modulation of siesta behavior could be surveyed in additional *Drosophila* populations and species to more precisely determine its degree of conservation. Intriguingly, in the distantly related fungus *Neurospora crassa* the central clock gene *frequency* undergoes temperature-dependent splicing to ensure a stable period length across a range of temperatures, suggesting that circadian clock circuits may have preferentially incorporated temperature-sensitive splicing mechanisms (Diernfellner et al., 2007).

In a broader context, the experiments presented in the paper point to multiple mechanisms besides dmpi8 splicing that seasonally modulate daily activity in *Drosophila*. First, photoperiod modulates the onset of evening activity in the absence of conditional *per* splicing at the last intron in *D. yakuba*. Second, temperature modulates siesta time in transgenic *D. melanogaster* lacking a temperature-sensitive dmpi8. While the explanation for the former observation likely involves clock-dependent synchronization of the evening activity peak relative to dusk, a combination of clock-dependent and clock-independent behavioral responses to temperature may help explain the latter observation. Regardless, it is clear that we can expect new discoveries concerning the molecular basis of seasonal behavior to keep appearing for quite some time.

REFERENCES

Chen, W.F., Majercak, J., and Edery, I. (2006). J. Biol. Rhythms *21*, 256–271.

Collins, B.H., Rosato, E., and Kyriacou, C.P. (2004). Proc. Natl. Acad. Sci. USA *101*, 1945–1950.

Diernfellner, A., Colot, H.V., Dintsis, O., Loros, J.J., Dunlap, J.C., and Brunner, M. (2007). FEBS Lett. *581*, 5759–5764.

Dubruille, R., and Emery, P. (2008). Mol. Neurobiol. *38*, 129–145.

Dunlap, J.C., Loros, J.J., and DeCoursey, P.J. (2004). Chronobiology: Biological Timekeeping (Sunderland, Mass: Sinauer Associates).

Hardin, P.E. (2005). Curr. Biol. *15*, R714–R722.

Low, K.H., Lim, C., Ko, H.W., and Edery, I. (2008). Neuron *60*, this issue, 1054–1067.

Majercak, J., Sidote, D., Hardin, P.E., and Edery, I. (1999). Neuron *24*, 219–230.

Majercak, J., Chen, W.F., and Edery, I. (2004). Mol. Cell. Biol. *24*, 3359–3372.

Stanewsky, R. (2003). J. Neurobiol. *54*, 111–147.

Wijnen, H., and Young, M.W. (2006). Annu. Rev. Genet. *40*, 409–448.

So Many Choices: What Computational Models Reveal about Decision-Making Mechanisms

Emilio Salinas[1,*]
[1]Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA
*Correspondence: esalinas@wfubmc.edu
DOI 10.1016/j.neuron.2008.12.011

Is decision making in the brain (a) optimal, (b) stochastic, (c) probabilistic, or (d) all of the above? Two papers in this issue of *Neuron* by Beck et al. and Furman and Wang address these questions by constructing model neural circuits capable of picking one option given multiple perceptual choices.

The neurobiological basis of decision-making has been intensely studied by systems neuroscientists over the last few decades. Much of this work has been based on paradigms in which, in each trial, a subject selects one of two possible choices. In parallel with the experimental work, a variety of theories and models have been proposed, and what stands out is that even highly simplified phenomenological models have been able to capture much of the psychophysical and neurobiological data in the two-alternative tasks (Smith and Ratcliff, 2004). Recently, however, Mike Shadlen's group added an important piece to this puzzle with a neurophysiological study that had a novel feature: it used a task with four alternatives (Churchland et al., 2008). These new experimental results, which on the surface seem deceptively similar to those obtained with two-alternative tasks, strongly disambiguate and constrain the models, and as a consequence, this issue of *Neuron* delivers something rather unprecedented in neuroscience: back-to-back theoretical papers addressing the same experimental data but starting from entirely different approaches (Beck et al., 2008; Furman and Wang, 2008). Both studies describe neural circuits that replicate psychophysical and neurophysiological results obtained during choice behaviors. However, Xiao-Jing Wang's group aimed to capture as much biophysical detail as possible, whereas Alex Pouget's group aimed to implement key mathematical principles that neural circuits should employ if they are to generate optimal choices.

One of the most popular paradigms for studying the neural basis of decision-making is a two-alternative forced-choice task in which a group of dots moving in random directions is displayed and the subject must decide whether the overall motion is to the left or to the right (or any opposite directions). The choice is indicated with an eye movement either to a left or to a right target. The difficulty of the task is controlled by the coherence parameter, which is the proportion of dots that move in the same direction, left or right. Psychophysical performance in this task is summarized by plotting two quantities, the percentage of correct choices and the subject's reaction time, as functions of coherence.

A lot is known about the neurophysiological underpinning of this task, which includes three key processing stages. First, neurons in cortical area MT respond to the random-dot stimuli and are thought to mediate the perceptual experience of visual motion (Salzman et al., 1992; Britten et al., 1993). Each MT neuron fires preferentially in response to motion in one particular direction.

Second, motion information encoded by the MT responses is then relayed to other cortical areas involved in the generation of motor commands for moving the eyes, such as area LIP. The perceptual decision is at least partially created at this point, because the activity of LIP neurons reflects the progressive accumulation of sensory evidence in favor of one or the other motor alternative (Shadlen and Newsome, 2001; Hanks et al., 2006). Think of two LIP cells with the left and right targets covering their receptive fields. When the subject starts viewing the random-dot display, the firing rates of both neurons are low. Then, as it becomes clear that the dots move, say, preferentially to the left, the activity of the neuron with the left target in its receptive field starts ramping up, and when it reaches a certain threshold level, an eye movement to the left target is triggered. This ramping activity is neither decidedly sensory nor decidedly motor. On one hand, it is tightly correlated with the subject's motor choice. For instance, at zero coherence the subject essentially picks one of the targets at random, and the chosen target is reliably signaled by the LIP firing rates. On the other hand, the slope of the climbing activity increases with coherence, so it is strongly modulated by the strength and direction of motion.

Third, LIP and other cortical areas project to downstream motor centers, such as the superior colliculus, which are thought to select one motor action unambiguously (McPeek and Keller, 2004; Lo and Wang, 2006). Activation of these higher-level motor areas, together with initiation of the eye movement, occurs once the accumulated sensory evidence reaches a certain threshold level (Hanes and Schall, 1996; Shadlen and Newsome, 2001; Lo and Wang, 2006).

The two new modeling studies in this issue of *Neuron* simulate the formation of a decision during the random-dot motion discrimination task in a circuit of interconnected LIP neurons. These models replicate the single-neuron activity recorded in monkeys, as well as the animals' performance and reaction-time curves as functions of coherence, in two- and multiple-alternative tasks. It may sound as if the two model circuits are very similar, and indeed, the networks have similar connectivity footprints and similar input and output representations—but they spring from extremely different mathematical approaches.

The model by Furman and Wang (2008) is closely related to "attractor networks," circuits that, due to strong recurrent interactions, are capable of generating multiple stable patterns of activity in the absence of specific inputs (Wang, 2001). Attractor networks have been traditionally used as models for memory because they produce self-sustained responses; however, only a subset of the neurons can be highly active at any given time, and it is this internal competition which makes them useful for decision making. In contrast, the model by Beck and colleagues (2008) is grounded on "Bayesian inference," which is the mathematics used for calculating the probability that a hypothesis is true. Bayesian inference is fundamental for many branches of science. In neuroscience, it is used to understand both the variability of neurons—their spontaneous fluctuations in activity—and how they represent uncertainty in the world (Pouget et al., 2003; Ma et al., 2006).

Furman and Wang (2008) took the bottom-up route. They constructed a network with roughly 2000 excitatory and 500 inhibitory neurons connected all-to-all, and included as much biophysical detail as possible given those numbers. The neurons in their network produce spikes and interact with each other through realistic synaptic conductances. This one is the descendent of an earlier model by Wang (2002) that had a much simpler architecture and which provided several key insights about the dynamics of spiking networks. In particular, it showed that a competition between two alternatives can be reproduced accurately by two populations of neurons that inhibit each other reciprocally but that contain excitatory recurrent interactions within each population. This organization creates a "slow reverberation" whereby the activity of one population progressively ramps up and dominates over the other, as observed in real data. In addition, this kind of model produces simulated spike trains with high variability, close to the variability observed in recorded cortical neurons. High variability is fundamental, because it allows a network to behave stochastically in the face of ambiguous or insufficient evidence: in trials with zero coherence, when the dots move in all directions, the model essentially flips a coin; it chooses one of the two options randomly, just like subjects do.

The new model by Furman and Wang (2008) generalizes these principles to the case of a larger circuit in which many populations compete with each other in a continuous way. Thus, the model is capable of slow reverberation and stochastic behavior but even when there are many possible choices. This is how it works (see Figure 1). The receptive fields of the model LIP neurons are spatially arranged along a circle around the fixation point. The choice targets are located on that same circle. The model neurons are driven by a background input, by an input that corresponds to the presence of a choice target inside the receptive field, and by simulated MT spikes. The neuron at, say, the 12 o'clock position receives input from MT neurons that prefer vertical motion, the neuron at the 6 o'clock position receives input from MT neurons that prefer downward motion, and so on. When the random dots are turned on in the simulation, the LIP units at the target positions slowly integrate their inputs until one of them dominates and suppresses the rest, due to the mutual inhibition across the circuit.

Beck and colleagues (2008), on the other hand, also constructed a competitive spiking network. Their description, however, is more abstract because the firing of each neuron is determined through an analog quantity that represents firing rate, or spiking probability. This is important, because it lets the authors control what the network is actually computing. How did Beck et al. (2008) generate this model? They took a decidedly top-down approach. Starting with Bayesian principles, they derived a number of conditions that a neural circuit should satisfy if it is to make optimal choices; optimal in the sense that the sensory information on which the choice is based is accumulated over time as efficiently as possible. Key to obtaining these conditions was an earlier study (Ma et al., 2006) showing that the variability of cortical neurons is of a special type, a type that drastically simplifies the representation of probabilities and their combinations through neural activity. As a consequence of that result, it turns out that the optimal computational strategy for LIP could not be simpler: each LIP unit should add over time the input that it receives from MT. What is surprising is that, according to the theory, such integration should be the same regardless of the task, the number of targets, the correlations in the input signals, and even if the parameters of the sensory signal (e.g., contrast or coherence) change during the course of a trial.

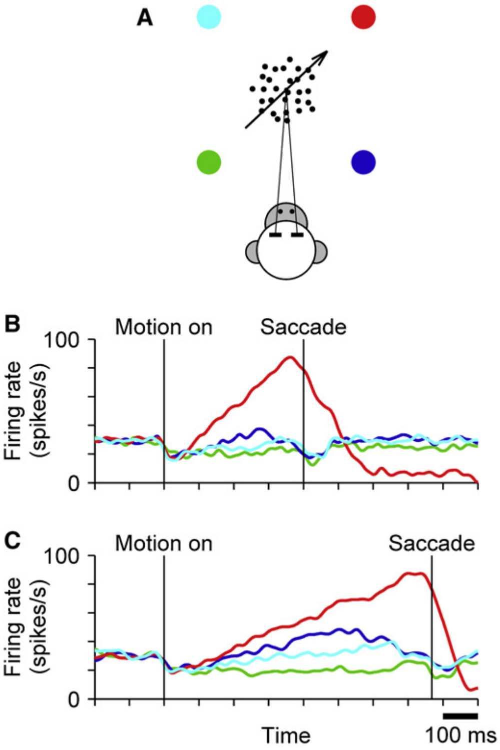

Figure 1. Hypothetical LIP Responses in the Four-Alternative Task

(A) Schematic of the display.

(B and C) Activity of four neurons with receptive fields at the target locations in two trials, one with high (B) and one with low (C) coherence.

This gives rise to an interesting prediction: if the coherence of the display changes in the middle of a trial, the slopes of the ramping responses should change accordingly, regardless of how much has been accumulated so far. Furman and Wang (2008) did not simulate this situation, and it will be important to see if their model—and real LIP neurons—show this effect. Similarly, their results predict that a particular pattern of errors should be seen when the choice targets are relatively close to each other, and it will be interesting to see if this also happens in reality and in the model of Beck and colleagues (2008).

What, then, are the key differences between the two LIP models? Which is better? I think these are the wrong questions. The biophysical model has many degrees of freedom that can be tweaked, and even more details can be added to it, so in a way its predictive power is limited by the very same level of detail that makes it convincing. On the other hand, the probabilistic model establishes a set of computational operations, rather than a specific circuit, so it requires some kind of translation into a biophysical substrate, which, again, is less constrained.

In my view, these models represent highly complementary approaches, and a better question is simply how to challenge and exploit them so that we learn something new about decision making. For example, according to the probabilistic model, the LIP responses should reflect not only the choice made but also its certainty, the probability that the decision is correct. This feature may be amenable to experimental verification through novel variants of the random-dot task. For instance, normally, certainty goes hand-in-hand with coherence—but this dependency can be disrupted. Suppose that there are four targets and the motion is clearly to the left; suddenly, the left target disappears and the subject has to choose one of the remaining targets. In this case, the LIP activity should encode the choice, as usual, but the subject will be quite certain that it was wrong. This certainty should have an impact on the LIP responses.

If such an effect is indeed found experimentally, then the next question will be, what biophysical mechanisms account for it? Turn, then, to Furman and Wang's level of description. Their model may reproduce the effect with little or no additional modification. Alternatively, to replicate the impact of certainty it might be necessary to simulate, for example, additional voltage-sensitive currents on the cell membranes, or the action of specific neuromodulators. In any case, with this tool it should be possible to identify at least candidate electrochemical mechanisms.

The broader point is that, ideally, there should be a back-and-forth dialog, not only between experimentalists and theoreticians, but also between the two levels of description, whereby the high-level model tells us what the circuit is computing and the low-level one tells us how. This type of synergy is much more likely to advance our understanding of decision making, and of neural function in general, than a single modeling framework alone, or no framework at all.

For now, we can say that decision making is based on stochastic neural dynamics and that, at least under certain conditions, it is optimal according to the principles of Bayesian inference. Therefore, regarding the question posed at the beginning of this preview, the evidence gathered so far indicates that the correct choice is (d), all of the above.

REFERENCES

Beck, J.M., Ma, W.J., Kiani, R., Hanks, T., Churchland, A.K., Roitman, J., Shadlen, M.N., Latham, P.E., and Pouget, A. (2008). Neuron *60*, this issue, 1142–1152.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1993). Vis. Neurosci. *10*, 1157–1169.

Churchland, A.K., Kiani, R., and Shadlen, M.N. (2008). Nat. Neurosci. *11*, 693–702.

Furman, M., and Wang, X.-J. (2008). Neuron *60*, this issue, 1153–1168.

Hanes, D.P., and Schall, J.D. (1996). Science *274*, 427–430.

Hanks, T.D., Ditterich, J., and Shadlen, M.N. (2006). Nat. Neurosci. *9*, 682–689.

Lo, C.C., and Wang, X.-J. (2006). Nat. Neurosci. *9*, 956–963.

Ma, W.J., Beck, J.M., Latham, P.E., and Pouget, A. (2006). Nat. Neurosci. *9*, 1432–1438.

McPeek, R.M., and Keller, E.L. (2004). Nat. Neurosci. *7*, 757–763.

Pouget, A., Dayan, P., and Zemel, R.S. (2003). Annu. Rev. Neurosci. *26*, 381–410.

Salzman, C.D., Murasugi, C.M., Britten, K.H., and Newsome, W.T. (1992). J. Neurosci. *12*, 2331–2355.

Shadlen, M.N., and Newsome, W.T. (2001). J. Neurophysiol. *86*, 1916–1936.

Smith, P.L., and Ratcliff, R. (2004). Trends Neurosci. *27*, 161–168.

Wang, X.-J. (2001). Trends Neurosci. *24*, 455–463.

Wang, X.-J. (2002). Neuron *36*, 955–968.