

## Decision making by urgency gating: theory and experimental support

David Thura, Julie Beauregard-Racine, Charles-William Fradet, and Paul Cisek

*Groupe de Recherche sur le Système Nerveux Central, Département de Physiologie, Université de Montréal, Montréal, Québec, Canada*

Submitted 28 November 2011; accepted in final form 15 September 2012

**Thura D, Beauregard-Racine J, Fradet CW, Cisek P.** Decision making by urgency gating: theory and experimental support. *J Neurophysiol* 108: 2912–2930, 2012. First published September 19, 2012; doi:10.1152/jn.01071.2011.—It is often suggested that decisions are made when accumulated sensory information reaches a fixed accuracy criterion. This is supported by many studies showing a gradual build up of neural activity to a threshold. However, the proposal that this build up is caused by sensory accumulation is challenged by findings that decisions are based on information from a time window much shorter than the build-up process. Here, we propose that in natural conditions where the environment can suddenly change, the policy that maximizes reward rate is to estimate evidence by accumulating only novel information and then compare the result to a decreasing accuracy criterion. We suggest that the brain approximates this policy by multiplying an estimate of sensory evidence with a motor-related urgency signal and that the latter is primarily responsible for neural activity build up. We support this hypothesis using human behavioral data from a modified random-dot motion task in which motion coherence changes during each trial.

sensory information processing; optimality; speed-accuracy trade off; reward rate; model; human

IMAGINE THAT you are driving somewhere and deciding on the best route. As you drive, your decision is informed by road signs, the advice of your passengers, information on a map, radio traffic reports, etc. Crucially, as you approach a potential turn, you are urged to make your decision even if you are not yet fully confident. Thus, the available information for making a choice is often changing continuously, and the urgency to choose one way or another is among many factors influencing the decision process.

Because such complexity cannot be easily addressed in a laboratory, research into the temporal aspects of decision making has primarily focused on simple perceptual choices. Most of these studies have provided support for a class of theories called “bounded integrator” or “drift diffusion” models (Bogacz and Gurney 2007; Carpenter and Williams 1995; Grossberg and Pilly 2008; Mazurek et al. 2003; Ratcliff 1978; Smith and Ratcliff 2004; Usher and McClelland 2001; Wong and Wang 2006). According to these, when a subject observes an informative stimulus, task-relevant variables are encoded in early sensory areas and fed to integrators that accumulate the total sensory evidence over time. A choice is made when the accumulated evidence in its favor reaches a threshold, and the setting of that threshold determines overall accuracy. This simple model produces a remarkably good match with the distribution of human reaction times (RTs) in a variety of decision-making tasks (Palmer et al. 2005; Ratcliff 2002; Ratcliff et al. 2004; Reddi and Carpenter 2000) and receives strong additional support from neurophysiological data showing

build-up activity related to the strength of sensory evidence in a large network of cortical and subcortical areas (Gold and Shadlen 2003, 2000; Hanes and Schall 1996; Kim and Shadlen 1999; Leon and Shadlen 2003; Ratcliff et al. 2007; Roitman and Shadlen 2002). In difficult tasks, such build up can last many hundreds of milliseconds, and its timing often predicts response latencies (Churchland et al. 2008; Roitman and Shadlen 2002). Finally, the process of bounded integration resembles the “sequential probability ratio test” (SPRT) (Wald 1945), a statistical test that minimizes the time required to reach a given accuracy criterion (Wald and Wolfowitz 1948). Taken together, these observations have led to the widespread acceptance of bounded integration as the mechanism underlying many kinds of decisions.

The studies described above have provided important insights into the neural mechanisms of decision making during “signal detection” tasks, in which subjects make perceptual judgments about stimuli whose informational content is constant during each trial. For example, a classic paradigm asks subjects to discriminate the direction of coherent motion within a field of randomly moving dots (Britten et al. 1992); although the stimulus is noisy and flickering, the signal within it (the percentage of coherently moving dots) is constant. However, during natural behavior, the environment can change without warning. Integrators are not well suited to such situations because they are slow to respond to changes in sensory information. To react quickly, animals must be very sensitive to novel information. Indeed, several studies have suggested that decisions are primarily based on information from a relatively short time window (Chittka et al. 2009; Cook and Maunsell 2002; Luna et al. 2005; Uchida et al. 2006; Yang et al. 2008). For simple color discrimination, this can be as short as 30 ms (Stanford et al. 2010), but even in much more difficult tasks, it appears to be on the order of 100–300 ms (Ghose 2006; Kiani et al. 2008; Ludwig et al. 2005; Price and Born 2010). But if decisions are determined by information from a short time window, then why should neural activity continue to build up for so much longer? Here, we propose a resolution to this paradox.

In a previous study (Cisek et al. 2009), we proposed that to trade off speed and accuracy in a changing environment, the nervous system should quickly estimate evidence and multiply that with a gradually growing “urgency” signal. We tested this model in a task in which subjects watched a set of tokens jumping from a central circle to one of two peripheral targets and were asked to guess which target would ultimately receive the most tokens. Our results strongly favored the urgency-gating model and were incompatible with models that integrate the stimulus. This led us to the conjecture that perhaps urgency gating is a general mechanism that even explains behavior during signal detection tasks and that the reason why neural activity grows in such tasks is due to urgency, not to stimulus integration (Ditterich 2006b). However, that conjecture could not be conclusive for several

Address for reprint requests and other correspondence: D. Thura, Département de Physiologie, Université de Montréal, 2960 chemin de la tour, Montréal, QC, Canada H3T 1J4 (e-mail: david.thura@umontreal.ca).

reasons. The results of Cisek et al. (2009) may have been relevant only to that particular task, which differed from previous experiments in four significant ways: 1) there was no noise in the stimulus, making integration less critical for filtering noise; 2) the tokens remained in their targets, providing a clear cue to the state of sensory information; 3) the depleting tokens in the central circle provided information on elapsing time that may have exaggerated the effect of urgency; and 4) the subjects' task was to make an inference about the future, not a judgment about the current stimulus, as in prior studies. For these reasons, one could argue that the results of the Cisek et al. (2009) experiment were of limited scope and could not be used to draw conclusions about decision making in general.

Here, we directly address these issues and show that the urgency-gating model applies to a wider range of decision tasks, including those that have been used to support bounded integration models. Furthermore, we go beyond our previous work by providing the mathematical foundation for why urgency-gating performs better than bounded integration in the sense that matters most for natural behavior: because it yields a higher reward rate.

Here, we present our proposal in three steps. First, we derived a decision policy for maximizing reward rate while taking into account the information conveyed by successive samples of the environment. While previous studies have suggested that optimal behavior is achieved by accumulating all sensory information to a fixed accuracy criterion (Bogacz et al. 2006), here we demonstrate that higher reward rates are achieved if what is accumulated is only novel information, and the result is compared with a decreasing accuracy criterion (Ditterich 2006a; Drugowitsch et al. 2012). Next, we suggest that the nervous system approximates that policy by detecting novel information and integrating it (or in simple tasks, just low-pass filtering the sensory stimulus) and then using an urgency signal to gradually bring the resulting neural activity closer to a given neural threshold. We show how this model accounts for behavioral and neural data previously explained with bounded integration models as well as recent results on how prior information is gated by elapsed time (Hanks et al. 2011). Finally, we describe a set of human behavioral experiments using a modified version of the classic random-dot motion discrimination task, in which the coherence of a noisy motion stimulus is changing over the course of the trial. This task was designed to test the urgency-gating model with a stimulus similar to those previously used. Moreover, we examined behavior both when subjects were instructed to make inferences about the future and when they were asked to detect perceived motion, as in previous studies. Finally, we examined how subjects modify their speed/accuracy trade off in two conditions of time pressure. Some of these experimental results have previously appeared in abstract form (Thura and Cisek 2009).

## MATERIALS AND METHODS

**Theoretical framework.** Here, we present a mathematical derivation of a decision-making policy that differs from widely used bounded integration models in two important ways: 1) to maximize rewards, it uses an accuracy criterion that decreases over time within each trial and 2) it takes into account the statistical dependence between sequential samples of the environment, yielding a policy that emphasizes novel information. We do not necessarily assume that the nervous system explicitly implements this policy using the particular steps shown here. Indeed, below, we show how a simple low-pass filter and time-dependent gain may provide animals with an adequate approximation.

Our derivation is partly based on the SPRT (Wald 1945), which is a statistical procedure for deciding to 1) accept a hypothesis, 2) reject it, or 3) continue the experiment by performing more observations. To achieve a given desired criterion of accuracy (e.g.,  $P < 0.05$ ), the SPRT requires fewer observations than any other procedure (Wald and Wolfowitz 1948). It is therefore optimal for any given desired level of accuracy in the sense of minimizing time. It has been suggested (Bogacz 2007; Bogacz et al. 2006; Gold and Shadlen 2007) that the bounded integration model approximates the continuous version of the SPRT and that it is therefore the optimal algorithm for making decisions in time. Note that in the bounded integration model, the setting of the threshold is associated with a given accuracy criterion, and both are fixed in time for a given trial.

However, while statistical tests, such as the SPRT, generally have a desired accuracy criterion (e.g., 95% or 99%) that is agreed upon by convention, this is not the case for natural behavior. In natural behavior, animals are not necessarily motivated to achieve a given level of accuracy and then reach that in the minimum time, but are instead motivated to maximize their reward rate (Balci et al. 2011). This creates a trade off between speed and accuracy. Below, we propose that to maximize reward rate, one needs an accuracy criterion that decreases over time within each trial; this can be implemented either through a decreasing value of a neural firing threshold or through an increasing gain of neural activity and a fixed firing threshold (Ditterich 2006a). This proposal is related to previous work suggesting that the accuracy criterion should decrease over time (Ditterich 2006a; Drugowitsch et al. 2012; Standage et al. 2011), but unlike that previous work, we do not assume the drift diffusion model as the basis of the decision-making process.

The second difference between our model and the bounded integration model concerns the quantity that is being integrated. Bounded integration of the sensory information from sequential samples is equivalent to the SPRT only in situations in which the samples are statistically independent, which is not the case in most natural situations. We suggest that in most situations, including most laboratory experiments, the first sample provides much more information than successive samples, which are increasingly redundant. Therefore, as we propose below, a good decision policy should only integrate sensory signals to the extent that they provide novel information.

**Maximizing rewards.** Suppose that you are in a situation in which you must make a correct guess to receive a reward, and after each guess there is a fixed period of time before you can try again. Suppose also that taking more time improves your chance of a correct guess. In terms of reward rate, what is the best trade off between taking an early guess versus waiting to make a better choice? For any given decision-making agent, one can quantify how the chance of success on a given trial ( $i$ ) increases as a function of time ( $t$ ); we denote this as  $P_i(t)$ . Note that this is not a "decision variable"; it is the probability that a given decision-making agent would make the right choice on a given trial if it were given a certain amount of time. If there are two choices, then  $P_i(t)$  must be between 0.5 and 1.0, but the exact value depends on many factors, including the difficulty of the trial (e.g., how informative is the stimulus, how much noise there is, etc.), the processing done by the agent, and the time that was so far spent in deliberation. For simplicity, we will first consider tasks in which the informational content of the sensory input is constant within each trial. We assume that each trial  $i$  has some level of difficulty that varies from trial to trial (intertrial variability) as well as sampling or processing noise (intra-trial variability). The agent can deal with intra-trial variability using a number of potential mechanisms, such as integrating or low-pass filtering the input. Importantly, these tend to be more effective if given more time. Because of this, even in constant information tasks,  $P_i(t)$  will increase over time as noise is being dealt with, until it reaches some asymptote (which may be below 1.0 for inherently probabilistic tasks such as gambling). Therefore, for the simple tasks considered here, we assume that the first derivative of  $P_i(t)$  is positive and the second derivative of  $P_i(t)$  is negative.

Now, let us assume that while each trial  $i$  has a certain level of difficulty, we don't know what the distribution of these trial difficulties is (we will consider some special cases below), and we don't even know whether a given trial is the only one or whether more trials will follow. In this condition, the best policy is to maximize the expected value of the time-discounted reward on each trial ( $R_i$ ). This is denoted as follows:

$$R_i = \frac{EV_i}{d_i + I} \quad (1)$$

where  $EV_i$  is the expected value of the reward (reward magnitude times the probability of success) on trial  $i$ ,  $d_i$  is the time spent in deliberation, and  $I$  is the intertrial interval. If the magnitude of rewards is constant, then the expected value is simply the probability of making a correct guess after a given deliberation time  $d_i$ . Thus, the expected value of the time-discounted reward is:

$$R_i(d_i) = \frac{P_i(d_i)}{d_i + I} \quad (2)$$

Because of the assumptions made above about  $P_i(t)$ , the function  $R_i$  has a single peak for each trial  $i$ , and our task is to find the decision policy that commits to a decision when that peak is reached.

On any given trial  $i$ , we don't know the exact shape of the function  $P_i$  because we don't know ahead of time the difficulty of this particular trial and we can't predict future sensory samples. However, we can calculate an estimate of our level of confidence in our current best guess on the basis of information that has already arrived (we will describe some methods for this below). Here, we will show that we can find the peak of  $R_i$  by comparing our estimate of confidence, at each moment in time, with a criterion level of accuracy that is a decreasing function of time. We demonstrate this for tasks in which  $P_i$  increases toward an asymptote and trials vary in difficulty. Again, we do not assume any particular mechanism for processing sensory information to make a guess or to calculate confidence and simply assume that it does better if it is given more time, if only because it can use some sensible strategy for dealing with noise (e.g., integration or low-pass filtering).

For any increasing function  $P_i(t)$ , we can find the maximum value of  $R_i(t)$  simply by looking for the time  $t$  (within a given trial) when the derivative of  $R_i(t)$  is zero and the second derivative is negative. The first derivative of  $R_i(t)$  is:

$$R_i'(t) = \frac{P_i'(t)}{t+I} - \frac{P_i(t)}{(t+I)^2} = \frac{P_i'(t)(t+I) - P_i(t)}{(t+I)^2} \quad (3)$$

and it is zero when:

$$P_i(t) = P_i'(t)(t+I) \quad (4)$$

To see that this is the point in time at which  $R_i(t)$  is maximized, note that its second derivative is:

$$R_i''(t) = \frac{P_i''(t)}{t+I} - \frac{2P_i'(t)}{(t+I)^2} + \frac{2P_i(t)}{(t+I)^3} \quad (5)$$

which, at the point of interest where  $P_i(t) = P_i'(t)(t+I)$ , reduces to

$$R_i''(t) = \frac{P_i''(t)}{t+I} \quad (6)$$

Because we assumed that the second derivative of  $P_i(t)$  is negative, then  $R_i''(t)$  is negative at this point. This means time  $t$  is a peak (as opposed to a trough), and therefore time  $t$  is the moment at which  $R_i(t)$  reaches its maximum value and is therefore the best moment to commit to the decision in trial  $i$ .

For a specific example, suppose we consider a family of functions:

$$P_i(t) = 1 - \frac{1}{2}e^{-a_i t} \quad (7)$$

where  $a_i$  is a parameter controlling trial difficulty that can vary from trial to trial (trials for which  $a_i$  is low are more difficult). This particular function is chosen simply for ease of differentiation, but it captures the assumptions we made above (it increases monotonically and reaches an asymptote). Figure 1A shows the points in time (red circles) when the function  $P_i(t)$  (black lines) intersects the function  $P_i'(t)(t+I)$  (blue lines) for different values of  $a_i$ . Note that these circles form a curve. Because this curve corresponds to the moments that maximize  $R_i$  in each trial, your decisions should be made whenever you estimate your confidence to cross this curve. On any given trial, your confidence in your best guess starts at a point (0.5) and grows at some rate. As long as you haven't crossed that red curve (i.e., are within the yellow shaded region), then you should continue to process information. However, as soon as you cross that red curve, you should commit. Crossing the lower arc of the red curve means that you should commit to just taking a nearly random guess, because confidence is low and rising too slowly to be worth the time. However, we will not consider these cases here; instead, we will focus on the upper arc of the red curve, which we will call the accuracy criterion function [ $C(t)$ ].

Note that  $C(t)$  is a decreasing function of time. Note also that if we increase the value of the intertrial interval  $I$ , the slope of  $C(t)$  becomes shallower, as expected (compare Fig. 1, A vs. B). Figure 1C shows the shape of the time-discounted reward function  $R_i$  for the same trials as in Fig. 1A. Note that the decision times (DTs) from Fig. 1A (red circles) fall precisely on the peaks of the individual  $R_i$  curves.

The observations made above can be proved formally by finding the equation for the curve  $C(t)$  by solving for the intersection of  $P_i(t)$  with  $P_i'(t)(t+I)$  to yield the following:

$$e^{a_i t} = \frac{1}{2}(a_i t + a_i I + 1) \quad (8)$$

We can approximate the exponential with a third-order Taylor expansion and use the cubic equation to solve for  $t$ , resulting in the following expression for  $t$  as a function of  $C$ :

$$t = -\frac{Iz}{A} \quad (9)$$

where  $z = \log(2 - 2C)$  and  $A = -\frac{1}{3}z^3 + z^2 - z + 1$ .

The above is an expression for the red curve in Fig. 1, A and B, in terms of how  $t$  changes as a function of  $C$ . However, what we want ideally is an expression for how  $C$  changes as a function of  $t$ . To find this, we compute the derivative of Eq. 9 with respect to  $C$  and then take the reciprocal to obtain the following:

$$\frac{dC}{dt} = \frac{A^2}{I} \left( \frac{1-C}{\frac{2}{3}z^3 - z^2 + 1} \right) \quad (10)$$

We now define  $w = \frac{2}{3}z^3 - z^2 + 1$  and note that Eq. 10 is undefined if  $w = 0$ . This occurs at time  $t_m$  (see Fig. 1A). We can find this point by solving for the root of  $w$ , which is  $z = -0.8065$  and corresponds to values of  $C = 0.7768$  and  $t_m = 0.3065I$ . This is the maximum time that one should ever wait, and it is linearly related to intertrial interval  $I$ . We can prove that until this moment, the upper arc of the red curve in Fig. 1A is dropping. We do so by computing the derivative of  $w$  with respect to  $C$

$$\frac{dw}{dC} = \frac{dw}{dz} \frac{dz}{dC} = (2z^2 - 2z) \frac{dz}{dC} = 2z \left( \frac{z-1}{C-1} \right) \quad (11)$$

and observing that this is always negative (because  $z < 0$  and  $0 < C < 1$ ). This confirms that Eq. 10 is negative when  $C$  is above the critical

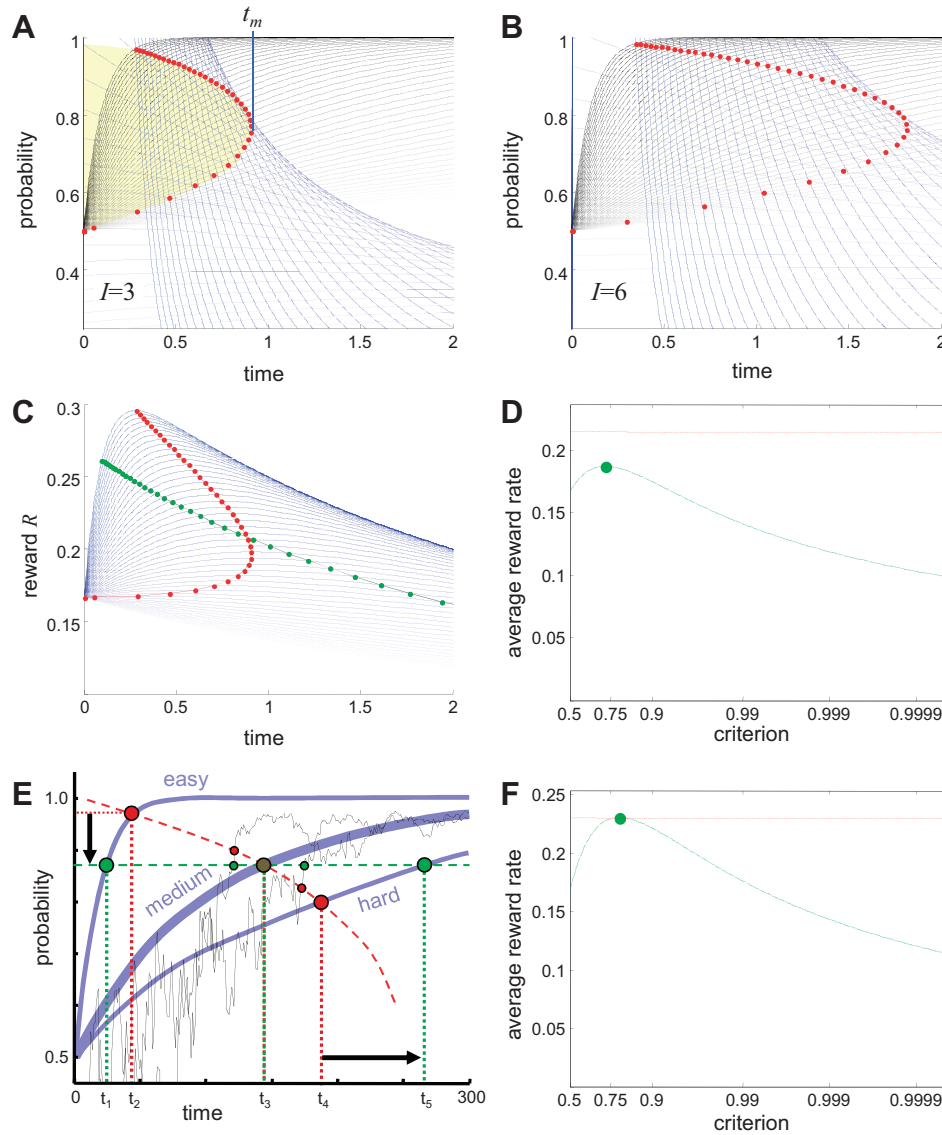


Fig. 1. Maximization of rewards with a dropping accuracy criterion [ $C(t)$ ]. **A**: the function  $P_k(t) = 1 - 1/2e^{-at}$  is plotted in black for different values of  $a$ . The corresponding function  $P'(t)(t + I)$  is plotted in blue for  $I = 3$  (where  $t$  is time and  $I$  is intertrial interval). Red circles indicate the intersection of each pair of plots, which corresponds to the moment at which decisions maximize rewards. The upper arc of the curve created by these circles is what we refer to as  $C(t)$ . The yellow shading indicates the region in which one should continue waiting, and the vertical blue line indicates the maximum time ( $t_m$ ) one should ever wait to make the decision. **B**: same but for  $I = 6$ . **C**: The expected time-discounted reward ( $R$ ) is plotted for the same condition as in **A**, and the red circles indicate the moments that  $P(t)$  intersects  $P'(t)(t + I)$ . Note that, in each case, this is the peak of  $R$ . **D**: the average reward rate obtained with different values of a fixed criterion (green line) compared with the maximum possible if decisions are always taken at the intersection of  $P(t)$  and  $P'(t)(t + I)$  (dashed red line). The best fixed-criterion value (0.72) is indicated by the circle, and its behavior is indicated by the green line in **C**. **E**: schematic demonstration of performance of a dropping criterion model (red) versus a fixed criterion model with the best setting of the fixed criterion (green). Three trial types are considered (easy, medium, and hard), and blue lines show the  $P_i(t)$  function for trials belonging to each group. The jagged black lines show the noisy estimate of momentary confidence that a decision-making system computes on two example trials of medium difficulty [here, we use a diffusion model with input sampled from a normal distribution with a mean of 0.1 and SD of 0.8, and the decision variable ( $x$ ) is converted into probability space using the following relation:  $p = e^x/(1 + e^x)$ ]. Note that these estimates gradually become more accurate as noise is being dealt with. Small circles indicate the moments at which the confidence estimate crosses the fixed criterion (green) or the dropping criterion (red) for the two example trials. Large circles show the moments when the decision will be made on average using the dropping (red) or fixed criterion (green). Note that if all three kinds of trials exist, then the fixed criterion will cause decisions to be too early on easy trials (at time  $t_1$  instead of  $t_2$ , so the probability of success is lower, as indicated by the downward arrow) and too late on hard trials (at time  $t_5$  instead of  $t_4$ , so time is wasted, as indicated by the rightward arrow). However, if only medium trials exist, then both models can achieve the same average performance. **F**: average reward rate obtained with different values of a fixed criterion (green line) compared with the dropping criterion model (dashed red line) in a condition where there is no intertrial variability. Note that in this condition, one can find the best fixed-criterion value (0.79; circle) that achieves the same reward rate as the dropping criterion model (see **E**, medium trials).

value of 0.7768 and thus that the accuracy criterion  $C(t)$  that maximizes rewards is a decreasing function of time. Furthermore, because  $I$  appears in the denominator of Eq. 10, we can conclude that the rate at which  $C$  drops is inversely related to  $I$ .

Figure 1D shows a plot of the reward rates obtained if decisions are made using different values of a fixed criterion. Note that the best-performing fixed criterion model (green circle at 0.72) does not yield the same reward rate that can be obtained with the

time-dependent criterion  $C(t)$ . The reason for this is explained the data shown in Fig. 1E. If there is intertrial variability [some trials are easy, some are hard, etc., as shown by the three different  $P_i(t)$  functions plotted as blue dashed curves], then the best-performing fixed criterion model (green dashed line) will cause the decision to be made too early in easy trials ( $t_1$  instead of  $t_2$ ), causing a loss of accuracy, as indicated by the black downward arrow, and too late in hard trials ( $t_5$  instead of  $t_4$ ), causing loss of time, as indicated by the black rightward arrow. In contrast, the dropping criterion (red dashed line) ensures that the decision is made at the peak of  $R_i$  for each trial, as shown in Fig. 1C. See Ditterich (2006a) and Standage et al. (2011) for related simulation results.

However, it is worth noting that for conditions in which difficulty is constant both within and across trials, a model with a fixed accuracy criterion can yield the same performance as our dropping accuracy criterion model. The reason for this is also shown in Fig. 1E. Suppose that all trials are of the medium difficulty and differ only due to intratrial variability (e.g., sampling noise). Because we assume that our decision-making system does not know the distribution of difficulties (it doesn't know that all trials are the same), it still uses the same dropping criterion (red dashed line). On average, this criterion is reached at time  $t_3$ . However, this does not imply that the decision will always be made at the same time, because intratrial variability may sometimes cause the animal to overestimate its confidence and guess slightly earlier (black trial ending slightly before  $t_3$ ) and sometimes to underestimate it and wait slightly longer (black trial ending slightly after  $t_3$ ). On average, these will cancel out yielding a reward rate that is nearly equivalent to a model that uses a single fixed criterion that does not change with time (green dashed line). Thus, for the special condition where difficulty is constant both within and across trials, the best fixed accuracy criterion model (with criterion = 0.79) reaches the same average reward rate as our dropping criterion model (see Fig. 1F).

Although a general proof was beyond the goals of this study, we conjecture that for most natural conditions and most forms of  $P_i(t)$ , the decision policy that maximizes rewards involves a dropping criterion  $C(t)$ . We expect that animals can learn the best way to decrease  $C$  on the basis of past experience with trials in a given condition (given task, given intertrial interval), but we do not address that here. Instead, the question we address next is the following: what is a good algorithm for estimating the current chance of guessing correctly given the sensory information that has so far arrived during a trial?

**Making decisions with nonindependent samples.** Here, we derive a strategy for estimating the current chance of guessing correctly [denoted as  $p(t)$ ] at a particular moment in time in a given trial by sequentially sampling relevant information from the environment. Instead of explicitly computing  $p(t)$ , we will instead compute a "decision variable" ( $x$ ), which is related to  $p(t)$  as follows:

$$x(t) = \log \frac{p(t)}{1 - p(t)} \quad (12)$$

Suppose that you are deciding between two mutually exclusive options ( $A$  vs.  $B$ ) and have a desired criterion of accuracy,  $C$ . Thus, you make choice  $A$  when  $p(A | s_1 \dots s_n) > C$ , where  $s_1 \dots s_n$  is a set of  $n$  samples of relevant information that you receive from the world at rate  $\Delta$  (thus,  $t = n \times \Delta$ ). Since  $A$  and  $B$  are mutually exclusive,  $p_B(t) = 1 - p_A(t)$ , and if both  $A$  and  $B$  choices yield the same rewards, then we can write our decision variable  $x_A$  as follows:

$$x_A(n) = \log \frac{p(A | s_1 \dots s_n)}{p(B | s_1 \dots s_n)} \quad (13)$$

Now, at each time step, we compare our current estimate of  $x_A$  to a criterion ( $K$ ), which is related to  $C$  as follows:  $K = \log(C/1 - C)$ .

If we have not yet received any information from the environment, then

$$x_A(0) = \log \frac{p(A)}{p(B)}$$

i.e., it is the logarithm of the ratio of prior probabilities that  $A$  or  $B$  is correct. Let us consider some ways to update  $x_A$  as each new sensory sample arrives. From Bayes' rule,

$$p(a|b) = \frac{p(a)p(b|a)}{p(b)}$$

we can see that after the first sample, we have

$$x_A(1) = \log \frac{p(A | s_1)}{p(B | s_1)} = \log \frac{p(A)p(s_1|A)}{p(B)p(s_1|B)} = \log \frac{p(A)}{p(B)} + \log \frac{p(s_1|A)}{p(s_1|B)} \quad (14)$$

The first term in Eq. 14 is the log ratio of priors, and the second term is the logarithm of the likelihood of seeing sample  $s_1$  in cases where  $A$  is the correct choice divided by the likelihood of seeing that sample if  $B$  is correct; this is the "log-likelihood ratio" (LogLR). Suppose now that we observe a second sample,  $s_2$ , which is statistically independent from the first. We can update our decision variable to obtain the following:

$$x_A(2) = \log \frac{p(A | s_1, s_2)}{p(B | s_1, s_2)} = \log \frac{p(A)}{p(B)} + \log \frac{p(s_1|A)}{p(s_1|B)} + \log \frac{p(s_2|A)}{p(s_2|B)} \quad (15)$$

More generally, we can say that we can calculate our decision variable as follows:

$$x_A(n) = \log \frac{p(A)}{p(B)} + \sum_{i=1}^n \log \frac{p(s_i|A)}{p(s_i|B)} \quad (16)$$

Equation 16 demonstrates that  $x_A(n)$  should start off at an initial value related to the log ratio of priors and then increase by the LogLR of each new sample. Note that the LogLR is positive for samples that are more likely given  $A$  and negative for ones more likely given  $B$ . Therefore, any individual sample provides evidence for one choice and against another, and a sequence of samples will generate a random walk of the variable  $x_A(n)$ . This process continues until  $x_A(n)$  crosses one of two thresholds,  $K$  and  $-K$ . This is equivalent to the diffusion model and, as shown by Bogacz et al. (2006), to a large family of related bounded integrator models such as the leaky competing accumulator (Usher and McClelland 2001).

However, the above derivation of how probabilities are calculated made a critical assumption: that each piece of information was statistically independent from previous ones. This is almost never true in natural situations, where animals often sample the same stimulus information repeatedly. To derive the more general case, we need to take into account the dependence between successive samples. For example, to calculate the decision variable after two samples requires an extension of Bayes' rule to three variables ( $a$ ,  $b$ , and  $c$ ), as follows:

$$p(a|b, c) = \frac{p(a)p(b|a)(c|a, b)}{p(b)p(c|b)}$$

yielding the following more general version of Eq. 15:

$$x_A(2) = \log \frac{p(A)}{p(B)} + \log \frac{p(s_1|A)}{p(s_1|B)} + \log \frac{p(s_2|s_1, A)}{p(s_2|s_1, B)} \quad (17)$$

As before, the first term is the log ratio of priors and the second term

is the LogLR of the first sample,  $s_1$ . However, the third term is different; it is the log of the ratio of the likelihoods of seeing the second sample given  $A$  or  $B$  and given the first sample. This is more complicated than the simple LogLR. From the definition of conditional probabilities, we have:

$$p(s_2 | s_1, A) = p(s_2 | A) \left[ \frac{p(s_1, s_2 | A)}{p(s_1 | A)p(s_2 | A)} \right] \tag{18}$$

where the first factor on the righthand side is the likelihood of the sample given  $A$  and the second factor is inversely related to the mutual information between  $s_1$  and  $s_2$  given  $A$ .

Now let us consider two extreme cases. First, suppose that the samples are, in fact, statistically independent. From the definition of independence, we know that  $p(s_1, s_2 | A) = p(s_1 | A)p(s_2 | A)$ , so the second factor in Eq. 18 is equal to 1, and we obtain  $p(s_2 | s_1, A) = p(s_2 | A)$ . Therefore, our equation for  $x_A(2)$  (Eq. 17) reduces back to Eq. 15.

In contrast, let us suppose that  $s_2$  is entirely predicted by  $s_1$ , i.e., that it is a redundant sample of the same sensory stimulus. This means that  $p(s_1, s_2 | A) = p(s_1 | A)$ , so Eq. 18 now becomes  $p(s_2 | s_1, A) = p(s_2 | A)[1/p(s_2 | A)] = 1$ . Therefore, our general equation for  $x_A(2)$  (Eq. 17) now becomes the following:

$$x_A(2) = \log \frac{p(A)}{p(B)} + \log \frac{p(s_1 | A)}{p(s_1 | B)} + \log \frac{1}{1} \tag{19}$$

It is clear that the last term is zero. In other words, if a given sample is already entirely predicted by previous samples, then it conveys no new information and should not cause the variable  $x_A$  to grow. In the intermediate cases where the first sample gives partial but incomplete information about the second, we have

$$x_A(2) = \log \frac{p(A)}{p(B)} + \log \frac{p(s_1 | A)}{p(s_1 | B)} + \log \left[ \frac{I_{12A}}{I_{12B}} \times \frac{p(s_2 | A)}{p(s_2 | B)} \right] \tag{20}$$

where

$$I_{12X} = \frac{p(s_1, s_2 | X)}{p(s_1 | X)p(s_2 | X)}$$

and is related to the conditional mutual information between the samples. As the mutual information increases, the difference between the numerator and denominator of the last term in Eq. 20 decreases, and so the logarithm goes toward zero.

This procedure generalizes to additional samples ( $n > 2$ ) after Bayes' rule is extended to additional variables, yielding the following conclusion: the amount by which  $x_A$  should increase after each sample is equal to the log of that sample's likelihood scaled by a factor related to the mutual information between it and all previous samples. If a sample is entirely independent of previous ones, then it contributes the entire LogLR. However, if the sample is already completely predicted by previous ones, then it conveys nothing, and our decision variable is unchanged. This latter situation is the case in all static evidence tasks without noise. In intermediate cases (such as static evidence tasks with noise), the value that is added by additional samples is nonzero but significantly smaller in magnitude than the LogLR of the first sample. In most situations, any given sensory sample is partially predicted by previous ones (i.e., there is mutual information), and so the earliest cues relevant to a choice convey significantly more information than later cues. Therefore, for any task in which the stimulus information is constant, the growth of the variable  $x_A$  will be brief because samples taken later in time are increasingly redundant and provide less and less new information.

Now let us suppose that at each moment  $x_A$  is compared with an accuracy criterion ( $K$ ) and that to maximize reward rate,  $K$  decreases with time (as discussed above). This can be described as follows:  $K = T/t$ , where  $T$  is a constant neural firing threshold and  $u(t)$  is some

function of time that is not related to evidence for any option. Now we can rewrite  $x_A(n)$  as follows:

$$x_A(n) = \log \frac{p(A)}{p(B)} \times u(t) + \sum_{k=1}^n \log \left[ \frac{p(s_k | A, s_{1..k-1})}{p(s_k | B, s_{1..k-1})} \right] \times u(t), \tag{21}$$

$- T < x_A(n) < T$

In other words, the decision variable  $x_A$  is computed as the sum of two terms: one that represents the prior and a second that represents the sum of the novel information favoring one choice over another, and each of these terms is multiplied by elapsed time, and the result compared with a constant neural firing threshold. We propose that this simple policy approximates the optimal algorithm for maximizing reward rate.

To summarize, we suggest that the classic bounded integration model is the optimal policy for making decisions only if samples are statistically independent and there is a preset desired accuracy criterion. In natural situations, these conditions usually do not hold. First, animals have the flexibility to trade off speed versus accuracy, motivating a dropping accuracy criterion. Second, successive samples of the environment are partially redundant, motivating a mechanism that only integrates novel information.

*The urgency-gating model.* While it seems unlikely that animals can precisely implement the policy outlined above, they can do very well with a highly simplified approximation. Here, we describe a mechanism that accomplishes this, called the urgency-gating model [see Cisek et al. (2009) for an earlier and still simpler version]. Figure 2 shows a schematic of the model.

The policy described above (Eq. 21) consists of three steps: 1) initialize the decision variable on the basis of prior information, 2) add the novel information favoring a given choice over other choices, and 3) multiply the result by an "urgency" signal that grows with time. If the resulting quantity exceeds a constant neural firing threshold, then the decision is made. Note that the first two steps result in an estimate of a quantity related to  $p(t)$ , whereas the last step implements what is effectively a dropping accuracy criterion. This general algorithm may be implemented in a variety of ways, and here we present just one possible set of equations and parameters.

How can the brain compute the extent to which a sensory sample is novel, i.e., not predicted by previous samples? One approach is to calculate the difference between the actual sensory signal and a

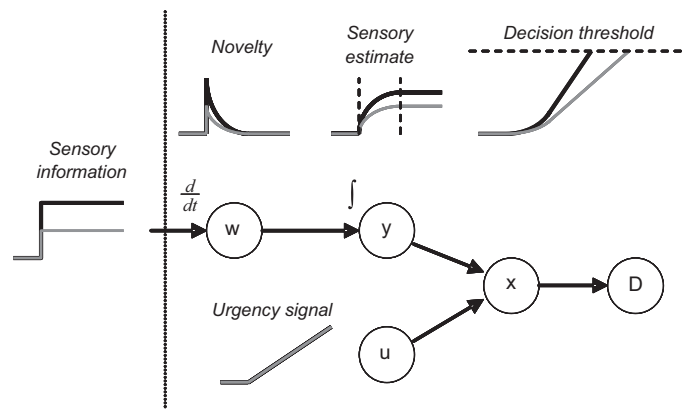


Fig. 2. New proposal for the temporal mechanism of decision making. In the model, sensory information (black: strong motion; gray: weak motion) is first differentiated and low-pass filtered ( $w$ ) to exaggerate novel events, and the resulting signal is then integrated ( $y$ ) to recover a filtered estimate of the sensory information. In a constant-evidence task such as our constant coherence motion detection (CMD) task, novel information only appears at the beginning, producing a burst in the  $w$  stage and transient growth in the  $y$  stage (in many conditions, these steps could be approximated with a low-pass filter). When this is combined with the stimulus-independent urgency signal ( $u$ ), the result is growth of activity ( $x$ ) to a neural firing threshold.

prediction of it. In other words, if the brain can compute sensory predictions, then anything that violates those predictions is novel and informative. The simplest, first-order prediction is to assume that the sensory signal stays the same. The difference between the actual sensory signal and this kind of crude first-order prediction is simply a derivative. We can implement this as follows:

$$N_A(t) = \frac{S_A(t) - S_A(t - dt)}{dt} \quad (22)$$

where  $S_A(t)$  is the difference in sensory evidence in favor of  $A$  versus  $B$  at time  $t$  [in the case of decisions about visual motion,  $S_A(t)$  is the difference of motion signals in direction  $A$  vs.  $B$ ] and  $N_A(t)$  is its time derivative. However, assuming that sensory information is noisy, one does not simply want to compute derivatives from moment to moment but instead apply some low-pass filtering, with a cutoff adjusted to respect the frequency at which relevant information changes. We implement the low-pass filter as follows:

$$\tau \frac{dw_A}{dt} = -w_A + N_A(t) + \theta \quad (23)$$

where  $\tau$  is a time constant (100 ms) and  $\theta$  is neural noise (normally distributed with a mean of 0 and a SD of 0.3). According to Eq. 23,  $w_A(t)$  is a low-pass-filtered version of the time derivative of the motion signal, i.e., it is a crude but simple estimate of novelty.

Next, we integrate  $w_A$  as follows:

$$y_A(t) = g \int w_A \quad (24)$$

where  $y_A$  is a neural variable that integrates  $w_A$  with gain ( $g$ ) = 0.005 and therefore recovers a filtered version of the original stimulus information  $S_A$ .

Because Eqs. 22–24 are all linear operations, their order can be rearranged and they can be equivalently described simply as a low-pass filter. This raises the possibility that in many simple decision-making tasks, the brain approximates the computation described by Eq. 21 with a low-pass filter gated by urgency (equivalent to being satisfied with a first-order estimate of novelty). This seems plausible for signal detection tasks such as random-dot motion discrimination, in which the relevant signal is available in the activity of area MT. For more complex tasks (e.g., Yang and Shadlen 2007), the brain may learn a more accurate computation of the novel information conveyed by successive sensory events and sum them explicitly.

Regardless of the steps used to compute  $y_A(t)$ , we compute our final decision variable as follows:

$$x_A(t) = z_A \times u(t) + y_A(t) \times u(t) \quad (25)$$

where  $z_A$  is the log of the priors (first term in Eq. 21) and  $u(t)$  is a growing urgency signal that is independent of any given choice. For simplicity, let us assume a linear urgency of  $u(t) = \beta t$ , where  $\beta = 2$  is a scalar gain. Choice  $A$  is made if  $x_A(t)$  exceeds a constant threshold,  $T = 0.2$ . An analogous computation is made for choice  $B$ , and because the two receive inverse stimulus information, they evolve as mirror images of each other.

In summary, Eqs. 22–25 can be written as follows:

$$x_A(t) = z_A \times t + f[S_A(t)] \times t \quad (26)$$

where  $f$  is a low-pass filter with a cutoff frequency dependent on  $\tau$ .

Let us now compare this to a bounded integration model, expressed as follows:

$$x_A^I(t) = z_A \times t + \int_0^t [S_A(\tau) + \theta] dt \quad (27)$$

where  $z_A$  is the log of the priors, which is multiplied by elapsed time (as proposed by Hanks et al. 2011), and the second term indicates the integration of sensory information favoring  $A$  over  $B$  (plus noise).

Note that if the information contained in the stimulus is constant over the course of each trial, then  $S_A(t) = E_A$ , where  $E_A$  is the constant evidence for  $A$  over  $B$ . Therefore, we can see that:

$$x_A^I(t) = z_A \times t + \int_0^t [E_A + \theta] d\tau = z_A \times t + E_A \int_0^t d\tau + \int_0^t \theta = z_A \times t + E_A \times t + \int_0^t \theta \quad (28)$$

We can compare this to the urgency-gating model (Eq. 26) with  $u(t) = t$ . Because a low-pass filter passes constant inputs through and integrates high-frequency noise, then after a short time (e.g.,  $t > 200$ ms), we can approximate Eq. 28 as follows:

$$x_A(t) = z_A \times t + f[E_A + \theta] \times t = z_A \times t + \left( E_A + \int_0^t \theta \right) \times t = z_A \times t + E_A \times t + t \times \int_0^t \theta \quad (29)$$

Note that this is identical to Eq. 28 except for the last term, which quickly goes to zero in both cases because the noise has a mean of zero. Thus, we suggest that any situation where the information contained in the stimulus is constant, the models behave nearly identically (Cisek et al. 2009).

We believe that the only way to distinguish between the bounded integration and urgency-gating models is through experiments in which the sensory information used to make a decision is varied over the course of each trial. This was done in the study of Cisek et al. (2009), but because subjects were provided with noiseless information that remained on the screen (potentially obviating the need for an integrator) and given a salient cue to elapsing time (potentially exaggerating urgency), that result may have been task dependent. For this reason, here we designed an experiment using a stimulus that is much more closely related to the coherent motion discrimination tasks used previously to study the temporal process of decision making.

**Experimental design.** Thirty-two subjects (19 women and 13 men, 31 right handed and 1 left handed, age: 21–51 yr, mean  $\pm$  SD: 28.4  $\pm$  8.3 yr) participated in this study. Each subject gave informed consent before the experiment, and the procedure was approved by the Ethics Committee of the University of Montréal.

Subjects made planar reaching movements using a digitizing tablet (CalComp), which recorded the position (125 Hz with 0.013-cm accuracy) of a cordless stylus embedded within a vertical plastic cylinder held in the hand. Target stimuli and cursor feedback were projected by an LCD monitor (60-Hz refresh rate) onto a half-silvered mirror suspended 16 cm above and parallel to the digitizer plane, creating the illusion that targets floated on the plane of the tablet.

In each experimental session of the main experiment, 24 subjects completed two tasks. The first was a RT version of a constant coherence motion detection (CMD) task (~150 choices) followed by a variable coherence motion discrimination (VMD) task (Fig. 3A). In both tasks, each trial began when subjects placed the cursor in a small starting circle (1-cm diameter). Next, a random-dot kinematogram consisting of 200 dots (in either a 6-cm-diameter circle or 6-cm-sided square aperture) appeared in the center of the workspace with two target circles (3.5-cm diameter) placed 180° apart at a distance of 6 cm of the center.

In the CMD task, after 200 ms, a certain fraction of the dots (defined as the coherence of the stimulus) began moving coherently in one of two potential directions (left or right), whereas the rest of the dots moved randomly. Note that in each successive stimulus frame it was not the same dots that moved coherently, but their proportion was constant. The coherence was varied randomly from trial to trial and was chosen from one of five possible levels (0%, 3%, 6%, 25.5%, and 51%). The subjects' task was to detect the direction of motion and to indicate it by moving the stylus into the target in that direction. They

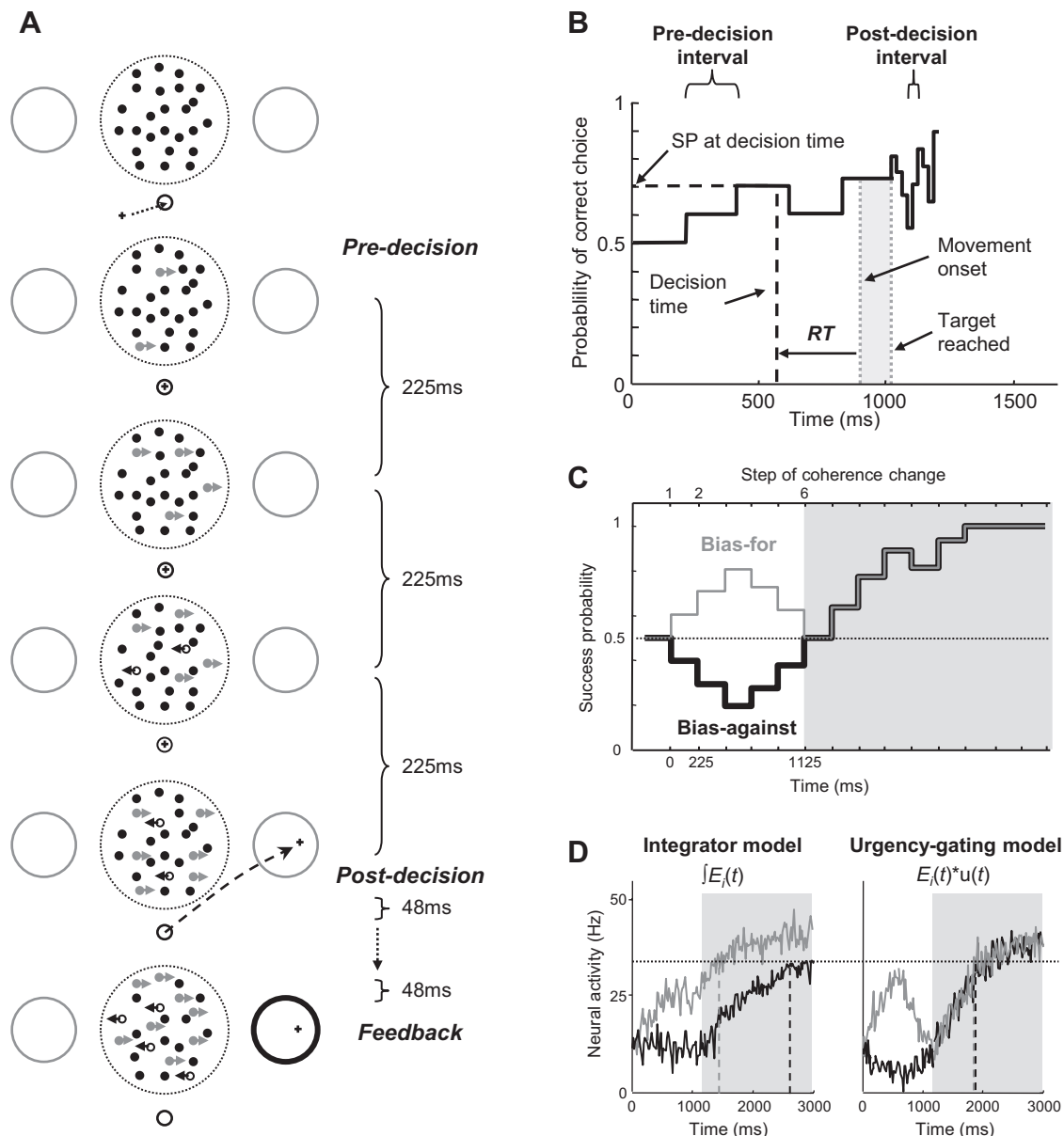


Fig. 3. Experimental design. *A*: variable coherence motion discrimination (VMD) task. See text for details. *B*: temporal profile of success probability (SP)  $[p_i(t)]$  associated with choosing target  $i$ . The movement onset is detected, and each subject's mean reaction time (RT; from the CMD task) is subtracted to estimate the "decision time" (DT). The profile of  $p_i(t)$  is used to compute the SP at DT. *C*: temporal profiles of SP for "bias-for" (gray) and "bias-against" (black) trials. *D*: schematic evolution of neural activity of a hypothetical cell that prefers the correct target during bias-for (gray) and bias-against (black) trials, as predicted by the "integrator" model (left) and the "urgency-gating" model (right). For any decision made after the sixth coherence step (gray area), the integrator model predicts that neural activity will reach the decision threshold (horizontal dotted line) faster in bias-for trials than in bias-against trials, resulting in shorter DTs in the former (vertical dotted gray line) than in the latter (vertical dotted black line). The urgency-gating model predicts no difference in behavior for the two trial types when decisions are made after the sixth coherence step.

were allowed to make their choice at any time. For each 0% coherence trial, the correct target was assigned randomly. Subjects who would later perform the "time pressure" version of the VMD task, as described below, had  $\sim 3$  s to make their choice, and subjects performing the "no time pressure" version had  $\sim 8$  s. When subjects entered one of the two possible targets, the motion stimulus was extinguished, and the correct target turned green. The subjects' mean RT in the 51% coherence trials was later used to estimate their DTs in the following VMD task.

In the VMD task, the 200 dots initially moved purely randomly. Next, after 225 ms, 6 dots began to move coherently to the left or right, whereas 194 dots continued moving randomly. Next, after another 225 ms, another six of the randomly moving dots all began to

move coherently either in the same or opposite direction. Thus, at that point there were either 12 dots moving coherently in one direction or 6 dots moving right and 6 dots moving left, and the remaining 188 dots continued moving randomly. The same procedure then continued: every 225 ms another six of the randomly moving dots were assigned to either left or right; this was called a "coherence step." After 15 coherence steps, the stimulus remained at the resulting constant coherence until time ran out. The task for the subject was to choose the target corresponding to the direction of motion in which s/he predicted the dots will be moving at the end of the trial. Importantly, subjects were allowed to make their choice as soon as they felt confident enough to do it. Fourteen subjects performed the time pressure version of the task, in which they had to make their



decisions before the end of the 15th coherence step (~3,375 ms). Sixteen subjects performed the no time pressure version, in which they had an extra 5 s of time (a total of 8,375 ms) to make their decisions, and during this time the coherence remained constant. Six subjects completed both of the two time pressure conditions in this task. Once a target was chosen, the interval between coherence steps was reduced from 225 to 48 ms. Thus, subjects were presented with a trade off between maximizing accuracy by waiting toward the end versus taking an early guess, which risks errors but could save time. If subjects entered the target before the end of the 15th coherence step, visual feedback about success or failure (the chosen target turning either green or red, respectively) was not provided before the 15th step was completed. In the no time pressure condition, if subjects chose a target after 3,375 ms, the visual feedback appeared immediately. The intertrial interval was 1,500 ms. In each time pressure condition, subjects were asked to complete 100 correct choices before taking a break. Subjects who performed the task with time pressure were asked to accomplish 4 blocks of 100 correct decisions, whereas those who performed the task without time pressure had to successfully complete 5 blocks.

The design of the VMD task allowed us to calculate, at each moment in time, the success probability  $p_i(t)$  associated with choosing a target  $i$ . If, at a particular moment in time,  $N_R$  coherence steps favored the right target, whereas  $N_L$  coherence steps favored the left target, and there were  $N_C$  steps remaining, then the probability that the target on the right (R) will ultimately be the correct one (i.e., the success probability of a rightward guess) is:

$$p(R|N_R, N_L, N_C) = \frac{N_C!}{2^{N_C}} \sum_{k=0}^{\min(N_C, 7-N_L)} \frac{1}{k!(N_C-k)} \quad (30)$$

To estimate decision times on a trial-by-trial basis, we detected the time of movement onset and subtracted each subject's mean RT (from the CMD task), as described by Eq. 31:

$$DT = RT_{VMD} - RT_{CMD} \quad (31)$$

where  $RT_{VMD}$  is the RT in a given trial of the VMD task and  $RT_{CMD}$  is that subject's mean RT in the CMD task with 51% coherent motion. Finally, we used Eq. 30 to compute the success probability at the time of the decision (see Fig. 3B). Importantly, for all analyses, we defined the success probability as the probability that the target chosen by the subject will be the correct one. For all statistical tests, the significance level was set at 0.05.

In each time pressure condition, all subjects were presented with the same pseudorandom sequence of trials. The time pressure and no time pressure sequences contained a total of 540 and 660 trials, respectively. Among them, ~25% were fully random (each coherence step was randomly assigned). The other trials belonged to specific classes. In easy trials (~15%), the initial coherence steps consistently favored one of the targets, quickly driving the success probability for that target to 1. In ambiguous trials (~14%), the initial coherence steps were more balanced, keeping the  $p_i(t)$  function close to 0.5 until late in the trial. In bias-for trials (~10%), the first three coherence steps favored the correct target, whereas the next three ones favored the opposite one, and the remaining steps resembled an easy trial. Bias-against trials (~10%) were identical to bias-for trials except that the first six steps were reversed (Fig. 3C). The bias-for-ambiguous and bias-against-ambiguous trials (4% of trials in the no time pressure condition only) were identical to bias-for and bias-against trials, respectively, except that the last 9 coherence steps resembled an ambiguous trial. As a control, we also added bias-updown and bias-downup trials (~10% of trials; see Fig. 8A). In the bias-updown trials, the first four coherence steps favored the correct direction and the next two favored the opposite direction. In the bias-downup trials, the two first steps favored the wrong direction and the next four steps favored the correct direction. In both of these, the remaining steps were similar and resembled an easy trial. Finally, in misleading trials

(~5% of trials in the time pressure conditions and 1% in the no time pressure conditions), the first four coherence steps favored the wrong target (data not shown).

To test the effect of task instruction (i.e., prediction vs. detection; see RESULTS), nine subjects (one of whom also participated in the main experiment) performed an additional control experiment consisting of three tasks. First, each subject performed 60 trials of the CMD task with either 3% or 51% coherence. Next, the subject was asked to perform four blocks (150 correct trials each) of the no time limit VMD task in two conditions: 1) in the "prediction" condition, subjects were asked, as above, to predict the net motion direction at the end of the trial; and 2) in the "detection condition," the stimulus was the same but subjects were instructed to indicate the direction of any net coherent motion as soon as they detected one and to ignore any subsequent motion changes. In the latter condition, the correct choice was always based on the net motion at the time the subject made their decision. Because it is difficult to estimate that moment online, and thus classify a trial as correct or wrong, no visual feedback was provided at the end of the trial in both conditions. Subjects were presented with the same sequence of 1,050 trials, including 350 random ones and 700 special ones, among which there were 70 bias-for and bias-against trials.

## RESULTS

**Behavior in a classical CMD task.** Each subject began by performing a RT version of a CMD task (Britten et al. 1992) using trials with five different coherence levels (0%, 3%, 6%, 25.5%, and 51%). Analysis of subjects' behavior in the CMD task showed that the decision speed depended critically on both motion strength (i.e., coherence level) and maximum RT allowed to respond. Indeed, RTs were significantly longer with lower motion strength, and these differences were greatest in the no time pressure condition (Fig. 4A). Accuracy also depended on motion strength and was nearly perfect at high motion strength but fell toward chance levels with lower motion strengths regardless of the time pressure condition (Fig. 4B). The time pressure condition had a significant effect on performance only for the 3% coherence trials (82% vs. 92% in time pressure vs. no time pressure conditions, respectively,  $P < 0.05$  by  $t$ -test) despite the fact that RTs in the 0%, 3%, and 6% trials were significantly longer for subjects who performed the no time pressure condition. In the 51% coherence trials

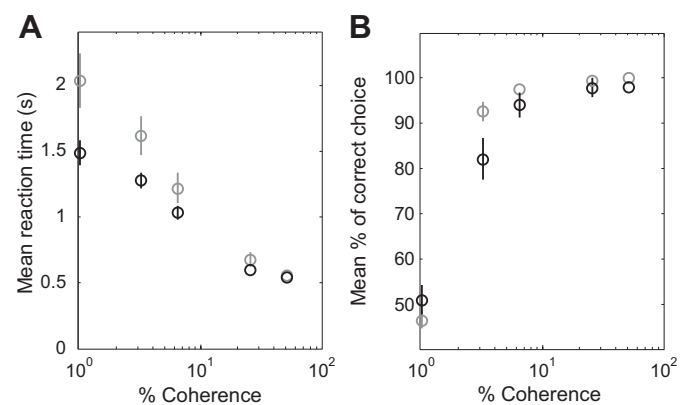


Fig. 4. Effect of coherence strength on subjects' mean RT and performance in the CMD task. RTs (left) and percentages of correct choices (right) were averaged across subjects during both the "time pressure" (black) and "no time pressure" (gray) conditions of the constant coherence RT task. Vertical lines denote SEs.

(those used to estimate subjects' DTs in the VMD task), RTs ranged from 417 to 734 ms (mean: 553 ms; SD: 83 ms), and we did not find any significant effect of the time pressure condition on either subjects' RTs or performance (546 vs. 557 ms and 98% vs. 100% in the time pressure vs. no time pressure conditions, respectively).

**VMD task: effects of time pressure.** Next, subjects performed a VMD task (see Fig. 3A and MATERIALS AND METHODS for details). Among our population, 14 subjects completed the time pressure version of the VMD task in which they had to make decisions before the last (15th) coherence step (3,375 ms). They performed an average of  $506 \pm 33$  trials to achieve the objective of 400 correct trials. The mean percentage of correct choices (performance) across subjects was 76.8%. For all trials, subjects' mean DT (measured from the first coherence step) was  $1,373 \pm 178$  ms (mean  $\pm$  SD), including both correct and error trials. Six of these subjects, along with ten other subjects, also performed a no time pressure version of the task in which they had an extra 5 s (thus, a total of 8,375 ms) to make their decisions and reach one of the two targets. On average, they needed a total of  $579 \pm 24$  trials to accomplish 500 correct trials. Thus, subjects' mean performance in the no time pressure condition (85.7%) of the VMD task was significantly better than in the time pressure condition ( $P < 0.001$  by *t*-test). Moreover, as expected, in the no time pressure condition, subjects' mean DT for both correct and error trials ( $2,474 \pm 596$  ms) was significantly longer compared with the time pressure condition ( $P < 0.001$  by *t*-test).

We analyzed the success rate as a function of the number of coherence steps that occurred before subjects made their decision. Across all trials (Fig. 5A), the success rate was quite low for very fast decisions, increased later in the trial, and then decreased again, especially under high time pressure. This was partially attributable to the fact that subjects generally waited for last steps only in trials that were more difficult and in which success was closer to chance levels. As expected, the success rate is clearly dependent on the pattern of coherence changes. For example, the success rate was higher for early decisions in easy trials compared with ambiguous trials (Fig. 5, B and C). In bias-for trials (Fig. 5D), most errors occurred between the third and seventh step of coherence change. This is particularly true in the time pressure condition. In contrast, in bias-against trials (Fig. 5E), most errors occurred before the seventh or sixth step of coherence change, depending on the time pressure condition. It was interesting to note that time pressure dramatically modified the subjects' strategy and exacerbated the observations described above. For instance, across all trials, the suc-

cess rate was lower in the time pressure condition for decisions made after the fourth coherence step. This makes good sense, since under time pressure, subjects had to make their decisions even if they were not completely confident, resulting in a lower level of performance. In contrast, without time pressure, it is more likely that subjects were willing to make their decisions only if they were confident enough (see the effects of time pressure on success probabilities during specific trial types at the population level in Figs. 7D, 8C, and 9D), yielding better performance.

It is also interesting to note that within a time pressure condition, subjects tended to adjust their speed-accuracy trade off over the time course of a session. Figure 6A shows the behavior of one subject who performed both time pressure conditions of the VMD task, and Fig. 6B shows results at the population level. As shown in Fig. 6, *top*, DTs significantly increased over the time course of the time pressure session (repeated-measures ANOVA across trial bins,  $F = 3.89$ ,  $P = 0.008$ ), leading to a weak (nonsignificant) but constant growth of success probabilities at DT across the session. In contrast, the subjects' strategy looked very different in the no time pressure condition. As expected, their mean DTs were significantly longer and their mean success probabilities significantly higher compared with the time pressure condition. Moreover, they also tended to reduce their mean DTs across the session (nonsignificant), and it was interesting to note that this time saving did not significantly affect their success probabilities.

To summarize, despite some obvious between-subject differences (some individuals more consistently made "fast and sloppy" decisions, whereas others were more meticulous and slow), these results tend to show that subjects performing the VMD task were pushed to modify their speed/accuracy trade off depending on time pressure (higher performance but more time spent in the no time pressure condition). They were more willing to tolerate lower success probability levels when urgency was increased. In a given time pressure condition, subjects also adjusted this trade off over the course of the session. The relatively weak performance of subjects doing the time pressure condition tended to push them to be slightly more conservative over time, whereas subjects doing the task without time pressure had very good performance and then tried to save some time across the session.

*The urgency-gating model explains behavior better than integrator models.* Among VMD trials with a random sequence of coherence steps, we interspersed several specific classes of trials in which the steps were designed to test specific hypoth-

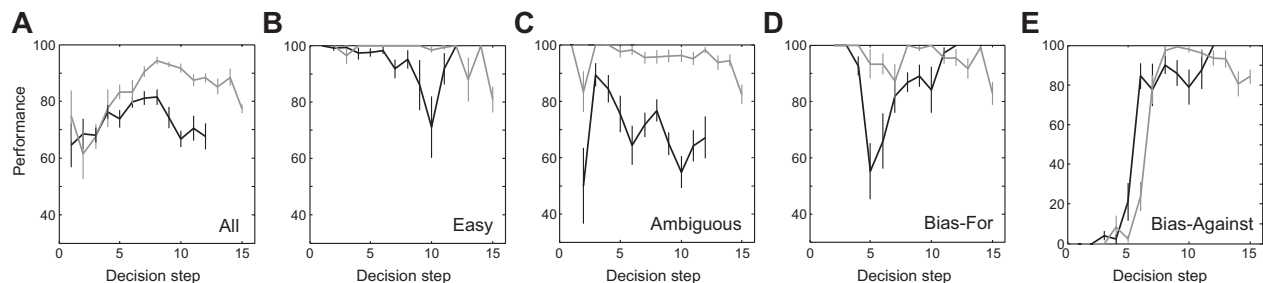
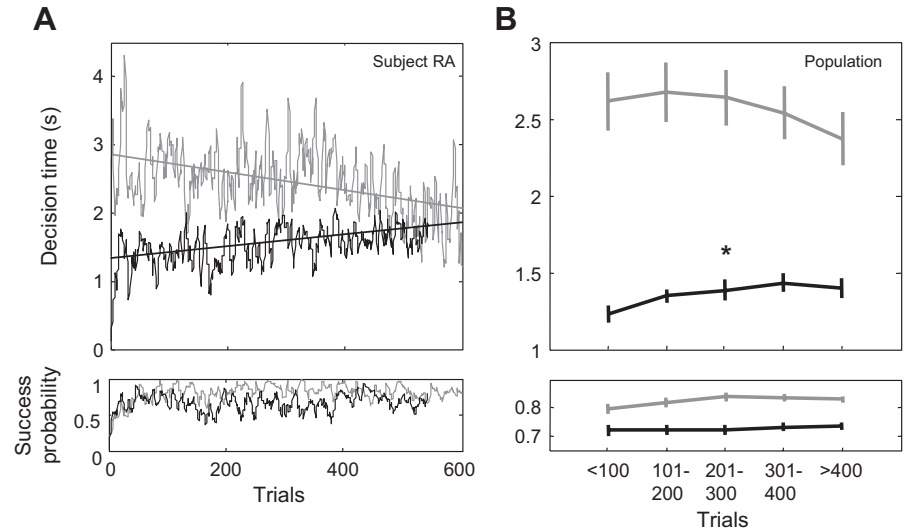


Fig. 5. Effect of DT on subjects' mean performance. We analyzed the success rate as a function of the number of coherence steps that occurred before subjects made their decision. A: the lines show subjects' mean success rate (means  $\pm$  SE, vertical lines) across all trials as a function of the number of coherence steps before the decision in the time pressure (black) and the no time pressure (gray) conditions. The success rate was quite low for very fast decisions, increased later in the trial, and then decreased again, especially under high time pressure. B: same as in A except only for easy trials. C: same as in A except for "ambiguous" trials. D: same as in A except for bias-for trials. E: same as in A except for bias-against trials.

Fig. 6. Effect of urgency on subjects' decision-making strategy. Among our population, six subjects performed both the time pressure and no time pressure versions of the task, in separate sessions (3 subjects did the time pressure condition first). *A*: evolution of DTs (*top*) and SPs (*bottom*) across the time course of a session for a subject who performed both the time pressure (black lines) and no time pressure (gray lines) versions of the task. *B*: population analysis (same convention as in *A*). We grouped trials according to five chronological epochs, defined as follows: *trials 0–100*, *trials 101–200*, *trials 201–300*, *trials 301–400*, and *trials >400*. Next, we calculated for each epoch and for all trials the mean DTs and SPs averaged over subjects (means  $\pm$  SE, vertical lines) in both time pressure and no time pressure conditions. In each time pressure condition, repeated-measures ANOVA was performed to test for an overall effect of trial bins on subjects' mean DT and SP. In the time pressure condition, subjects tended to become more conservative across the session ( $P = 0.008$ ).



eses about the temporal dynamics of decisions. Here, we focus on trials that helped us to distinguish between integrator and urgency models (Fig. 3C). In bias-for trials, the first three coherence steps favored the correct target, whereas the next three steps favored the opposite option, and the remaining steps again mostly favored the correct target. Bias-against trials were identical except that the first six steps were reversed. This comparison is critical because the two classes of models make distinct predictions about the timing of decisions in these trials (Fig. 3D). In particular, because integrator models retain a “memory” of previous coherence steps, they predict that after 1,125 ms (6 steps of novel information; see Fig. 3C), neural activity related to the correct target will be higher (and therefore closer to threshold) in bias-for trials than in bias-against trials, because during the first six steps of bias-for trials, the net motion is always in the correct direction. Consequently, these models predict faster decision times in bias-for than bias-against trials. In contrast, because the urgency-gating model integrates changes in motion information, it does not predict faster decisions in bias-for than bias-against trials. This is because after the sixth coherence step, the changes in sensory information are balanced in both kinds of trials.

To evaluate these predictions, we focused our analyses on trials in which correct decisions were made after 1,125 ms, i.e., excluding any early decisions whose outcome would be trivial (on average, ~20% and ~6% of bias-for/bias-against trials were removed in the time pressure and no time pressure conditions, respectively). For most subjects (10 of 14 subjects in the time pressure condition and 14 of 16 subjects in the no time pressure condition), there were no significant difference between DTs in the bias-for and bias-against trials [ $P > 0.05$  by Kolmogorov-Smirnov (KS) test]. Figure 7A shows the behavior of one subject, and Fig. 7C shows results at the population level. It was interesting to note that for those six cases in which we found a significant difference between DTs, decisions were faster in the bias-against trials—opposite to the predictions of models that integrate motion. Next, we analyzed the success probabilities at DT and found that these were also similar in the two classes of trials for most subjects (13 of 14 subjects and 14 of 16 subjects in time pressure and no time pressure conditions, respectively,  $P > 0.05$  by KS test). This result is shown for one subject in Fig. 7B and for the population

in Fig. 7D. Finally, we analyzed two more trial types related to bias-for and bias-against trials. These bias-for-ambiguous and bias-against-ambiguous trials were identical to bias-for and bias-against trials except that the last nine coherence steps were very ambiguous. Thus, subjects were motivated to take a best guess, and we were interested in how that guess reflected the early bias. For the same reasons as described above, integrator models suggest that neural activity accumulated in the early part of bias-for-ambiguous trials will result in better performance than in bias-against-ambiguous trials. In contrast, the urgency-gating model does not predict a significant difference in performance because no bias due to the early part of the trial should affect the way a subject will make a decision in the late and ambiguous period. Only subjects who performed the no time pressure version of the task ( $n = 16$ ) were tested with these trials. For decisions made after the sixth coherence step, there was no tendency for subjects to have better performance in the bias-for-ambiguous trials compared with the bias-against-ambiguous trials ( $P = 0.3$  by  $t$ -test; Fig. 7E), contradicting the predictions of integrator models.

The potential capacity for recognizing special trial classes may have some crucial implications regarding the interpretation of our results. For instance, it is possible to explain our results if we postulate that the integrators can get “reset” if subjects can recognize the condition of complete ambiguity (e.g., after 6 coherence steps in bias-for and bias-against trials, when the motion favoring each target is the same). For this reason, we embedded among the random trials variations of bias-for and bias-against trials in which the first few steps of coherence change were not three and three (see Fig. 8A and MATERIALS AND METHODS for details). Therefore, success probability in bias-updown trials never returned to the critical value of 0.5 that could potentially trigger a reset of the integrators. In both time pressure conditions, there was no significant difference between decision times in the two trial types (Fig. 8B). Moreover, the success probabilities at DT were similar in the two kinds of trials for most of subjects, in both the time pressure (13 of 14 subjects) and no time pressure (15 of 16 subjects) conditions (Fig. 8C).

We then examined our human data to see whether a leak could explain our results in bias-for versus bias-against trials. Indeed, it is possible that by the time the decision is made,

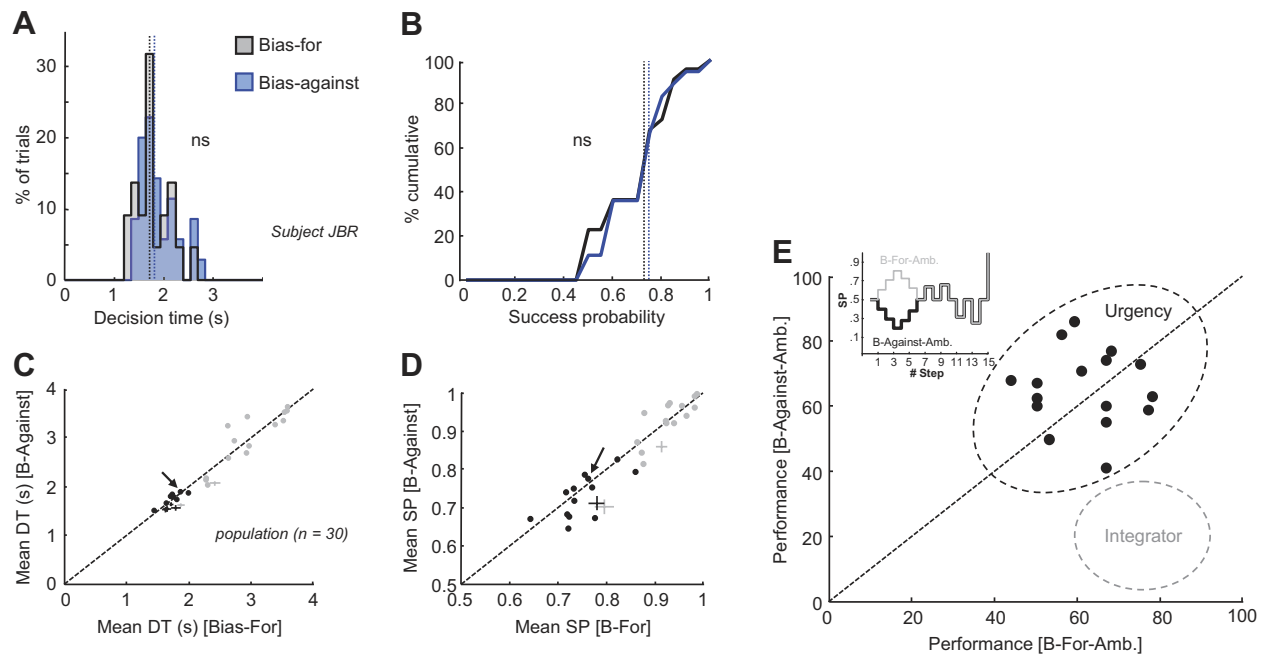


Fig. 7. Behavior in bias-for versus bias-against trials and “bias-for-ambiguous” versus “bias-against-ambiguous” trials. *A*: distribution of DTs for one subject in bias-for (gray,  $n = 82$ ) and bias-against trials (blue,  $n = 47$ ). The mean DT in bias-for trials was not significantly different from the mean DT in bias-against trials (dotted lines; 1,715 vs. 1,810 ms,  $P > 0.05$ ). *B*: cumulative distribution of SPs at the time of the decision during bias-for (black) and bias-against (blue) trials for the same subject. *C*: average DTs for all subjects (black: time pressure, gray: no time pressure). Solid pluses indicate means  $\pm$  SE of subjects for whom the difference was significant; circles represent subjects for whom it was not. There was no significant difference (ns) between DTs in bias-for and bias-against trials for 24 of 30 subjects. The arrow indicates the subject shown above. *D*: the SP at DT also was similar in the two trial types for most of subjects, in both the time pressure and no time pressure conditions. In all of these, we only used data from trials in which decisions were made after the sixth coherence step. *E*: performance (percentage of correct choices) for each subject in the bias-for-ambiguous (abscissa) versus bias-against-ambiguous (ordinate) trials (see inset). The dotted diagonal line represents the null hypothesis: that there is no difference between the two trial types. The integrator model predicts better performance in bias-for-ambiguous trials (gray dotted ellipse), whereas the urgency-gating model does not predict any performance difference (black dotted ellipse). Only subjects who performed the no time pressure condition were included ( $n = 16$ ).

differences between the accumulated activities in the early part of bias-for and bias-against trials would have decayed away, and behavior would be similar in both kinds of trials. In particular, we looked at DTs from a subset of bias-for and bias-against trials in which a subject made the decision within 450 ms after the sixth coherence step (Fig. 8*D*, shaded area). These early decisions might still retain some bias that has not leaked away. Eleven and one subjects who performed the time pressure and no time pressure conditions, respectively, made enough of these fast decisions (at least 5 trials of each type) to make the comparison possible. The distributions of DTs across all subjects showed that there were no significant differences between DTs for bias-for versus bias-against trials (Fig. 8*E*). Analyses of DTs in individual subjects showed that in only two cases, decisions were significantly different in the two trial types ( $P < 0.05$  by KS test), being faster in bias-against trials, contradicting again the predictions of models that integrate motion signals (Fig. 8*F*).

One remaining concern regarding these results is the possibility that subjects just ignored the first five to six coherence steps in each trial. However, this is highly unlikely. First, when subjects made decisions before 1,125 ms, they were correct 84.3% (96.7%) of the time in bias-for trials during the time pressure (no time pressure) condition, but only 9.7% (1.9%) of the time in bias-against trials. This confirms that they attended to the motion signal during those first 1,125 ms. Furthermore, when we tested them on other kinds of trials, they clearly were influenced by the overall profile of the success probability

function, even by the first few coherence steps. For example, Fig. 9 shows a comparison of behavior during easy trials, in which the coherence steps tended to consistently favor one of the targets, and ambiguous trials, in which the steps were more balanced between the two targets until late in the trial. As expected, most subjects made decisions significantly later in ambiguous trials than in easy trials, both in the time pressure (12 of 14 subjects) and no time pressure (16 of 16 subjects) conditions ( $P < 0.05$  by KS test; Fig. 9, *A* and *C*). Importantly, in easy trials, subjects often made decisions within the first four to five coherence steps. Still more interesting was the observation that almost all subjects (27 of 30 sessions) made decisions at a significantly lower level of success probability in ambiguous trials ( $P < 0.05$  by KS test; Fig. 9, *B* and *D*) than in easy trials. This held true for 14 of 14 subjects in time pressure conditions and 13 of 16 subjects in no time pressure conditions. Thus, subjects appeared more willing to guess in ambiguous trials than in easy trials, and, unsurprisingly, they had significantly lower performance in ambiguous trials than in easy trials (81% vs. 96% correct, respectively,  $P < 0.005$  by *t*-test). This is also compatible with the urgency-gating model, which effectively implements a dropping accuracy criterion over time.

*Effect of DT on subjects' confidence level.* One key prediction of the urgency-gating model is that the level of confidence at which the subjects will make decisions should decrease as a function of the time taken to make the decision (Cisek et al. 2009). According to the model, the confidence should be

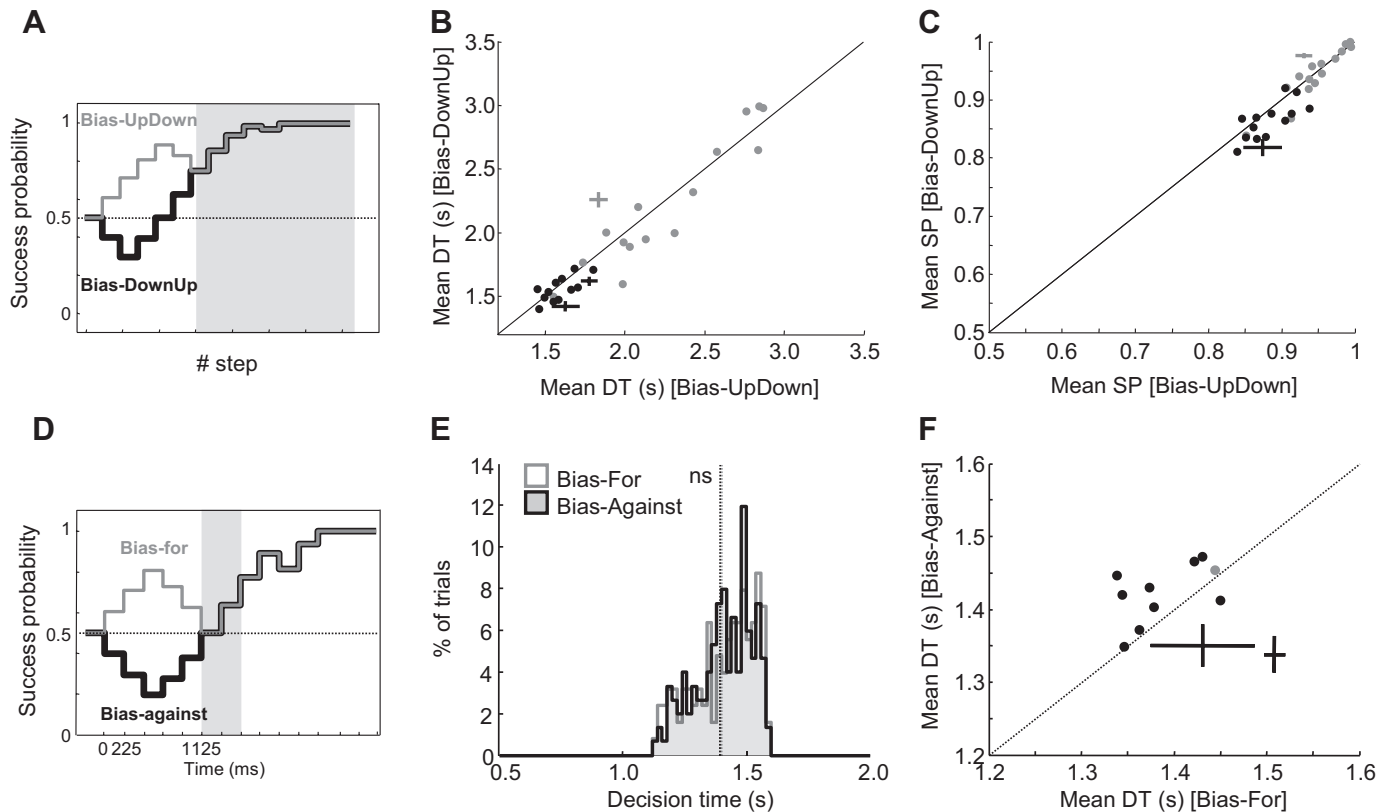


Fig. 8. Additional control analyses. **A**: temporal profiles of SP for “bias-updown” and “bias-downup” trials. In bias-updown trials, the first four coherence steps were toward the correct direction and the next two steps were in the opposite direction. In the bias-downup trials, the two first steps favored the wrong direction and the next four steps favored the correct direction. In both trial types, the remaining steps were similar and resembled an easy trial. **B**: average DTs for all subjects (same conventions as in Fig. 7C). We only focused on those trials in which decisions were made after the sixth coherence step (gray area in A). **C**: the SPs at DT (same conventions as in Fig. 7D) were similar in the two kinds of trials for most of subjects, in both the time pressure (13 of 14 subjects) and no time pressure (15 of 16 subjects) conditions. **D**: temporal profiles of SP for bias-for and bias-against trials. We examined at DTs from a subset of bias-for and bias-against trials in which a subject made the decision within 450 ms after the sixth coherence step (gray area). Eleven and one subjects who performed the time pressure and no time pressure conditions, respectively, made enough of these fast decisions (at least 5 trials of each type) to make the comparison possible. **E**: distributions of DTs across all subjects showed that there were no significant differences between DTs for bias-for versus bias-against trials. **F**: analyses of DTs in individual subjects showed that in only two cases, decisions were significantly different in the two trial types [ $P < 0.05$  by Kolmogorov-Smirnov (KS) test], being faster in bias-against trials, contradicting again the predictions of models that integrate motion signals.

related to the variable  $y(t)$  (Eq. 24), which approximates the current state of the sensory information. It is reasonable to assume that subjects’ confidence strongly relies on the sensory evidence provided by the stimulus at the moment of decision. Of course, we cannot know the exact form of the  $E(t)$  function used by our subjects to solve the task, but we can safely assume that they don’t calculate Eq. 30 at every coherence step. As an alternative, we propose that subjects estimate sensory evidence using a simple “first-order” estimate of sensory evidence. If this is computed simply by adding up novel information, then it will be related to the sum of LogLRs (SumLogLR) of individual coherence steps. To test this, we grouped trials according to the number of coherence steps that passed before the DT and calculated SumLogLR for the selected target at the time of the decision (see Cisek et al. 2009). Figure 10A shows the result of this analysis for one subject who did both time pressure and no time pressure versions of the task. In both conditions, SumLogLR decreased over time, but the dropping effect was stronger (i.e., significant negative regression and stronger negative slope) in the time pressure condition. Note that this trend existed despite the fact that, on average, the success probability increased over time. At the population level, we found a negative slope regression for all subjects

performing the time pressure version of the task (mean:  $-0.068$ ; SD: 0.036), among which 10 of 14 subjects (71%) were significant (Fig. 10B). In contrast, we found 12 and 4 regressions with a negative and a positive slope, respectively, in the no time pressure condition (mean:  $-0.029$ ; SD: 0.032). Among the 12 negative regressions, only 4 negative regressions were significant (25%; Fig. 10C). In summary, there was a trend for later decisions to be made at a lower level of SumLogLR than decisions made early in the trial, especially under time pressure, consistent with the predictions of the urgency-gating model.

**Effect of task instruction.** Despite the use of a noisy stimulus similar to the well-studied motion discrimination task, one important difference remains between our paradigm and those of previous studies. In studies that used constant evidence, subjects were asked to make a perceptual judgment about the current state of the stimulus. In contrast, here (and in Cisek et al. 2009), subjects were asked to use current perceptual judgments (e.g., detected changes in motion coherence) to infer a prediction about the future state of the stimulus. It is plausible that this produces a difference in the strategies used by the subjects. To test this, nine subjects performed an additional control experiment. In a prediction condition, these subjects

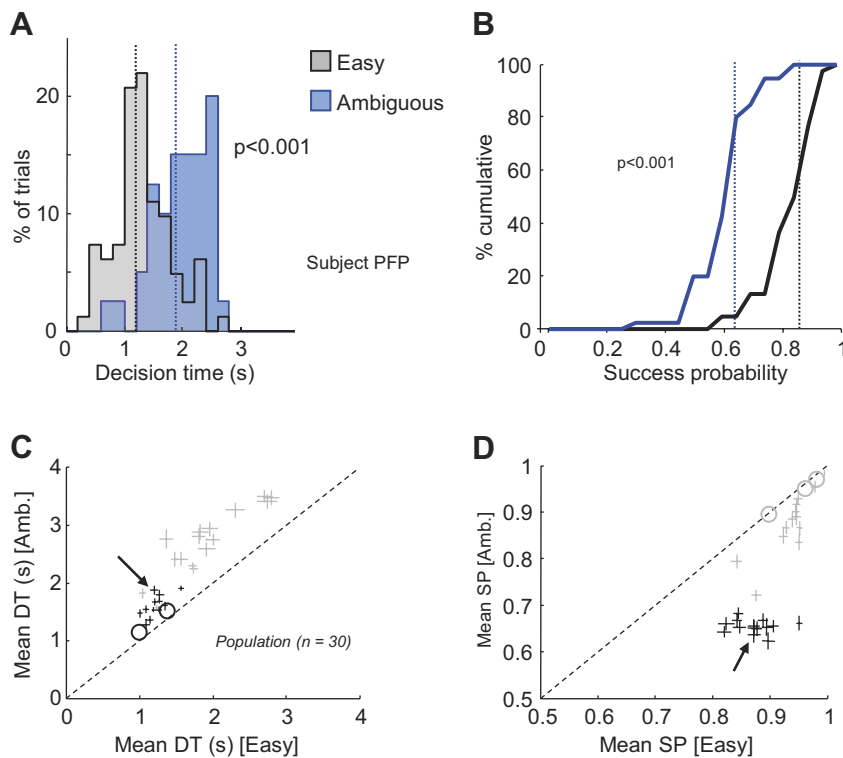


Fig. 9. Behavior in easy versus ambiguous trials. *A*: distribution of DTs for one representative subject in easy (gray,  $n = 82$ ) and ambiguous (blue,  $n = 40$ ) trials. The mean DT (black and blue dotted lines) was shorter in easy trials than in ambiguous trials (1,216 vs. 1,886 ms,  $P < 0.001$  by KS test). *B*: cumulative distribution of SPs at DT during easy (black) and ambiguous (blue) trials for the same subject. The mean SP in the easy trials (0.87) was significantly higher than in the ambiguous trials (0.65). *C*: average DTs for each subject during the time pressure (black) and no time pressure (gray) conditions (same conventions as in Fig. 7). Of 30 subjects, 28 subjects made decisions significantly later in ambiguous trials than in easy trials. *D*: average SPs for all subjects.

performed the VMD task and were instructed to predict the net motion direction at the end of the trial, just as in the experiments described above. In separate blocks, these same subjects also performed a detection condition, in which they were instructed to indicate the direction of motion as soon as they detected one and to ignore any subsequent motion changes. As shown in Fig. 11, the majority of subjects (8 of 9 subjects) did not show a significant difference between DTs (after 1,125 ms) for bias-for versus bias-against trials even in the detection condition. Again, this does not indicate that they ignored the early motion signals, because decisions made in the detection

condition before 1,125 ms were correct in 97.3% of bias-for trials and 90.3% of bias-against trials (note: a correct answer was defined on the basis of the current motion stimulus). This suggests that even when the subjects were asked to detect momentary motion, they were influenced by the motion signal but did not accumulate it for very long. In other words, their behavior was not governed by an integration process with a long time constant, but it was compatible with urgency gating.

**Model simulations.** Figure 12A shows the time course of the model variables during the CMD task. When the motion began ( $t = 500$  ms), there was a transient peak in the variable  $w$  (Fig.

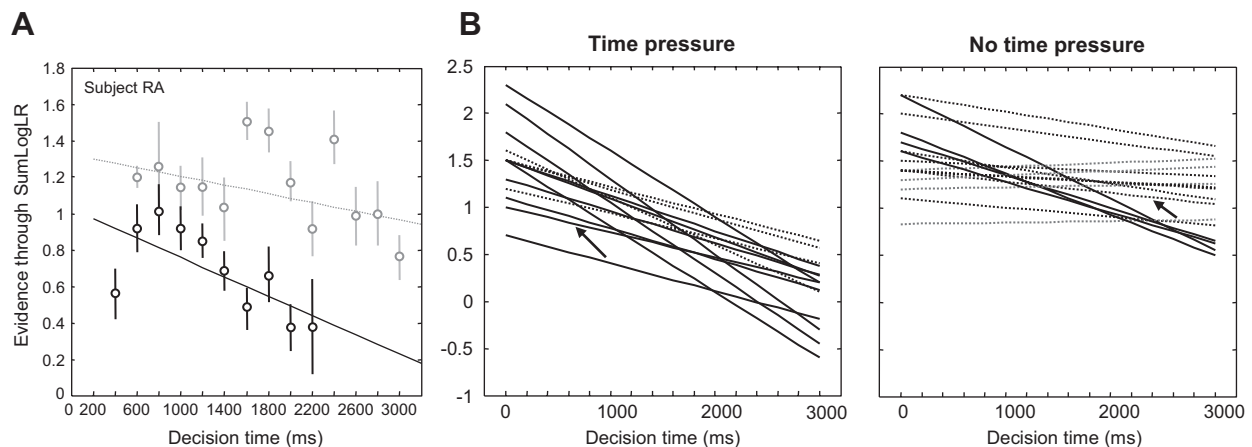


Fig. 10. Effect of decision time on confidence level. *A*: analysis of the confidence level [approximated as the sum of log-likelihood ratios at DT (see Cisek et al. 2009)] for decisions made at different times for an individual subject's data from time pressure (black circles) and no time pressure (gray circles) conditions. All trials were grouped according to the number of coherence steps that occurred before DT, and we show means  $\pm$  SE of the sum of log-likelihood ratio (vertical lines) calculated for the selected target at that time. Lines illustrate linear regressions through the data. For this subject, the linear regression was significant in the time pressure condition (black solid line, slope:  $-0.053$ ,  $r = -0.68$ ,  $P = 0.03$ ) but not in the no time pressure condition (gray dotted line, slope:  $-0.024$ ,  $r = -0.42$ ,  $P = 0.15$ ). *B*: at the population level, a significant regression (black solid lines) was found for 10 of 14 subjects in the time pressure condition (left), whereas for the no time pressure condition (right), we found only 4 significant regressions among the 12 negative slopes (dotted lines) as well as 4 regressions with a positive slope (gray dotted lines). In both *A* and *B*, black arrows indicate the subject shown in *A*.

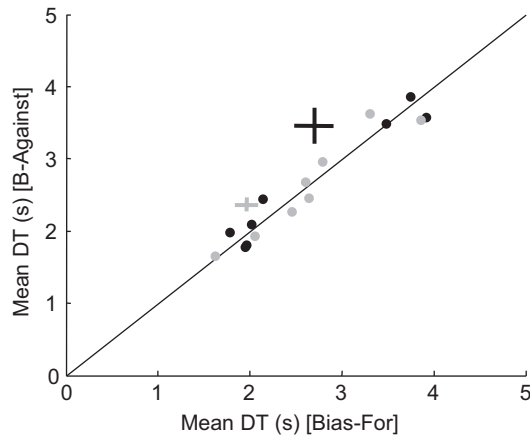


Fig. 11. Effect of task instruction on DTs in bias-for versus bias-against trials. In a control experiment, we tested nine subjects who performed both a “prediction” and a “detection” version of the VMD task (see text). Here, we show the mean DTs (same conventions as Fig. 7) in bias-for versus bias-against trials (using data only from trials in which decisions were made after the sixth coherence step). There were no significant differences between DTs in bias-for and bias-against trials for 8 of 9 subjects in both the prediction (black) and detection (gray) versions of the task.

12A, *left*), which caused a brief growth of the variable  $y$  (Fig. 12A, *middle*). Because of the growing urgency signal (not shown), the variable  $x$  (Fig. 12A, *right*) continued to grow until it reached the decision threshold. Figure 12B shows model behavior in the no time pressure version of the VMD task, using the same set of trials presented to our subjects. As shown in Fig. 12B, *middle*, the level reached by the variable  $y$  (which is a low-pass-filtered estimate of stimulus information) dropped gradually with time, as observed in the data (Fig. 10B). With time pressure (Fig. 12C), this drop was steeper. Fig. 12, *D* and *E*, shows model behavior during the VMD task for the same easy, ambiguous, bias-for, and bias-against trials used by our human subjects. As shown in Fig. 12D, the model made decisions faster but at a higher value of success probability in easy trials than in ambiguous trials. As shown in Fig. 12E, there were no significant differences of behavior between bias-for and bias-against trials when decisions were made after six coherence steps. Note, however, that there was a slight and nonsignificant tendency for the model to decide faster in bias-against trials (2,379 ms) than in bias-for trials (2,446 ms). This is similar to what was observed in the data (Fig. 7C). Cisek et al. (2009) simulated the behavior of several kinds of integrator models during an analogous changing evidence task and found that no variation of models that integrate sensory information can reproduce human behavior in such conditions. The only models that succeed are ones in which the sensory information (computed by low-pass filtering or by integrating the change in the information) is multiplied by a growing urgency signal before comparison with a neural firing threshold.

## DISCUSSION

Many recent models have proposed that decision making involves the temporal integration of sequential sensory samples until a fixed bound is reached (Bogacz et al. 2006; Bogacz and Gurney 2007; Carpenter and Williams 1995; Grossberg and Pilly 2008; Mazurek et al. 2003; Ratcliff et al. 2007; Roitman and Shadlen 2002; Smith and Ratcliff 2004; Usher and McClelland 2001; Wong and Wang 2006). However, slow inte-

gration is suboptimal if the environment changes, motivating animals to make perceptual judgments using a short temporal window (Chittka et al. 2009; Trimmer et al. 2008). Indeed, several studies have shown that decisions appear to be based primarily on information from a short time window (Cook and Maunsell 2002; Luna et al. 2005; Uchida et al. 2006; Yang et al. 2008), raising the question of what could be responsible for the much longer lasting build up of neural activity that appears to determine the timing of decisions.

Here, we propose an alternative to the classic model. In our model, what is integrated is not the state of evidence pertinent to a choice but rather the change in that state. This follows from considering the redundancy among successive samples, motivating animals to emphasize information that is novel. In a constant evidence task, integration of novel information will indeed be brief (<200 ms). Furthermore, we propose that the prolonged build up of neural activity beyond that initial time window is primarily due to an urgency signal that delays commitment to allow the environment to provide new information. This is motivated by a simple policy for achieving a trade off between speed and accuracy that maximizes what animals care about most: reward rate (Balci et al. 2011).

Our conclusions are in good agreement with those of an earlier study that presented subjects with changing evidence indicated by moving tokens (Cisek et al. 2009). That study also supported the urgency-gating model, but it was unclear whether the results were task dependent. In particular, the stimulus used in the Cisek et al. experiment may have favored the urgency-gating model because it required no memory, had no noise, and provided a cue of elapsing time. Here, however, subjects were presented with a stimulus whose perceptual properties were very similar to those previously used in many perceptual decision-making tasks. Even with this very noisy stimulus, subjects' behavior did not conform with the predictions of models in which the motion itself is integrated over time to a fixed decision bound (Mazurek et al. 2003). Finally, and most importantly, this holds true even when subjects were given the same instructions as those used in most of the perceptual decision tasks, i.e., to detect the current motion of the stimulus (Fig. 11). Although subjects were clearly influenced by biases in the motion signal (as evidenced by their early choices), that bias was quickly abandoned if they made their decisions later in time. Therefore, we propose that even during classic signal detection tasks, the underlying mechanism determining decisions may be analogous to our urgency-gating model.

In simple detection tasks, the brain may approximate the detection of novelty and its integration with a simple low-pass filter, whose output is then gated by urgency. This is equivalent to what Ditterich (2006a, 2006b) described as a leaky integrator with time-varying gain. Importantly, however, our data suggests that the time constant of integration is very short (e.g., 100 ms), much shorter than what is normally assumed by bounded integration models. This is in line with studies showing a short time window from which sensory information is gathered to make decisions (Cook and Maunsell 2002; Ghose 2006; Ludwig et al. 2005; Ratcliff 2002; Stanford et al. 2010; Uchida et al. 2006). It remains to be seen whether the time window can be longer in some conditions. Indeed, it appears to be relatively short even during noisy motion discrimination tasks. In a fixed-duration task, Kiani et al. (2008) found that

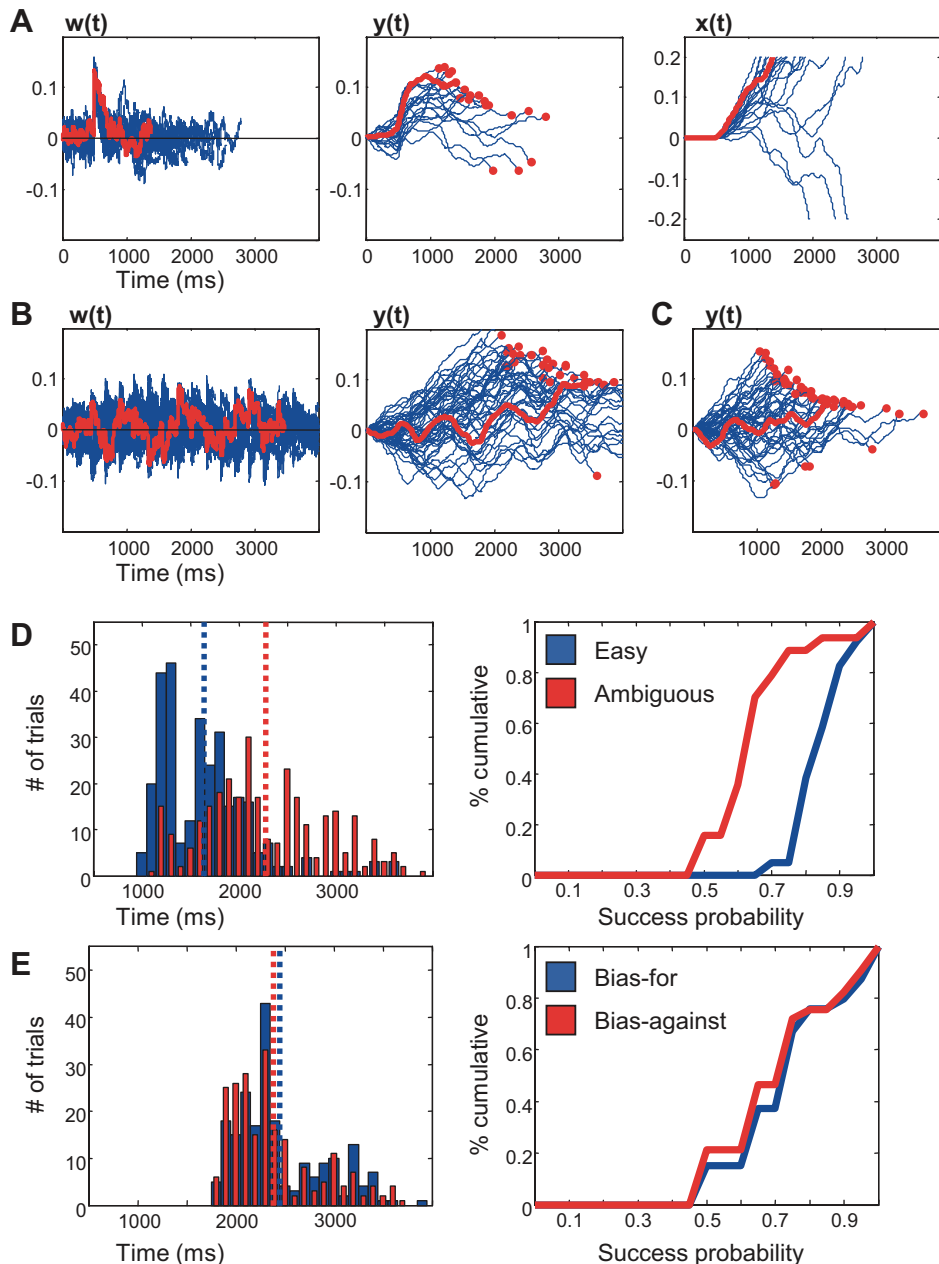


Fig. 12. Model simulations. *A*: simulations of model behavior in the constant-coherence task. Time courses of three of the model variables when a motion stimulus of strength  $S = 12$  appeared at time  $t = 500$  and remained for the rest of the trial are shown. In the *left*, *middle*, and *right* graphs, 20 individual trials are shown, with one highlighted in red. In the *middle* graphs, red circles indicate the value reached at the time of the decision. In this simulation, scalar gain ( $\beta$ ) was set to 2. In all of the trials, the correct answer corresponded to  $x$ , reaching the positive threshold ( $T = 0.2$ ), but in some cases, noise produced an error and  $x$  instead reached  $T = -0.2$ . *B*: behavior in 50 VMD trials performed in the no time pressure condition ( $\beta = 0.66$ ). *C*: behavior in the same 50 trials in the time pressure condition ( $\beta = 2$ ). *D*: comparison of DT distributions (*left*) and cumulative distributions of SP (*right*) of 300 easy trials (blue) versus 300 ambiguous trials (red). Vertical lines indicate means of DT distributions. *E*: comparison of 300 bias-for trials (blue) and 300 bias-against trials (red), in the same format as in *D*.

brief motion pulses had an effect on monkeys' performance only if they occurred within the first 400 ms of motion viewing. In a RT version of the same task, motion pulses had a long-lasting effect on neural activity, but this effect was much weaker for pulses that occurred  $>200$  ms after the motion onset (Huk and Shadlen 2005), as if the temporal integration used to quantify evidence was completed quickly.

Here and in our earlier study (Cisek et al. 2009), we proposed that the urgency-gating model cannot be distinguished from the bounded integration model using tasks in which the evidence is constant. However, Churchland et al. (2011) recently suggested a method for distinguishing such models at the neural level using analyses of high-frequency noise in neural spike trains. They concluded that the available data are only consistent with an integration process and not with urgency gating. However, they only examined a version of the urgency-gating model without any low-pass filtering.

We propose that the sensory signal is low-pass filtered before scaling by urgency, as in Eq. 26. If this occurs, then the model again cannot be distinguished from an integrator for the trivial reason that a low-pass filter and an integrator are mathematically equivalent with respect to high-frequency noise. Thus, to our knowledge, these models cannot be distinguished with such methods, including Fano factor analyses (Churchland et al. 2010; Nawrot et al. 2008), and, ultimately, the question can only be settled with tasks presenting at least some change in sensory evidence.

One observation that appears to argue for the presence of slow integration is that during fixed-duration sensory discrimination tasks, success rates tend to increase with viewing duration (Britten et al. 1992; Gold and Shadlen 2003; Palmer et al. 2005; Roitman and Shadlen 2002). Bounded integrator models explain this result because with more time, integration is more and more effective at canceling out noise (Ratcliff 2001) (this is also true of our model



because it also integrates high-frequency noise). Nevertheless, improvements after 400 ms tend to be very modest, even in very noisy trials (Gold and Shadlen 2003; Uchida et al. 2006). A brief period of improvement may result from the kind of brief integration that we describe above, and improvements with additional viewing time beyond that may be due to attentional fluctuations: given more viewing time, there is more chance of attending a stimulus, resulting in more accurate performance when averaged across a large number of trials. Finally, the time window of integration may be task dependent, potentially explaining why performance accuracy sometimes asymptotes quickly (Burr and Santoro 2001; Ratcliff 2002) and sometimes continues to improve with prolonged exposure to a stimulus (Palmer et al. 2005).

Our model is in agreement with the idea that decision making involves two processes that turn sensory evidence into action (Carpenter et al. 2009; Reddi 2001): the detection of the stimulus and the response commitment itself. A recent experiment by Bennur and Gold (2011) separated these by providing monkeys with a random-dot motion stimulus but indicating the mapping between motion direction and saccade target before, during, or after the motion stimulus. The main finding was that the activity of some cells in the lateral intraparietal area (LIP) reflected the direction of motion even if the mapping was not yet known. Interestingly, such motion-related activity reached a plateau quickly ( $\sim 300$  ms). A gradual build up was only observed when the mapping was known. Even more striking evidence that build up of activity may not be attributable to sensory integration is provided by the work of Stanford et al. (2010). In their “compelled-saccade task,” these authors dissociated sensory and motor contributions to decisions by varying the time at which the sensory stimulus was provided to guide the monkey’s choice, and it was sometimes provided after saccade initiation. In such cases, the motor plan and related build up of frontal eye field activity are initiated in the absence of the visual cue, but a choice (i.e., a guess) is still made. This suggests that the build up is attributable to an urgency signal (Churchland et al. 2008; Cisek et al. 2009; Drugowitsch et al. 2012; Standage et al. 2011), which is related to motor initiation (Janssen and Shadlen 2005; Renoult et al. 2006) rather than to an accumulation of sensory evidence.

Our proposal is also consistent with recent results showing that the extent to which human and monkey decisions are influenced by priors increases as a function of DT, in a nearly linear manner (Hanks et al. 2011). Hanks and colleagues separated the decision variable into two terms: one related to the priors and one to sensory evidence, which they assumed to involve an integrator. They showed that the relative contribution of the priors increases linearly with time. Our explanation for this is captured by *Eqs. 21* and *26*, in which both the term related to priors and the term related to evidence are scaled by the urgency signal, and we propose that this is a direct consequence of a mechanism for maximizing reward rate.

One question raised by our model concerns the site of the novelty detection and integration mechanisms shown in Fig. 2. For a motion discrimination task, one would expect this to involve area MT (Britten et al. 1993). However, MT cells respond in a manner proportional to motion strength (albeit sometimes with a transient burst) (Britten et al. 1993; Cook and Maunsell 2002), suggesting that they lie either at the output of  $y(t)$  or serve as input to the system. Although our mathematical derivation suggests separate stages of novelty detection, low-

pass filtering, and integration, it is possible that in the brain these processes may simply be approximated with a low-pass filter. Our data cannot distinguish these alternatives. An interesting line of future research would be to record MT activity in conditions in which the motion stimulus changes in a manner that is either more or less predicted by prior stimuli.

Another question raised by our model concerns the origin of the urgency signal. In monkeys, time-dependent neural activity has been reported in several cortical areas known to be involved in both sensorimotor control of movement and decision making. These notably include the LIP (Churchland et al. 2008; Hanks et al. 2011; Janssen and Shadlen 2005; Leon and Shadlen 2003; Maimon and Assad 2006), supplementary motor area (SMA) and pre-SMA (Mita et al. 2009), and prefrontal cortex (Genovesio et al. 2006). Although we cannot exclude a role of these areas in the generation of the urgency signal, an alternative and appealing option is that they all receive a common urgency signal coming from the basal ganglia (BG). Parts of the BG are in direct communication with the thalamus and cortex, creating functional loops (Alexander et al. 1990). Recent data have suggested that the output of the BG regulates the speed and size (the “vigor”) of movement. In monkeys, inactivation of the internal segment of the globus pallidus reduces movement velocity and acceleration (Desmurget and Turner 2010; Horak and Anderson 1984), and a major deficit of Parkinson’s disease is the inability to move rapidly (Mazzoni et al. 2007). If we consider that decision and movement are ultimately aimed at yielding rewards, the total elapsed time between stimulus appearance and final movement offset can be seen as a *temporal cost* that discounts the value of the reward. In this view, the duration of action or, more generally, of any neural process carries a penalty related to elapsed time. Recently, Shadmehr et al. (2010) showed a relationship between reward discounting and movement speed, raising the possibility that both are modulated by a common signal. Our results are in agreement, proposing that the timing of decisions may also be influenced by that same signal, potentially originating in the BG, whose overarching role is to achieve the highest average reward rate.

Finally, although the notion of urgency in decision making has been raised many years ago (Reddi and Carpenter 2000), it is still an open question how a hypothetical urgency signal is incorporated into the decision process. Churchland et al. (2008) suggested that urgency is a time-varying signal added to the decision variable coded by LIP neurons. In contrast, computational models have suggested that urgency multiplicatively gates sensory information (Cisek et al. 2009; Ditterich 2006a; Standage et al. 2011). In the present study, as in Cisek et al. (2009), we propose a multiplicative urgency signal because this successfully explains previous data from constant evidence tasks and recent data on the influence of priors (Hanks et al. 2011). The “accelerated race to threshold” model developed by Stanford et al. (2010) to describe data in their compelled-saccade task is also compatible with a multiplicative interaction between sensory evidence and a time-varying motor urgency signal. In each trial, activities related to each option start growing linearly to a threshold. When a sensory cue is provided, it causes the acceleration of these build-up signals—consistent with the multiplication of each with a briefly growing signal produced by a low-pass filter of the sensory cue. However, we cannot exclude the possibility that urgency is

additive but that current paradigms, including ours, have not yet adequately dissociated these possibilities. Additional neurophysiological studies may be needed to resolve this question conclusively.

In summary, considerations of reward rate maximization and redundancy among samples have led us to propose a novel model of decision making that differs from classic models in two important ways. First, we propose that if computation of evidence does involve integration, it is integration of only novel evidence. Therefore, it is brief in situations when novel information is only presented at the beginning of a trial and subsequent samples are increasingly redundant. Indeed, it is possible that the brain implements the integration of novel evidence simply with a low-pass filter. Second, we propose that the build up of neural activity observed in many decision-making tasks is not primarily attributable to the integration of sensory information but that it is caused by an urgency signal, related to motor preparation, that implements a simple policy for achieving a speed-accuracy trade off. We propose that our model is compatible with all previous results favoring integration models because it is nearly equivalent in the conditions previously tested. It provides an explanation for the seemingly paradoxical finding that even when a long time is taken for a decision, the choice is influenced primarily by information from a very short and early time window (Cook and Maunsell 2002; Luna et al. 2005; Uchida et al. 2006; Yang et al. 2008). Finally, we believe that the model may provide a promising theoretical link between the mechanisms of temporal decisions and one of the major motivations in animal behavior: the maximization of reward.

#### ACKNOWLEDGMENTS

The authors are grateful to Andrea Green, Erik P. Cook, John Kalaska, and Francois Rivest for discussions and comments on the manuscript.

#### GRANTS

The work was supported by a Fondation Fyssen Fellowship (to D. Thura), Canadian Institutes of Health Research Grant MOP-102662 and EJLB Foundation grants (to P. Cisek), and a Fonds de la Recherche en Santé du Québec infrastructure grant.

#### DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

#### AUTHOR CONTRIBUTIONS

Author contributions: D.T. and P.C. conception and design of research; D.T., J.B.-R., and C.-W.F. performed experiments; D.T., J.B.-R., and C.-W.F. analyzed data; D.T., J.B.-R., C.-W.F., and P.C. interpreted results of experiments; D.T. prepared figures; D.T. drafted manuscript; D.T. and P.C. edited and revised manuscript; D.T. and P.C. approved final version of manuscript.

#### REFERENCES

- Alexander GE, Crutcher MD, DeLong MR.** Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog Brain Res* 85: 119–146, 1990.
- Balci F, Simen P, Niyogi R, Saxe A, Hughes JA, Holmes P, Cohen JD.** Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Atten Percept Psychophys* 73: 640–657, 2011.
- Bennur S, Gold JI.** Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. *J Neurosci* 31: 913–921, 2011.
- Bogacz R.** Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn Sci* 11: 118–125, 2007.
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD.** The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* 113: 700–765, 2006.
- Bogacz R, Gurney K.** The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput* 19: 442–477, 2007.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA.** The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA.** Responses of neurons in macaque MT to stochastic motion signals. *Vis Neurosci* 10: 1157–1169, 1993.
- Burr DC, Santoro L.** Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Res* 41: 1891–1899, 2001.
- Carpenter RH, Reddi BA, Anderson AJ.** A simple two-stage model predicts response time distributions. *J Physiol* 587: 4051–4062, 2009.
- Carpenter RH, Williams ML.** Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377: 59–62, 1995.
- Chittka L, Skorupski P, Raine NE.** Speed-accuracy tradeoffs in animal decision making. *Trends Ecol Evol* 24: 400–407, 2009.
- Churchland AK, Kiani R, Chaudhuri R, Wang XJ, Pouget A, Shadlen MN.** Variance as a signature of neural computations during decision making. *Neuron* 69: 818–831, 2011.
- Churchland AK, Kiani R, Shadlen MN.** Decision-making with multiple alternatives. *Nat Neurosci* 11: 693–702, 2008.
- Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV.** Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13: 369–378, 2010.
- Cisek P, Puskas GA, El-Murr S.** Decisions in changing conditions: the urgency-gating model. *J Neurosci* 29: 11560–11571, 2009.
- Cook EP, Maunsell JH.** Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5: 985–994, 2002.
- Desmurget M, Turner RS.** Motor sequences and the basal ganglia: kinematics, not habits. *J Neurosci* 30: 7685–7690, 2010.
- Ditterich J.** Evidence for time-variant decision making. *Eur J Neurosci* 24: 3628–3641, 2006a.
- Ditterich J.** Stochastic models of decisions about motion direction: behavior and physiology. *Neural Netw* 19: 981–1012, 2006b.
- Drugowitsch J, Moreno-Bote R, Churchland AK, Shadlen MN, Pouget A.** The cost of accumulating evidence in perceptual decision making. *J Neurosci* 32: 3612–3628, 2012.
- Genovesio A, Tsujimoto S, Wise SP.** Neuronal activity related to elapsed time in prefrontal cortex. *J Neurophysiol* 95: 3281–3285, 2006.
- Ghose GM.** Strategies optimize the detection of motion transients. *J Vis* 6: 429–440, 2006.
- Gold JI, Shadlen MN.** The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J Neurosci* 23: 632–651, 2003.
- Gold JI, Shadlen MN.** The neural basis of decision making. *Annu Rev Neurosci* 30: 535–574, 2007.
- Gold JI, Shadlen MN.** Representation of a perceptual decision in developing oculomotor commands. *Nature* 404: 390–394, 2000.
- Grossberg S, Pilly PK.** Temporal dynamics of decision-making during motion perception in the visual cortex. *Vision Res* 48: 1345–1373, 2008.
- Hanes DP, Schall JD.** Neural control of voluntary movement initiation. *Science* 274: 427–430, 1996.
- Hanks TD, Mazurek ME, Kiani R, Hopp E, Shadlen MN.** Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J Neurosci* 31: 6339–6352, 2011.
- Horak FB, Anderson ME.** Influence of globus pallidus on arm movements in monkeys. I. Effects of kainic acid-induced lesions. *J Neurophysiol* 52: 290–304, 1984.
- Huk AC, Shadlen MN.** Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J Neurosci* 25: 10420–10436, 2005.
- Janssen P, Shadlen MN.** A representation of the hazard rate of elapsed time in macaque area LIP. *Nat Neurosci* 8: 234–241, 2005.

- Kiani R, Hanks TD, Shadlen MN.** Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci* 28: 3017–3029, 2008.
- Kim JN, Shadlen MN.** Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2: 176–185, 1999.
- Leon MI, Shadlen MN.** Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38: 317–327, 2003.
- Ludwig CJ, Gilchrist ID, McSorley E, Baddeley RJ.** The temporal impulse response underlying saccadic decisions. *J Neurosci* 25: 9907–9912, 2005.
- Luna R, Hernandez A, Brody CD, Romo R.** Neural codes for perceptual discrimination in primary somatosensory cortex. *Nat Neurosci* 8: 1210–1219, 2005.
- Maimon G, Assad JA.** A cognitive signal for the proactive timing of action in macaque LIP. *Nat Neurosci* 9: 948–955, 2006.
- Mazurek ME, Roitman JD, Ditterich J, Shadlen MN.** A role for neural integrators in perceptual decision making. *Cereb Cortex* 13: 1257–1269, 2003.
- Mazzoni P, Hristova A, Krakauer JW.** Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J Neurosci* 27: 7105–7116, 2007.
- Mita A, Mushiake H, Shima K, Matsuzaka Y, Tanji J.** Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat Neurosci* 12: 502–507, 2009.
- Nawrot MP, Boucsein C, Rodriguez Molina V, Riehle A, Aertsen A, Rotter S.** Measurement of variability dynamics in cortical spike trains. *J Neurosci Methods* 169: 374–390, 2008.
- Palmer J, Huk AC, Shadlen MN.** The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J Vis* 5: 376–404, 2005.
- Price NS, Born RT.** Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *J Neurosci* 30: 14036–14045, 2010.
- Ratcliff R.** A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychon Bull Rev* 9: 278–291, 2002.
- Ratcliff R.** Putting noise into neurophysiological models of simple decision making. *Nature Neuroscience* 4: 336, 2001.
- Ratcliff R.** A theory of memory retrieval. *Psychol Rev* 85: 59–108, 1978.
- Ratcliff R, Gomez P, McKoon G.** A diffusion model account of the lexical decision task. *Psychol Rev* 111: 159–182, 2004.
- Ratcliff R, Hasegawa YT, Hasegawa RP, Smith PL, Segraves MA.** Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *J Neurophysiol* 97: 1756–1774, 2007.
- Reddi BA.** Decision making: the two stages of neuronal judgement. *Curr Biol* 11: R603–606, 2001.
- Reddi BA, Carpenter RH.** The influence of urgency on decision time. *Nat Neurosci* 3: 827–830, 2000.
- Renoult L, Roux S, Riehle A.** Time is a rubberband: neuronal activity in monkey motor cortex in relation to time estimation. *Eur J Neurosci* 23: 3098–3108, 2006.
- Roitman JD, Shadlen MN.** Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22: 9475–9489, 2002.
- Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY.** Temporal discounting of reward and the cost of time in motor control. *J Neurosci* 30: 10507–10516, 2010.
- Smith PL, Ratcliff R.** Psychology and neurobiology of simple decisions. *Trends Neurosci* 27: 161–168, 2004.
- Standage D, You H, Wang DH, Dorris MC.** Gain modulation by an urgency signal controls the speed-accuracy trade-off in a network model of a cortical decision circuit. *Front Comput Neurosci* 5: 7, 2011.
- Stanford TR, Shankar S, Massoglia DP, Costello MG, Salinas E.** Perceptual decision making in less than 30 milliseconds. *Nat Neurosci* 13: 379–385, 2010.
- Thura D, Cisek P.** Human perceptual decisions in noisy, changing conditions (Abstract). *Soc Neurosci Abstr* 576.3: 2009.
- Trimmer PC, Houston AI, Marshall JA, Bogacz R, Paul ES, Mendl MT, McNamara JM.** Mammalian choices: combining fast-but-inaccurate and slow-but-accurate decision-making systems. *Proc Biol Sci* 275: 2353–2361, 2008.
- Uchida N, Kepecs A, Mainen ZF.** Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nat Rev Neurosci* 7: 485–491, 2006.
- Usher M, McClelland JL.** The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev* 108: 550–592, 2001.
- Wald A.** Sequential tests of statistical hypotheses. *Ann Math Stat* 16: 117–186, 1945.
- Wald A, Wolfowitz J.** Optimum character of the sequential probability ratio test. *Ann Math Stat* 19: 326–329, 1948.
- Wong KF, Wang XJ.** A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci* 26: 1314–1328, 2006.
- Yang T, Shadlen MN.** Probabilistic reasoning by neurons. *Nature* 447: 1075–1080, 2007.
- Yang Y, DeWeese MR, Otazu GH, Zador AM.** Millisecond-scale differences in neural activity in auditory cortex can drive decisions. *Nat Neurosci* 11: 1262–1263, 2008.