multiple-self models are largely unsupported by the bulk of our existing data. Of course, emotions do influence decision-making and choosers do show varying levels of self-control; that is beyond doubt. The question is, how do emotions and circuits related to self-control effect this influence? The amygdala, to take one example, may provide an answer. The amygdala projects strongly to the ventral striatum and there is physiological and anatomical evidence that activity in the amygdala strongly influences activity in the ventral striatum. That does argue that the amygdala, and perhaps the emotions that it encodes, can influence valuation-related activity in this area, but it does not make a compelling case for a Freudian multiple-self model of neural decision-making.

p0910

## CONCLUSION

What emerges from a review of the available human and animal data on decision-making is evidence of a two-stage model for choice. The first (or valuation) stage learns and represents the values of both actions and goods. Within this stage, at least

three learning mechanisms distributed in the basal ganglia and frontal cortex contribute to the construction of what we refer to as subjective value. These areas are hypothesized to learn subjective values, at a biophysical level, through the well-studied process of synaptic plasticity. These learning processes operate both during choice and during the passive receipt of rewards, effecting a disassociation between choice and valuation.

Our available evidence makes it clear that subjective value is a stochastic quantity, effectively drawn from a spiking distribution dependent on these synaptic strengths. It is also a reference-dependent quantity, as indicated by the Tom *et al*. (2007) study. In this regard, subjective value is most closely allied to a reference-dependent random-utility model in economic theory. I propose that SV is encoded specifically in the activity of the medial prefrontal cortex and the ventral striatum. I note that while SV is responsible for preferences, it can violate the axioms of expected utility theory; indeed, it must if it is to account for true preferences. Some of these violations doubtless reflect the influence of emotion-related brain structures on medial prefrontal cortical and ventral striatal activity.

p0890

Choice, I propose, is accomplished in a network that includes the posterior parietal cortex and a number of movement-related areas subsequent to it in the motor control stream. In these areas, the SVs of objects within a single choice set are normalized to RSVs. These RSVs are further modified by the addition of a variable noise term, of Poisson-distributional origin, prior to a winner-takes-all operation that accomplishes choice itself. This is a feature reminiscent of the trembling hand of economic theory in some important ways. Let me stress that the winner-takes-all choice operation must be broadly distributed, and involves structures that range from the superior colliculus to the orbitofrontal cortex.

p0870

Of particular interest are several features of the model that remain unspecified. While there are many candidate pathways by which information from the medial prefrontal cortex and the ventral striatum may influence activity in the posterior parietal cortex, which of these pathways is critical for choice has not yet been determined. It has also been noted (see Chapter 29) that much of the posterior parietal cortex encodes SV with regard to actions, while neurons in the orbitofrontal cortex (Padoa-Schioppa and Assad, 2006), and perhaps the medial prefrontal cortex, encode SV with regard to goods. We do not know how a transformation between these representations occurs, although we do know that it does occur. We also have only limited information about the systems that "decide to choose." In some tasks animals have to be trained to make a choice as

s0200

p0880

soon as possible, and under these conditions one can observe the parietal and frontal networks converging towards choice. In other situations, however, the time-courses of valuation and choice are separable, as is more typically the case in human economic behavior. This suggests the existence of a circuit that can essentially force the parietal networks towards convergence. Such a system would almost necessarily involve cortical networks of inhibitory connections, but the features of this process that decides when to choose remain completely absent from this standard back-pocket model.

Over the course of the past decade an extraordinary amount of progress has been made in identifying the basic features of the primate mechanism for choice, and there is remarkable consensus about much of this mechanism. This is a device that can be the subject of economic study, and the existing neurobiological data clearly identify some areas of economic theory as more relevant to the study of this device than others. The existing theory also identifies questions that must be answered by neurobiology. That, of course, is the whole point of this endeavor.

## Acknowledgments

## References

Allais, M. (1953). Le comportment de l'homme rationnel devant le risqué: critique des postulates et axioms de l'ecole americaine. *Econometrica* 21, 503–546.

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 1–13.

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Dahlström, A. and Fuxe, K. (1964). Evidence for the existence of monoamine-containing neurons in the central nervous system. I. Demonstration of monoamines in the cell bodies of brain stem neurones. *Acta Physiol. Scand.* 62(Suppl. 1), 1–55.

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

Delgado, M.R., Nystrom, L.E., Fissell, C. *et al*. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072–3077.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the subjective desirability of an action. *Neuron* 44, 365–378.

Elliott, R., Friston, K.J., and Dolan, R.J. (2000). Disassociable neural responses in human reward systems. *J. Neurosci* 20, 6159–6165.

Elliott, R., Newman, J.L., Longe, O.A., and Deakin, J.F.W. (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J. Neurosci.* 23, 303–307.

Glimcher, P.W. (2003). The neurobiology of visual saccadic decision making. *Annu. Rev. Neurosci.* 26, 133–179.

Glimcher, P.W. (2005). Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25–56.

Glimcher, P.W. and Sparks, D.L. (1992). Movement selection in advance of action in the superior colliculus. *Nature* 355, 542–545.

Glimcher, P.W., Kable, J.W., and Louie, K. (2007). Neuroeconomic studies of impulsivity: now or just as soon as possible? *Am. Econ. Rev.* 97, 142–147.

Gold, J.I. and Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.

Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. In: A. Caplin and A. Schotter (eds), *The Foundations of Positive and Normative Economics: A Handbook*. Oxford: Oxford University Press, (forthcoming).

Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* 9, 81–198.

Hollerman, J.R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Iyengar, S. and Lepper, M. (2000). When choice is demotivating: can one desire too much of a good thing? *J. Pers. Social Psych.* 79, 995–1006.

Iyengar, S.S., Wells, R.E., and Schwartz, B. (2006). Doing better but feeling worse: looking for the "best" job undermines satisfaction. *Psychol. Sci.* 17, 143–150.

Janssen, P. and Shadlen, M.N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* 8, 234–241.

Kable, J.W. and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.

Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). fMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* 12, 20–27.

Knutson, B., Fong, G.W., Bennett, S.M. *et al*. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related FMRI. *NeuroImage* 18, 263–272.

Knutson, B., Taylor, J., Kaufman, M. *et al*. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Krug, K., Cumming, B.G., and Parker, A.J. (2004). Comparing perceptual signals of V5/MT neurons in two binocular depth tasks. *J. Neurophysiol.* 92, 1586–1596.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Economics* May, 443–477.

Lau, B. and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–643.

Lee, C., Rohrer, W.H., and Sparks, D.L. (1989). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332, 357–360.

Levy, I., Schluppeck, D., Heeger, D.J., and Glimcher, P.W. (2007). Specificity of human cortical areas for reaches and saccades. *J. Neurosci.* 27, 4687–4696.

Louie, K. and Glimcher, P.W. (2006). Temporal discounting activity in monkey parietal neurons during intertemporal choice. *Soc. Neurosci. Abstr.*, 605.5.

Louie, K., Grattan, L., and Glimcher, P.W. (2007). Relative reward encoding and cortical normalization in parietal area LIP. *Soc. Neurosci. Abstr.*, 645.7.

McClure, S.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2004a). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

McClure, S.M., Li, J., Tomlin, D. *et al.* (2004b). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379–387.

McFadden, D. (1974). Conditional Logit analysis of qualitative choice behavior. In: P. Zarembka (ed.), *Frontiers in Econometrics*. New York, NY: Academic Press, pp. 105–142.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1997). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.

O'Doherty, J., Deichmann, R., Critchley, H.D., and Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron* 33, 815–826.

O'Doherty, J.P., Buchanan, T.W., Seymour, B., and Dolan, R.J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49, 157–166.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Padoa-Schioppa, C. and Assad, J.A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes in menu. *Nat. Neurosci.* 11, 95–102.

Pagnoni, G., Zink, C.F., Montague, P.R., and Berns, G.S. (2002). Activity in human ventral striatum locked to errors in reward prediction. *Nat. Neurosci.* 5, 97–98.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Platt, M.L., Lau, B., and Glimcher, P.W. (2003). Situating the superior colliculus within the gaze control network. In: W.C. Hall and A. Moschovakis (eds), *The Oculomotor System: New Approaches for Studying Sensorimotor Integration*. Boca Raton, FL: CRC Press, pp. 1–34.

Romo, R. and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J. Neurophysiol.* 63, 592–606.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Savage, L. (1954). *Foundations of Statistics*. New York, NY: Wiley.

Schultz, W. and Romo, R. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* 63, 607–624.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* 13, 900–913.

Schwartz, O. and Simoncelli, E.P. (2001). Natural signal statistics and sensory gain control. *Nat. Neurosci.* 4, 819–825.

Selten, R. (1975). A reexamination of the perfectness concept for equilibrium points in extensive games. *Intl J. Game Theory* 4, 25–55.

Shadlen, M.N. and Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proc. Natl Acad. Sci. USA* 93, 628–633.

Stevens, S. (1970). Neural events and the psychophysical law. *Science* 170, 1043–1050.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.

Tolhurst, D.J., Movshon, J.A., and Dean, A.F. (1983). Statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* 23, 775–785.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.