

# Decision Making in Recurrent Neuronal Circuits

Xiao-Jing Wang<sup>1,\*</sup>

<sup>1</sup>Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA

\*Correspondence: [xjwang@yale.edu](mailto:xjwang@yale.edu)

DOI 10.1016/j.neuron.2008.09.034

Decision making has recently emerged as a central theme in neurophysiological studies of cognition, and experimental and computational work has led to the proposal of a cortical circuit mechanism of elemental decision computations. This mechanism depends on slow recurrent synaptic excitation balanced by fast feedback inhibition, which not only instantiates attractor states for forming categorical choices but also long transients for gradually accumulating evidence in favor of or against alternative options. Such a circuit endowed with reward-dependent synaptic plasticity is able to produce adaptive choice behavior. While decision threshold is a core concept for reaction time tasks, it can be dissociated from a general decision rule. Moreover, perceptual decisions and value-based economic choices are described within a unified framework in which probabilistic choices result from irregular neuronal activity as well as iterative interactions of a decision maker with an uncertain environment or other unpredictable decision makers in a social group.

## Introduction

Decision making is a cognitive process of choosing an opinion or an action among a set of two or more alternatives, with several defining characteristics. First, choice alternatives are not merely reflexive responses but involve goal-directed actions for which the expected outcomes can be assessed to some degree and taken into account in a decision process. Second, a hallmark of controlled decisions is the process of information accumulation and deliberate consideration. Third, risk is inherent in virtually all interesting decisions; indeed, one can say that the essence of decision making is to make a right choice in the face of uncertainty about its long-term consequences.

Aside from momentous decisions, such as those on war and peace, marriage, or judicial verdict, decision making pervades all aspects of flexible behavior in our daily lives. We decide on a goal, then make a series of choices in order to achieve that goal. Voluntary selective attention, in the sense of purposefully directing sensory processing, relies on decisions about what in the external world are the most relevant, behaviorally, at any moment. Perception relies on judgments about the sensory scene, where conflicting and ambiguous input signals need to be detected, identified, and discriminated. Given the sensory information, an organism is faced with the task of selecting a course of action among available options, based on expected outcomes and associated risks of these actions. Choice preference and strategies must be flexibly adaptive when the environment changes or when the outcome depends on all the choices of interacting decision makers in a social setting.

In spite of a central role of decision making in cognition, little was known about its neuronal underpinning until recently. The current decade has witnessed a surge of interest and activity in this area, thanks to a confluence of animal behavioral physiology, human brain imaging, theory, and neural circuit modeling. In particular, neurophysiologists have begun to undertake studies of behaving nonhuman primates in a variety of decision tasks, including perceptual discrimination (Shadlen and Newsome, 1996,

2001; Romo and Salinas, 2001; Roitman and Shadlen, 2002; Romo et al., 2004; Heekeren et al., 2008), target selection (Hanes and Schall, 1996; Schall, 2001, 2004; Cisek and Kalaska, 2005; Scherberger and Andersen, 2007), economic choice behavior (Platt and Glimcher, 1999; Sugrue et al., 2004, 2005; Padoa-Schioppa and Assad, 2006), and competitive games (Barraclough et al., 2004; Dorris and Glimcher, 2004; Glimcher, 2003; Lee, 2008). These experiments have uncovered neural signals at the single-cell level that are correlated with specific aspects of decision computation. Yet, in the mammalian brain, a decision is not made by single cells, but by the collective dynamics of a neural circuit. How are the observed neural signals generated? What are the properties of a local cortical area (e.g., in the prefrontal or posterior parietal cortex) that enable it to subserve decision computations, in contrast to early processing in primary sensory areas? How can one establish the chain of causation linking molecules, circuits to decision behavior?

In close interaction with experiments, realistic neural circuit modeling provides a valuable tool to address these fundamental issues. Biophysically based models can help bridge different levels of description, probing cellular and network mechanisms that underlie the observed neural spiking activity on one hand and account for the performance at the behavioral level on the other hand. Moreover, decision computations depend on cortical circuits endowed with an abundance of positive and negative feedback loops, the behavior of which is not readily predictable by intuition alone. Theory of nonlinear dynamical systems offers a mathematical framework for describing and predicting the behavior of such strongly recurrent neural systems.

Cellular-level modeling has proven tremendously useful for understanding the behavior of single synapses, single neurons, and sensory processing such as the mechanism of orientation selectivity in primary visual cortex. On the other hand, cognitive processes like decision making have largely been described by abstract mathematical models. The situation has been changing in recent years. Biophysically based spiking network models

have been developed and applied to various experimental paradigms, including perceptual tasks that involve both decision making and working memory, action selection and preparation, learning flexible sensorimotor associations, and reward-based economic choice behaviors such as foraging or interactive games. These models are similar in their basic assumptions. Recurrent synaptic excitation is assumed to be sufficiently strong to generate multiple self-sustained stable states of neural populations, which are mathematically referred to as "attractor states." Reverberating excitation is instantiated by a slow cellular process, giving rise to long ramping of neural activity over time. Therefore, the network's behavior is not necessarily dominated by steady states (representing categorical choices), but slow transient dynamics provide a neural mechanism for temporal accumulation of information. On the other hand, feedback inhibition implements competitive dynamics underlying the formation of a categorical choice. Furthermore, highly irregular spiking activity of neurons plays a key role in generating stochastic choice behavior. Finally, reward-dependent synaptic plasticity implements learning that reflects outcomes of past choice history, leading to choice adaptation in a changing environment or in interaction with other decision makers in a social setting. Because of their commonalities, these models will be collectively referred to as "the recurrent neural circuit model."

This article reviews recent electrophysiological findings from this computational perspective. The focus will be on basic computations: (1) accumulation of evidence (what is the cellular basis of temporal accumulation of information?), (2) formation of a categorical choice (what is the termination rule for a deliberation process in neuronal terms?); (3) reward-based adaptation (are values of alternative responses learned by neurons or synapses; what may be the underlying plasticity process?); (4) stochasticity inherent in choice behavior (how is the uncertainty about the world represented in the brain? what are the intrinsic neuronal sources of randomness in choice behavior?). These computations are at the core of many decision processes, regardless of their diversity and complexity; therefore, understanding their neuronal underpinnings is essential for a biological foundation of decision making.

### Neuronal Processes in the Frontoparietal Circuitry Underlying Accumulation of Information and Categorical Choice

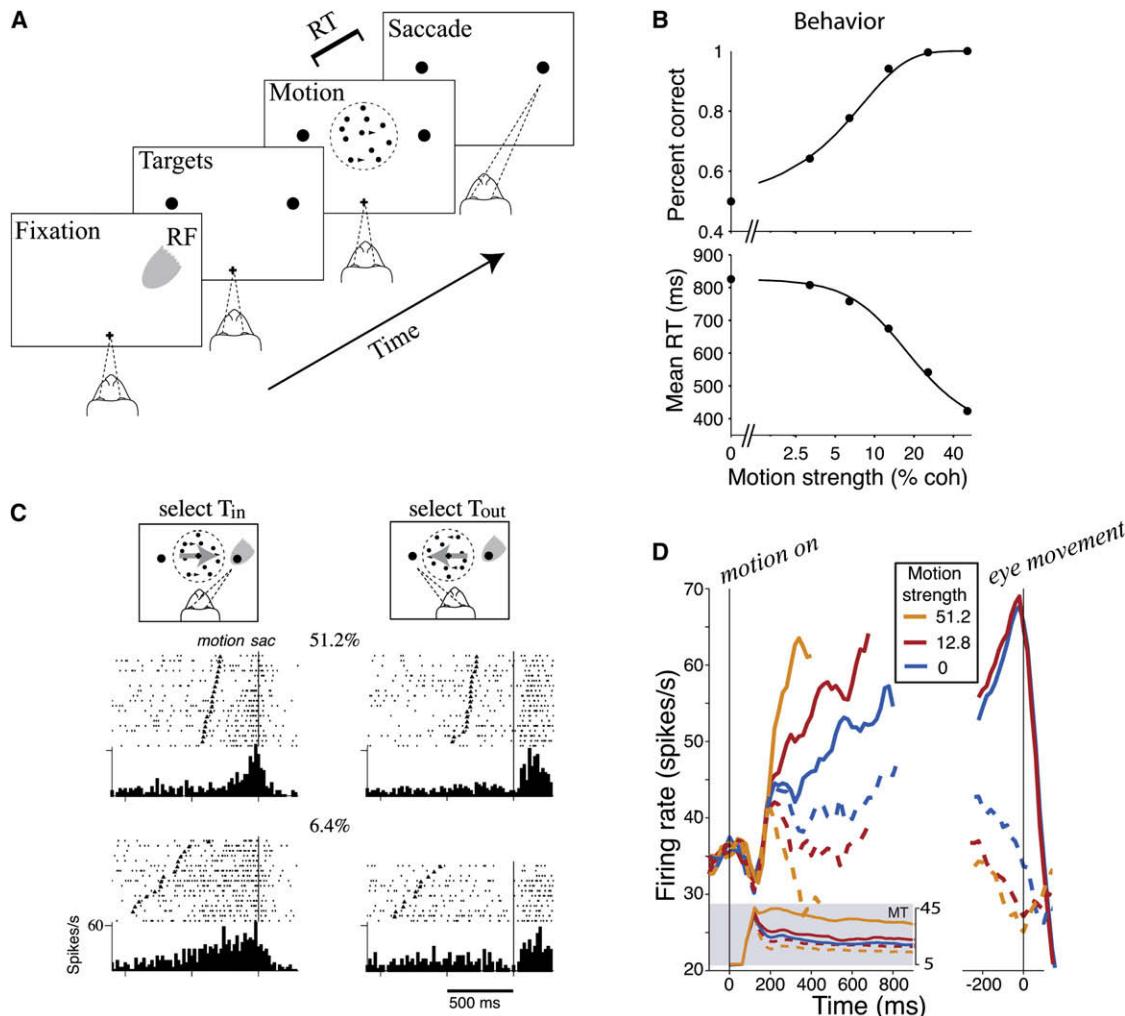
A hallmark of deliberate decision making is time integration, a process that enables us to accumulate evidence in favor of or against alternative propositions and mull over choice options. Although we are capable of producing rapid responses, rushed decisions may yield ill effects. There is often a tradeoff between speed and accuracy: performance improves with slower response times (Wickelgren, 1977). Moreover, we typically take a longer time to ponder a more difficult decision, when information provided by the external world is conflicting or when there are numerous options to consider (Hick, 1952; Vickers, 1970).

At the behavioral level, reaction time (RT) measurements have provided a powerful tool for probing time integration in perception, memory, and cognitive processes (Donders, 1969; Posner, 1978; Luce, 1986; Meyer et al., 1988). RT studies have led to the development of accumulator models, which implement in vari-

ous ways the idea of stochastic integration of input signals to a fixed decision threshold. In a race model, accumulators representing different choice options build up their activities, and whichever is the first to reach a prescribed threshold produces the choice (Logan and Cowan, 1984). In a drift diffusion model for two-alternative forced choices, an accumulator adds evidence in favor of one alternative and subtracts evidence in favor of the other; a decision is made when it reaches either a positive threshold or a negative threshold (Stone, 1960; Laming, 1968; Ratcliff, 1978; Smith and Ratcliff, 2004). A linear leaky competing accumulator (LCA) model, which mimics a neural network, takes into account a leakage of integration and assumes competitive inhibition between accumulators selective for choice alternatives (Usher and McClelland, 2001). This model is easily extended to decisions with multiple alternatives (Usher and McClelland, 2001; McMillen and Holmes, 2006; Bogacz et al., 2007), which is not straightforward for the diffusion model (Niwa and Ditterich, 2008; Churchland et al., 2008). For the two-alternative tasks, the LCA model is reduced to the diffusion model in the special case when the leak and inhibition cancel out each other (Usher and McClelland, 2001). The diffusion model is popular because of its simplicity yet proven success with fitting behavioral data in numerous human studies (Ratcliff, 1978; Busemeyer and Townsend, 1993; Smith and Ratcliff, 2004).

Although the concept of time integration is appealing, it is not obvious what types of choice behavior engage such accumulation process (a characteristic of deliberate decision making) and on what timescale (Uchida et al., 2006). Selection among a set of possible actions is a form of choice that can occur quickly, when speed is at a premium. This is illustrated by examples from simple organisms (Real, 1991): a toad produces a prey-catching or an avoidance behavior, depending on whether an ambiguous moving object is perceived as prey or a predator (Ewert, 1997); or an archerfish, by watching a prey's initial condition, quickly (within 100 ms) decides on the course of action in order to catch the prey (Schlegel and Schuster, 2008). In rodents, a study reported that olfactory discrimination is fast (~300 ms) in rats. Performance varied from chance level to near 100% correct, as the task difficulty was varied by adjusting the relative proportions of two odorants in a binary mixture, but the RTs were changed only slightly (by ~30 ms) (Uchida and Mainen, 2003). In contrast, another study found that, in mice, olfactory discrimination performance was high (~95% correct), regardless of discrimination difficulty, while RTs increased by ~80 ms from the easiest to the hardest (Abraham et al., 2004), suggesting that in these tasks rodents exhibit speed-accuracy tradeoff on a timescale of less than 100 ms. In human studies, mean RTs typically range from tens of milliseconds to about a second in simple perceptual tasks (Luce, 1986; Usher and McClelland, 2001).

What are the neural processes underlying time integration? Recently, electrophysiological studies with behaving monkeys have revealed that reaction times can be related to neural activity at the single-cell level. In a two-alternative forced-choice visual random-dot motion (RDM) direction discrimination task, monkeys are trained to make a binary judgment about the direction of motion of a near-threshold stochastic random dot visual motion stimulus and to report the perceived direction with a saccadic eye movement. The task difficulty can be parametrically



**Figure 1. Neural Mechanism of a Decision in a Monkey Random-Dot Motion Direction Discrimination Experiment**

(A) Reaction time (RT) version of the task. The subject views a patch of dynamic random dots and decides the net direction of motion. The decision is indicated by an eye movement to one of two peripheral targets (representing the two forced choices). In the RT task, the subject controls the viewing duration by terminating each trial with an eye movement whenever ready. The gray patch shows the location of the response field (RF) of an LIP neuron.

(B) Monkey's performance (top) and mean RT (bottom) as a function of the motion strength.

(C) Response of a single LIP neuron. Only correct choices at two motion strengths (6.4% and 51.2%) are shown. Spike rasters and response histograms are aligned to the beginning of the monkey's eye movement response (vertical line). Carets denote the onset of random-dot motion. Trial rasters are sorted by RT.

(D) Average response of LIP neurons during decision formation, for three levels of difficulty. Responses are grouped by motion strength and direction of choice, as indicated. (Left) The responses are aligned to onset of random-dot motion. Averages are shown during decision formation (curves truncated at the median RT or 100 ms before the eye movement). Shaded insert shows average responses from direction-selective neurons in area MT to motion in the preferred and antipreferred directions. After a transient, MT responds at a nearly constant rate. (Right) The LIP neural responses are aligned to the eye movement.

(A), (B), and (D) are reproduced with permission from Gold and Shadlen (2007) (insert from online database used in Britten et al. [1992]); (C) is reproduced from Roitman and Shadlen (2002).

varied by the fraction of dots moving coherently in the same direction, called the motion strength or percent coherence  $c'$ . Extensive physiological and microstimulation studies have shown that while direction-sensitive neurons in the area MT encode the motion stimulus (Newsome et al., 1989; Britten et al., 1992, 1993, 1996), the decision process itself occurs downstream of MT, potentially in the posterior parietal cortex and/or prefrontal cortex. Shadlen and Newsome found that activity of neurons in the lateral intraparietal area (LIP) was correlated with monkey's perceptual choice in both correct and error trials (Shadlen and Newsome, 1996, 2001). Moreover, in a reaction time version of

the task (Roitman and Shadlen, 2002; Huk and Shadlen, 2005), the subject's response time increased by ~400 ms from the easiest (with  $c' = 100\%$ ) to the hardest (with  $c' = 3.2\%$ ) (Figures 1A and 1B). LIP cells exhibited characteristic firing time courses that reflected the monkey's response time and perceptual choice (Figures 1C and 1D). From the onset of a random-dot motion stimulus until the time the monkey produced a choice response by a rapid saccadic eye movement, spike activity of LIP neurons selective for a particular saccadic target increased for hundreds of milliseconds. The ramping slope was larger with a higher  $c'$  (a higher quality of sensory information). Furthermore,

it was observed that the decision choice (as indicated by a saccade) was made when the firing rate of LIP neurons (selective for that choice response) reached a threshold that was independent of  $c'$  and the response time. Therefore, these LIP neurons display stochastic ramping to a set level, as expected for a neural integrator.

Another physiological study reported that, while monkeys performed a task of detecting the presence of a visual motion stimulus, neurons in ventral intraparietal (VIP) area of the posterior parietal cortex exhibited ramping activity that was correlated with the subjective judgment (a higher activity in hit and false-alarm trials than in miss and correct rejection trials) and response time (ranging from 400 to 700 ms) (Cook and Maunsell, 2002). MT neural responses were larger in hits (with successful detection) and misses (failure), implying that the strength of signals provided by MT neurons to higher areas was stronger in trials with ultimately successful detection than in failed trials. In VIP but not MT, neuronal activity was significantly larger than baseline in false-alarm trials, suggesting that the subjective judgment was computed downstream of MT. This study thus supports the notion that the parietal cortex is part of the brain system underlying accumulation of information and subjective judgment in visual perception.

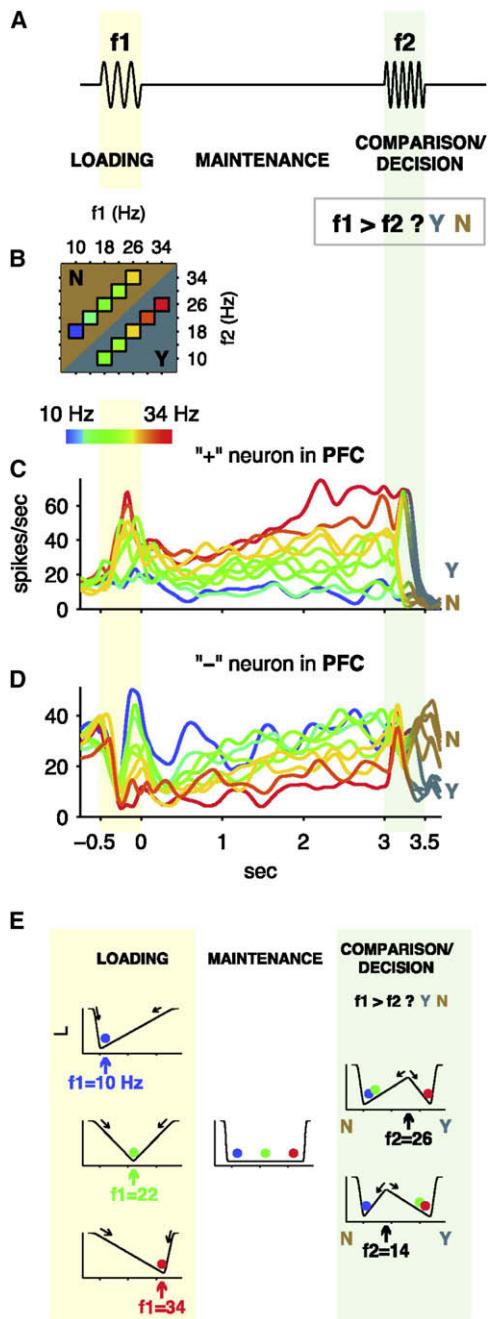
Regarding ramping neural activity, so far most recordings have been limited to one cell at a time, and ramping activity is usually reported as trial-averaged neural activity (but see Roitman and Shadlen, 2002). This leaves the question open as to whether a spike train in a single trial indeed displays a quasilinear ramp of firing rate. Alternatively, neurons could actually undergo a sudden jump from one rate to another, but the jumping time varies from trial to trial in such a way that the trial average shows a smooth ramp (Okamoto et al., 2007). Additional experiments, perhaps with multiple single-unit recording, would help to resolve this issue. Moreover, it is still unclear whether the observed neural activity in the parietal cortex is generated locally or reflects inputs from elsewhere.

The posterior parietal cortex plays a major role in selective visual attention (Colby and Goldberg, 1999; Corbetta and Shulman, 2002; Sereno and Amador, 2006; Ganguly et al., 2008), but it seems unlikely that neural signals observed in these perceptual decision experiments can be solely attributed to attentional effects. Attention presumably should be focused on the RDM stimulus, not the targets, until a choice is made. Even with divided attention between the motion stimulus and targets, it is unclear how an attentional account can explain the time course of neural dynamics, namely target-specific ramping activity that reflects the gradual formation of a decision and marks the end of the process. Another potential interpretation of the ramping activity is motor intention, as LIP is specifically involved with planning saccades (Andersen and Buneo, 2002). One way to test this possibility is to ascertain whether decision-correlated activity of parietal neurons remains the same regardless of the modality of the ultimate behavioral response.

Decision-related activity has also been observed in the prefrontal cortex, a high-level cognitive area that plays a major role in time integration and a gamut of other cognitive functions (Fuster, 2008; Miller and Cohen, 2001; Wang, 2006a). In a fixed-duration version of the RDM direction discrimination task, the

monkey was required to maintain fixation through a 1 s viewing period, the stimulus offset was followed by a delay period, then the monkey signaled its choice by a saccadic eye movement. In contrast to the RT version, now the activity of recorded neurons correlated with decision (during motion viewing) and that with motor response (after the delay) could be temporally dissociated. Neurons in the lateral prefrontal cortex (Kim and Shadlen, 1999) and LIP (Shadlen and Newsome, 1996, 2001; Roitman and Shadlen, 2002) showed a similar activity pattern: their activity reflected the monkey's choice, but, while it was a graded function of the motion strength  $c'$  during motion viewing, persistent activity during the delay became insensitive to  $c'$ . The implication was that the subject made the decision during stimulus viewing and maintained actively the binary choice across the delay to guide a later behavioral response. These results suggest that decision making and working memory can be subserved by the same cortical mechanism, presumably residing in the parietofrontal circuit. It should be noted that recordings from prefrontal neurons have not yet been done in the reaction time version of the RDM discrimination task. Prefrontal neurons often display ramping activity during a mnemonic delay period, but this could reflect anticipation and timing of an upcoming response rather than decision computation (Fuster, 2008; Chafee and Goldman-Rakic, 1998; Quintana and Fuster, 1999; Brody et al., 2003; Miller et al., 2003; Watanabe and Funahashi, 2007). Physiological recordings from the prefrontal cortex using RT paradigms are needed to assess whether prefrontal neurons display ramping activity in a way similar to that observed in parietal neurons and, if so, whether ramping activity is generated locally in one area (and reflected in another) or through a reciprocal loop between the two areas.

Romo and collaborators carried out a series of experiments, using a different task paradigm and sensory modality, that provided ample evidence for the involvement of the prefrontal cortex in perceptual decisions. In a somatosensory delayed discrimination task, monkeys report a decision based on the comparison of two mechanical vibration frequencies  $f_1$  and  $f_2$  applied sequentially to the fingertips, separated in time by a delay of 3–6 s. Therefore, the behavioral response (signaling whether  $f_1$  is perceived as larger or smaller than  $f_2$ ) requires the animal to hold in working memory the frequency of the first stimulus across the delay period (Figure 2A). It was found that neurons in the inferior convexity of the prefrontal cortex and the premotor cortex showed persistent activity during the delay, with the firing rate of memory activity increasing (a “plus cell”) or decreasing (a “minus cell”) monotonically with the stimulus frequency (Figures 2C and 2D). During comparison decision, neural activity became binary: a “plus neuron” showed high firing in trials when the monkey's choice was  $f_1 > f_2$ , low firing in trials when the monkey's choice was  $f_1 < f_2$ ; a “minus neuron” showed the opposite trend. In this task, working memory precedes decision making, but again the same circuit is engaged in both processes. In a modified version of the task, the decision report is postponed a few seconds after the comparison period, and medial premotor neurons were found to retain the monkey's choice and past sensory information in the form of persistent activity across the second delay (Lemus et al., 2007), reinforcing the point that the same circuit is involved in both working memory and decision making. A



**Figure 2. Delayed Vibrotactile Discrimination Task and Neuronal Responses in the Prefrontal Cortex**

(A) Schematic diagram of the task, where two mechanical vibration stimuli with frequencies  $f_1$  and  $f_2$  are applied sequentially (separately by a delay) to the tip of a monkey's finger, and the subject has to decide whether  $f_1$  is larger than  $f_2$ . (B) Typical stimulus set used in the neurophysiological studies. Each colored box indicates a  $(f_1, f_2)$  stimulus pair. For each pair, monkeys made the correct response more than 91% of the time. (C) and (D) show smoothed firing rates of two different PFC neurons recorded over many trials. (C) shows a positively monotonic (plus) neuron and (D) shows a negatively monotonic (minus) neuron.

remarkable accomplishment of Romo's work is to systematically explore neural activity during the same task across a number of cortical areas (primary and secondary somatosensory areas, premotor and lateral prefrontal areas), which yielded a rich picture of neural dynamics in these areas as the process unfolds in time (Hernández et al., 2002; Romo et al., 2004). Additional evidence for a role of the frontal lobe in subjective decisions was reported in a detection task using near-threshold vibrotactile stimuli. de Lafuente and Romo showed that activity of premotor neurons in the frontal lobe, but not that of primary somatosensory neurons, covaried with trial-by-trial subjective reports (whether a stimulus was present or absent) (de Lafuente and Romo, 2005). Similar detection psychometric functions were obtained with premotor cortex microstimulation or mechanical stimulation, suggesting that the stimulated frontal site may be causally related to this decision behavior.

Human studies on the physiological correlates of RTs in perceptual discrimination tasks began in the 1980s, with the development of event-related potential measurements (Gratton et al., 1988). Interestingly, it was found that the time-to-peak of the P300 component recorded in the parietal cortical area increased with the difficulty of stimulus discriminability and RT but was indifferent to the stimulus-response compatibility (Kutas et al., 1977; McCarthy and Donchin, 1981). Electroencephalographic recordings, however, lack sufficient spatial resolution to localize brain areas critically involved in time integration. Recently, human functional magnetic resonance imaging is beginning to be applied to studies of evidence accumulation. In one study (Binder et al., 2004), a subject was asked to identify a speech sound in a noisy auditory stimulus, blood-oxygen-level-dependent (BOLD) functional magnetic resonance imaging (fMRI) of the auditory cortex was found to reflect the signal-to-noise ratio (the quality of sensory information), whereas that of the inferior frontal cortex increased linearly with the reaction time (an index of decision process). In another study (Ploran et al., 2007), as a subject viewed pictures that were revealed from a blank screen gradually in eight steps (2 s each), the areas that exhibited a gradual buildup in activity peaking in correspondence with the time of recognition were the parietal and frontal areas as well as the inferior temporal area. Moreover, a recent study showed that a free motor decision could be decoded from BOLD signals in the frontal and parietal cortex long (seconds) before it entered awareness (Soon et al., 2008).

Therefore, growing evidence from human neuroimaging and monkey single-neuron physiology suggests that the parietal and frontal cortices form a core brain system for temporal accumulation of data and categorical choice formation in perceptual judgments. These areas may provide top-down signals to

(E) One-dimensional dynamical algorithm for two-stimulus interval discrimination. Abcissa: The state variable (e.g., the difference in the firing rates of the plus and minus neurons shown in [C] and [D]). Ordinate: A computational energy function (with minima corresponding to stable attractor states). During the loading period, the first stimulus creates a unique attractor state located at a point along the horizontal axis that encodes the  $f_1$  value. The energy landscape is flat during the delay period, so the memory of  $f_1$  is maintained internally in the form of parametric persistent activity. During the comparison period, the system is again reconfigured, the second stimulus  $f_2$  in interplay with the internal memory state gives rise to a categorical decision  $f_1 > f_2$  or  $f_1 < f_2$ . Reproduced with permission from Machens et al. (2005).

sensory neurons, whose spiking activity commonly displays weak trial-to-trial correlated variability with monkeys' choices (Britten et al., 1996; Shadlen and Newsome, 1996; Parker and Krug, 2003; Law and Gold, 2008).

A challenge for future work is to elucidate precisely how the parietal and frontal circuits work together, potentially playing differential and complementary roles, in decision making. Rather than proceeding strictly in serial stages, a decision is likely to involve parallel processing across brain regions. Such a scenario has been advocated in a model where action selection and response preparation take place simultaneously in diverse cortical areas (Cisek, 2006). Empirical evidence, however, is scarce. One promising approach to this outstanding issue is to examine interareal interactions by recording simultaneously single units and local field potentials from two or more areas in behaving animals (Pesaran et al., 2008).

### Recurrent Cortical Circuit Mechanism

How are decision computations instantiated in a cortical circuit? A clue came with the observation that decision-related neural activity has been reported in cortical areas that typically exhibit mnemonic persistent activity during working memory maintenance. For instance, in an oculomotor delayed response task, neurons in both LIP and prefrontal cortex display directionally tuned persistent activity (Gnadt and Andersen, 1988; Funahashi et al., 1989). Motivated by this observation, it has been proposed that this is not a mere coincidence, but suggests a common circuit mechanism underlying decision making and working memory (Wang, 2002). A leading candidate mechanism for the generation of persistent activity is strong recurrent excitation in a local cortical circuit that gives rise to stimulus-selective attractor states—self-sustained population activity patterns of a neural network (Amit, 1995; Goldman-Rakic, 1995; Wang, 2001). Can such an attractor network model also account for decision-making computations?

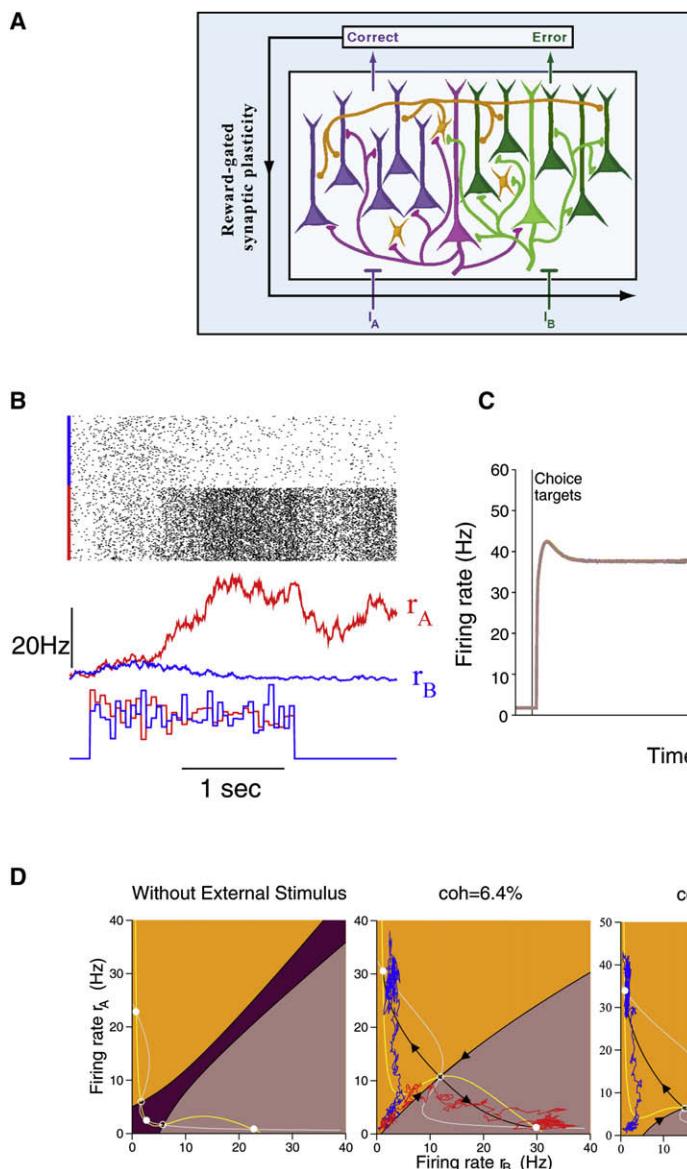
To address this question, a biophysically based model originally developed for working memory (Brunel and Wang, 2001) was applied to simulate the RDM discrimination experiment (Wang, 2002). It is worth emphasizing that this local circuit model stresses shared characteristics of the prefrontal and parietal areas and does not speak to the issue of whether memory or decision-related neural activity is generated in one of these areas or both. "Biophysically based models" generally refer to models with an anatomically plausible architecture, in which not only single spiking neurons are described biophysically with a reasonable level of accuracy but also synaptic interactions are calibrated by quantitative neurophysiology (which turned out to be critically important).

Figure 3 illustrates such a recurrent neural circuit model (Wang, 2002; Wong and Wang, 2006; Wong et al., 2007). In a two-pool version of the model, subpopulations of spiking neurons are selective for two-choice alternatives (e.g., A = left motion, B = right motion). Within each pyramidal neural group, strong recurrent excitatory connections can sustain persistent activity triggered by a transient preferred stimulus. The two neural groups compete through feedback inhibition from interneurons. Conflicting sensory inputs are fed into both neural pools in the circuit, with the motion strength  $c'$  implemented as the rel-

ative difference in the inputs (Figure 3A). Figure 3B shows a simulation with zero motion strength. At the stimulus onset, the firing rates ( $r_A$  and  $r_B$ ) of the two competing neural populations initially ramp up together for hundreds of milliseconds before diverging from each other when one increases (by virtue of recurrent excitation within that neural pool) while the other declines (due to winner-take-all competition mediated by feedback inhibition). The perceptual choice is decided based on which of the two neural populations wins the competition. With a varying  $c'$ , the ramping activity is faster when the quality of sensory data is higher (Figure 3C). The model captures important features of activity of LIP cells recorded from behaving monkeys. First, neural activity is primarily correlated with the decision choice (even in error trials or when the motion strength is zero). Second, the neural decision process proceeds in two steps: sensory data are first integrated over time in a graded fashion, followed by winner-take-all competition leading to a binary choice. Third, after the stimulus is withdrawn, the network stores the decision choice in working memory, in the form of persistent activity that is insensitive to  $c'$ .

The "attractor landscape" can be illustrated in the decision space, where  $r_A$  is plotted against  $r_B$  (Figure 3D). In this example, the sensory evidence is in favor of the choice A, so attractor A has a larger basin of attraction (orange) than that of attractor B (brown). The system is initially in the spontaneous state, which falls in the basin of attraction A, and evolves toward the decision state A in a correct trial (blue). However, at low  $c'$  the bias is not strong, and noise can induce the system's trajectory to travel across the boundary of the two attraction basins, in which case the system eventually evolves to the decision state B in an error trial (red). The crossing of a boundary between attraction basins is slow, which explains why the reaction times are longer in error trials than in correct trials, as was observed in the monkey experiment (Rothman and Shadlen, 2002). This decision-space analysis hammers home the point that the system is not rigid but is flexible in response to external signals. Attractor states can be created or destroyed by inputs; hence, the same network can subserve different functions, such as decision making during stimulus presentation followed by active memory of the choice across a delay in the fixed-duration RDM task. This conclusion is supported by other recent modeling studies of the RDM experiment (Roxin and Ledberg, 2008; Grossberg and Pilly, 2008).

As illustrated by the above example, an attractor network is not just limited to steady-state behavior but can use long transients to perform interesting computations. As initially proposed for working memory, one candidate cellular substrate for slow reverberation is the NMDA receptors at local recurrent excitatory synapses (Wang, 1999). A simple estimate of the network's time constant is given by  $\tau_{\text{network}} = \tau_{\text{syn}} / (|1 - w_{\text{rec}}|)$  (where  $\tau_{\text{syn}}$  is the synaptic time constant, and  $w_{\text{rec}}$  is the strength of recurrent connections), which is longer than  $\tau_{\text{syn}}$  in the presence of  $w_{\text{rec}}$  (Seung, 1996; Wang, 2001). For instance, if  $\tau_{\text{syn}} = 100$  ms for the NMDA receptor-mediated synapses (Hestrin et al., 1990), and  $w_{\text{rec}} = 0.9$ , then  $\tau_{\text{network}} = 1$  s. Thus, the network displays transient dynamics of ramping activity on a timescale of up to 1 s, and this ability critically depends on the NMDA receptors (Wang, 2002; Wong and Wang, 2006). Note that this mechanism relies on the slow kinetics of NMDA receptor-mediated channel



correspond to the basins of attraction for the coexisting attractor states. In the absence of noise, the system starting in one of the basins converges to the corresponding attractor state. Note that the basin for the correct choice state is much larger at a high (right panel) than a low (middle panel) motion strength.  
(A) and (B) were reproduced with permission from Wang (2002), (C) from Wong et al. (2007), (D) was computed using the model of Wong and Wang (2006).

and emphasizes the importance of NMDA receptors for online cognition (rather than its well-known role in long-term synaptic plasticity). Other slow biophysical mechanisms, such as short-term synaptic facilitation (Abbott and Regehr, 2004) or calcium-dependent processes in single cells (Major and Tank, 2004), may also contribute to time integration. These candidate scenarios, all positive-feedback mechanisms, can be experimentally tested with behaving monkeys (using pharmacological means) as well as rodent and other simpler animal systems.

Is the recurrent neural circuit model simply an implementation of the diffusion model? Interestingly, in contrast to the one-dimensional diffusion model, a “decision-space” analysis (Figure 3D) showed that the dynamics of the attractor neural network

**Figure 3. A Cortical Circuit Model of Spiking Neurons for Two-Alternative Forced-Choice Tasks**

(A) Model scheme. There are two (purple and green) pyramidal cell subpopulations, selective to the two directions (A or B), respectively, of random moving dots in a visual motion discrimination experiment. A third (orange) neural subpopulation represents inhibitory interneurons. Each of the three subpopulations consists of a few hundreds of spiking neurons. The circuit is endowed with strong recurrent excitation (mediated by AMPA and NMDA receptors) among pyramidal cells within each selective subpopulation and competitive inhibition (mediated by GABA<sub>A</sub> receptors) through shared feedback inhibition. The motion coherence is expressed as  $c' = (I_A - I_B)/(I_A + I_B)$ , where  $I_A$  and  $I_B$  are the mean inputs. For nonzero  $c'$ , one of the choices is correct, the other is erroneous, the resulting outcome may lead to reward-dependent plastic changes of some input synapses.

(B) A sample simulation of the spiking network model with zero coherence. Top to bottom: Network spiking raster, population firing rates  $r_A$  and  $r_B$ , stochastic inputs. Note the initial slow ramping (time integration) and eventual divergence of  $r_A$  and  $r_B$  (categorical choice).

(C) Trial-averaged neural activity for different motion strengths, with the inclusion of target inputs. Solid curves: Winning neural population. Dashed curves: Losing neural population. Note the transient dip at the onset of the RDM stimulus.

(D) Phase-plane plot for the two selective neural populations in a fixed-duration version of the task, without external input (left panel), in the presence of a motion stimulus with  $c' = 6.4\%$  (middle panel) or  $51.2\%$  (right panel). In the absence of stimulation (left panel), three attractors coexist (white circles): a spontaneous state (when both  $r_A$  and  $r_B$  are low), and two persistent activity states (with a high  $r_A$  and a low  $r_B$  or vice versa). Upon the presentation of a stimulus (middle panel with  $c' = 6.4\%$ ), the attractor landscape is altered, and the spontaneous steady state disappears, so that the system is forced to evolve toward one of the two active states that represent perceptual decisions (A or B), as shown by the network's trajectory in two individual trials (blue and red). After the offset of the stimulus, the system's configuration reverts back to that in the left panel. Because a persistently active state is self-sustained, the perceptual choice (A or B) can be stored in working memory for later use, to guide behavior. Colored regions

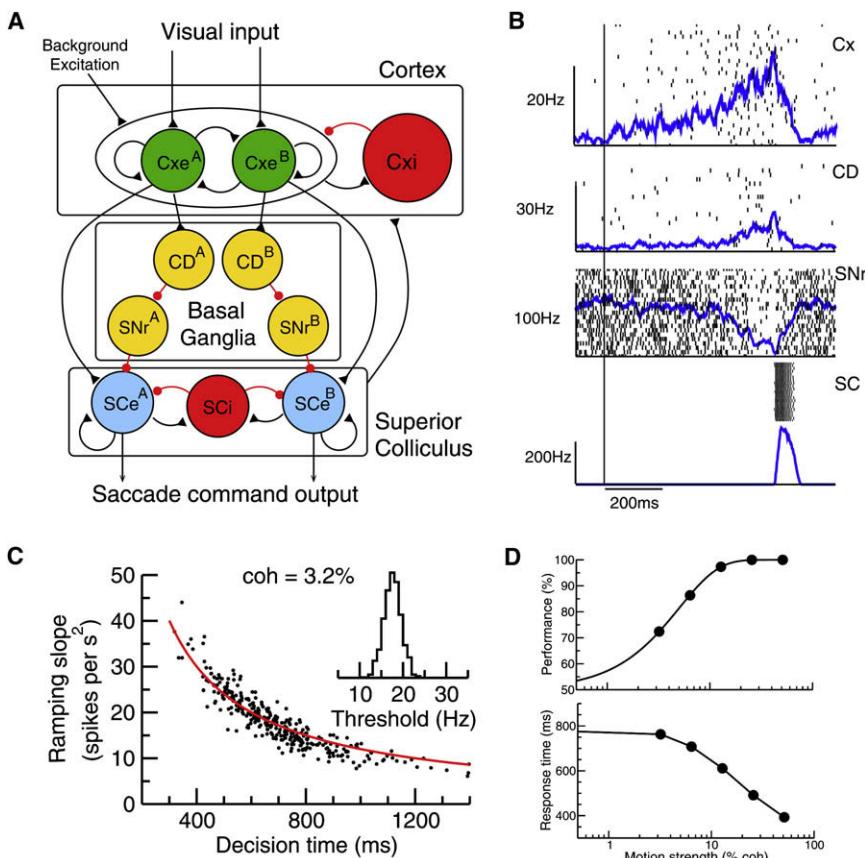
is inherently two dimensional (Wong and Wang, 2006). This is consistent with the finding that, in the LIP data recorded from the RDM experiment (Roitman and Shadlen, 2002), the dynamics within each of the two selective neural pools is dominated by a slow mode (Ganguli et al., 2008); thus, the description of two competing neural pools requires two dynamical variables. A two-variable model is needed to explain the observation that LIP neuronal activity displays a biphasic time course, with neurons selective for the two opposite targets first ramping up together before diverging away from each other (Roitman and Shadlen, 2002; Huk and Shadlen, 2005). The same type of behavior was also observed in a free motor decision task (Scherberger and Andersen, 2007).

Furthermore, importantly, the diffusion model and the recurrent neural circuit model have distinct predictions at the behavioral level. First, the recurrent circuit model produces longer response times in error trials than in correct trials (Wong and Wang, 2006), consistent with the monkey experiment (Roitman and Shadlen, 2002). By contrast, a neural implementation of the diffusion model yields the opposite effect (Mazurek et al., 2003). Longer RTs in error trials can be realized in the diffusion model with the additional assumption that the starting point varies stochastically from trial to trial (Ratcliff and Rouder, 1998). Second, the diffusion model never reaches a steady state and predicts that performance can potentially improve indefinitely with a longer duration of stimulus viewing, e.g., by raising the decision bound. In the recurrent circuit model, ramping activity eventually stops as an attractor state is reached (Figure 3D). Consequently performance plateaus at sufficiently long stimulus-viewing times (Wang, 2002). This model prediction is confirmed by a recent monkey experiment (Kiani et al., 2008). Third, the attractor network model has been shown to be able to subtract negative signals as well as add positive evidence about choice alternatives, but the influence of newly arriving inputs diminishes over time, as the network converges toward one of the attractor states representing the alternative choices (Wang, 2002). This prediction is also confirmed by the monkey experiment, which showed that the impact of a brief motion pulse in addition to the random-dot stimulus was greater with an earlier onset time (Huk and Shadlen, 2005; Wong et al., 2007). This violation of time-shift invariance cannot be accounted for by the inclusion of a leak. In fact, in contrast to the recurrent circuit model, the LCA model actually predicts that later, not earlier, signals influence more the ultimate decision, because an earlier pulse is gradually “forgotten” due to the leak and does not affect significantly the decision that occurs much later (Wong et al., 2007).

Recurrent excitation must be balanced by feedback inhibition (Brunel and Wang, 2001). The diffusion model assumes that a difference signal about the conflicting inputs is computed occurs at the input level (Ratcliff, 1978; Mazurek et al., 2003). This idea has been taken seriously in a human fMRI experiment, in which the task was to discriminate whether an ambiguous image is a face or a house, and the BOLD signal in the dorsolateral prefrontal cortex was found to covary with the difference signal between the face- and house-selective regions in the ventral temporal cortex (Heekeren et al., 2004). This work suggests that some brain region(s) may encode difference signals in discrimination of categorically distinct signals. The situation is likely to be different for discrimination between options in the same dimension, such as left versus right motion direction, which is likely to occur within a local network. In the recurrent circuit model, competition between neural pools selective for choice alternatives is instantiated by lateral synaptic inhibition (Wang, 2002; Wong and Wang, 2006). This feedback mechanism, not the feed-forward subtraction, is supported by the observation that microstimulation of one neural pool in LIP not only sped up the decisions in its preferred direction but also slowed down the decisions in the antipreferred direction (Hanks et al., 2006). In another relevant analysis, Ditterich (2006) found that a diffusion model produced reaction time histograms with long right tails (reflecting unusually long RTs), inconsistent with the monkey ex-

periment. The inclusion of lateral inhibition worsened the problem, resulting in even longer right tails, especially at low coherence levels. This is not the case in the recurrent neural circuit model, which produces decision-time distributions that do not show pronounced right tails and are similar to those observed in the monkey experiment (X.-J. Wang, 2006, Soc. Neurosci., abstract). A distinguishing feature of the nonlinear attractor model is strong recurrent excitation, which is absent in linear accumulator models. The positive-feedback mechanism ultimately leads to an acceleration of the ramping neural activity toward a decision bound, preventing excessively long decision times. Indeed, Ditterich showed that monkey's reaction-time distributions could be well fitted by an accumulator model, with the additional assumption that the decision bound decreased over time. This is functionally similar to a hypothesized “urgency signal” that grows over time (T.D. Hanks et al., 2007, Soc. Neurosci., abstract). Equivalently, the desired effect can be accomplished by a temporally increasing ramping slope, which naturally occurs in the recurrent circuit model without additional assumptions. On the other hand, human studies commonly report skewed RT distributions with a long right tail, which is well captured by the diffusion model (Ratcliff, 1978; Luce, 1986; Ratcliff and Rouder, 1998; Usher and McClelland, 2001; Sigman and Dehaene, 2005) but not the existing neural circuit model. It will be interesting to identify, in animal as well as human studies, conditions under which RT distributions do or do not display a prominent tail, and to come up with a neural mechanistic account of this phenomenon.

Recurrent circuit models have also been developed for the somatosensory discrimination experiment (Romo et al., 2002, 2004). Miller et al. (Miller et al., 2003) showed that fine-tuning of connectivity in this model yields a line attractor capable of parametric working memory, similar to the one in the gaze-control system (Seung et al., 2000). Moreover, two such neural populations coupled by reciprocal inhibition exhibit persistent activity that is positively and negatively, respectively, monotonically tuned to the first frequency  $f_1$ , as observed in prefrontal neurons in this task (Romo et al., 1999; Brody et al., 2003; Miller et al., 2003). This circuit is thus capable of storing  $f_1$  across the delay period. Machens et al. (2005) showed that such a circuit could also perform discrimination computation ( $f_1 > f_2$  or  $f_1 < f_2$ ) during the comparison period (Figure 2E), provided that a switching mechanism was posited for the afferent stimuli, such that the second stimulus reached the memory/decision circuit with the opposite sign of tuning to the first stimulus. In an alternative scenario, an integral feedback mechanism instantiates comparison computation without requiring input switching (Miller and Wang, 2006a). Using a phase-plane analysis, Machens et al. showed elegantly how the attractor landscape is differentially reconfigured by external signals for each of the task epochs (cue loading, memory maintenance, comparison); hence, the same circuit can subserve both working memory and decision computation (Machens et al., 2005). The model predicts positive trial-by-trial correlation between neural pairs of the same (positively or negatively monotonic) tuning type, and negative correlation between neural pairs of opposite tuning. This prediction is confirmed by neural data recorded from the prefrontal cortex. Interestingly, in contrast to the fixed-duration version of the RDM



discrimination task, where decision precedes a delayed response that requires working memory of a binary chosen option, in the vibrotactile discrimination task, parametric working memory of an analog quantity ( $f_1$ ) precedes a two-choice decision process. Nevertheless, models for these two kinds of behavioral tasks display striking similarities (including phase-plane plots). The near-threshold detection experiment (de Lafuente and Romo, 2005) has also been modeled by a similar attractor network model (Deco et al., 2007a). Taken together, these results further support slow reverberating dynamics as a general mechanism for both working memory as well as decision making.

#### Termination Rule for a Decision Process

How do we know precisely when a graded accumulation process ends and a categorical decision is formed? In an RT task, a decision time can be deduced from measured reaction time (minus motor response latency), which has been shown to be correlated with threshold crossing of LIP neuronal activity (Roitman and Shadlen, 2002). If so, what would be the biological substrate of such a decision threshold? The answer may lie downstream. A plausible scenario is that, when decision neurons integrate inputs and reach a particular firing rate level, this event triggers an all-or-none response in downstream neurons and leads to the generation of a behavioral output. In the case of oculomotor tasks, a natural candidate is movement neurons in the frontal eye field (FEF) and superior colliculus (SC), which are brain regions essential for selecting, preparing, and initiating

**Figure 4. A Multiple-Module Network Mechanism for Two-Alternative Forced-Choice Tasks**

(A) Schematic model architecture. Neural pools in the cortical network integrate sensory information and also compete against each other. They project to both the superior colliculus (SC) and the caudate nucleus (CD) in the basal ganglia. CD sends inhibitory projection to the substantia nigra pars reticulata (SNr), which through inhibitory synapses connect with movement neurons in the SC. Each population consists of noisy spiking neurons.

(B) A single trial simulation of the model, showing spike trains from single cells and population firing rates of Cxe, SNr and CD, and Sce. A burst of spikes in movement neurons (Sce) is triggered when their synaptic inputs exceed a threshold level, which results from both direct excitation by cortical neurons and disinhibition from SNr via the cortico-striatal projection. Time zero corresponds to stimulus onset.

(C) The ramping slope of Cxe firing rate is inversely related to decision time on a trial-by-trial basis (each data point corresponds to an individual trial). The red curve is  $12,000/(decision\ time)$ .

(D) Performance (percentage of correct choices) and mean response time as a function of the motion coherence  $c'$ . Reproduced with permission from Lo and Wang (2006).

saccadic eye movements. These neurons are selective for saccade amplitude and direction and fire a stereotypical burst of spikes immediately before a saccade is initiated (Hanes and Schall, 1996; Munoz and Fecteau, 2002). While here we focus on response execution, FEF and SC, like LIP, are also involved in other aspects of oculomotor decision and response selection.

To test this scenario for a decision threshold, we considered an extended, multiple-circuit model (Lo and Wang, 2006). Decision neurons in the cortex (as described above) project to movement neurons in the SC (Figure 4A). This model also includes a direct pathway in the basal ganglia, with an input layer (caudate, CD) and an output layer (substantia nigra pars reticulata, SNr), which is known to play a major role in controlling voluntary movements (Hikosaka et al., 2000). As a neural pool in the cortex ramps up in time, so do their synaptic inputs to the corresponding pool of SC movement neurons as well as CD neurons. When this input exceeds a well-defined threshold level, an all-or-none burst of spikes is triggered in the SC movement cells, signaling a particular (A or B) motor output. In this scenario, a decision threshold (as a bound of firing rate of decision neurons) is instantiated by a hard threshold of synaptic input for triggering a special event in downstream motor neurons. Figure 4B shows a sample trial of such a model simulation for the visual motion direction discrimination experiment. The rate of ramping activity fluctuates from trial to trial, as a result of stochastic firing dynamics in the cortex, and is inversely related to the decision time (as defined by the time when a burst is triggered in the SC) on a trial-by-trial basis (Figure 4C). Moreover, when the task is more difficult (with a lower motion coherence), ramping activity is slower, leading to longer reaction times. However, the threshold of cortical firing

activity that is read out by the downstream motion system has the same narrow distribution (insert in Figure 4C), regardless of the ramping speed or reaction times. Therefore, the variability of reaction times is mostly attributed to the irregular ramping of neural activity itself rather than trial-to-trial variability of the decision bound. This model reproduced the monkey's behavioral performance and reaction times quantitatively (Figure 4D).

Can a decision threshold be adaptively tuned in this circuit? In a speed-accuracy tradeoff, too low a threshold leads to quicker responses but more errors, whereas too high a threshold improves the accuracy but prolongs response times. Neither of these yields maximal rewards. A commonly held idea is that optimality can be achieved by adaptively tuning the decision threshold (Gold and Shadlen, 2002; Bogacz et al., 2006). Since in the neural circuit model the decision threshold is defined as the minimum cortical firing needed to induce a burst response in the downstream SC neurons, one would expect that this threshold could be adjusted by plastic changes in the cortico-collicular pathway: with an enhanced synaptic strength, the same level of cortical input to the superior colliculus could be achieved with less firing of cortical neurons. Interestingly, this is not the case when the system is gated by the basal ganglia. This is because neurons in SNr normally fire tonically at a high rate (Figure 4B) and provide a sustained inhibition to SC movement neurons (Hikosaka et al., 2000). This inhibition must be released (as ramping activity in the cortex activates CD neurons, which in turn suppresses the activity in the SNr), in order for SC neurons to produce a burst output. This highly nonlinear disinhibition mechanism implies that the decision threshold is much more readily adjustable by tuning the synaptic strength of cortico-striatal pathway than by changes of the cortico-collicular synaptic strength (Lo and Wang, 2006). This finding is particularly appealing in light of the fact that cortico-striatal synapses represent a prominent target of innervations by dopamine neurons. Given that dopamine neurons signal rewards or reward-prediction errors (Schultz, 1998) and that long-term potentiation and depression of the cortico-striatal synapses depend on the dopamine signals (Reynolds et al., 2001; Shen et al., 2008), our work suggests that dopamine-dependent plasticity of cortico-striatal synapses represents a candidate neural locus for adaptive tuning of a decision threshold in the brain, a prediction that is testable experimentally. More generally, it remains to be seen whether a decision threshold (defined neurophysiologically in terms of a neural activity bound), or some other attributes like the onset time or ramping slope of putative decision neurons, is actually dynamically adjusted in a speed-accuracy tradeoff.

Although decision threshold is a critical element in reaction-time tasks, it should not be equated to a general decision rule that terminates an accumulation process. This is clearly illustrated in the fixed-duration version of the RDM task in which the viewing time is controlled externally and the subject is required to refrain from making an overt response until either at the stimulus offset (Britten et al., 1992; Kiani et al., 2008) or after a mnemonic delay period (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002). The subjects, and LIP neurons, do not appear to integrate sensory information through the whole stimulus presentation period (provided it is sufficiently long), which was suggested to indicate that threshold crossing may still

be the rule for terminating the decision process in these situations (Kiani et al., 2008). Such a scenario would require a readout system (that detects the event of threshold crossing) to send a feedback signal to stop the integration process in decision neurons. In contrast, in the recurrent neural circuit model, integration stops naturally when the system has reached a steady state. This can occur without triggering an overt behavioral response, presumably because downstream movement neurons (in FEF and SC) are inhibited by virtue of an external cue (e.g., the fixation signal) and/or internally generated control signals. Conceptually, the recurrent neural network suggests that categorical choice is determined naturally by which of the alternative attractors wins the competition. The response time is interpreted in terms of the time at which neural signals from a decision circuit are read out by the motor system, which can be flexibly adjusted and differently controlled in a reaction-time task or a fixed-duration task.

This general idea not only applies to perceptual decisions but also action control. Indeed, a recurrent circuit approach has been used to build models for action selection and movement preparation (Wilimzig et al., 2006; Cisek, 2006; Heinze et al., 2007). The timing of a movement, or even whether a response is ultimately produced, is potently controlled by inhibitory processes, such as suppressive gating of movement neurons by "holding" neurons (Hanes and Schall, 1996; McPeek and Keller, 2002; Narayanan and Laubach, 2006; Boucher et al., 2007), and by the basal ganglia (Hikosaka et al., 2000). Therefore, selection of an action may not be casually linked to the reaction time, at least under some circumstances. In a popular model for inhibitory control of action, there is a race between a GO process and a STOP process, whichever crosses a threshold first wins the race and determines whether a response is inhibited or not (Logan and Cowan, 1984; Boucher et al., 2007). In a recurrent neural circuit model of countermanding action, the decision is described as a bistable dynamics, similar to the two forced-choice perceptual decision, except that now the attractor states correspond to cancelled versus noncancelled response (C.C. Lo and X.-J. Wang, 2007, Soc. Neurosci., abstract). In this view, the outcome of a decision process is to some extent insensitive to the precise value of the decision threshold, e.g., raising the threshold beyond a certain level does not necessarily improve the performance. Hence, again, categorical choice can be dissociated from a decision threshold.

### Value-Based Economic Choice

To make the "right" decision is ultimately about achieving a behavioral goal. In laboratory experiments, the goal is often to garner maximum rewards. This type of decision making relies on the brain's ability to evaluate the desirabilities of available options as a prerequisite to choosing and to adaptively change decision strategies when choice outcomes do not meet expectations. This field, a fusion of reinforcement learning theory and neuroeconomics, has been the topic of several recent reviews (Sugrue et al., 2005; Rushworth and Behrens, 2008; Loewenstein et al., 2008; Soltani and Wang, 2008).

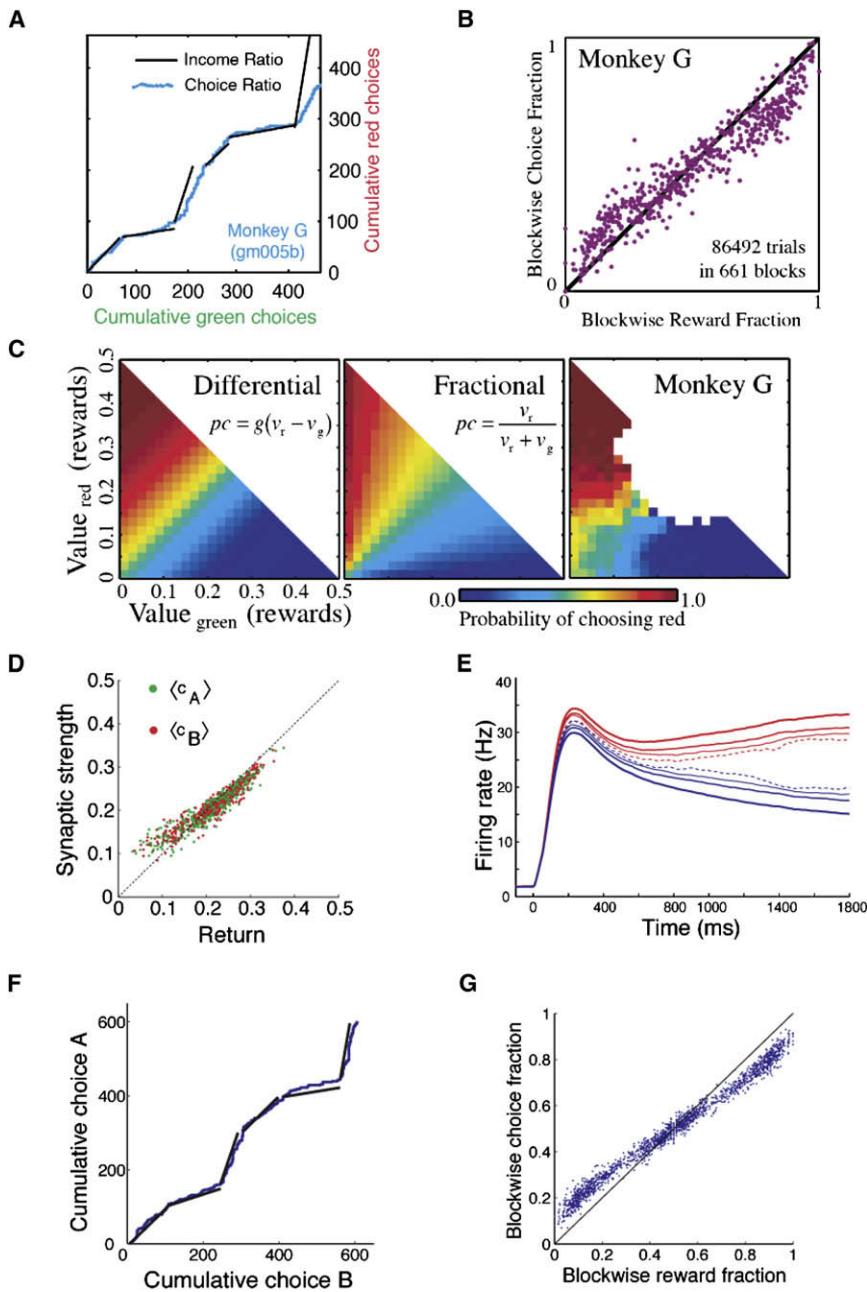
Significant progress has been made on neural representation of reward signals and reward expectations. A seminal finding was that phasic activity of dopamine neurons in the ventral

tegmental area (VTA) encodes a reward prediction error (the difference between the expected and actual reward) (Schultz et al., 1997; Schultz, 1998; Roesch et al., 2007). Consistent with a prediction error signal, spiking activity of dopamine neurons increases with both reward magnitude and probability (Fiorillo et al., 2003; Tobler et al., 2005; Roesch et al., 2007). Moreover, a recent study found that neurons in the primate lateral habenula reflect reward-prediction errors with an opposite sign from dopamine neurons (exhibiting a strong increase in spiking activity when the actual reward is smaller than the expected outcome), suggesting that the lateral habenula is a source for negative prediction error signal (Matsumoto and Hikosaka, 2007). On the other hand, neural signals correlated with reward expectation have been consistently observed in the striatum, amygdala, orbitofrontal cortex (OFC), and anterior cingulate cortex (ACC) in single-unit recording from behaving animals (reviewed in Rushworth and Behrens, 2008). The expected value is often characterized as a leaky integrator of experienced rewards. For instance, neural firing in ACC of behaving monkeys has been described as a temporal filter of past rewards, on a timescale of several trials (or tens of seconds) (Kennerley et al., 2006; Seo and Lee, 2007). In reinforcement learning theory, the error signal is postulated to update the reward expectation, which in turn is used to compute the error signal (Sutton and Barto, 1998; Bayer and Glimcher, 2005; Rushworth and Behrens, 2008). Thus, the reward expectation and the prediction error depend on each other and must be computed iteratively, possibly in different brain regions connected in a reciprocal loop. For instance, through a learning process that depends on dopaminergic inputs, reward expectation may be evaluated in a circuit including the striatum and frontal cortex. This signal is then fed back to midbrain dopamine cells to be compared with the actual reward to yield a prediction error. However, reward expectation and prediction error signals are mixed in multiple brain areas and are often difficult to disentangle.

It is useful to distinguish a brain system for reward valuation and those neural circuits that use this information to guide choice behavior. Brain structures activated in decision making and modulated by reward signals include caudate (Samejima et al., 2005; Hikosaka et al., 2006), lateral parietal cortex area LIP (Platt and Glimcher, 1999; Sugrue et al., 2004), and prefrontal cortex (Watanabe, 1996; Roesch and Olson, 2003). Reinforcement learning models suggest that action values are learned at synapses onto neurons in a decision circuit, thereby influencing choice behavior (Seung, 2003; Wörgötter and Porr, 2005). To illustrate this point, consider a neural network shown in Figure 2A. Recall that the network behavior is described by a softmax decision criterion, that is, the probability of choosing A versus B is a sigmoid function of the difference in the inputs ( $\Delta I$ ) to the two competing neural pools (Figure 4D, upper panel). Suppose that the strengths of the two synaptic connections  $c_A$  and  $c_B$  are plastic, then synaptic modifications will alter the network decision behavior over time. Specifically, we used binary synapses that undergo a stochastic Hebbian learning rule, namely that synaptic plasticity depends on coactivation of presynaptic and postsynaptic neurons and takes place probabilistically (Fusi, 2002; Fusi et al., 2007). In addition, it is assumed that synaptic learning depends on reward signals, based on the observation that dopa-

mine signal is known to gate synaptic plasticity in the striatum (Wickens et al., 2003; Shen et al., 2008) and prefrontal cortex (Otani et al., 2003; Matsuda et al., 2006). For instance, synapses for inputs to decision neurons are potentiated only if the choice is rewarded, and depressed otherwise. Therefore, in a learning process, synapses acquire information about reward outcomes of chosen responses, i.e., action-specific values. As a result of synaptic modifications, the input strengths for the competing neural groups of the decision network vary from trial to trial, leading to adaptive dynamics of choice behavior.

Such a model was tested by applying it to a foraging task in which a subject makes a sequence of choices adaptively in an unpredictable environment. In a monkey experiment (Sugrue et al., 2004; Lau and Glimcher, 2005, 2008), rewards were delivered to two (A and B) response options stochastically at baiting rates  $\lambda_A$  and  $\lambda_B$ , respectively, according to a concurrent variable-interval reinforcement schedule, in which choice targets are baited with rewards probabilistically and remain baited until the subject chooses the target and collects the reward. The studies found that monkey's behavior conformed to the matching law, which states that a subject allocates her or his choices in a proportion which matches the relative reinforcement obtained from these choices (Herrnstein et al., 1997) (Figures 5A and 5B). Moreover, neural activity of LIP neurons selective for a saccadic response was modulated by a representation of the outcome value, which was defined behaviorally as a leaky integration of past rewards on that target (Sugrue et al., 2004). Interestingly, the monkey's choice behavior is fit well by a softmax function of the difference in the two incomes (Corrado et al., 2005) (Figure 5C). These behavioral and neurophysiological observations were reproduced in the neural circuit model of decision making endowed with reward-dependent plasticity (Soltani and Wang, 2006) (Figures 5D–5G). It turns out that, in the model, the synaptic strengths ( $c_A$  and  $c_B$ ) are proportional to the returns (the average reward per choice) rather than the incomes (the average reward per trial) of the two targets (Figure 5D). Note that matching implies that the two returns are equalized, thus encoding reward values in terms of returns is especially suited for matching computation. Moreover, because synapses are potentiated or weakened stochastically over time, they are forgetful and behave like a leaky integrator of past choice outcomes, with a time constant determined by the learning rate as well as the reward statistics in the environment (Soltani and Wang, 2006). Hence, the decision behavior is influenced by past rewards harvested locally in time, in agreement with the observed monkey's behavior (Sugrue et al., 2004; Lau and Glimcher, 2005). As observed in LIP, in the model, neurons are modulated by the values of the response options (Figure 5E), even though they are not directly responsible for valuation itself. The model reproduces the matching behavior: as the reward rate  $\lambda_A/\lambda_B$  varies from one block of trials to the next block, the model's behavior changes quickly, so that the probability of choosing A versus B matches approximately  $\lambda_A/\lambda_B$  (Figure 5F). Further, the model also accounts for the observation that, in the monkey experiment, matching is not perfect, and the relative probability of choosing the more rewarding option is slightly smaller than the relative reward rate ("undermatching") (Figure 5G). A model analysis showed that undermatching is a natural consequence



(G) Each point shows the blockwise choice fraction as a function of the blockwise reward fraction for a block of trials on which the baiting probabilities are held constant. The model reproduces the matching behavior as well as the undermatching phenomenon.

(A) is reproduced with permission from Sugrue et al. (2004), (B) and (C) from Corrado et al. (2005), and (D)–(G) from Soltani and Wang (2006).

of fluctuating network dynamics (Soltani and Wang, 2006). Without neural variability, decision behavior tends to get stuck with the more rewarding alternative; stochastic spiking activity renders the network more exploratory and produces undermatching as a consequence.

The same type of model is also applicable to competitive games, where several decision makers interact according to a payoff matrix. Specifically, model simulations have been car-

ried out for the experiment of Barraclough et al. (2004), in which monkeys played matching pennies with a computer opponent. The model reproduced salient behavioral observations (Soltani et al., 2006). Similar to monkey's behavior, when the opponent is fully interactive according to the rules of matching pennies, the model behavior becomes quasirandom. For instance, if initially  $c_A$  is larger than  $c_B$ , and the system chooses target A more frequently, it would be exploited by the opponent, and

**Figure 5. Neural Basis of Matching Law in Foraging Behavior**

(A) Dynamic matching behavior of a monkey during a single experimental session. Continuous blue curve shows cumulative choices of the red and green targets. Black lines show average ratio of incomes (red:green) within each block (here, 1:1, 1:3, 3:1, 1:1, 1:6, and 6:1). Matching predicts that the blue and black curves are parallel.

(B) Block-wise matching behavior. Each data point represents a block of trials with the baiting probabilities for each target held constant. Reward and choice fractions are shown for the red target (those for the green target are given by one minus the fraction for the red target). Perfect matching corresponds to data points along the diagonal line. Deviations (undermatching) are apparent, as the choice probability is lower than reward probability when the latter is larger than 0.5.

(C) In a linear-nonlinear model, past rewards are integrated across previous trials with a filter time constant of approximately five to ten trials, yielding estimated values for the two targets  $v_r$  and  $v_g$ . Choice probability as a function of  $v_r$  and  $v_g$  is modeled as either a softmax rule (left panel) or a fractional rule (middle panel). Monkey's behavioral data are fitted better by the softmax (sigmoid) decision criterion (right panel).

(D) In a recurrent neural circuit model endowed with reward-dependent plasticity (Figure 3A) applied to the foraging task, the average synaptic strength is a linear function of the return from each choice (the reward probability per choice on a target). Red and green data points are for the synaptic strengths  $c_A$  (for red target) and  $c_B$  (for green target), respectively.

(E) Graded activity of neurons in the two selective neural populations. The activity of decision neurons shows a graded pattern if single-trial firing rates are sorted and averaged according to the choice and the difference between synaptic strengths. Activity is aligned by the onset of two targets, and it is shown separately for the choice that is the preferred (red) or nonpreferred (blue) target of the neurons. In addition, trials are subdivided into four groups according to the difference between the values encoded by the synaptic strength onto the two competing neural populations ( $c_A - c_B = -0.05$  to  $-0.14$  [dashed], 0 to  $-0.05$  [thin], 0 to  $0.05$  [normal], 0.05 to  $0.14$  [thick]).

(F) For one session of the model simulation of the foraging experiment, the cumulative choice on target A is plotted versus the cumulative choice on target B (blue). The black straight lines show the baiting probability ratio in each block. The same baiting probability ratios are used as in the monkey's experiment (A).

the unrewarded outcomes from choosing A induce depression of the synapses to the neural pool A, so that the difference  $c_A - c_B$  decreases over time, and the system gradually chooses B more frequently (Soltani et al., 2006; Lee and Wang, 2008).

Therefore, although activity of decision neurons depends on values of response options, valuation may occur elsewhere, perhaps at the synaptic level. It remains to be seen how such a learning rule works when the outcome (reward or not) is revealed only long after the behavioral response, by incorporating either persistent neural activity that bridges the temporal gap between the two events or an “eligibility trace” in the synapses (Sutton and Barto, 1998; Seung, 2003; Izhikevich, 2007). Virtually nothing is known empirically on this important issue, and new experiments in this direction would be highly desirable. Another key factor is cost (punishment, loss, and effort), the flip side of reward, which is poorly understood at the neural level. Furthermore, for the sake of simplicity, most biophysically based models have so far been limited to considerations of a local network and remain agnostic about the actual site of synaptic plasticity underlying valuation. Candidate loci include the corticostriatal connections in the basal ganglia or synaptic pathways within the orbitofrontal cortex, which have been explored in connectionist neural network models (Cohen et al., 1996; Frank and Claus, 2006) and reinforcement learning models (Samejima and Doya, 2007). Thus, it is likely that reward-dependent synaptic plasticity occurs in specific brain areas (or subpopulations of neurons in those areas) dedicated to signaling action values, whereas others are more directly involved with the generation of behavioral choice (Samejima and Doya, 2007; Rushworth and Behrens, 2008). Elucidation of the inner working of such large-scale decision circuits represents a major challenge in the field.

### Uncertainty and Stochastic Neural Dynamics

Decisions are often fraught with risk because the sensory world and choice outcomes, as well as intentions of interactive decision agents, are known only with varying levels of probability. This is illustrated by the aforementioned monkey experiments: sensory information is meager and conflicting in a near-threshold discrimination task; whereas in a foraging task, the possible outcomes of response options are given by reward probabilities that change unpredictably over time. And in a matching pennies game task, agents must decide without knowing each other’s intended actions, but the outcome depends on all the agents’ choices. Uncertainty is considered a key factor for explaining economic choice behavior (Kahneman, 2002), and decision making under risk represents a central point of converging interest for economists and neuroscientists. Recently, human studies using a combination of gambling tasks and functional neuroimaging and neurophysiological studies with behaving animals have been carried out to examine neural representations of uncertainty. Some studies aimed at identifying distinct brain systems that are recruited by different types of uncertainty (Yu and Dayan, 2005; Hsu et al., 2005; Huettel et al., 2006; Behrens et al., 2007). Other studies quantified uncertainty-correlated neural signals in terms of probabilities (Fiorillo et al., 2003) or the variance of a probability distribution (Tobler et al., 2007). Yet others examined estimation of confi-

dence about a probabilistic decision (Grinband et al., 2006; Kepes et al., 2008).

The origin of randomness in decision making has been an issue pertaining to the debate on whether the same core mechanisms could underlie perceptual decisions and valued-based choice behavior (Glimcher, 2005; Gold and Shadlen, 2007). Gold and Shadlen (2007) described decision making as a process in which “the decision variable (DV) represents the accrual of all sources of priors, evidence, and value into a quantity that is interpreted by the decision rule to produce a choice.” Thus, in accumulator models of perceptual decision, randomness originates from noise in the external input, so the DV is stochastic but the decision rule (the bound) is fixed. The computational benefit of time integration is understood in terms of the signal-to-noise ratio, which increases over time ( $\sim\sqrt{t}$  in the diffusion model). On the other hand, in reinforcement models for reward-dependent choice (Barracough et al., 2004; Sugrue et al., 2004, 2005; Lau and Glimcher, 2005), the DV is defined by values of response options which are deterministically updated according to past rewards, whereas the choice is generated by a probabilistic decision rule (e.g., a softmax criterion) based on the DV. The source of stochasticity is thus interpreted as internal. Glimcher (2005) argued that intrinsic indeterminacy may be essential for unpredictable behavior. For example, in interactive games like matching pennies or rock-paper-scissors, any trend that deviates from random choice by an agent could be exploited to his or her opponent’s advantage.

The recurrent neural circuit model offers a way to reconcile these two seemingly contrasting views. In this model, there is no fundamental distinction between the DV and the decision rule, insofar as the same recurrent neural dynamics instantiate the accrual process as well as categorical choice. We interpret neural activity in a putative decision network as the DV in both reward-based choice tasks and perceptual tasks. Reward or value signals modulate neural firing through synaptic inputs, just like sensory stimuli, in consonance with the view of Gold and Shadlen (2007). The neural dynamics give rise to stochastic decisions, with the aggregate behavior characterized by a softmax function of the difference  $\Delta I$  in the inputs to the competing decision neurons. This softmax is simply a description of behavioral statistics, not the decision criterion used to produce individual choices in single trials. The smaller is the absolute value of  $\Delta I$ , the more random is the network behavior. In a foraging task, the serial responses and outcomes lead to changes in the synaptic strengths so that  $\Delta I$  reflects the difference in the values of the choice options. When the amount of reward uncertainty is varied,  $\Delta I$  is adjusted through synaptic plasticity so that the system behaves more, or less, randomly in compliance with the matching law (Soltani and Wang, 2006). In a competitive game, the interplay with the opponent induces reward-dependent synaptic plasticity that forces  $\Delta I$  to be close to zero, resulting in random behavior. Therefore, a decision maker does not have a goal to play randomly, but simply tries to play at its best, given the environment or other decision agents in an interactive game (Soltani et al., 2006; Lee and Wang, 2008). This conclusion is consistent with behavioral studies demonstrating an indispensable role of feedback in producing random patterns of responses (Rapoport and Budescu, 1992; Camerer, 2003; Glimcher, 2005). In this

view, a decision circuit produces random-choice behavior, not necessarily because the system relies on a “random number generator,” but because the trial-to-trial interplay between a decision maker with a volatile environment or with other decision makers leads to adaptive and seemingly random decision patterns.

This perspective, emphasizing intrinsic stochastic neural dynamics, also applies to perceptual decisions, where  $\Delta/\sigma$  measures the relative input strength (such as the coherence of an RDM stimulus). To appreciate this point, it is worth noting that even determinist choices are associated with some behavioral variability, notably the trial-to-trial variability of response times. Consider the simple act of saccading (a ballistic eye movement) to a suddenly appearing visual target. There is no sensory uncertainty, and the behavioral response is always the same. However, saccade response time (the time between target onset and saccadic eye movement) fluctuates considerably from trial to trial (Carpenter, 1981). In monkey physiological experiments, Hanes and Schall found that a saccade was initiated when the firing activity reached a threshold level of movement neurons in the FEF (Hanes and Schall, 1996; Schall, 2001). Trial-to-trial variability of saccade response time was shown to be inversely correlated with the slope of the buildup activity of movement-related neurons, whereas the threshold level remained constant independent of the response time (Hanes and Schall, 1996). There is also evidence for a trial-to-trial correlation between the response latency and the preparatory activity of cortical movement neurons, before the target onset, in saccade and other sensory-motor tasks (Dorris et al., 1997; Churchland et al., 2006; Nakahara et al., 2006). In situations when there is a conflict, for instance when the subject has to inhibit a planned saccade by a stop signal introduced with a short delay after the target onset, the behavior becomes probabilistic (the saccade is suppressed on some trials, but not on other trials) (Logan and Cowan, 1984; Boucher et al., 2007). Therefore, in some sense, the stochasticity inherent in a neural system reveals itself by external uncertainty.

The signal-detection theory of perception explains behavioral indeterminacy in terms of noisy input (Green and Swets, 1966). However, precious little is known physiologically about the identities and relative weights of various sources that contribute to randomness in a decision process. In a monkey RDM discrimination experiment, there are at least three components of noise that influence a decision circuit like LIP: the stochastic spatio-temporal dot pattern presented in a trial, the trial-to-trial stimulus variation, and fluctuating dynamics intrinsic in the nervous system. Interestingly, it was found that the trial-to-trial stimulus variation had no discernible effect on the trial-to-trial variance of firing activity in MT neurons (Britten et al., 1993), nor on the relationship between MT neural responses and behavioral choice (Britten et al., 1996). Similarly, in the recurrent neural circuit model, the probabilistic decision behavior, measured by the psychometric function and the variability of response time, was found to be the same in simulations when the external inputs varied from trial to trial or remained fixed across trials, suggesting that the main source of variability may not be the sensory stimulus but within the neural system itself (Wang, 2002). In addition, Deco et al. applied this model to the monkey somatosensory discrimination experiment and showed that the intrinsically sto-

chastic decision circuit dynamics could account for Weber’s law, which states that the ratio of just-noticeable input difference to absolute input intensity is constant (Deco and Rolls, 2006; Deco et al., 2007b). These examples illustrate the potential of how statistical behavioral laws can ultimately be explained in neuronal terms.

Both for models of perceptual decisions and value-based choice behavior, intrinsic stochasticity arises from highly irregular neural firing (Amit and Brunel, 1997; Brunel and Wang, 2001), a characteristic of cortical neuronal firing (Softky and Koch, 1993; Shadlen and Newsome, 1994). Evidence suggests that this randomness is inherent within cortical circuits. For instance, Poisson-like statistics is a characteristic of delay-period persistent activity of prefrontal neurons recorded from behaving monkeys, even in the absence of external stimulus during working memory (Compte et al., 2003). Theoretically proposed mechanisms for highly irregular neural activity in the cortex posit recurrent cortical circuit dynamics endowed with balanced synaptic excitation and inhibition (van Vreeswijk and Sompolinsky, 1996; Amit and Brunel, 1997; Mattia and Del Giudice, 2004; Renart et al., 2007; Barbieri and Brunel, 2008). Unlike the diffusion model in which the input is given and integrated by the DV, in a recurrent circuit a substantial component of the synaptic input comes “from within” and builds up over time in parallel with the spiking activity. Therefore, time integration is over the total (external and recurrent) input, rather than sensory stimulus alone. If neurons in a putative decision circuit like LIP exhibit Poisson statistics through a ramping time course, then the signal-to-noise ratio improves over time simply as a consequence of the increased mean firing rate (true for any Poisson process), rather than because noise in the external stimulus is averaged out. Furthermore, it is not a foregone conclusion that the signal-to-noise ratio indeed decays over time for decision neurons, because a neural accumulator, due to its lack of a significant leak, is expected to display unusual fluctuations, e.g., the Fano factor of spike trains (the ratio of the variance over the mean of the spike count) may actually grow over time (Miller and Wang, 2006b).

The phase-plane analysis (Figure 3D) offers a new look at the issue of signal-to-noise ratio. At any given time, define signal as the distance  $d(t)$  between the current network state (given by the two neural pool firing rates  $r_A$  and  $r_B$ ) and the boundary that separates the basins of attraction of the two choice attractors. Noise can be quantified by the trial-to-trial standard deviation [ $\sigma(t)$ ] of the network state. The signal-to-noise ratio is  $d(t)/\sigma(t)$ . At the onset of a RDM stimulus, the initial signal  $d(t = 0)$  depends on the motion strength  $c'$ ; it is zero if  $c' = 0$ , but is positive for nonzero  $c'$  because the network already starts inside the basin of the correct choice attractor (cf. Figure 3D). However, the network remains “undecided” as long as  $d(t)/\sigma(t)$  is small. As the network evolves further into one of the basins of attraction, both  $d(t)$  and  $\sigma(t)$  may increase over time, but the ratio  $d(t)/\sigma(t)$  grows, therefore it becomes increasingly unlikely that noise can “bring back” the network across the boundary to the other, alternative, attraction basin. In this sense, one may say that a categorical choice is reached when  $d(t)$  becomes much larger than  $\sigma(t)$ , even though the network may be still far away from the actual attractor state or neither of the two firing rates have reached

a prescribed decision threshold. According to this state dynamics perspective on signal-to-noise ratio, there is no need to separately treat external noisy stimulus and internal neuronal stochasticity. The computational benefit of time integration is understood through the network dynamics in an attractor landscape, rather than in terms of a traditional time domain analysis. Note that the network does not directly “see” the motion coherence  $c'$ , only the RDM stimulus, but the decision-space landscape is sensitive to the mean input that reflects  $c'$ . For a higher  $c'$  (Figure 3D,  $c' = 51.2\%$  versus  $6.4\%$ ), the network starts out deeper in the territory of the correct choice attractor,  $d$  is larger at the stimulus onset, and the time point of a categorical choice is earlier. With a sufficiently large  $c'$ , the performance is 100% because, as soon as the stimulus is presented,  $d(t)/\sigma(t)$  is already so large that switching to the alternative choice attractor (an error) is impossible. One can say that the system has already “made up its mind” at the stimulus onset, even though it takes sometime for neural activity to reach a threshold level. Note that each state-based depiction corresponds to a (sufficiently long) stationary input. It does not mean that the decision is irreversible; a change in the external input (e.g., reversing the direction of the motion stimulus) can radically alter the attractor landscape, leading to a different choice.

It has been proposed that in a single trial, neural population activity patterns explicitly represent probability density functions (Ma et al., 2006). Applied to the RDM discrimination experiment, ramping spiking activity of LIP neurons has been interpreted as a temporal summation of the logarithm of the likelihood ratio (Gold and Shadlen, 2001; Jazayeri and Movshon, 2006), or of the posteriors (that combine evidence with prior information) (Ma et al., 2006; Jazayeri and Movshon, 2006), about the two alternatives. Such models require a Bayesian decoder that uses a nonlinear process to readout the categorical choice. The recurrent neural circuit model offers a different perspective in which the same decision circuit performs both temporal integration of data and categorical choice by attractor dynamics. Furthermore, decision making in single trials is based on random sampling of fluctuating neural network activity; probabilistic distributions appear only in the aggregated statistics across trials. Future research will help us to understand these different modes of operation, possibly deployed differentially in distinct brain regions.

### Concluding Remarks

Decision making has recently attracted increasing attention not only in the neurobiological studies of cognition but also in psychiatry with the recognition that impaired decision making is prominently associated with various mental disorders (Fellows, 2004; Sachdev and Malhi, 2005). In this review, I have marshaled experimental findings on the basis of which a recurrent neural circuit theory for decision making has been developed. As it has now become possible to investigate decision making across species, from flies, rats, and monkeys to human subjects, the time is ripe to investigate the underlying mechanisms in terms of the biophysics of single neurons (Llinás, 1988; Magee et al., 1998), the dynamics of synaptic connections (Abbott and Regehr, 2004), and the microcircuit wiring connectivity (Somogyi et al., 1998; Douglas and Martin, 2004). An insight from a nonlin-

ear dynamical systems perspective is that quantitative differences give rise to qualitatively different functions. Thus, while the posterior parietal and prefrontal cortex may have qualitatively the same architecture layout as sensory cortices, sufficiently strong synaptic recurrence (provided that it is slow) can naturally lead to the generation of persistent activity and ramping activity suitable for subserving cognitive-type computations (Wang, 2006a). Conversely, relatively modest reductions of recurrent excitation and inhibition could produce marked impairments of cognitive functions (Wang, 2006b; Durstewitz and Seamans, 2008; Rolls et al., 2008).

A key neural computation in both working memory and decision making can be conceptualized as the time integral of inputs: working memory relies on neurons that convert a transient input pulse into a self-sustained persistent activity, whereas decision making involves quasilinear ramping activity in response to a constant input for accumulation of information. Perceptual discrimination (Pasternak and Greenlee, 2005) and action selection (Tanji and Hoshi, 2008) tasks often also depend working memory, in order to retain information useful for a future decision or to remember a choice made previously. However, it is still unclear whether the underlying circuit mechanism is necessarily the same for stimulus-selective persistent activity in working memory and accumulation of evidence in a decision process. In RDM discrimination experiments, recorded LIP neurons were preselected using the criterion that they displayed directionally tuned mnemonic activity in a delayed oculomotor response task (Shadlen and Newsome, 1996, 2001; Roitman and Shadlen, 2002; Huk and Shadlen, 2005). It would be interesting to examine whether other neurons that do not show persistent activity also display slow ramping activity in the RDM discrimination task. Furthermore, the cellular and synaptic mechanisms of neural ramping activity remain to be elucidated experimentally. Whether LIP indeed acts as an attractor network has also been questioned on the ground that certain aspects of neural responses in LIP during selective attention have not been reproduced by existing attractor models (Ganguli et al., 2008).

Choice behavior is commonly formulated in terms of value-based optimization. One challenge is thus to elucidate how various dimensions of valuation are represented in the brain (Sugrue et al., 2005; Rushworth and Behrens, 2008). Another is to understand the neural metrics of uncertainty and risk and how it influences decision making. In an uncertain world, reinforcement sometimes needs to be counterbalanced by exploratory decisions. It will be interesting to study how, according to behavioral demands, the brain can exploit a known environment or explore optimal options in a volatile world (Daw et al., 2006; Behrens et al., 2007).

Although the present article is centered on simple behavioral tasks, elemental building blocks of decision computation and their neural mechanisms are likely to be relevant to more complex cognitive processes as well. Even in spoken-language processing, there is evidence that a spoken word elicits multiple lexical representations, and spoken-word recognition proceeds from real-time integration of information sources to categorical choice among phonological competitors (Spivey et al., 2005). It is thus hoped that understanding higher-level decisions can benefit from detailed neurophysiological studies of simpler perceptual and economic decisions.

## ACKNOWLEDGMENTS

I am grateful to my collaborators for their contributions: Nicolas Brunel, Albert Compte, Stefano Fusi, Alexander Huk, Daeyeol Lee, Feng Liu, Chung-Chuan Lo, Earl K. Miller, Paul Miller, Alfonso Renart, Michael N. Shadlen, Alireza Soltani, and Kong-Fatt Wong. I also thank Peter Dayan, Josh Gold, and the anonymous referees for their helpful comments on the manuscript; to Kong-Fatt Wong, Alireza Soltani, Carlos Brody, and Josh Gold for their help with figures. This work was supported by NIH grant 2-R01-MH062349 and the Kavli Foundation.

## REFERENCES

- Abbott, L.F., and Regehr, W.G. (2004). Synaptic computation. *Nature* 431, 796–803.
- Abraham, N., Spors, H., Carleton, A., Margrie, T., Kuner, T., and Schaefer, A. (2004). Maintaining accuracy at the expense of speed: stimulus similarity defines odor discrimination time in mice. *Neuron* 44, 865–876.
- Amit, D.J. (1995). The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behav. Brain Sci.* 18, 617–626.
- Amit, D.J., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252.
- Andersen, R.A., and Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189–220.
- Barbieri, F., and Brunel, N. (2008). Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex? *Frontiers in Neurosci.* 2, 114–122.
- Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.
- Bayer, H., and Glimcher, P. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.
- Behrens, T., Woolrich, M., Walton, M., and Rushworth, M. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.
- Binder, J., Liebenthal, E., Possing, E., Medler, D., and Ward, B. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295–301.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.
- Bogacz, R., Usher, M., Zhang, J., and McClelland, J. (2007). Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1655–1670.
- Boucher, L., Palmeri, T., Logan, G., and Schall, J. (2007). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376–397.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* 10, 1157–1169.
- Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13, 87–100.
- Brody, C., Hernández, A., Zainos, A., and Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cereb. Cortex* 13, 1196–1207.
- Brunel, N., and Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63–85.
- Busemeyer, J., and Townsend, J. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* 100, 432–459.
- Camerer, C.F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton, NJ: Princeton University Press).
- Carpenter, R.H.S. (1981). Oculomotor procrastination. In *Eye Movements: Cognition and Visual Perception*, D.F. Fisher, R.A. Monty, and J.W. Senders, eds. (Hillsdale, NJ: Lawrence Erlbaum), pp. 237–246.
- Chafee, M.V., and Goldman-Rakic, P.S. (1998). Neuronal activity in macaque prefrontal area 8a and posterior parietal area 7ip related to memory guided saccades. *J. Neurophysiol.* 79, 2919–2940.
- Churchland, M., Afshar, A., and Shenoy, K. (2006). A central source of movement variability. *Neuron* 52, 1085–1096.
- Churchland, M., Kiani, R., and Shadlen, M. (2008). Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702.
- Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: a computational model. *J. Neurosci.* 26, 9761–9770.
- Cisek, P., and Kalaska, J. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* 45, 801–814.
- Cohen, J., Braver, T., and O'Reilly, R. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1515–1527.
- Colby, C.L., and Goldberg, M.E. (1999). Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349.
- Compte, A., Constantinidis, C., Tegnér, J., Raghavachari, S., Chafee, M.V., Goldman-Rakic, P.S., and Wang, X.-J. (2003). Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J. Neurophysiol.* 90, 3441–3454.
- Cook, E.P., and Maunsell, J.H.R. (2002). Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat. Neurosci.* 5, 985–994.
- Corbetta, M., and Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Corrado, G., Sugrue, L., Seung, H., and Newsome, W. (2005). Linear-nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.
- Daw, N., O'Doherty, J., Dayan, P., Seymour, B., and Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.
- de Lafuente, V., and Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* 8, 1698–1703.
- Deco, G., and Rolls, E. (2006). Decision-making and Weber's law: a neurophysiological model. *Eur. J. Neurosci.* 24, 901–916.
- Deco, G., Prez-Sanagustín, M., de Lafuente, V., and Romo, R. (2007a). Perceptual detection as a dynamical bistability phenomenon: a neurocomputational correlate of sensation. *Proc. Natl. Acad. Sci. USA* 104, 20073–20077.
- Deco, G., Scarano, L., and Soto-Faraco, S. (2007b). Weber's law in decision making: integrating behavioral data in humans with a neurophysiological model. *J. Neurosci.* 27, 11192–11200.
- Ditterich, J. (2006). Evidence for time-variant decision making. *Eur. J. Neurosci.* 24, 3628–3641.
- Donders, F.C. (1969). On the speed of mental processes. [Translation of Die Schnelligkeit Psychischer Prozesse, first published in 1868]. *Acta Psychologica* 30, 412–431.
- Dorris, M., Paré, M., and Munoz, D. (1997). Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements. *J. Neurosci.* 17, 8566–8579.
- Dorris, M.C., and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

- Douglas, R.J., and Martin, K.A.C. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451.
- Durstewitz, D., and Seamans, J. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 10.1016/j.biopsych.2008.05.015.
- Ewert, J. (1997). Neural correlates of key stimulus and releasing mechanism: a case study and two concepts. *Trends Neurosci.* 20, 332–339.
- Fellows, L.K. (2004). The cognitive neuroscience of human decision making: a review and conceptual framework. *Behav. Cogn. Neurosci. Rev.* 3, 159–172.
- Fiorillo, C., Tobler, P., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.
- Frank, M., and Claus, E. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113, 300–326.
- Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349.
- Fusi, S. (2002). Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.* 87, 459–470.
- Fusi, S., Asaad, W.F., Miller, E.K., and Wang, X.-J. (2007). A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron* 54, 319–333.
- Fuster, J.M. (2008). The Prefrontal Cortex, Fourth Edition (New York: Academic Press).
- Ganguli, S., Bisley, J.W., Roitman, J.D., Shadlen, M.N., Goldberg, M.E., and Miller, K.D. (2008). One-dimensional dynamics of attention and decision making in LIP. *Neuron* 58, 15–25.
- Glimcher, P.W. (2003). Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics (Cambridge, MA: MIT Press).
- Glimcher, P.W. (2005). Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25–56.
- Gnadt, J.W., and Andersen, R.A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216–220.
- Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16.
- Gold, J.I., and Shadlen, M.N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.
- Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
- Goldman-Rakic, P.S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485.
- Gratton, G., Coles, M., Sirevaag, E., Eriksen, C., and Donchin, E. (1988). Pre- and poststimulus activation of response channels: a psychophysiological analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 331–344.
- Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (New York: John Wiley and Sons).
- Grinband, J., Hirsch, J., and Ferrera, V.P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron* 49, 757–763.
- Grossberg, S., and Pilly, P. (2008). Temporal dynamics of decision-making during motion perception in the visual cortex. *Vision Res.* 48, 1345–1373.
- Hanes, D.P., and Schall, J.D. (1996). Neural control of voluntary movement initiation. *Science* 274, 427–430.
- Hanks, T.D., Ditterich, J., and Shadlen, M.N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* 9, 682–689.
- Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature* 431, 859–862.
- Heekeren, H., Marrett, S., and Ungerleider, L. (2008). The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467–479.
- Heinzle, J., Hepp, K., and Martin, K. (2007). A microcircuit model of the frontal eye fields. *J. Neurosci.* 27, 9341–9353.
- Hernández, A., Zainos, A., and Romo, R. (2002). Temporal evolution of a decision making process in medial premotor cortex. *Neuron* 33, 959–972.
- Herrnstein, R.J., Rachlin, H., and Laibson, D.I. (1997). The Matching Law: Papers in Psychology and Economics (Cambridge: Harvard University Press).
- Hestrin, S., Sah, P., and Nicoll, R. (1990). Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron* 5, 247–253.
- Hick, W.E. (1952). On the rate of gain of information. *Q. J. Exp. Psychol.* 4, 11–26.
- Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953–978.
- Hikosaka, O., Nakamura, K., and Nakahara, H. (2006). Basal ganglia orient eyes to reward. *J. Neurophysiol.* 95, 567–584.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.
- Huettel, S., Stowe, C., Gordon, E., Warner, B., and Platt, M. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.
- Huk, A.C., and Shadlen, M.N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.
- Izhikevich, E. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* 17, 2443–2452.
- Jazayeri, M., and Movshon, J.A. (2006). Optimal representation of sensory information by neural populations. *Nat. Neurosci.* 9, 690–696.
- Kahneman, D. (2002). Nobel prize lecture: Maps of bounded rationality: a perspective on intuitive judgment and choice. In Nobel Prizes 2002: Nobel Prizes, Presentations, Biographies, & Lectures, T. Frängsmyr, ed. (Stockholm: Almqvist & Wiksell Int), pp. 416–499.
- Kennerley, S., Walton, M., Behrens, T., Buckley, M., and Rushworth, M. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940–947.
- Kepecs, A., Uchida, N., Zariwala, H., and Mainen, Z.F. (2008). Neural correlations, computation and behavioral impact of decision confidence. *Nature* 455, 227–231.
- Kiani, R., Hanks, T., and Shadlen, M. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.
- Kim, J.-N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2, 176–183.
- Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197, 792–795.
- Laming, D. (1968). Information Theory of Choice Reaction Times (New York: Academic Press).
- Lau, B., and Glimcher, P.W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.
- Lau, B., and Glimcher, P. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463.
- Law, C., and Gold, J. (2008). Neural correlates of perceptual learning in a sensorymotor, but not a sensory, cortical area. *Nat. Neurosci.* 11, 505–513.

- Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404–409.
- Lee, D., and Wang, X.-J. (2008). Mechanisms for stochastic decision making in the primate frontal cortex: Single-neuron recording and circuit modeling. In *Neuroeconomics: Decision Making and the Brain*, E.F.P.W. Glimcher, C.F. Camerer, and R.A. Poldrack, eds. (New York: Academic Press).
- Lemus, L., Hernández, A., Luna, R., Zainos, A., Nácher, V., and Romo, R. (2007). Neural correlates of a postponed decision report. *Proc. Natl. Acad. Sci. USA* 104, 17174–17179.
- Llinás, R. (1988). The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* 242, 1654–1664.
- Lo, C.C., and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.
- Loewenstein, G., Rick, S., and Cohen, J. (2008). Neuroeconomics. *Annu. Rev. Psychol.* 59, 647–672.
- Logan, G.D., and Cowan, W.B. (1984). On the ability to inhibit thought and action: a theory of an act of control. *Psychol. Rev.* 91, 295–327.
- Luce, R.D. (1986). *Response Time: Their Role in Inferring Elementary Mental Organization* (New York: Oxford University Press).
- Ma, W., Beck, J., Latham, P., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.* 9, 1432–1438.
- Machens, C.K., Romo, R., and Brody, C.D. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 318, 1121–1124.
- Magee, J., Hoffman, D., Colbert, C., and Johnston, D. (1998). Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons. *Annu. Rev. Physiol.* 60, 327–346.
- Major, G., and Tank, D. (2004). Persistent neural activity: prevalence and mechanisms. *Curr. Opin. Neurobiol.* 14, 675–684.
- Matsuda, Y., Marzo, A., and Otani, S. (2006). The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex. *J. Neurosci.* 26, 4803–4810.
- Matsumoto, M., and Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447, 1111–1115.
- Mattia, M., and Del Giudice, P. (2004). Finite-size dynamics of inhibitory and excitatory interacting spiking neurons. *Phys. Rev. E Stat. Nonlin. Soft Matter Physiol.* 70, 052903.
- Mazurek, M.E., Roitman, J.D., Ditterich, J., and Shadlen, M.N. (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex* 13, 1257–1269.
- McCarthy, G., and Donchin, E. (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science* 211, 77–80.
- McMillen, T., and Holmes, P. (2006). The dynamics of choice among multiple alternatives. *J. Math. Psychol.* 50, 30–57.
- McPeek, R., and Keller, E. (2002). Saccade target selection in the superior colliculus during a visual search task. *J. Neurophysiol.* 88, 2019–2034.
- Meyer, D.E., Osman, A.M., Irwin, D.E., and Yantis, S. (1988). Modern mental chronometry. *Biol. Psychol.* 26, 3–67.
- Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Miller, P., and Wang, X.-J. (2006a). Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli. *Proc. Natl. Acad. Sci. USA* 103, 201–206.
- Miller, P., and Wang, X.-J. (2006b). Power-law neuronal fluctuations in a recurrent network model of parametric working memory. *J. Neurophysiol.* 95, 1099–1114.
- Miller, P., Brody, C.D., Romo, R., and Wang, X.-J. (2003). A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb. Cortex* 13, 1208–1218.
- Munoz, D.P., and Fecteau, J.H. (2002). Vying for dominance: dynamic interactions control visual fixation and saccadic initiation in the superior colliculus. *Prog. Brain Res.* 140, 3–19.
- Nakahara, H., Nakamura, K., and Hikosaka, O. (2006). Extended LATER model can account for trial-by-trial variability of both pre- and post-processes. *Neural Netw.* 19, 1027–1046.
- Narayanan, N., and Laubach, M. (2006). Top-down control of motor cortex ensembles by dorsomedial prefrontal cortex. *Neuron* 52, 921–931.
- Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.
- Niwa, M., and Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435–4445.
- Okamoto, H., Isomura, Y., Takada, M., and Fukai, T. (2007). Temporal integration by stochastic recurrent network dynamics with bimodal neurons. *J. Neurophysiol.* 97, 3859–3867.
- Otani, S., Daniel, H., Roisin, M., and Crepel, F. (2003). Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cereb. Cortex* 13, 1251–1256.
- Padoa-Schioppa, C., and Assad, J. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
- Parker, A., and Krug, K. (2003). Neuronal mechanisms for the perception of ambiguous stimuli. *Curr. Opin. Neurobiol.* 13, 433–439.
- Pasternak, T., and Greenlee, M. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6, 97–107.
- Pesaran, B., Nelson, M., and Andersen, R. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453, 406–409.
- Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
- Ploran, E., Nelson, S., Velanova, K., Donaldson, D., Petersen, S., and Wheeler, M. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J. Neurosci.* 27, 11912–11924.
- Posner, M.I. (1978). *Chronometric Exploration of Mind* (Hillsdale, NJ: Lawrence Erlbaum Associates).
- Quintana, J., and Fuster, J.M. (1999). From perception to action: temporal integrative functions of prefrontal and parietal neurons. *Cereb. Cortex* 9, 213–221.
- Rapoport, A., and Budescu, D.V. (1992). Generation of random series in two-person strictly competitive games. *J. Exp. Psychol. Gen.* 121, 352–363.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.
- Ratcliff, R., and Rouder, J.F. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.
- Real, L. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science* 253, 980–986.
- Renart, A., Moreno-Bote, R., Wang, X.-J., and Parga, N. (2007). Mean-driven and fluctuation-driven persistent activity in recurrent networks. *Neural Comput.* 19, 1–46.
- Reynolds, J.N.J., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67–70.
- Roesch, M., and Olson, C. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* 90, 1766–1789.
- Roesch, M., Calu, D., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.

- Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.
- Rolls, E., Loh, M., Deco, G., and Winterer, G. (2008). Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat. Rev. Neurosci.* 9, 696–709.
- Romo, R., and Salinas, E. (2001). Touch and go: Decision-making mechanisms in somatosensation. *Annu. Rev. Neurosci.* 24, 107–137.
- Romo, R., Brody, C.D., Hernández, A., and Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470–474.
- Romo, R., Hernández, A., Zainos, A., Lemus, L., and Brody, C. (2002). Neuronal correlates of decision-making in secondary somatosensory cortex. *Nat. Neurosci.* 5, 1217–1225.
- Romo, R., Hernández, A., and Zainos, A. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron* 41, 165–173.
- Roxin, A., and Ledberg, A. (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4, e1000046.
- Rushworth, M., and Behrens, T. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389–397.
- Sachdev, P., and Malhi, G. (2005). Obsessive-compulsive behaviour: a disorder of decision-making. *Aust. N. Z. J. Psychiatry* 39, 757–763.
- Samejima, K., and Doya, K. (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. *Ann. N Y Acad. Sci.* 1104, 213–228.
- Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of actionspecific reward values in the striatum. *Science* 310, 1337–1340.
- Schall, J.D. (2001). Neural basis of deciding, choosing and acting. *Nat. Neurosci.* 2, 33–42.
- Schall, J. (2004). On building a bridge between brain and behavior. *Annu. Rev. Psychol.* 55, 23–50.
- Scherberger, H., and Andersen, R. (2007). Target selection signals for arm reaching in the posterior parietal cortex. *J. Neurosci.* 27, 2001–2012.
- Schlegel, T., and Schuster, S. (2008). Small circuits for large tasks: high-speed decision-making in archerfish. *Science* 319, 104–106.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Seo, H., and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.
- Sereno, A., and Amador, S. (2006). Attention and memory-related responses of neurons in the lateral intraparietal area during spatial and shape-delayed matchto-sample tasks. *J. Neurophysiol.* 95, 1078–1098.
- Seung, H.S. (1996). How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. USA* 93, 13339–13344.
- Seung, H. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* 40, 1063–1073.
- Seung, H.S., Lee, D.D., Reis, B.Y., and Tank, D.W. (2000). Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26, 259–271.
- Shadlen, M.N., and Newsome, W.T. (1994). Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* 4, 569–579.
- Shadlen, M.N., and Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proc. Natl. Acad. Sci. USA* 93, 628–633.
- Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.
- Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848–851.
- Sigman, M., and Dehaene, S. (2005). Parsing a cognitive task: a characterization of the mind's bottleneck. *PLoS Biol.* 3, e37.
- Smith, P.L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.
- Softky, W.R., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* 13, 334–350.
- Soltani, A., and Wang, X.-J. (2006). A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* 26, 3731–3744.
- Soltani, A., and Wang, X.-J. (2008). From biophysics to cognition: reward-dependent adaptive choice behavior. *Curr. Opin. Neurobiol.* 18, 209–216.
- Soltani, A., Lee, D., and Wang, X.-J. (2006). Neural mechanism for stochastic behavior during a competitive game. *Neural Netw.* 19, 1075–1090.
- Somogyi, P., Tamás, G., Luján, R., and Buhl, E.H. (1998). Salient features of synaptic organization in the cerebral cortex. *Brain Res. Brain Res. Rev.* 26, 113–135.
- Soon, C., Brass, M., Heinze, H., and Haynes, J. (2008). Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543–545.
- Spivey, M., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. Natl. Acad. Sci. USA* 102, 10393–10398.
- Stone, M. (1960). Models for choice reaction time. *Psychometrika* 25, 251–260.
- Sugrue, L.P., Corrado, G.C., and Newsome, W.T. (2004). Matching behavior and representation of value in parietal cortex. *Science* 304, 1782–1787.
- Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.
- Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (Cambridge, MA: MIT Press).
- Tanji, J., and Hoshi, E. (2008). Role of the lateral prefrontal cortex in executive behavioral control. *Physiol. Rev.* 88, 37–57.
- Tobler, P., Fiorillo, C., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.
- Tobler, P., O'Doherty, J., Dolan, R., and Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621–1632.
- Usher, M., and McClelland, J. (2001). On the time course of perceptual choice: the leaky competing accumulator model. *Psychol. Rev.* 108, 550–592.
- Uchida, N., and Mainen, Z. (2003). Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* 6, 1224–1229.
- Uchida, A., Kepecs, A., and Mainen, Z.F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nat. Rev. Neurosci.* 7, 485–491.
- van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. In *Contemporary Problems in Perception: Ergonomics*, A.T. Welfare and L. Houssiaud, eds. (London: Taylor & Francis), pp. 37–58.
- Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* 19, 9587–9603.
- Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455–463.

- Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.
- Wang, X.-J. (2006a). A microcircuit model of prefrontal functions: ying and yang of reverberatory neurodynamics in cognition. In *The Prefrontal Lobes: Development, Function and Pathology*, J. Risberg, J. Grafman, and F. Boller, eds. (New York: Cambridge University Press), pp. 92–127.
- Wang, X.-J. (2006b). Toward a prefrontal microcircuit model for cognitive deficits in schizophrenia. *Pharmacopsychiatry* 39 (*Suppl 1*), 80–87.
- Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.
- Watanabe, K., and Funahashi, S. (2007). Prefrontal delay-period activity reflects the decision process of a saccade direction during a free-choice ODR task. *Cereb. Cortex* 17 (*Suppl 1*), 88–100.
- Wickelgren, W.A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol. (Amst.)* 41, 67–85.
- Wickens, J.R., Reynolds, J.N.J., and Hyland, B.I. (2003). Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* 13, 685–690.
- Wilimzig, C., Schneider, S., and Schöner, G. (2006). The time course of saccadic decision making: dynamic field theory. *Neural Netw.* 19, 1059–1074.
- Wong, K.F., and Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328.
- Wong, K.F., Huk, A.C., Shadlen, M.N., and Wang, X.-J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision-making. *Frontiers in Comput. Neurosci.* 1, 6. 10.3389/neuro.10/006.2007.
- Wörgötter, F., and Porr, B. (2005). Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Comput.* 17, 245–319.
- Yu, A.J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.