

Oxford Handbooks Online

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Shih-Wei Wu and Paul W. Glimcher

The Oxford Handbook of Computational Economics and Finance

Edited by Shu-Heng Chen, Mak Kaboudan, and Ye-Rong Du

Print Publication Date: Feb 2018

Subject: Economics and Finance, Economic Theory and Mathematical Models, Econometrics, Experimental and Quantitative Methods

Online Publication Date: Feb 2018 DOI: 10.1093/oxfordhb/9780199844371.013.45

Abstract and Keywords

The standard neurobiological model of decision making has evolved, since the turn of the twenty-first century, from a confluence of economic, psychological, and neuroscientific studies of how humans make choices. Two fundamental insights have guided the development of this model during this period, one drawn from economics and the other from neuroscience. The first derives from neoclassical economic theory, which unambiguously demonstrated that logically consistent choosers behave “as if” they had some internal, continuous, and monotonic representation of the values of any choice objects under consideration. The second insight derives from neurobiological studies suggesting that the brain can both represent, in patterns of local neural activity, and compare, by a process of interneuronal competition, internal representations of value associated with different choices.

Keywords: neuroeconomics, revealed-preference theory, prospect theory, decision under risk, probability weighting function, dopamine, reward prediction error, reinforcement learning, medial prefrontal cortex, striatum

23.1 Overview

In the 1930s Samuelson famously demonstrated that consistent human choosers behave as if they had an internal representation of an idiosyncratic subjective value, or the *utility*, of choice objects under current consideration and selected from these internal representations the single choice object that had the highest utility. Taking that proof as a

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

starting point, many neuroeconomists have argued for a very literal and mechanistic reinterpretation of Samuelson's insight: That consistent choosers behave as they do *because* they have an internal representation of subjective value encoded in units of physical action potentials per second in their brains (e.g., Dorris and Glimcher 2004; Kim et al. 2008; Hayden et al. 2009; Glimcher 2010, 2011). They hypothesize that the brain performs an *argmax* operation on this internal representation based on the ordering of these action potential rates to select the most desirable option from a choice set. Of course, it is critical to note that these action potential rates are physical objects that are fully cardinal and unique, a property that makes them quite different from an economist's notion of utility. For that reason, and for others related to the causal relation between this activity and choice, this physical correlate of utility is typically referred to as *subjective value*.

This basic construct has led naturally to the notion that the human choice mechanism can be usefully divided into two subcomponents. The first is presumed to learn, represent, and store the values of goods and actions. The network in the brain involved in these computations is generally referred to as the *valuation network*. It is this (p. 689) mechanism that explains, for any given choice set, how humans assign values to choice objects that are unique to the individual decision maker. The second of these subcomponents is presumed to allow the direct comparison of two or more valued objects and results in the selection of the option associated with higher levels of neural activation through a "winner-take-all" computational process—an algorithmic instantiation of the mathematician's *argmax*. The brain network that performs this algorithmic operation is typically referred to as the *choice network*. Although our current evidence suggests that these processes, and the networks that embody them, cannot be seen as entirely separate, there is good evidence that these processes are instantiated as, at least in part, separable and sequentially executed algorithms (for an alternative view see Padoa-Schioppa 2011). It is this mechanism that explains how humans select the best option from any given choice set based on the values computed, stored, and represented in the antecedent valuation mechanism. Our goals in this chapter are to provide a more detailed overview of these two components and to discuss the strengths and weaknesses of this general two-stage model.

23.2 Stage 1: The Valuation Mechanism

23.2.1 Ordinal Utility to Cardinal Subjective Value

Perhaps the first critical challenge faced by any theory which assumes that humans choose the way they do because of an underlying utilitylike representation in the nervous system is that of ordinality. Since Pareto, nearly all economists have acknowledged that measurements of utility are largely ordinal. Although we can say that a chooser prefers

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

apples to oranges based on, for example, the Strong Axiom of Revealed Preference (Houthakker 1950), we cannot meaningfully say either that an apple produces twice as much utility as an orange for some chooser or that an apple specifically produces 2 utils and an orange 1 util. Von Neumann and Morgenstern (1944) elaborated on this issue when they introduced the independence axiom, but even their approach only specifies utilities to within a linear transform. Ordinality (or perhaps weak cardinality in the case of vNM utilities) is a fundamental feature of economic utility derived by choice.

This is a challenge because the measurements neurobiologists make are fundamentally cardinal and necessarily unique. When neurobiologists measure activity in the nervous system, they typically employ one of two techniques: direct measurements of the times of occurrence of electrochemical action potentials in single nerve cells or indirect measurements of this activity using functional magnetic resonance imaging (fMRI). In either case, neurobiologists measure (with error) a unique and fully cardinal object. If, as all neurobiologists believe, all of human behavior is generated through transformations of this cardinally specified activity, then measurements in the nervous (p. 690) system cannot be measurements of utility itself. In practice, neuroeconomists address this issue by searching for neural signals that are linearly correlated with economically specified expected utilities or which correlate ordinally with less cardinal systems of utility. These signals are typically referred to as subjective values and are defined as real numbers ranging from zero to one thousand (the range of physically possible action potential rates). Mean subjective values are the mean firing rates of specific populations of neurons and are linearly proportional to fMRI measurements of these activities (Heeger and Ress 2002).¹ Note that mean subjective values predict choice stochastically, which reflects the fact that these action potential rates are stochastic.

Notice as well that the features of this stochasticity are reasonably well understood (e.g., Tolhurst et al. 1983; Glimcher 2005). This indicates that subjective value theory will be most closely allied with random utilitytype models from economics, a parallel now being carefully explored by a number of economists. Finally, subjective values, because of their causal relation to action, should be *always* consistent with choice, though stochastically, even when choice is not consistent with utility theory. This is, of course, a critical point that may turn out to have profound implications for welfare theory.

23.2.2 Primary Sensory Transformation and Subjective Value

Whence do these subjective values arise? To some degree they must arise from the algorithmic mechanisms that transform physical events in the outside world into neural activities that guide choice. The fact that all human choosers prefer sugar to quinine, to take one obvious example, must reflect innate properties of the mechanisms by which we sense the external world.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

(p. 691) In fact, these processes have now been widely studied, and the transformations that relate such external properties as sugar or quinine concentration to the internal (or endogenous) representations of these quantities tend to be strictly concave functions that bear a striking resemblance to utility functions in some regards (Fechner 1912; Stevens 1957; Glimcher 2011). This set of observations has led quite naturally to the suggestion that at least one source of concavity in subjective values involves processes that simply learn, through repeated sampling, the action potential rates associated with repeatedly consumed goods.

23.2.3 Learning and Storing the Values of Actions

To understand how neurobiologists think about this process, consider what is typically called a Pavlovian conditioning task. In such a situation a cue, for example a visual stimulus, is presented to a subject and is followed by the delivery of a reward (a positively valued good). After experiencing several times this cue-reward pairing, the subjects begin to exhibit a direct response to the cue, suggesting (sometimes indirectly) that he or she views it as a positive utility shock. This is, of course, the famous salivating dog of Pavlov's experiments (Pavlov 1927).

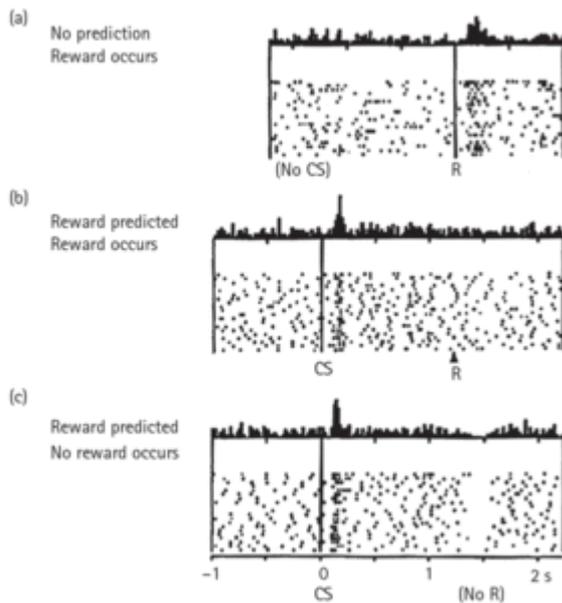
In 1997 Wolfram Schultz and his colleagues measured the mean action potential rates of a class of nerve cells in the base of the brain, midbrain dopamine (DA) neurons, while this process unfolded and found the following: When there was no cue associated with reward, there was a burst of DA firing at the time of reward delivery (figure 23.1). When a cue consistently preceded the reward, the activity of DA neurons at reward delivery would remain at their unique baseline (or zero) firing rate. But under these conditions, the DA neurons would fire at the time of cue presentation. Perhaps even more interesting, they observed a decrease in action potential rates (a uniquely negative number) when an apparently expected reward was omitted. These observations (figure 23.1) suggested that the dopamine signal could be seen as encoding a kind of "utility shock" that related expectations about future positive outcomes to the properties of directly sensed rewards.

This work has triggered enormous interest in the neuroscientific (and neuroeconomic) community. The primary question of interest is how to model the dynamics of neuronal activity that change with experience and how to relate these changes to choice behavior. A number of models have emerged, but the dominant class appears to be the temporal-difference (TD) learning model developed by computer scientists Richard Sutton and Andrew Barto (Sutton and Barto 1981, 1988). This is an algorithmic model that provides clear ties to normative theories of learning.

Consider modeling the dynamics of DA activity in the Pavlovian conditioning task using the TD model. In this model, an agent computes an estimate of value separately at each moment in time within a trial (a trial being a stereotyped multi-period learning problem of

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

finite duration that is repeatedly encountered). Suppose that there are n time points within each trial. The consumption value available at time t within a trial (p. 692)



[Click to view larger](#)

Figure 23.1 Activity of midbrain dopamine neurons in the Pavlovian conditioning task: (A) The DA neurons exhibit an increase in activity immediately after delivery of a reward (unconditioned stimulus, US) when no conditioned stimulus (CS) is presented prior to the reward. (B) After repeated CS-US pairings, when a CS is paired with a US, DA neurons start to show an increase in activity when CS is presented. At US delivery, DA activity remains at baseline level, when US was delivered. (C) At US delivery, DA activity goes below baseline when US is not delivered. Adapted from Schultz et al. (1997).

$$V(t) = E[\gamma^0 r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \dots]$$

where $E[\cdot]$ denotes expectation, $r(t)$ is the physical reward experienced at time t , and $\gamma = [0,1]$ is the discount parameter. By definition, $V(t)$ can be further written as the sum of expected reward at time t and the value at $t+1$ weighted by γ :

$$(23.2)$$

$$V(t) = E[r(t)] + \gamma V(t+1)$$

Hence, the expected reward at time t is the difference between the estimate of value at time t , or $V(t)$, and the estimate of value at $t+1$, $V(t+1)$. The learning agent updates the value estimate by computing the prediction error $\delta(\cdot)$ —the difference between actual reward received at t and expected reward at t , $r(t) - E[r(t)]$ —by the following equation

$$(23.3)$$

$$V_{\text{updated}}(t) = V_{\text{current}}(t) + \alpha \delta(t),$$

is defined by the sum of expected, temporally discounted future rewards during one entire trial:

$$(23.1)$$

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

(p. 693) where the updated value estimate at time t , $V_{updated}(t)$, is the current value estimate at time t , $V_{current}(t)$, plus the weighted prediction error $\alpha\delta(t)$. The weight $\alpha = [0,1]$ assigned to the prediction is a parameter that determines how fast the agent learns and is often referred to as the learning rate. Some normative conditions can be placed on this term, but a discussion of those details lies outside the scope of this presentation (see Sutton and Barto 1981 for details). Given equation (23.2), we can rewrite equation (23.3) as

(23.4)

$$V_{updated}(t) = V_{current}(t) + \alpha[r(t) + \gamma V_{current}(t+1) - V_{current}(t)]$$

The attractive property of this model lies in the fact that the expected reward at any given point in time within a trial is the difference between the value estimate at that time point and the discounted value estimate *at the next time point* per equation (23.2). Because of this property, an update on the value estimate at the time of reward delivery would subsequently affect the value estimate of the preceding time point. Eventually (after a sufficient number of repeated trials), the value of any moment in time propagates back to the point where that future reward can first be anticipated. This backward propagation of expectation thus causes the agent (in at least some environments) to form correct (rational) expectations about all future reward deliveries with a nonzero probability of occurrence on cue presentation. The model thus learns what an economist might call “consumption paths” and responds to any event that signals a change in current or future consumption path with a learning signal. Of course, the goal of this learning is to develop a policy for choosing among possible consumption paths the one that maximizes the discounted sum of future rewards, but the details of the policy element would take us too far from the dopamine neurons that form our principle subject.

What is striking about the TD model is that it quite accurately describes the dynamics of DA neuron activity and how this activity changes over time in the Pavlovian learning tasks. These were the subjects of early DA studies. To summarize those empirical findings, DA neurons at the beginning of an experimental session do not fire when a visual cue is presented that, unknown to the subject, signals a future reward. Instead, these neurons fire when a reward is delivered. After repeated trials in which the cue-reward association consistently happens, DA neurons come to fire at the time of cue presentation, that is, at the time of the utility shock. And this is, of course, exactly what is predicted by TD-class algorithms.

Perhaps it is not surprising that different learning models have been proposed that vary in some ways from this basic template but produce quantitatively similar results. For a discussion of TD-class algorithms and their limitations in explaining DA activity, see Niv and Montague (2008) and Daw and Tobler (2013). For recent advances in modeling reinforcement learning (RL) and, in particular, on dissociating the contributions of model-free RL (e.g., the TD model) and model-based RL to choice and neural activity, see Gläscher et al. (2010) and Daw et al. (2011). Using methods from neoclassical economics, Caplin and Dean (2008) proposed an axiomatic description of (p. 694) this class of

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

learning algorithm which has been more widely influential in economic circles and that has been tested empirically by Caplin et al. (2010). They identified a set of axioms that are necessary and sufficient conditions for representing utility shocks and found that brain activity (mean action potential rates measured with fMRI) in at least one brain region, the ventral striatum, met the axiomatic conditions in a way that could drive this kind of near-normative learning of the subjective values of consumable rewards. This brain area is rich in the neurotransmitter dopamine and receives direct projections from the midbrain DA system.

23.2.4 An Overview of the Network for Valuation

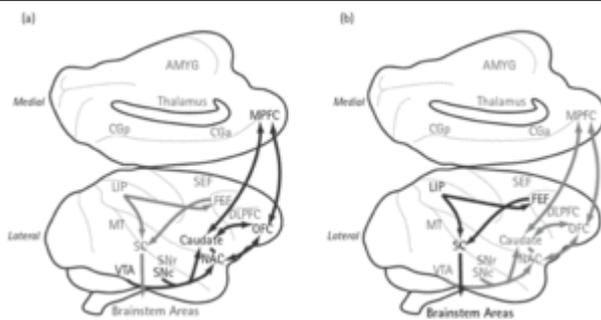
There is now accumulating evidence that neural circuitry including the midbrain DA neurons, mentioned above, and a series of other brain areas including the striatum, the ventromedial prefrontal cortex (vmPFC), and the orbitofrontal cortex (OFC) are involved in the representation of the subjective values of consumable goods and monetary rewards (Padoa-Schioppa and Assad 2006; Lau and Glimcher 2008; Plassmann et al. 2007; Chib et al. 2009; Levy and Glimcher 2011). Neuronal action potential rates in these areas have been widely shown to be both linearly proportional to the utilities (or expected utilities in probabilistic lotteries) and predictive of choice even when subjects behave inconsistently (e.g., Kable and Glimcher 2009). Moreover, a critical feature of these brain areas is that activity elicited by a given option, although stochastic in nature, is independent of what the other available options are (Padoa-Schioppa and Assad 2008). For reviews and meta-analysis on the valuation system see Bartra et al. (2013) and Clithero and Rangel (2014).

23.3 Stage 2: The Choice Mechanism

The choice stage refers to the algorithmic processes that compare the subjective values associated with different objects in a choice set so as to guide the chooser. In principle, the neural circuits involved in the choice process should be able to represent the subjective value associated with each available option in any given choice set. Hence, the choice circuit should receive information about subjective value from the valuation circuit, but in a way restricted to the current choice set (figure 23.2). However, one should remain cautious when thinking algorithmically about the interaction between valuation and choice circuits as purely “feed forward” in the sense that subjective-value signals are passed unidirectionally from the valuation circuit to the choice circuit. In fact, these two systems are heavily and reciprocally interconnected, suggesting that as we come to understand the algorithmic process more completely, the logical separability of these two systems will come to be reduced. Indeed, there is already evidence that

(p. 695)

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions



[Click to view larger](#)

Figure 23.2 Neural circuitry of valuation and choice. (A) Valuation circuitry. (B) Choice circuitry. Structures in these networks and the directions of their connections are highlighted in black. Adapted from Kable and Glimcher (2009).

choice and valuation circuits may interact algorithmically. Although the implications of this for reduced-form models are currently unclear, several models have been proposed. In one model, Padoa-Schioppa (2011) proposed that subjective value is being computed and compared in the space of “goods” in the OFC and

that vmPFC and that these computations are done in a fashion that is independent of the sensorimotor contingencies of choice. After a choice is made, a transformation that maps the chosen good onto the appropriate course of action originates in the OFC and vmPFC and culminates in the planning and execution of motor action. Other models, such as the one proposed by Glimcher (e.g., Louie et al. 2011) and by Shadlen and colleagues (Gold and Shadlen 2007), emphasize value coding in the space of motor actions that are required to obtain the desirable goods. The latter view is reviewed in detail in the next section.

23.3.1 An Overview of the Network for Choice

Our current understanding of the value comparison process at the theoretical, algorithmic, and circuit levels is largely based (for technical reasons) on studies of a well-understood model system of decisionmaking in monkeys. In these *awake-behaving monkey electrophysiology studies*, monkeys choose between two lotteries by making an eye movement (saccade) to one of two possible visual targets that vary in the magnitude or probability of reward, sometimes under conditions of partial information. This model system consists of a heavily interconnected network of brain areas that participate in both the encoding of the subjective value of the lotteries under consideration (p. 696) and the winner-take-all, or argmax, process. The brain areas that participate in this process include the lateral intraparietal area (LIP), the frontal eye field (FEF), and the superior colliculus (SC) (figure 23.2). There is now accumulating evidence that this circuitry is involved in representing the *relative subjective value* (RSV) associated with different options (Platt and Glimcher 1999; Gold and Shadlen 2001; Louie et al. 2011). The current data suggest that at any moment in time neurons in the LIP represent the instantaneous RSV of each lottery (e.g., Dorris and Glimcher 2004; Rorie et al. 2010), a representation that is believed to be derived (algorithmically) from the representation of SV localized in the valuation network, particularly in the vmPFC, OFC, and ventral striatum.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Note that RSV would serve to map SV onto the limited dynamic range of the LIP neurons. Such neurons are limited in number, fire over a roughly 100Hz dynamic range, and have (errors that are drawn from) Poisson-like distributions. 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 failures to deterministically identify the preferred option when selecting among large numbers of possible movements (Louie et al. 2013). In summary, 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 past a fixed numerical threshold, triggering the physical action that instantiates choice.

23.4 Future Directions

Although our current model incorporates many existing data and provides a useful framework for thinking about the neurobiological mechanisms of decision making, the model still lacks descriptions of certain concepts that have been identified by economists and psychologists as critical in the decisionmaking processes. In this section, we seek to expand the current model in two directions. The first direction is motivated by the notion of reference dependence, which, as many psychologists have argued (e.g., Kahneman and Tversky 1979), is a core feature of the valuation process. It has been observed not only in economic decision making but also in a wide variety of judgment tasks. Economists have also begun to incorporate this concept into newly developed models of decision making (Sugden 2003; Koszegi and Rabin 2006). This concept has not received much attention in the neuroeconomics community, but as we mention later, there is a close tie between what dopamine neurons encode and reference dependence. Our goal must therefore be to incorporate reference dependence to the computational algorithm implemented during valuation.

(p. 697) The second direction is motivated by the fact that in many decision scenarios we face, information about probability associated with potential outcomes is not explicitly given and often needs to be estimated by the decision maker. This feature makes these decisions unlike the classical lottery tasks studied in a typical economic laboratory, where probability information is explicitly revealed to the subjects in numerical or graphical form. We introduce recent studies concerning the way information about probability appears to be distorted (more formally: how the independence axiom is violated) in classical economic lottery tasks and in mathematically equivalent “motor” and “perceptual” lottery tasks. Our goal is to expand the standard model to include the

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

violations of the independence axiom in different tasks and to search for the algorithmic sources of this distortion at a neural level.

23.4.1 Incorporating Reference Dependence into Value Computation

Kahneman and Tversky (1979) defined the choice-related “value” (a utilitylike construct) of potential outcomes as gains or losses relative to a reference point. The reference point, as the authors put it, can be viewed as an adaptation level, status quo level, or expectation level defined by the past and present experiences of the decision maker. They argued that the evaluation of monetary changes from this reference point shares many of the mathematical properties of perceptual judgments about such things as sugar concentration, temperature, or brightness, and they noted that many of these perceptual experiences show shifting unique zero-levels that impact perception. For example, it is well known that in a room that is 15°C, it is easier to discriminate 17°C from 19°C than it is to discriminate 27°C from 29°C but the reverse is true when the room is 25°C. In more economic terms, the discriminability of temperature change decreases as the distance from the reference point increases (Weber 1850). This is relevant to economic choice because the neural mechanisms that underlie these phenomena are now fairly well understood and turn out to be ubiquitous. The second feature of Kahneman and Tversky’s reference-dependent value function is that it captures simultaneous risk aversion in the gain domain and risk-seeking in the loss domain (although there are, of course, other ways to capture this, e.g., Friedman and Savage 1948). The third feature of the value function is that it captures aversion to losses. As these authors often put it, losses loom larger than gains, for “the aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount (Kahneman and Tversky 1979, p. 279).” To understand loss aversion, consider a lottery (0.5,\$x; 0.5,-\$x) with a 50-50 chance of gaining \$x or losing \$x. Empirically, it has been observed that most people find this lottery very unattractive. Furthermore, for $x > y \geq 0$, (0.5,\$y; 0.5,-\$y) is often preferred to (0.5,\$x; 0.5,-\$x), according to Kahneman and Tversky. They used these two observations to motivate a value function for losses that is steeper than the value function for gains.

(p. 698) Thus the value function they proposed was

(23.5)

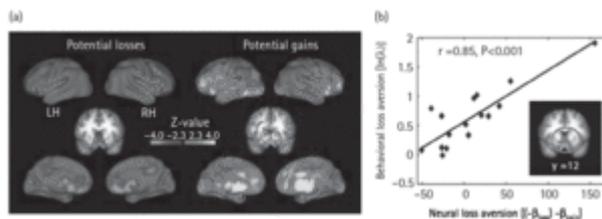
$$v(x) = \begin{cases} x^\alpha, & x > 0 \\ -\lambda(-x)^\beta, & x < 0 \end{cases}$$

where x denotes outcomes relative to the reference point, α and β characterize the curvature of the function in the gain domain and loss domain respectively, and λ is used to represent the degree of loss aversion.

In a seminal paper, Tom et al. (2007) attempted to study the neural basis of loss aversion using fMRI in humans. In their experiment, on each trial the subjects had to decide whether to accept or reject a mixed lottery (0.5,\$x; 0.5,-\$y), a 50-50 chance of winning

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

\$x or losing \$y. The amounts of gain and loss were independently manipulated throughout the experiment. This is critical because in the fMRI analysis, gains and losses could be implemented as separate and uncorrelated parametric regressors of interest. The authors found that regions including the ventromedial prefrontal cortex and the ventral striatum encode both the gains and the losses associated with any given lottery (figure 23.3a). Activity in these regions was positively correlated with gains and negatively correlated with losses. This result was consistent with Kahneman and Tversky's value function if one assumes that the reference point was fixed at zero throughout the experiment for each subject and remained so across all subjects. When treating the value function as linear and only modeling the loss aversion parameter, Tom and colleagues found that their behavioral measure of λ was highly correlated with the neural measure of λ (the asymmetry of the gain and loss regression slopes) in regions including the ventral striatum (figure 23.3b). This pointed out the possibility of a neural representation of a simplified version of the value function in the valuation circuitry



[Click to view larger](#)

Figure 23.3 Neural correlates of gains, losses, and loss aversion - Neural correlates of gains, losses, and loss aversion. **A.** Neural correlates of gains and losses. Two separate, thresholded statistical parametric maps showing regions whose activity was negatively correlated with losses (blue), and regions whose activity was positively correlated with gains (orange/yellow). **B.** Neural measure of loss aversion obtained from ventral striatum (red) was correlated with behavioral measure of loss aversion. Adapted from Tom et al. (2007).

(p. 699) that respects both the unique zero imposed by Kahneman and Tversky's model and the gain-loss asymmetry that they hypothesized.

These advances at the neural-algorithmic level aside, Koszegi and Rabin (2006, 2007) have, in related work, begun to develop a behavioral model of reference-dependent preferences in which they have defined

the reference point as the decision maker's rational beliefs about future outcomes. Instead of a fixed point or status quo, such rational beliefs are stochastic. Let $r = [r_1, \dots, r_k] \in R^k$ denote the set of possible reference-level consumptions and $G(r)$ denote the probability distribution over r . For an option with a set of potential consumption outcomes $c = [c_1, \dots, c_k] \in R^k$ with support $F(c)$, the utility of this option is given by

(23.6)

$$u(F|G) = \int \int u(c|r) dG(r) dF(c)$$

where $u(c|r)$ denotes the utility of a consumption given a reference level. What is unique about these authors' definition is that of $u(c|r)$. It is defined as the sum of two components: the consumption utility $m(c)$ —the pure and absolute pleasure derived from

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

consuming an outcome—and a reference-dependent gain-loss utility $n(c|r)$ similar to Koszegi and Rabin's value function

(23.7)

$$u(c|r) = m(c) + n(c|r).$$

This implies that the utility of consuming an outcome has both absolute and relative components. The absolute component is the consumption utility $m(c)$. The relative component is the reference-dependent gain-loss utility $n(c|r)$. Specifically, $n(c|r)$ is derived from comparing $m(c)$ with the consumption utility of a reference level

(23.8)

$$n(c|r) = u(m(c) - m(r)).$$

Here, $u(\cdot)$ is the gain-loss utility function. In this definition, gain or loss associated with an outcome is based on comparing its consumption utility with that of a reference level. The utility of gain or loss is equivalent to saying “how a person feels about gaining or losing depends in a universal way on the changes of consumption utility associated with such gains or losses” (Koszegi and Rabin 2006, p. 1139).

Several intriguing features in Koszegi and Rabin's formalization of the reference-dependent utility deviate from Kahneman and Tversky's reference-dependent value function. First, unlike Kahneman and Tversky's model where reference point is a fixed value, reference point in Koszegi and Rabin is defined as probability distribution over potential outcomes. Specifically, such probability distribution reflects rational beliefs established on the basis of recent experience. Second, Koszegi and Rabin defined the utility associated with an outcome as the sum of its consumption utility, which is absolute and context-independent, and its gain-loss utility, which is relative and reference-dependent. Third, gain-loss utility is strictly defined on the basis of consumption utility as in equation (23.8). Finally, gain-loss utility associated with a (p. 700) potential outcome depends on the deviation of the consumption utility of that outcome from the consumption utility of the reference level.

These different models thus provide exciting and interesting hypotheses for testing reference-dependent value computations performed at the neural level. We outline several questions of interest and testable hypotheses below.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

1. What is the neurobiological nature of consumption utility? Consumption utility is not required in the Kahneman and Tversky definition, but it is the core of the Koszegi and Rabin utility function. Hence, to be able to test the two models at the neural level, one needs to first establish neurobiological evidence for the presence of consumption utility. The notion that it is absolute, strictly increasing, and represents the hedonic value of consuming a good or reward provides important search criteria for relevant neural signals. However, neurobiologists and economists need to work together to axiomatize the necessary and sufficient conditions for the neural consumption utility. For example, one key question is whether consumption utility necessarily and sufficiently requires one to measure neural signals at the time of consumption. Can neural signals associated with a cue that predicts a reward or neural signals right before consumption of the reward be labeled as consumption utility of that reward?

2. What are the neurobiological natures of gains and losses? Here, the two models provide different hypotheses because of their differences in defining the reference point. Kahneman and Tversky defined gain and loss with respect to something like a fixed status quo that is based on animals' current state. This state can be the current wealth level (e.g., monetary rewards) or the satiation or consumption level (e.g., a primary reward such as food or juice). In contrast, Koszegi and Rabin defined a reward as a gain or loss by comparing its consumption utility with the expectations about rewards established by recent experience with the environment. Such expectations are summarized by a probability distribution on possible rewards. To test the two models, one needs to independently manipulate status quo and expectation. For example, status quo manipulation can be achieved via endowment (for monetary rewards) or selective satiation (for food or juice rewards) prior to an experimental session, while one can manipulate expectation by systematically varying rewards in terms of, (e.g., quantity, size, or type) that the subjects would experience in an experimental session.

3. If the answer to the previous question supports an expectation-based reference point, then one needs to further investigate whether the neural representation of the reference point is stochastic and how reference-dependent value computations take into account such a feature. Koszegi and Rabin provided a specific algorithm for such computations in equation (23.6). In summary, the utility of a potential reward is computed by the sum of the gain-loss utility associated with each possible reward weighted by its probability of occurrence.

4. Finally, a key feature of Koszegi and Rabin's notion of "utility" is that it is composed of two components: the consumption utility of a potential reward and the (p. 701) gain-loss utility of that reward. To test this conjecture, one needs to independently manipulate these two components. For example, we can manipulate consumption utility on a trial-by-trial basis by varying the amount of juice reward associated with an option and, independent of this manipulation, we can manipulate the gain-loss utility by changing the reference-point distribution across different experimental sessions.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

The neurobiological evidence for valuation and choice accumulated since the beginning of the twentyfirst century has only provided some very preliminary results for resolving the questions outlined above. For example, we now know that the valuation system represents the expected value of a cue that predicts the delivery of a reward (e.g., Schultz et al. 1997; Tremblay and Schultz 1999). But do such signals reflect consumption utility? At the time of reward consumption, dopamine neurons and the ventral striatum encode prediction error. This seems to support the notion that the valuation network does not represent consumption utility-like signals. Rather, it encodes the deviation of actual outcome from expectation at the time of consumption. In contrast, there is evidence suggesting that the OFC might be a candidate region for representing consumption utility. Neurons in the OFC are known to selectively respond to different rewards. For example, a population of OFC neurons may respond only to grape juice, while another population may respond only to apple juice. Moreover, these neurons respond (1) at the time when the cue that predicts the specific reward is presented, (2) prior to reward delivery, and (3) immediately after reward delivery (Schultz et al. 2000). These results indicate distinct computation roles in the subcomponents of the valuation system and provide key insights into the formation of the reference point and how it relates to subjective value and choice at the neural algorithmic level. Another example of the challenge we face is the stochastic, expectation-based reference-point hypothesis. Although we know that the valuation network represents the expected value of a stimulus that predicts a reward at the time of stimulus presentation, we know very little about how such expected-value signals would be modulated by the expected value of other stimuli present in the recent past. What we know is that valuation signals in the OFC are context-dependent. It adapts to the range of subjective values associated with rewards experienced in the recent past (Tremblay and Schultz 1999; Padoa-Schioppa 2009).

23.4.2 Decision Under Risk: Neural Representations and Distortions of Different Sources of Probability Information

In this section we focus on the representation of probability information, another key variable when people are engaged in decision making under risk. We begin with a review of past research on probability distortion in standard lottery tasks. Then we review recent behavioral and neural studies investigating how differences in the way information about probability is revealed to choosers affects how it is distorted. Finally, we discuss recent fMRI studies that employ these behavioral approaches to shed light (p. 702) on the neural representation of probability and the neural mechanisms underlying probability distortion (FitzGerald et al. 2010; Wu et al. 2011).

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

23.4.2.1 The Allais Paradox and the Common Consequence Effect

Traditionally, in decision under risk, people choose between lotteries, expressed as

$(p_1, x_1; \dots; p_n, x_n)$, $\sum_{i=1}^n p_i = 1$ where x denotes the outcome and p denotes the probability of

occurrence associated with the outcome (e.g., Von Neumann and Morgenstern 1944). In expected utility theory (EUT), the desirability of a lottery is, of course, specified by its expected utility, the sum of the utility associated with each outcome weighted by its objective probability of occurrence. Formally, von Neumann and Morgenstern (1944) expressed this with their independence axiom: for any lottery L_1 and L_2 , if L_1 is preferred to L_2 , then $\alpha L_1 + (1-\alpha)L_3$ should be preferred to $\alpha L_2 + (1-\alpha)L_3$, and vice versa, $0 < \alpha \leq 1$. yet a wealth of empirical evidence now suggests that human choosers often violate this axiom (e.g., Allais 1953). Consider the following violation of independence from Kahneman and Tversky (1979):

1. Problem 1: Choose between A: (0.33, \$2,500; 0.66, \$2,400) and B: (1, \$2,400)
2. Problem 2: Choose between C: (0.33, \$2,500) and D: (0.34, \$2,400)

In problem 1, 82 percent of subjects chose B , while 83 percent of subjects chose C in problem 2. This violates the independence axiom because adding a common consequence of a 66 percent chance of winning \$2400 to C and D in problem 2 to construct A and B in problem 1 should not alter subjects' choice. If subjects prefer A in problem 1, then they should prefer C in problem 2, and vice versa. If subjects prefer B in problem 1, then they should prefer D in problem 2 and vice versa.

In order to interpret the violation of the independence axiom in decision under risk, Kahneman and Tversky hypothesized that people assign nonlinear weights to probabilities when making risky decisions. The function that characterizes these nonlinear weights is often referred to as the weighting function $\pi(\cdot)$ (e.g., Kahneman and Tversky 1979; Tversky and Kahneman 1992) or the probability weighting function (Wu and Gonzalez 1996). Note that the weighting function is at least somewhat conceptually distinct from the notion of subjective probability advanced by Savage (Savage 1954). In the original prospect theory (Kahneman and Tversky 1979), the decision value (or prospect) over a lottery was hypothesized to be derived by choosers at an algorithmic level by the sum of the value associated with each outcome in the lottery (defined in equation [23.5] in the preceding section) weighted by its decision weight. Note that the decision weight associated with an outcome might not be the weight associated with its probability of occurrence. In the original prospect theory, this is the case. In the subsequent version, cumulative prospect theory, however, it is not. For example, in the original theory (Kahneman and Tversky 1979), the prospect of a lottery with two nonzero outcomes $(p, \$x; q, \$y)$, $x > y \geq 0$, $p + q = 1$ is $v(y) + \pi(p)[v(x) - v(y)]$ where $\pi(\cdot)$ is the weighting function and $v(\cdot)$ is the value function. In cumulative prospect theory, Tversky and Kahneman (1992) incorporated (p. 703) rank-dependence into this framework such that the prospect of a lottery became $\pi(p)v(x) + [\pi(p+q) - \pi(p)]v(y)$.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Broadly speaking, when models of this type are parameterized in any number of ways (e.g., Tversky and Kahneman 1992; Gonzalez and Wu 1999; Wu and Gonzalez 1996), $\pi(\cdot)$ is found to be well described by an inverse-S-shaped function, that is, a function that is concave at small probabilities and convex at moderate-to-large probabilities. Prelec (1998) developed an axiomatic foundation for these functions, deriving axioms that were necessary conditions for the probability weighting function proposed by Kahneman and Tversky (Kahneman and Tversky 1979; Tversky and Kahneman 1992). Since that time a host of studies have examined human choice behavior with prospect theory and the Prelec function and almost universally found that parameterizations indicate this inverse S-shaped structure for the Prelec function.

Notice, however, that nearly all of these parameterizations have been based on data gathered from human subjects in what might be called classical lottery tasks. In these kinds of experiments information about probability distributions on possible outcomes are explicitly described in numerical or graphical form to subjects who then express their preferences. What is worth noting is that this kind of decision making scenario describes only a subset of the risky decision making scenarios we face in everyday life. What is surprising is that a growing body of evidence now suggests that the parameterized probability weighting function extracted outside classical lottery tasks looks quite different from that extracted in these more classical situations.

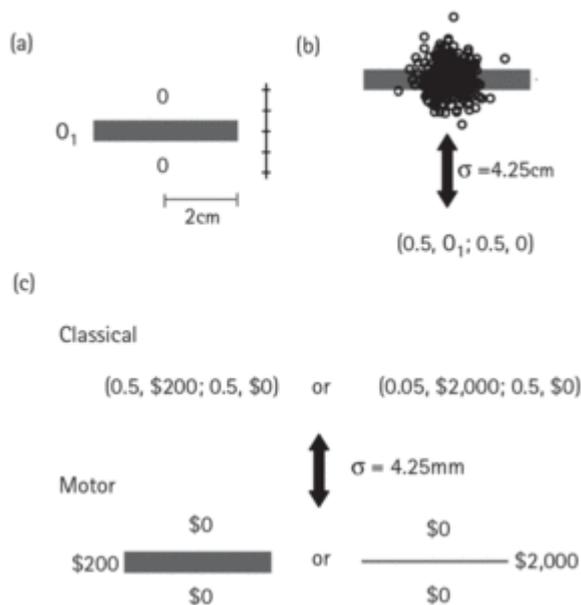
23.4.2.2 Motor Decision Making and Probability Distortion

A baseball player deciding whether to swing a bat at an incoming ball is not given explicit numerical estimates of the probability of producing a base hit, a home run, or a miss. In situations like these, decision makers typically estimate probability based on experience. And of course these estimates must take into account a number of sources of variance including errors in neurobiologically derived estimates of the speed and position of the ball and estimates of the movement error associated with a plan to swing the bat toward a fixed location in space and time. How humans estimate probability in these situations, and how well they do it, are currently under intense investigation. There is accumulating evidence that in these domains humans achieve near-normative performance, taking into account noise coming from the perceptual and motor systems in a way that seems to obey the independence axiom (Geisler 1989; Trommershäuser et al. 2003b,a; Körding and Wolpert 2004; Najemnik and Geisler 2005; Tassinari et al. 2006; Battaglia and Schrater 2007; Dean et al. 2007). These findings present a sharp contrast to results from economic decision under risk, in which information about probability is explicitly stated. Few studies, however, directly compared decision making in classical lottery tasks with perceptual or motor tasks.

To formally compare decision making under different modalities, Wu et al. (2009) developed a method for translating a classical lottery to a mathematically equivalent “motor” (or movement-based) lottery (figure 23.4). They then asked the subjects to perform identical sets of incentive-compatible classical and motor lotteries. Information

(p. 704)

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions



[Click to view larger](#)

Figure 23.4 Construction of a motor lottery task. (A) In a rapid point task, the subjects were trained to hit a single target within a very short time window (usually < 0.7 s). (B) Here we superimposed the distribution of movement end points from an actual subject. We verified that the distribution of movement end points is bivariate Gaussian and characterized motor noise by the estimated standard deviation (σ) of the distribution. For this subject ($\sigma = 4.25$ mm), this target was equivalent to a lottery $(0.5, O_1; 0.5, 0)$. (C) Given the motor noise separately estimated for each subject, we could translate a binary classical lottery task, for example, choosing between $(0.5, \$200; 0.5, \$0)$ and $(0.05, \$2,000; 0.5, \$0)$, to a mathematically equivalent motor lottery task. In a later, decision-making phase of the experiment, we asked the subjects to choose between lotteries in classical tasks and in motor tasks. We emphasized that subjects during the motor lottery task only indicated which lottery they preferred. They did not execute any pointing movement during the decision tasks. Adapted from Wu et al. (2011).

about the probability of winning in the motor lotteries depended on both the size of the target that the subjects had to successfully hit and the intrinsic variability of the subject's movement (which subjects had to learn from experience).

In an initial training session aimed at teaching subjects about their movement variability, the subjects were asked to repeatedly and quickly (within < 0.7 seconds) hit with their finger a rectangular target that appeared on a computer touchscreen. The size of that target was varied independently to control the probability of "winning" a given lottery (or trial). Hitting the target (p) resulted in a small monetary gain (v), and hitting anywhere else on the screen ($1 - p$) won nothing (0).² The critical idea of these lotteries,

(p. 705) what makes them lotteries, is that subjects do not have perfect control over their movements owing to intrinsic noise in the motor system introduced by the very short time window. After extensive training under the same time constraint, the motor noise often becomes stable at a within-subject level (Trommershäuser et al. 2003a,b). For the experimenter, this means that any binary lottery can be constructed once the movement variability, or "motor noise" has been measured, although the approach assumes that subjects can estimate the probability of their hitting a target given knowledge of their own motor noise.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

The questions raised by this line of research are whether decisions made in this way show different patterns of rationality, particularly with regard to the independence axiom, and whether they show different risk preferences. In fact, Wu and colleagues found that subjects violated the independence axiom in motor lotteries just as much as they violated the axiom in the classical lottery task. The pattern of violation was markedly different, however. Their parametric analysis suggested that this difference could be attributed to a change in the probability weighting function. Rather than the typical overweighting of small probabilities and underweighting of moderate-to-large probabilities, subjects in the motor task tend to *underweight small probabilities and overweight moderate-to-large ones*.

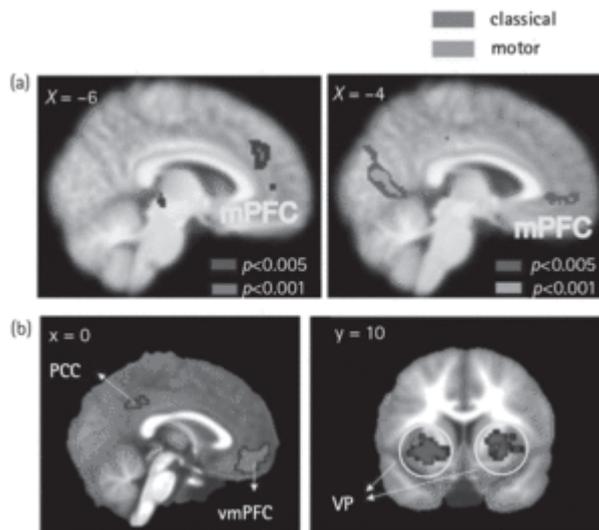
This pattern of inferred probability distortion is of particular interest because the ability of subjects to estimate the probability of reward in the motor lottery task depends on the subjects' previous experience hitting targets on the touchscreen. Hence, it is an experience-based lottery task in which knowledge about probability associated with hitting motor targets was established by experience. Behavioral studies have begun to reveal that, as opposed to overweighting rare events when probability information is revealed explicitly, people tend to underweight rare events when information about probability associated with rare monetary gains is acquired by sampling experience (Hertwig et al. 2004; Jessup et al. 2008; Ungemach et al. 2009); for a review, see Hertwig and Erev (2009). This difference is often called the description-experience gap. Despite accumulating evidence suggesting the existence of such difference at the behavioral level, very few studies (FitzGerald et al. 2010; Wu et al. 2011) have directly compared the neural representation of probability in decision under risk when information about probability comes from different sources, for example, when it is described explicitly versus when it is learned from experience. That seems important because the neural measurements might give insight into the algorithmic constraints that shape these two classes of decision.

Neurobiological studies of decisions involving risk and uncertainty have identified the neural systems that correlate with these economic variables (Platt and Huettel 2008). In reinforcement learning tasks, dopamine neurons have been shown to represent the probability of both reward and risk (defined as the variance) associated with reward-predicting stimuli (Fiorillo et al. 2003). In humans, fMRI studies have reported that the striatum, the anterior insula, the medial prefrontal cortex, lateral prefrontal cortex, and posterior parietal cortex represent these variables as well (FitzGerald et al. (p. 706) 2010; Hsu et al. 2009; Huettel et al. 2005, 2006; Knutson et al. 2005; Paulus et al. 2003; Preusschoff et al. 2006; Tobler et al. 2008; Wu et al. 2011).

Unfortunately, the neural results available today are not entirely consistent. Wu and colleagues (2011) found that the medial prefrontal cortex (mPFC) encodes "probability weight" in a classical lottery task and in a motor lottery task. In that study, mPFC activity only showed correlation with probability of reward in the motor lottery but was not correlated with physical size of the target in a size judgment task in which the physical properties of the stimuli were identical to those in the motor lottery task (figure 23.5a).

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Together, the results suggest a convergence of two mechanisms for probability encoding and pushes neuroeconomists to search upstream (in the algorithmic sense) for these two probability encoding mechanisms. Others, however, have found that activity in the dorsolateral prefrontal cortex (Tobler et al. 2008) and in the striatum (Hsu et al. 2009) is correlated with probability distortion in similar tasks. FitzGerald and colleagues (2010) had subjects choose between lotteries for which the probability of reward associated with one lottery was revealed explicitly and the probability of reward associated with the other was acquired by sampling experience. They found that at the time when the subjects were asked to choose between the lotteries, activity in the medial



[Click to view larger](#)

Figure 23.5 Comparisons of neural correlates of different sources of probability in described and in experienced lottery tasks. (A) Neural correlates of probability weight in classical (described) lottery tasks (red) and in motor lottery task (green); mPFC: medial prefrontal cortex. (B) Neural correlates of subjective value (estimated probability of reward) in classical (described) lottery tasks (red) and in experienced or learned lottery tasks (blue); vmPFC: ventromedial prefrontal cortex; PCC: posterior cingulate cortex; VP: ventral putamen. Adapted from Wu et al. (2011) and FitzGerald et al. (2010).

(p. 707) prefrontal cortex and another area (the posterior cingulate cortex) was correlated with the decision-value of the lottery learned by experience, while activity in yet another brain area, the ventral putamen, was correlated with the decision-value of a classically described lottery (figure 23.5b).

In summary, since Allais (1953), converging evidence at the behavioral level suggests that humans distort information about probability in decision under risk in a highly characteristic manner. That is, people tend to overweight small

probabilities and underweight moderate-to-large ones. However, such a pattern of distortion has been found primarily in situations where information about outcomes and their associated probabilities of occurrence are described explicitly in numerical or graphical form to the subjects. This line of results was recently challenged, however, by two other lines of research. The first pointed out the description-experience gap in decision making and found that the pattern of distortion was markedly different in situations where people acquired information about probability by sampling experience (Barron and Erev 2003; Hertwig et al. 2004; Jessup et al. 2008; Ungemach et al. 2009). The second challenge came from research about perception and action suggesting that people are near-optimal EUT maximizers in perceptual and motor tasks that are formally equivalent to decision making under risk (Geisler 1989; Trommershäuser et al. 2003b,a;

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Körding and Wolpert 2004; Najemnik and Geisler 2005; Tassinari et al. 2006; Battaglia and Schrater 2007; Dean et al. 2007). A direct comparison of the classical descriptive lottery tasks against mathematically equivalent motor lottery tasks (Wu et al. 2009) suggested that, similar to the findings in the description-experience gap, people tend to underweight small probabilities and overweight large probabilities in motor tasks. Together, these results pointed out that in more naturalistic settings where probability information has to be inferred, either from experience or from knowledge about the environment or the nervous system, the way such information is distorted is markedly different from descriptive scenarios. Neuroimaging studies have just begun to address these phenomena observed in behavior by investigating the underlying neural correlates of probability distortion. We believe that the ultimate success or failure of the contribution of neuroeconomics to understanding probability representations would depend on identifying the algorithms involved in probability distortion at the neural level and how their neural implementations would tie to current understanding about reward learning systems, particularly the midbrain dopamine system.

In addition, fundamental questions such as how the reference-dependent utility of a reward is integrated with probability weight at the time of choice and the computational algorithm for such integration also need to be addressed. In doing so it is important to first separately examine the neural systems involved in computing the reference-dependent utility and probability, in either choice or nonchoice situations. Once the neural representations in these domains have been established, one can then create a decision task that independently manipulates rewards and probability. This will then allow for the possibility of investigating the integration computations and will allow for an assessment of how neural systems involved in either kind of computation (p. 708) (reference-dependent utility or probability weight) contribute to the systems that represent the integrated products.

23.5 Concluding Remarks

Since Samuelson (1938), neoclassical economics has largely relied on revealed-preference theory in conducting economic analysis. What is central in this approach is that the relation between internal representations of subjective value and choice behavior is established on the basis of axioms about revealed preference. We began the chapter by arguing that subjective value is represented by neuronal activity in units of action potential and that choosers behave as they do *because*, not *as if*, there is an internally consistent cardinal representation of subjective value. Armed with this assumption, neuroeconomists have accumulated a wealth of neurobiological evidence related to decision computations. As a result of this rapid accumulation of findings, a standard neurobiological model emerged. This emerging model proposed that the neurobiological mechanisms of decision making could be subdivided into two components, valuation and choice. Valuation refers to computing subjective value for each option available in a

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

choice set. The network involved in this component includes the dopamine neurons in the midbrain, the medial prefrontal cortex, the orbitofrontal cortex, and the striatum. We reviewed evidence suggesting SV computation and algorithms for learning and representing SV in the valuation circuitry. The second component, referred to as choice, is involved in integrating SV computed in the valuation circuitry so as to compute relative subjective value associated with different options. The RSV serves to guide action planning for the animal to obtain the desired option. In the framework of saccadic decision making, the choice network involves the lateral intraparietal area, the frontal eye field, and the superior colliculus. To summarize, this emerging model placed a particular emphasis on identifying algorithms involved in decision-related computations at different stages and their implementations at the system level.

As part of an attempt to expand the model at both the algorithmic and the neural-implementation levels, we identified two novel future directions for value computations raised by economic theorists and psychologists. The first direction was related to the notion of the reference point. Different models in economics and psychology for how to formally model reference points have been proposed. The models in essence provided precise algorithms for reference-dependent value computations that could be tested and compared at the neural level. The theoretical notion of reference point proposed by Koszegi and Rabin (2006) drew particular attention to neurobiologists. In their view, the reference point is the rational belief, represented by probability distribution over possible outcomes, established by experience from the recent past. This is closely related to what neurobiologists found in midbrain dopamine neurons in reward learning tasks, and several laboratories have begun to investigate the algorithmic form of reference-dependent computations in the valuation network.

(p. 709) The second direction was related to different sources of probability information and how they might differentially affect choice in decision under risk. Classical results indicated that humans tend to distort probability information, but these results tended to be replicated in situations where information about probability was explicitly revealed to the chooser. More recently, accumulating evidence suggests that the way probability information is distorted is different from descriptive scenarios in motor decisions and in scenarios where knowledge of probability was acquired via sampling experience. However, the neural correlates associated with different sources of probability information and the precise neural mechanisms responsible for differences in probability distortion remain largely unknown and are under investigation. We believe that the differences lie in the systems representing probability information and the algorithms they are capable of in computing probability information. Hence, the key contribution that neuroeconomics could possibly make to this field would be to identify the systems involved in representing different sources of probability information and the potential differences in the algorithms implemented in those systems.

References

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Allais, M. (1953). Le comportement de l'homme rationnel devant le risqué: Critique des postulats et axiomes de L'école Américaine. *Econometrica* 21, 503–546.

Barron, G., and I. Erev (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making* 16, 215–233.

Bartra, O., J. McGuire, and J. Kable (2013). The valuation system: A coordinate-based meta-analysis of bold fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76, 412–427.

Battaglia, P., and P. Schrater (2007). Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *Journal of Neuroscience* 27, 6984–6994.

Caplin, A., and M. Dean (2008). Axiomatic neuroeconomics. In P. Glimcher, C. Camerer, E. Fehr, and R. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*, pp. 21–31. Academic Press.

Caplin, A., M. Dean, P. Glimcher, and R. Rutledge (2010). Measuring beliefs and rewards: A neuroeconomic approach. *Quarterly Journal of Economics* 125, 923–960.

Chib, V., A. Rangel, S. Shimojo, and J. O'Doherty (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience* 29, 12315–12320.

Clithero, J., and A. Rangel (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social, Cognitive and Affective Neuroscience* 9, 1289–1302.

Daw, N., S. Gershman, B. Seymour, P. Dayan, and R. Dolan (2011). Model-based influences on humans choices and striatal prediction errors. *Neuron* 69, 1204–1215.

Daw, N., and P. Tobler (2013). Value learning through reinforcement: The basics of dopamine and reinforcement learning. In P. Glimcher and E. Fehr (Eds.), *Neuroeconomics: Decision Making and the Brain*, (2nd ed.) pp. 283–298, Elsevier.

Dean, M., S.-W. Wu, and L. Maloney (2007). Trading off speed and accuracy in rapid, goal-directed movements. *Journal of Vision* 7, 1–12.

(p. 710) Dorris, M., and P. Glimcher (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Fechner, G. (1860/1912). Elemente der psychophysik. In B. Rand (Ed.), *The Classical Psychologists*, pp. 562–572. Houghton Mifflin.

Fiorillo, C., P. Tobler, and W. Schultz (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

FitzGerald, T., B. Seymour, D. Bach, and R. Dolan (2010). Differentiable neural substrates for learned and described value and risk. *Current Biology* 20, 1823–1829.

Friedman, M., and L. Savage (1948). The utility analysis of choices involving risk. *Journal of Political Economy* 56, 279–304.

Geisler, W. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review* 96, 267–314.

Gläscher, J., N. Daw, P. Dayan, and J. O'Doherty (2010). State versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595.

Glimcher, P. (2005). Indeterminacy in brain and behavior. *Annual Review of Psychology* 56, 25–56.

Glimcher, P. (2010). *Foundations of Neuroeconomic Analysis*. Oxford University Press.

Glimcher, P. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences USA* 108, 15647–15654.

Gold, J., and M. Shadlen (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences* 5, 10–16.

Gold, J., and M. Shadlen (2007). The neural basis of decision making. *Annual Review of Neuroscience* 30, 535–574.

Gonzalez, R., and G. Wu (1999). On the shape of the probability weighting function. *Cognitive Psychology* 38, 129–166.

Hayden, B., J. Pearson, and M. Platt (2009). Fictive reward signals in the anterior cingulate cortex. *Science* 324, 948–950.

Heeger, D., and D. Ress (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience* 3, 142–151.

Hertwig, R., G. Barron, E. Weber, and I. Erev (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science* 15, 534–539.

Hertwig, R., and I. Erev (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences* 13, 517–523.

Houthakker, H. (1950). Revealed preference and the utility function. *Economica* 17, 159–174.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Hsu, M., I. Krajbich, C. Zhao, and C. Camerer (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *Journal of Neuroscience* 29, 2231–2237.

Huettel, S., A. Song, and G. McCarthy (2005). Decisions under uncertainty: Probabilistic context influences activity of prefrontal and parietal cortices. *Journal of Neuroscience* 25, 3304–3311.

Huettel, S., C. Stowe, E. Gordon, B. Warner, and M. Platt (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.

Jessup, R., A. Bishara, and J. Busemeyer (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science* 19, 1015–1022.

Kable, J., and P. Glimcher (2009). The neurobiology of decision: Consensus and controversy. *Neuron* 63, 733–745.

(p. 711) Kahneman, D., and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47, 263–291.

Kim, S., J. Hwang, and D. Lee (2008). Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* 59, 161–172.

Knutson, B., J. Taylor, M. Kaufman, R. Peterson, and G. Glover (2005). Distributed neural representation of expected value. *Journal of Neuroscience* 25, 4806–4812.

Körding, K., and D. Wolpert (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.

Koszegi, B., and M. Rabin (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics* 121, 1133–1166.

Koszegi, B., and M. Rabin (2007). Reference-dependent risk attitudes. *American Economic Review* 97, 1047–1073.

Lau, B., and P. Glimcher (2008). Value representations in the primate caudate nucleus. *Neuron* 58, 451–463.

Levy, D., and P. Glimcher (2011). Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience* 31, 14693–14707.

Logothetis, N. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878.

Logothetis, N., J. Pauls, M. Augath, T. Trinath, and A. Oeltermann (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Louie, K., L. Grattan, and P. Glimcher (2011). Reward value-based gain control: Divisive normalization in parietal cortex. *Journal of Neuroscience* 31, 10627–10639.

Louie, K., M. Khaw, and P. Glimcher (2013). Normalization is a general mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences USA* 110, 6129–6144.

Najemnik, J., and W. Geisler (2005). Optimal eye movement strategies in visual search. *Nature* 434, 387–391.

Niv, y., and P. Montague (2008). Theoretical and empirical studies of learning. In P. W. Glimcher, C. F. Camerer, E. Fehr, and R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*, pp. 331–351. Academic.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: A good-based model. *Annual Review of Neuroscience* 34, 333–359.

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

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

Paulus, M., C. Rogalsky, A. Simmons, J. Feinstein, and M. Stein (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* 19, 1439–1448.

Pavlov, I. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press.

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

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

(p. 712) Platt, M., and S. Huettel (2008). Risky business: The neuroeconomics of decision making under uncertainty. *Nature Neuroscience* 11, 398–403.

Prelec, D. (1998). The probability weighting function. *Econometrica* 66, 497–527.

Preuschoff, K., P. Bossaerts, and S. Quartz (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Rorie, A., J. Gao, J. McClelland, and W. Newsome (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS ONE* 5(e9308).

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Samuelson, P. (1938). A note on the pure theory of consumer's behavior. *Economica* 5, 61-71.

Savage, L. (1954). *The Foundation of Statistics*. Wiley.

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

Schultz, W., L. Tremblay, and R. Hollerman (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex* 10, 272-283.

Stevens, S. (1957). On the psychophysical law. *Psych Rev* 64, 153-181.

Sugden, R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory* 111, 172-191.

Sutton, R., and A. Barto (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review* 88, 135-170.

Sutton, R., and A. Barto (1988). *Reinforcement Learning: An Introduction*. MIT Press.

Tassinari, H., T. Hudson, and M. Landy (2006). Combining priors and noisy visual cues in a rapid pointing task. *Journal of Neuroscience* 26, 10154-10163.

Tobler, P., G. Christopoulos, J. O'Doherty, R. Dolan, and W. Schultz (2008). Neuronal distortions of reward probability without choice. *Journal of Neuroscience* 28, 11703-11711.

Tolhurst, D., J. Movshon, and A. Dean (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research* 23, 775-785.

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

Tremblay, L., and W. Schultz (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704-708.

Trommershäuser, J., L. Maloney, and M. Landy (2003a). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A* 20, 1419-1433.

Trommershäuser, J., L. Maloney, and M. Landy (2003b). Statistical decision theory and tradeoffs in the control of motor response. *Spatial Vision* 16, 255-275.

Tversky, A., and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

Ungemach, C., N. Chater, and N. Stewart (2009). Are probabilities overweighed or underweighted, when rare outcomes are experienced (rarely)? *Psychological Science* 20, 473–479.

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

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.

Weber, E. (1850). Der tastsinn und das gemeingefühl. In W. R. Ernst Weber (Ed.), *Handwörterbuch der Physiologie*, Vol. 3, pp. 481–588. Vieweg, Brunswick, Germany.

(p. 713) Wu, G., and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science* 42, 1676–1690.

Wu, S.-W., M. Delgado, and L. Maloney (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences USA* 106, 6088–6093.

Wu, S.-W., M. Delgado, and L. Maloney (2011). The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *Journal of Neuroscience* 31, 8822–8831.

Notes:

⁽¹⁾ A brief introduction to two principal techniques often used to measure activity in the brain and the signals they measure is in order here. (A) “Single-neuron recording” measures activity from a single neuron by placing a tiny electrical probe very near to a targeted neuron. This technique measures the electrochemical state of a single neuron. Because the probe must be inserted into the brain, however, the technique cannot be applied to humans except in rare surgical environments. (B) Functional magnetic resonance imaging typically measures a physicochemical signal called the “blood-oxygen-level-dependent (BOLD) response” using an MRI scanner. The BOLD signal reflects changes in blood flow, blood volume, and blood oxygenation caused by changes in the metabolic demands neurons. Because changes in metabolic demand closely parallel the electrochemical states of nearby neurons, the BOLD signal is an indirect measure of neuronal activity. This measurement technique is entirely noninvasive and thus has revolutionized the study of brain and behavior in humans. However, the precise mapping between neural activity and the BOLD signal has not yet been specified with complete accuracy. To a first approximation, fMRI yields a measurement that is a linear transform of mean action potential rates across a population of nerve cells over a spatial extent of several millimeters and over a period of several seconds (Heeger and Ress 2002). We caution, however, that the precise mapping of the fMRI signal to underlying activity is a subject of intense current scrutiny. There is no significant doubt that this

The Emerging Standard Neurobiological Model of Decision Making: Strengths, Weaknesses, and Future Directions

signal is monotonic with mean action potential rates, but it may well be that it maps more linearly to aggregate membrane depolarization than to the action potential rates derived physically from this quantity. See Logothetis et al. (2001); Heeger and Ress (2002); Logothetis (2008) for more about this issue.

(²) We note that, during training, hitting the screen after the 0.7-second time limit resulted in a monetary loss five times greater than the gain. This manipulation served to train the subjects to respond within 0.7 seconds. In practice, the probability of this occurring in a trained subject is negligible.

Shih-Wei Wu

Shih-Wei Wu, Institute of Neuroscience, National Yang-Ming University, Taipei, Taiwan

Paul W. Glimcher

Paul W. Glimcher, Center for Neural Science, New York University, New York, USA
Institute for the Interdisciplinary Study of Decision Making, New York University, New York, USA

