75 Neuroeconomics and the Study of Valuation

PAUL W. GLIMCHER

ABSTRACT Just over a decade ago neurobiologists knew almost nothing about the neural mechanisms of voluntary choice. In contrast, economists and psychologists working at that time had well-developed frameworks for describing the many hidden processes that must underlie choice, but these frameworks had very little impact in neurobiological circles. The last decade, however, has seen a revolution in the neurobiological understanding of choice that has been driven by an integration of economics and psychology into mainstream neuroscience. Today, the basic outlines of the primate system for decision making are emerging from studies on humans and monkeys that rely on techniques ranging from single-neuron electrophysiology to functional magnetic resonance imaging (fMRI). Indeed, since the last edition of this book was published a new field for the study of decision making has emerged, *neuroeconomics*, and an edited volume has been published that surveys the field (Glimcher, Camerer, Fehr, & Poldrack, 2008). This chapter provides an outline of the primate mechanism for choice as we understand it today. In broad strokes, we now believe that choice involves a two-stage neural process. The first stage, largely resident in the frontal cortex and the basal ganglia, learns and represents the value of our actions. The second stage, largely resident in a frontoparietal network, selects the option that has the highest subjective value from among the options before us at any moment in time.

Introduction

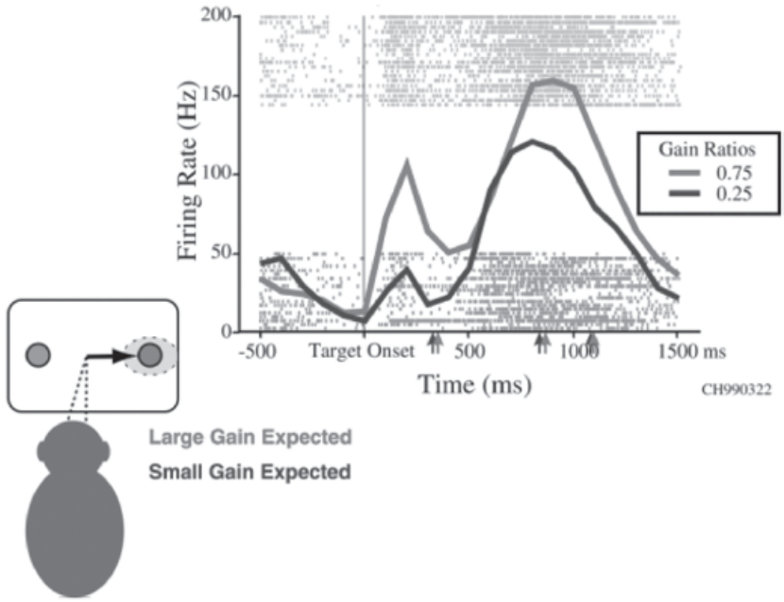
Our existing data now suggest that when we make a choice we employ a two-step neurobiological process with some remarkable similarities to both psychological and economic process models of decision making. The first step in the neurobiological processes that guides decision making places idiosyncratic valuations on the options before a chooser. These valuations involve the activation of many frontocortical and basal ganglia circuits. The second step chooses, based on those valuations, a single action for execution. Although less well understood than the valuation processes, these choice processes involve both frontal and parietal circuits. What follows is an overview of the valuation and choice mechanisms as they are understood today. Without a doubt, this understanding is fragmentary, and some of the conclusions made here will be somewhat controversial, but the presentation captures the state of the field today and suggests just how much has been accomplished since the third edition of this volume was published only five years ago.

PAUL W. GLIMCHER Professor of Neural Science, Economics, and Psychology; Director, Center for Neuroeconomics; New York University, New York, New York.

The two-stage model

The neurobiological evidence for a two-stage model emerged initially from studies of decision making in awake behaving monkeys conducted throughout the 1990s by two groups of researchers. The first of these groups was concerned with understanding how animals engaged in traditional psychophysical tasks that required the evaluation of visual stimuli reached a perceptual decision (Newsome, Britten, Salzman, & Movshon, 1990; Gold & Shadlen, 2007). The second emerged from the study of movement control and was concerned with understanding how changes in the magnitude or probability of reward influenced decision making (Platt & Glimcher, 1999; Glimcher, 2002). Both groups converged, however, to the view that neurons of the posterior parietal cortex participated in the actual process of deciding (selecting one action from a finite set of alternatives) and that these neurons received inputs that encoded something about the magnitude or likelihood of future rewards (associated with each of those alternatives) that originated from signals generated elsewhere in the brain.

In Platt and Gimcher's (1999) study, the authors recorded from neurons in the posterior parietal cortex while thirsty monkeys participated in a simple forced-choice task. In that task, monkeys fixated a central yellow target while two eccentric visual stimuli (one red and one green) were coilluminated (figure 75.1). One of those targets was located within the response field of a parietal neuron under study. After a brief delay, the central target then switched color to either red or green, indicating which of the two eccentric stimuli the animal should fixate in order to receive a reward. What the authors varied, across blocks of about 100 trials, was either the magnitude of reward associated with each of the targets or the probability that the fixation target would turn red. They found that, *immediately after target onset*, if the magnitude of reward associated with the target



22 FIGURE 75.1 Same movement, different values. (From Glimcher, 2003.)

inside the response field was increased, the firing rates of neurons encoding an eye movement to that target increased at the very beginning of the trial. They also showed that if the probability that an eye movement toward the response field would be reinforced was high, the units responded more strongly than if a movement toward that target was unlikely to yield a reward. *Immediately before eye movement onset*, however, the neuronal firing rate indicated whether or not the animal had chosen to produce the saccade encoded by that neuron.

In interpreting this result, they noted that all economic theories of choice predict that valuation should always be influenced by both the probability and the magnitude of
1 reward. This theory suggested that the early activity observed in these neurons might well encode the subjective value of the eye movements to the monkeys. The late activity, in contrast, appeared to encode choice, the output of an operation performed on the set of eye movements available to the animal. The suggestion, then, was that the inputs to these parietal circuits might well encode an idiosyncratic subjective valuation of the kind described by economic theories of choice and that parietal (and related extraparietal) circuits might use these valuation inputs as part of a winner-take-all computation to choose actions for execution.

At the same time that these studies were being conducted, a number of lines of evidence began to suggest that portions of the striatum and the frontal cortex both learn and represent the values of goods and actions—a finding suggesting that these areas might serve as the source of the valuation signals identified in parietal cortex. The critical first step toward 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 toward a broader understanding of valuation. In the early 1990s, Wolfram Schultz and his colleagues (e.g., Romo & Schulz, 1990; Schultz & Romo, 1990; Schultz, Apicella, & Ljungberg, 1993) demonstrated that midbrain dopaminergic neurons encode a *reward prediction error*. Montague, Dayan, and Sejnowski (1997) 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 that could be used to guide choice. What followed were 10 years of work that established the existence of at least three interrelated subsystems in these brain areas that employ distinct mechanisms for learning and representing value and that interact to produce the valuations that guide choice (Dayan & Balliene, 2002; Balliene, Daw, & O'Doherty, 2008; Niv & Montague, 2008).

In a similar way, studies of the movement control systems of the brain strengthened the conviction of many that a discrete choice mechanism used these valuation signals to select and execute actions. Our current evidence indicates that the choice system involves large portions of the parietal cortex, among other areas. 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 of the frontal cortex and basal ganglia participate directly in the choice process with these parietal areas. We now know that specific neurons in the orbitofrontal cortex (Padoa-Schioppa & Assad, 2006, 2008) and the dorsal striatum (Samejima, Ueda, Doya, & Kimura, 2005; Lau & 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 an action for execution; 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.

The basic structure of the valuation system

The critical breakthrough that allowed modern studies of valuation to crystallize were insights into the function of the midbrain dopaminergic pathways. In 1993, Schultz and colleagues measured the spiking activity of single dopamine neurons while monkeys passively received rewards during a classical conditioning task. 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 but rather appeared to encode something more closely related to learning itself. 2 This revelation led Montague, Dayan, and Sejnowski (1996) to propose that dopamine neurons encoded the difference between expected and obtained rewards: the *reward prediction error* of reinforcement learning theory. The critical idea that emerged over the next several years was that dopamine spike rates communicated to frontocortical and striatal circuits the degree to which rewards actually obtained by the subject matched previously learned predictions of reward magnitude. This explained why dopamine neurons responded to unconditioned rewards (which the animals did not expect) while remaining silent when animals received conditioned rewards (which the animals expected).

More formally, what these studies suggested was that dopamine neurons coded a term from reinforcement learning theory that had been previously developed within psychological circles. In 1972, Rescorla and Wagner had proposed that the associative strength between a stimulus and a reward during classical conditioning could be described by the rule

3 $$AssStr_{new} = AssStr_{old} + \alpha(\mathrm{R}eward - AssStr_{old})$$

where *AssStr*, or "associative strength," is thus incremented (or decremented) by the difference between the reward obtained and the reward expected (the *old* associative strength) until the prediction matches the experience and learning is thus complete. In this formulation, α is a number between 0 and 1 that controls how gradually learning shifts the prediction across trials from old values to new values. (In fairness, this is not exactly the form in which Rescorla and Wagner made their proposal. Their proposal employed an additional term associated with stimulus salience and also made predictions about how two stimuli competed to predict a single reward. The form shown here is much closer to a model originally proposed in 1951 by Bush and Mosteller that served as the basis of Rescorla and Wagner's later model.)

Subsequent studies of the dopamine neurons and many of their targets have largely validated this early conclusion of Montague's and extended these insights into the domain of operant conditioning in animals. In 1992, Schultz and his colleagues (Ljungberg, Apicella, & Schultz, 1992) showed, for example, that even in a classical conditioning task dopamine neurons encoded a signal that closely paralleled the reward prediction error term of theory. Subsequent studies using more sophisticated computational methods (Bayer & Glimcher, 2005; Morris, Nevet, Arkadir, Vaadia, & Bergman, 2006) have also validated this hypothesis. Together, these data demonstrate unequivocally that dopamine neurons carry a signal to the striatum and frontal corticies that is sufficient to account for how animals learn the values of their actions, clear evidence that a valuation signal of some kind could be constructed and stored in these areas (or their targets) within the brains of monkeys.

Fortunately, there is also clear evidence that these dopaminergic neurons behave in the same manner in humans as they do in monkeys. 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, Deichmann, Critchley, & Dolan, 2002; Pagnoni, Zink, Montague, & Berns, 2002) demonstrated simultaneously that activity in the dopaminergic terminal fields of the striatum and the frontal cortex during both gustatory and monetary reward tasks behaved exactly as predicted. This result indicated that there existed dopaminergic signals appropriate for value learning in humans.

Contemporary studies of these neurons continue to extend and refine these conclusions in important ways. We now have reason to believe that the actual algorithm computed by the dopamine neurons or their sources is a more refined version of the Rescorla and Wagner model known as the *temporal difference model* (Sutton & Barto, 1988). This model 4

explains not just how expected rewards are encoded, but how a dopamine-based system could develop associations between stimuli and rewards that are separated in time. The temporal difference model, like the dopamine neurons, is able to connect the ringing of a bell with a food reward that follows it seconds later (Schultz et al., 1997). This is an important advance, but one that lies beyond the scope of this brief review.

What remains, then, is to understand where and how these dopamine activations are used to mechanistically compute and store the values of actions. Two lines of evidence contribute to our contemporary understanding of these issues: Neuronal recording studies in animals and fMRI studies in humans. The recording studies in animals have now established that the basal ganglia (and in particular the striatum) 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 (Hikosaka, Takikawa, & Kawagoe, 2000; Samejima et al., 2005; Lau & Glimcher, 2008), neurons that encode the actions that have just been executed (Samejima et al.; Lau & Glimcher, 2007), and neurons with firing rates dependent on the current state of the environment (Hikosaka, 2007), among 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 projects in turn 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 (Padoa-Schioppa & Assad, 2006, 2008). Functional MRI studies in humans tell a similar story (Knutson, Westdorp, Kaiser, & Hommer, 2000; Knutson, Taylor, Kaufman, Peterson, & Glover, 2005; O'Doherty et al., 2002; O'Doherty, Buchanan, Seymour, & Dolan, 2006), suggesting that frontal and basal ganglia circuits form the core of the human mechanism for learning and representing value.

There is, however, evidence for other learning mechanisms in these same structures that interact with this well-studied Rescorla-Wagner-style learning mechanism. The details of these other learning systems are still being worked out, but 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 (Balliene et al., 2008).

For neuroeconomists, 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 these values that guide choice, likely computed by several interaction neural circuits organized around the frontal cortex and the basal ganglia to the choice system, are acted on by the choice system that guides action?

One way to begin to answer this question is to look at the existing fMRI data and to ask, Are there a small number of areas that are actively correlated with subjective value 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 values we infer humans and animals place on their own actions.

Activity in the ventral striatum has been shown to be correlated with both rewards and punishments (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000), the magnitude of cumulative rewards (Elliot, Friston, & Dolan, 2000), the anticipation of reward (Knutson, 2001, 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, Newman, Longe, & Deakin, 2003), monetary expected values (Knutson et al., 2005), behavioral preference rankings among 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 six months (Kable & Glimcher, 2007). Single-unit recording studies in the dorsal striata of monkeys, both in the caudate (Lau & Glimcher, 2006) and the putamen (Samejima et al., 2005), tell a similar story. Neurons that clearly code action values have been identified in these areas. All 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.

A similar correlation 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., 2001, 2003), preference ordering among primary rewards (McClure, Laibson, Loewenstein, & Cohen, 2004), the expected value of a lottery (Knutson et al., 2005), the subject-specific valuation of gains and losses (Tom et al., 2007), subject-specific discounted reward value (Kable & Glimcher, 2007), and willingness to pay (Plassman et al., 2007). Activity in this area appears to be correlated with valuation under all these conditions. These data have led to the proposal that mean activity in the medial prefrontal cortex and the ventral striatum serves as a final common path for encoding the values of actions (Glimcher, Dorris, & Bayer, 2005).

It should be noted, however, that this conclusion remains somewhat controversial. An alternative hypothesis active in the literature proposes that the valuations we infer from behavior reflect the interaction of two or more largely independent neural systems that compete to govern behavior, the so-called *multiple-self* models. 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 most of the existing corpus of neurobiological data. Still, it is germane to ask whether the existing evidence supports a two-agent model of decision making of the type proposed by Laibson and colleagues (e.g., Laibson, 1997; McClure et al., 2004). In that model, it is argued that the basal ganglia and medial prefrontal cortex form an emotional decision-making module that 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 *competing* emotional and rational subgroups can be supported by the available data. 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 & Shadlen, 2005; Louie & Glimcher, 2006), for large rewards and for small rewards (Platt & Glimcher, 1999; Dorris & Glimcher, 2004), for high-probability and low-probability rewards (Shadlen & Newsome, 1996; Platt & Glimcher, 1999). The data from animals seem to be unambiguous—lateral interparietal area (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 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 here, 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 previously, 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. Together, these data paint a picture of structures globally involved in valuation driven by all mental states—not a structure driven exclusively by immediacy, fear, or emotionality.

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 influence decision making, and choosers show varying levels of self-control; those conclusions are beyond doubt. The question is, How do neural circuits related to emotions influence decision making? 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 evidence argues that the amygdala, and thus perhaps the emotions to which it is related, 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.

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. That emphasis 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 nonhuman primates.

Initial studies of choice in monkeys evolved almost simultaneously from studies of sensory-perceptual systems (e.g., Newsome, Britten, & Movshin, 1989) and movement-control studies (e.g., Glimcher & Sparks, 1992), as noted earlier. 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 so by leveraging an extensive preexisting literature on the structure of the visual and eye movement systems to search for the decision-making circuits that connected them in these tasks (Glimcher, 2005). Subsequent work has generalized many, but not all, of these findings to arm movement control systems and to studies of humans.

We have to begin, therefore, with a review of the basic structure of the saccadic control system (figure 75.2). The 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 be required to align gaze with those objects (for a review see Glimcher, 2005). Thus a particular location on the map (or more precisely the neurons on the map at that location) activate when a visual stimulus appears 10 degrees to the right of fixation, *and* that region might become particularly active milliseconds before an eye movement that shifts gaze 10 degrees. 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, Lau, & Glimcher, 2003), a fact which is probably important for understanding choice. Finally, the colliculus is connected to brain stem circuits that actually govern eye movements in real time. The connection between these brain stem systems and the colliculus are mediated by a class of collicular neurons called *burst neurons*. Burst neurons have

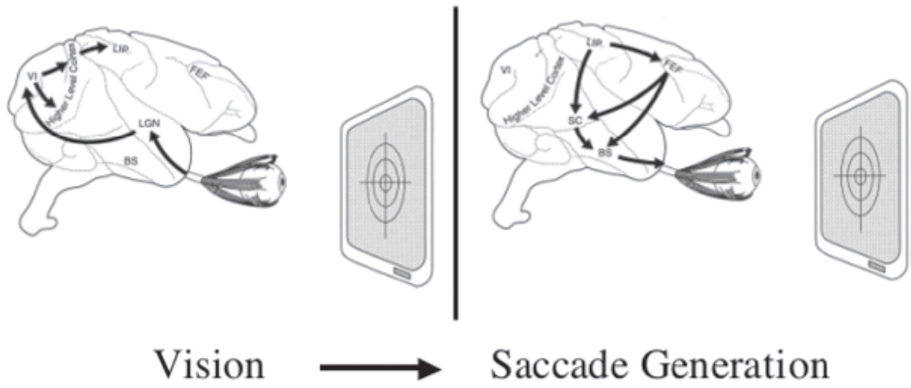


FIGURE 75.2 Saccadic control system: The visual-saccadic brain.

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 burstlike activity occurring at more than one location at a time, suggesting that the collicular architecture allows only a single movement to be executed at a time. Studies in area LIP, the frontal eye fields (FEF), and the superior colliculus (SC) 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 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 (Dorris & Glimcher, 2004; Gold & Shadlen, 2007).

Together, these data suggest the following model for eye movement generation. At any moment in time neurons in LIP represent the instantaneous subjective value of each movement in the saccadic repertoire. Movements that have nonzero values are thus each represented by local activity on the map. One might even hypothesize that the representation of subjective value localized in the medial prefrontal cortex and the ventral striatum serve as the initial source of this signal.

In summary then, the available data suggest that at all three of these areas, LIP, FEF, and SC, carry signals encoding subjective value 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 (reviewed in Glimcher, 2005; Gold & Shadlen, 2007) 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 subjective value project to these areas, probably through LIP. These signals propagate recursively through these networks while reflecting value 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 that occurs when the collicular neurons are driven above their burst threshold.

Two questions, however, immediately arise: How does this system achieve choice among more abstract objects that do not have specific movements associated with them? Does this model generalize to humans and non-eye-movement conditions? A limited amount of data exist that do suggest that this general class of system operates under conditions in which choices are made between more abstract objects. Shadlen and Gold (2000; see also Sugrue, Corrado, & Newsome, 2004), 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. This finding clearly indicates that the saccadic choice circuit has access to instantaneous mapping information relating abstract properties to actions. It cannot tell us, however, how 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 (2008) have observed choice neurons in the dorsal striatum. At the very least, this finding suggests that the choice circuit can send information about decisions frontally, 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 (Andersen & Buneo, 2002) 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
14 (cf. Levy & Glimcher, 2007). But it does seem clear that the general theories of movement control advanced for the monkey 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 result of course raises the possibility that the weaker fMRI signal reflects the temporal dynamics of choice. Because subjective value is only represented until a decision is made, in these areas the magnitude of the subjective value signal, integrated over an entire trial, may be much less than in areas located more frontally where subjective value is represented throughout a trial.

Summary

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. A network, which includes the posterior parietal cortex and a number of movement-related areas subsequent to it in the motor control stream, appears to perform a winner-take-all operation on these values that accomplishes choice itself. Let me stress that the winner-take-all choice operation must be broadly distributed and involves structures that range from the superior colliculus to the orbitofrontal cortex.

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. 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 soon as possible, and under these conditions one can observe the parietal and frontal networks converging rapidly toward choice. In other situations, however, the time courses of valuation and choice are separable. This possibility suggests the existence of a circuit that can essentially force the parietal networks toward convergence, the circuits that "decide to choose." Such circuits almost necessarily involve cortical networks of inhibitory connections, but the features of this process that decides when to choose remain completely absent from our standard model.

Over the course of the past decade a remarkable amount of progress has been made in identifying the basic features of the primate mechanism for choice. While many critical questions remain, progress in the last decade has marked this as an exciting and innovative area in cognitive neuroscience.

REFERENCES

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in the posterior parietal cortex. *25*, 189–220.

Balliene, B. W., Daw, N. D., & O'Doherty, J. (2008). Multiple forms of value learning. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain*. New York: Elsevier.

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

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychol. Rev.*, *58*, 313–323.

Dayan, P, & Balliene, B. W. (2002). Reward, motivation and reinforcement learning. *Neuron*, *36*, 285–298.

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

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

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

Elliott, R., Newman, J. L., Longe, O. A., & 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. (2002). Decisions, decisions, decisions: Choosing a biological science of choice. *Neuron*, *36*, 323–332.
Glimcher, P. W. (2003). The neurobiology of visual saccadic decision making. *Annu. Rev. Neurosci.*, *26*, 133–179.
Glimcher, P. W., Camerer, C. F., Fehr, E., & Poldrack, R. A. (2008). *Neuroeconomics: Decision making and the brain.* New York: Elsevier.
Glimcher, P. W., Dorris, M. C., & Bayer, H. M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.*, *52*, 213–256.
Glimcher, P. W., & Sparks, D. L. (1992). Movement selection in advance of action in the superior colliculus. *Nature*, *355*, 542–545.
Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, *30*, 535–574.
Hikosaka, O. (2007). Basal ganglia mechanisms of reward-oriented eye movement. *Ann. NY Acad. Sci.*, *1104*, 229–249.
Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive eye movements. *Physiol. Rev.*, *80*, 953–978.
Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.*, *1*, 304–309.
Janssen, P., & 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., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.*, *10*, 1625–1633.
Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. S., & Hommer, D. (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., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *J. Neurosci.*, *25*, 4806–4812.
Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage*, *12*, 20–27.
Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Econ.*, 443–477.
Lau, B., & Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate nucleus. *J. Neurosci.*, *27*, 14502–14514.
Lau, B., & Glimcher, P. W. (2008). Value representations in the primate caudate during a matching-law task. *Neuron*. In press.
Levy, I., Schluppeck, D., Heeger, D. J., & Glimcher, P. W. (2007). Specificity of human cortical areas for reaches and saccades. *J. Neurosci.*, *27*, 4687–4696.
Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.*, *67*, 145–163.
McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503–507.
Montague, P. R., Dayan, P., & Sejnowski, T. J. (1997). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.*, *16*, 1936-1947.
Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future actions. *Nat. Neurosci.*, *9*, 1057–1063.
Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, *341*, 52–54.
Newsome, W. T., Britten, K. H., Salzman, C. D., & Movshon, J. A. (1990). Neuronal mechanisms of motion perception. *Cold Spring Harb. Symp. Quant. Biol.*, *55*, 697–706.
Niv, Y., & Montague, P. R. (2008). Theoretical and empirical studies of learning. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the Brain.* New York: Elsevier.
O'Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*, *49*(1), 157–166.
O'Doherty, J., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, *33*, 815–826.
Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*, 223–226.
Padoa-Schioppa, C., & 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., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors in reward prediction. *Nat. Neurosci.*, *5*(2), 97–98.
Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*, 233–238.
Platt, M. L., Lau, B., & Glimcher, P. W. (2003). Situating the superior colliculus within the gaze control network. In W. C. Hall, & A. Moschovakis (Eds.), *The oculomotor system: New approaches for studying sensorimotor integration.* Boca Raton, FL: CRC Press.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black, & W. Prokasy (Eds.), *Classical conditioning II* (pp. 64-99). Appleton-Century-Crofts.
Romo, R., & 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., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science*, *310*, 1337–1340.
Schultz, W., Apicella, P., & Ljungberg, T. (1993). Respones of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.*, *13*, 900-913.
Schultz, W., & Romo, R. (1990). Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.*, *63*, 607–624.
Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. *Proc. Natl. Acad. Sci. USA*, *93*, 628–633.
Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science*, *304*, 1782–1787.
Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. Cambridge, MA: MIT Press.
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*, 1642–1645.

AUTHOR QUERY FORM

Dear Author

During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.

Many thanks for your assistance.

Query References	Query	Remarks
1.	Author: "This theory" OK? If not, then this what?	
2.	Author: We don't have Montague, Dayan, & Sejnowski '96, We have '97. (cited previously)	
3.	Author: "Re *ward*"? Should this be "*Reward*"?	
4.	Author: We don't have Sutton & Barto '88. We have '98. We don't have Schultz et al. '97	
5.	Author: Are Delgadro '00 and Elliot … '00 OK as edited?	
6.	Author: We don't have Knutson '01, '03, we have et al.'s for '00, '03, and '05. See also last line of p. 75-13.	
7.	Author: We don't have Breiter et al. '01, Tom et al. '07, as Lau & Glimcher '06.	
8.	Author: Only one McClure et al., right?	
9.	Author: We don't have Tom et al. '07 or Plassman et al. '07	
10.	Author: We don't have Glimcher '05. Should this be et al.?	
11.	Author: Glimcher et al., '05?	
12.	Author: Glimcher et al., '05?	
13.	Author: We don't have Shadlen & Gold '00	
14.	Author: We don't have Levy & Glimecher '07	
15.	Author: Andersen …: Pls give journal name (abbreviated)	
16.	Author: Hollerman & Schultz '98 wasn't cited. OK to delete?	
17.	Author: Knutson, Fong … '03 wasn't cited as such. See p. 75-12.	
18.	Author: Laibson … '97: Plas give vol.	
19.	Author: Lau & Glimicher '08: Pls give full data.	
20.	Author: Rescorla … '72: Pls give place of publication.	
21.	Author: Tobler … '05 wasn't cited. OK to delete?	
22.	Author: Are these caption OK? Or do you wish to provide fuller ms.?	