J Neurophysiol 114: 99–113, 2015.
First published May 6, 2015; doi:10.1152/jn.00793.2014.

CALL FOR PAPERS | *Decision Making: Neural Mechanisms*

Confidence estimation as a stochastic process in a neurodynamical system of decision making

Ziqiang Wei[1,2] and Xiao-Jing Wang[3,4]

[1]*Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia;* [2]*The Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, Maryland;* [3]*Center for Neural Science, New York University (NYU), New York, New York; and* [4]*NYU-ECNU Institute of Brain and Cognitive Science, NYU Shanghai, Shanghai, China*

Submitted 8 October 2014; accepted in final form 6 May 2015

Wei Z, Wang XJ. Confidence estimation as a stochastic process in a neurodynamical system of decision making. *J Neurophysiol* 114: 99–113, 2015. First published May 6, 2015; doi:10.1152/jn.00793.2014.—Evaluation of confidence about one's knowledge is key to the brain's ability to monitor cognition. To investigate the neural mechanism of confidence assessment, we examined a biologically realistic spiking network model and found that it reproduced salient behavioral observations and single-neuron activity data from a monkey experiment designed to study confidence about a decision under uncertainty. Interestingly, the model predicts that changes of mind can occur in a mnemonic delay when confidence is low; the probability of changes of mind increases (decreases) with task difficulty in correct (error) trials. Furthermore, a so-called "hard-easy effect" observed in humans naturally emerges, i.e., behavior shows underconfidence (underestimation of correct rate) for easy or moderately difficult tasks and overconfidence (overestimation of correct rate) for very difficult tasks. Importantly, in the model, confidence is computed using a simple neural signal in individual trials, without explicit representation of probability functions. Therefore, even a concept of metacognition can be explained by sampling a stochastic neural activity pattern.

decision confidence; lateral intraparietal cortex; line-attractor neural model

A KEY TO MONITORING OF COGNITION (metacognition) is our ability to evaluate the degree of confidence that we have about a decision, a strategy to tackle the problem at hand, a newly acquired piece of knowledge, etc. Confidence estimation has been an important subject of research in cognitive and developmental psychology (Flavell 1979; Vickers 1979). In laboratory studies, confidence can be measured using postdecision wagering (PDW), where subjects first perform a perceptual decision and then make a high-low bet between a risky option (associated with a high reward if the first-order choice is correct, a loss otherwise) and a safe option (associated with a low reward regardless of the first-order choice). Consequently, if subjects have less confidence about their choice, they would be more likely to bet on the low but certain reward option (Dienes and Seth 2010; Fleming and Dolan 2010; Fleming et al. 2010; Kepecs and Mainen 2012; Persaud et al. 2007). Recently, researchers have begun to use PDW and other task designs with behaving animals to explore the neural basis of confidence estimation (Kepecs et al. 2008; Lak et al. 2014; Middlebrooks and Sommer 2011, 2012; Smith et al. 2003).

In a monkey experiment, Kiani and Shadlen (2009) extended a well-known discrimination task to examine neural signals correlated with confidence. In this task, a subject is required to decide between two possible directions (indicated by two directional targets) of a random-dots motion stimulus. Specifically, Kiani and Shadlen used a fixed-duration (FD) version of the task (Shadlen and Newsome 2001), where the visual motion stimulus is followed by a delay period, and monkeys must indicate the decision at the end of the delay by a saccadic response to one of the directional targets. In a random subset of trials, they offered a third target (T_s) during the delay period, and monkeys could opt to T_s for a certain but small amount of reward. Interestingly, monkeys selected T_s more often when motion strength was weaker or stimulus duration became shorter, under which conditions the error rate was higher and selecting T_s gave rise to an improvement of performance across trials. The probability of choosing T_s (P_{sure}) thus reflected a degree of choice uncertainty. Importantly, P_{sure} was found to modulate single-neuron activity in lateral intraparietal (LIP) area that was correlated with accumulating decision evidence of a choice. This finding supports the intuitive idea that confidence signal is an integral part of a decision process (Vickers 1979) and reflected in a neural decision variable (Gold and Shadlen 2007).

Computational schemes have been proposed for the study of confidence (Kepecs and Mainen 2012; Kepecs et al. 2008; Kiani and Shadlen 2009; Moreno-Bote 2010; Pleskac and Busemeyer 2010; Ratcliff and Starns 2009; Rolls et al. 2010a, 2010b; Vickers 1979). In particular, with the use of drift diffusion model (DDM) (Ratcliff and Smith 2004), confidence has been defined in terms of the log posterior ratio for the two choices given the decision variable at the time of behavioral response (Kiani and Shadlen 2009). This DDM, nevertheless, has some limitation to account for the complexity of confidence (Van Zandt 2000), e.g., a decision variable that terminates at a fixed threshold may not present a graded confidence across trial (Kiani et al. 2014).

In this work, to uncover neural circuit mechanisms underlying confidence estimation, we took a different approach and employed a biophysically realistic cortical network model of spiking neurons, which was previously shown to successfully

Address for reprint requests and other correspondence: X.-J. Wang, Center for Neural Science, New York Univ., 4 Washington Place, New York, NY 10003 (e-mail: xjwang@nyu.edu).

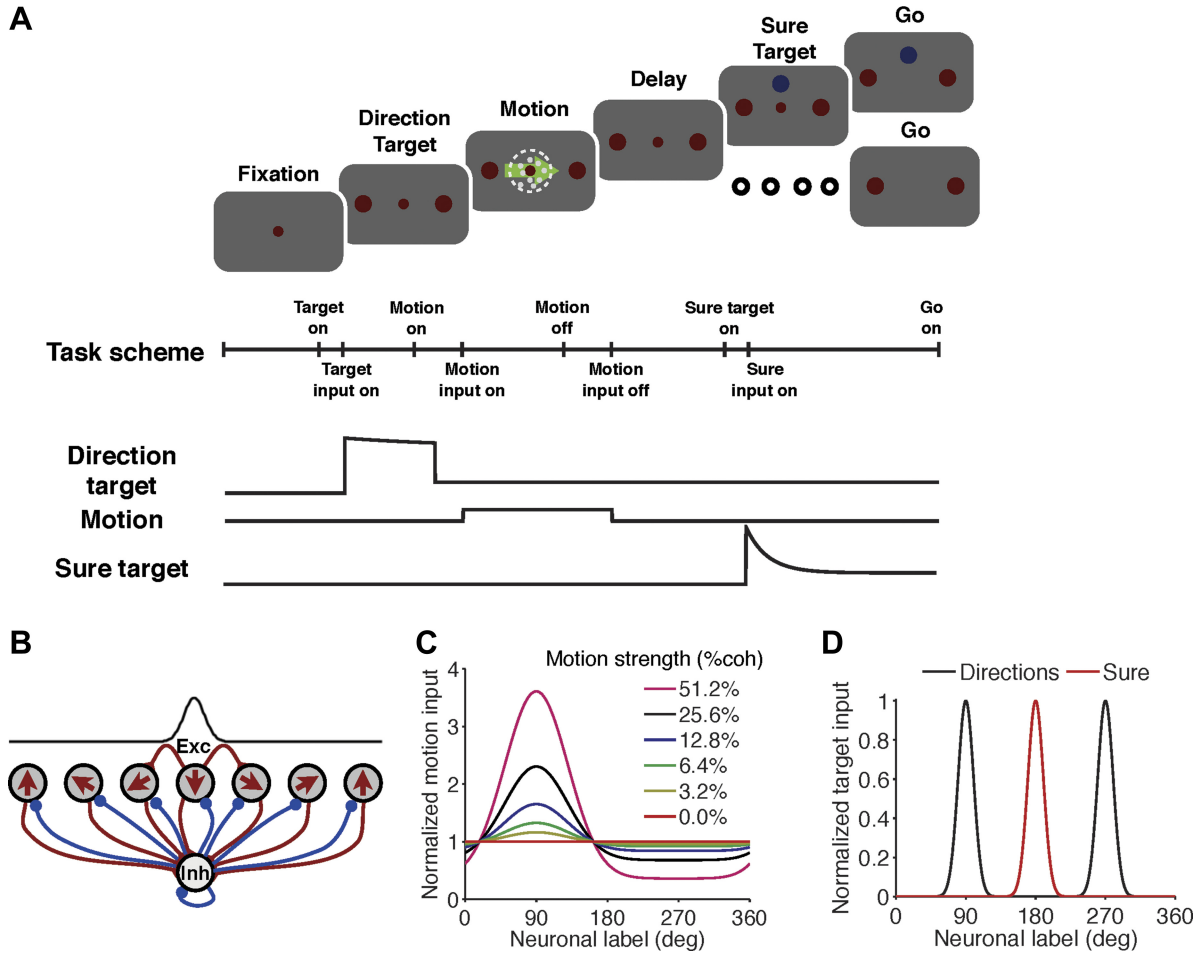

Fig. 1. Schematic description of the decision task and model architecture. *A*: procedure of a simulated fixed-duration discrimination task. Following a fixation period, two targets (large red circles) appear, indicating the alternative choices. A random-dots motion stimulus is presented, followed by a delay period. A saccade to one of the alternatives indicates the decision at the end of the delay. In some trials, a sure target (blue circle) is shown after the motion offset, and choosing it leads to a certain but small amount of reward. *Bottom*: detailed task and input schemes. *B*: neural network structure. The network consists of excitatory pyramidal cells (Exc) and inhibitory interneurons (Inh). The pyramidal cells are uniformly placed on a continuous ring, and each neuron is labeled by its preferred motion direction (shown as the arrow in pyramidal cells). The excitatory-to-excitatory connections between pyramidal cells are structured as a Gaussian function of the difference in their preferred directions (upper black curve), and the connections from and onto interneurons are broad. *C*: motion input (centered at 90°) with different motion strengths, the integral of which is identical for all motion strengths. *D*: normalized input of 2 directional targets (namely T_A at 90° and T_B at 270°) and a sure target (namely T_s at 180°).

simulate the random-dots motion-direction discrimination experiment (Furman and Wang 2008). We investigated whether the same model could accurately reproduce the salient findings from Kiani and Shadlen (2009). The model is endowed with a continuous network of neurons that can represent any direction; therefore, it can be readily extended to incorporate the presentation of a sure target (T_s) during a delay period. Notably, such a model of decision-making and memory processes was not originally designed for the Kiani-Shadlen experiment to account for confidence estimation. It is thus surprising that the model can capture both a range of behavioral performance data and physiological observations from LIP single neurons. We noted that neurons selective for T_s win the competition (thus T_s was chosen) when the differential activity of neurons selective for the two alternative choices is small. Quantitatively, we found that confidence could be estimated, at any time, as a sigmoid function of the differential firing activity of the two competing neural pools selective for the alternative choices. Therefore, choice confidence is computed simultaneously when a decision is made, and a trial-by-trial variation of choice is generated by sampling of stochastic neural dynamics (Wang 2008).

MATERIALS AND METHODS

Network model. We employed a spiking neural network model, which has been previously used to simulate a categorical decision of an analog feature, like motion direction (Furman and Wang 2008; Liu and Wang 2008). This model consists of 2,048 pyramidal cells and 512 interneurons. Both pyramidal cells and interneurons are modeled as integrate-and-firing neurons; excitatory postsynaptic currents from pyramidal cells are mediated by α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and *N*-methyl-D-aspartate (NMDA) receptors, whereas inhibitory postsynaptic currents from interneurons are mediated by GABA receptors. Pyramidal cells are uniformly placed on a ring according to their preferred motion directions and continuously span 360° of possible motion directions (Fig. 1*B*), while the interneurons constitute a nonselective neural pool. The recurrent connectivity strength between two pyramidal cells is a Gaussian function of the difference between their preferred motion directions, whereas those from and onto the interneurons are broad and uniform (Fig. 1*B*). All the cells receive a background noise mediated by

AMPA receptors, which is modeled as uncorrelated Poisson spike train. We used the neuronal and synaptic parameters from Furman and Wang (2008), which are fully specified therein, with a change of the background noise rate to 2,200 Hz, which ensures a choice is generated even if the motion strength is weak and stimulus duration is short. With these parameters, the network is endowed with winner-take-all competition so that only one of the neural pools wins (reaching an average population firing rate >50 Hz for at least 50 ms), and the decision is maintained in the form of a bell-shaped persistent activity pattern ("bump attractor") in a delay period.

Simulation protocol of FD discrimination-decision task. Our model assumes that neurons in area LIP incorporate sensory evidence (Cook and Maunsell 2002; Hanks et al. 2006; Roitman and Shadlen 2002) and reward signals (Platt and Glimcher 1999; Sugrue et al. 2004; Tobler et al. 2005). For simplicity, we assumed that the amount of reward for each target (i.e., two directional targets and a sure target) is associated with the instantaneous input strength of its current at the moment of the target onset (Soltani and Wang 2006). That is to say, the amplitude of the sure target input does not correspond to its physical properties (like the luminance) in the experiment; instead it is related to the behavioral significance of the sure target that a monkey learned, i.e., the amount of reward it received by choosing the sure target (Fig. 8).

In an FD version of the two-alternative direction-discrimination task (Fig. 1*A*), two directional targets, T_A (90°) and T_B (270°), are first presented to the network. A random-dot stimulus with net motion to T_A is presented at 500 ms after target onset. The difficulty of the task is modulated by the stimulus duration (randomly chosen from 110, 130, 152, 178, 208, 244, 289, 348, 439, and 627 ms) and the percentage of coherently moving dots (the motion strength). We modeled the external input to pyramidal cell i (at θ_i) as a sum of two target signals, $I^{tar}_i(t)$ ($tar = \{A, B\}$; Fig. 1*D*, black line), and the motion stimulus, $I^m_i(t)$ (Fig. 1*C*). The target inputs to T_A and T_B are identical:

$$I^{tar}_i(t) = I^{tar}(t)\exp[-(\theta_i - \theta_{tar})^2/2\sigma^2_{tar}]$$

where $\theta_A = 90°$; $\theta_B = 270°$; $\sigma_{tar} = 10°$.

$$I^{tar}(t) = I_1 + I_2\exp[-(t - t_d - 200)/\tau_d],\ t_d + 100 < t \le t_m + 80$$

$$I^{tar}(t) = I_3 + (I_1 - I_3)\exp[-(t - t_m - 80)/\tau_{ss}],\ t_m + 80 < t$$

where $t_d = 400$ ms and $t_m = 800$ ms are the onset times of targets and motion, respectively; $\tau_d = 500$ ms and $\tau_{ss} = 15$ ms are the time constants of the adaption and suppression, respectively; $I_1 = 250$ pA, $I_2 = 50$ pA, and $I_3 = 60$ pA. Specially, the target (motion, respectively) input onset time is 100 (200, respectively) ms after the target (motion, respectively) onset time, and the target input is suppressed by the motion stimulus with a latency of 80 ms (Roitman and Shadlen 2002); with the high intensities of the target inputs, winner-take-all competition between the two targets does not take place before the motion stimulus onset (Furman and Wang 2008; Liu and Wang 2008; Wong et al. 2007).

In simulation, motion input is modeled to imitate the neural response in the middle-temporal area (MT) to the random-dot stimuli. We constructed such a population activity as a Gaussian function with a tuning width independent of motion strength while motion presented ($t_m + 200 < t < t_{mo}$, t_{mo} is the moment of motion input offset)

$$I^m(i) = m_0 + coh\ \{-m_1 + m_2\exp[-(\theta_i - \theta_m)^2/2\sigma^2_m]\}$$

where the motion strength $0 \le coh \le 1$; net direction $\theta_m = 90°$; $\sigma_m = 40°$. We kept the activity normalized, i.e., $<I^m(i)> = m_0 = 4$ pA; $m_1 = 4.93$ pA; $m_2 = 25$ pA.

In trials with T_s ($\theta_s = 180°$), where there was the opt-out safe target presented (Fig. 1*D*, red line), we modeled its time-dependent current, $I^s(i)$, as:

$$I^s(i) = I^s(t)\exp[-(\theta_i - \theta_s)^2/2\sigma^2_{tar}]$$

which is added to the external input. We used:

$$I^s(t) = I_3 + I_4\exp[-(t - t_s - t_{mo})/\tau_s]$$

for $t > t_s + t_{mo}$, where t_s is T_s input onset time after the motion input offset, t_{mo}, with a latency of 100 ms to the network after T_s onset. In simulations, we used $t_s = 575$ ms; $\tau_s = 90$ ms; $I_4 = 240$ pA (see Fig. 8 for a discussion on the choice of I_4), expect Fig. 5 (t_s is equal to 575, 750, or 925 ms).

The network model is taken from Furman and Wang (2008), with a few parameter changes, i.e., background noise that ensures a choice is generated even if the motion strength is weak and stimulus duration is short, and a different set of parameter values for the choice target input, motion input, and sure target input that are adopted to the new experimental protocol of the Kiani-Shadlen experiment. Although the network was not originally designed for a confidence-estimation experiment, unexpectedly it can reproduce the behavioral and neurophysiological observations that are similar to those in the Kiani-Shadlen experiment [Figs. 1, *B* and *C*, and 2, *A* and *B*; Fig. 5, *B* and *C* in Kiani-Shadlen study (2009)]: *1*) neurons inside T_A and T_B display firing activities with small differences when T_s is chosen, while their firing activity is divergent when T_s is shown but waived [Fig. 2, *B* and *D*, *top*, vs. Fig. 2, *A* and *B*, in Kiani and Shadlen (2009)]; *2*) neurons inside T_s response field have weak and uninformative spontaneous activity before T_s onset and then exhibit a fast ramping followed by a decay after T_s onset [Fig. 2*D*, *bottom*, vs. Fig. 5, *B* and *C*, in Kiani and Shadlen (2009)]. In the model, we assumed that the input strength of each target goes to the same level attributable to the adaptation. For simplicity, we fixed I_4, while adjusting τ_s to match a majority of points on the performance curves of P_{sure} from Kiani and Shadlen (2009) [Fig. 3*B*, vs. Fig. 1, *B* and *C*, in Kiani and Shadlen (2009)].

We also studied the choice confidence in a reaction time (RT) version of the task. In this task, a network can generate choice at any time after the motion onset and, at the same time, report directly its choice confidence. We followed the same simulation protocols as those in the FD task without T_s, except that the motion input is terminated when one of the activity bumps crosses the decision threshold, 60 Hz for at least 50 ms. We measured the corresponding time, t_r, and calculated RT as $t_R = t_r - t_m + 80$, where 80 ms is the latency period for implementation of saccadic eye movement (Roitman and Shadlen 2002). In the FD version of the task, an initial choice is assumed to be made when one of the two competing neural pools reaches a decision threshold of 50 Hz for at least 50 ms after motion onset because the decision threshold of an FD task was experimentally observed lower than that of an RT task (Roitman and Shadlen 2002).

In FD tasks, we performed 1,500 trials at each motion strength and stimulus duration level, where T_s was not presented, and 3,500 trials at each motion strength and stimulus duration level, where T_s was presented. In Fig. 5, we simulated 1,500 trials for each condition. In RT task, we carried out 3,000 trials for each motion strength level. All the simulated behavioral data reported were computed using all trials for each simulation set. The integration method was a modified second-order Runge-Kutta algorithm with firing-time interpolation (Hansel et al. 1998) and a time step $dt = 0.02$ ms.

Measurements of activity trajectories. We calculated the average response of the population of units associated with targets T_A, T_B, and T_s, namely R_A, R_B, and R_s, as the average firing activity of the neurons within 8.4° around each target with a time window of 100 ms preceding the time point (e.g., the moment of decision and onset of T_s) for R_A, R_B, and R_s in analysis, except for Figs. 2, *B* and *D*, and 7*A*. In Figs. 2, *B*, *D*, *E*, *F*, *G*, and *H*, and 7*A*, each trajectory or point represents the activity of a single neuron at each target. We applied a 100-ms Gaussian sliding window to smooth the peristimulus time histograms for the temporal evolutions of the firing rates of R_A, R_B, and R_s in Figs. 2, *B*, *D*, *E*, *F*, *G*, and *H*, and 7*A*.

Choice confidence assessment. In the monkey experiment, as well as in our model simulations, the introduction of a sure target only serves as a probe examining the confidence of the system. That is to say, with carefully choosing the ratio of sure target reward to that of

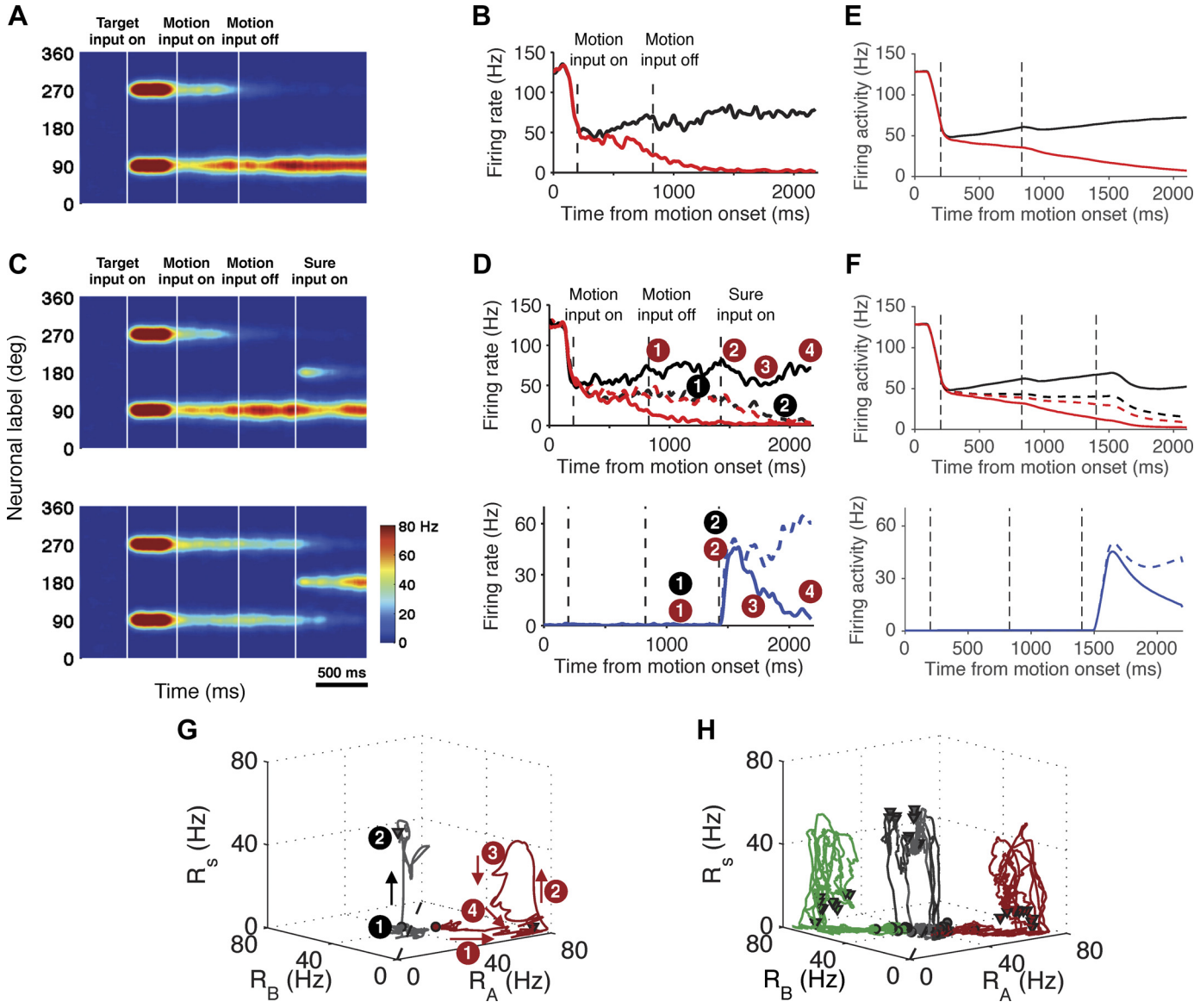

Fig. 2. Neuronal activity of sample trials at 0 motion strength. Spatiotemporal activity pattern of pyramidal cells in trials when the motion strength is 0. *A*: sample trial where T_s was not presented (stimulus duration is 627 ms). Neural pools centered around the 2 directional targets eventually diverge from each other; that near T_A wins the competition, and its activity persists during the delay in the form of a bell-shaped "bump attractor." *B*: average firing rate of the neural pools at T_A (black line) and T_B (red line) of the trial in *A*. *C*: 2 sample trials where T_s was presented (stimulus duration is 627 ms). *Top*: sure target induced a transient response that was suppressed because of feedback inhibition within the circuit, and the neural pool at T_A preserves similar activity to that in *A*; therefore T_s was waived. *Bottom*: neural pool around T_s fires at a sufficiently high rate that it overcomes the competition with the other neural pools, which in turn is suppressed by feedback inhibition; therefore T_s was selected. Note that in this trial the neural activities of 2 competing bumps are indistinguishable before T_s onset and gradually decay to a low level after T_s onset. *D*: neural activities at T_A (black lines), T_B (red lines), and T_s (blue lines) of the trials in *C*. Dashed curves: the trial where T_s was selected; solid curves: the trial where T_s was waived. Note that the stimulus condition was identical for the 2 sample trials; whether the sure target was chosen or waived was completely determined by network dynamics that fluctuated from trial to trial. *E* and *F*: average activities of R_A, R_B, and R_s across different motion strengths (100 trials for each motion strength), which follow the same conventions as those in *B* and *D*. *G* and *H*: network dynamics underlying a trial-by-trial variation of choice in a 3D (R_A, R_B, R_s) decision state space. *G*: neural activity trajectories of the 2 sampling trials in *C* from 150 ms after motion onset, the starting (end, respectively) points of which are marked by circles (triangles, respectively); time sequence of the trial selecting T_s (black line, 1–2 steps) follows that network walks around $R_A = R_B$ before T_s onset (Step 1 in *D*; black circles) and then converges to R_s after T_s onset (Step 2 in *D*; black circles); time sequence of the trial waiving T_s (red line, 1–4 steps) follows that network goes toward R_A before T_s onset (Step 1 in *D*; red circles) and then walks along R_s direction after T_s onset (Step 2 in *D*; red circles) and then converges back to R_A again (Step 3–4 in *D*; red circles). *H*: neural activity trajectories of the other 18 sampling trials at the same stimulus condition. For the trials waiving T_s, the network first converges to a choice attractor T_A (red lines) or T_B (green lines) preceding T_s onset; it then moves along the direction parallel to R_s axis because of the presentation of T_s and finally converges back to the initial choice attractor. For the trials choosing T_s (gray lines), the network first walks around the diagonal line $R_A = R_B$ ($R_s \approx 0$) and then converges to the sure attractor T_s after its presentation. Neural dynamics therefore acts as a 3-way competition.

choice targets (i.e., I_4 in our simulation), one can access the choice confidence across trials. In our simulation, the probability of opting for the sure target is bounded (Fig. 3*B*); it thus represents a good choice for estimating confidence. Furthermore, we will show later in RESULTS that the probability of choosing the sure target, P_{sure}, reflects the uncertainty of a choice in an opt-out task. Here, we defined choice confidence, *cc*, as the probability of waiving a sure target, i.e., $cc = 1 - P_{sure}$, using the trials in which T_s is presented (binned by 0.5 Hz;

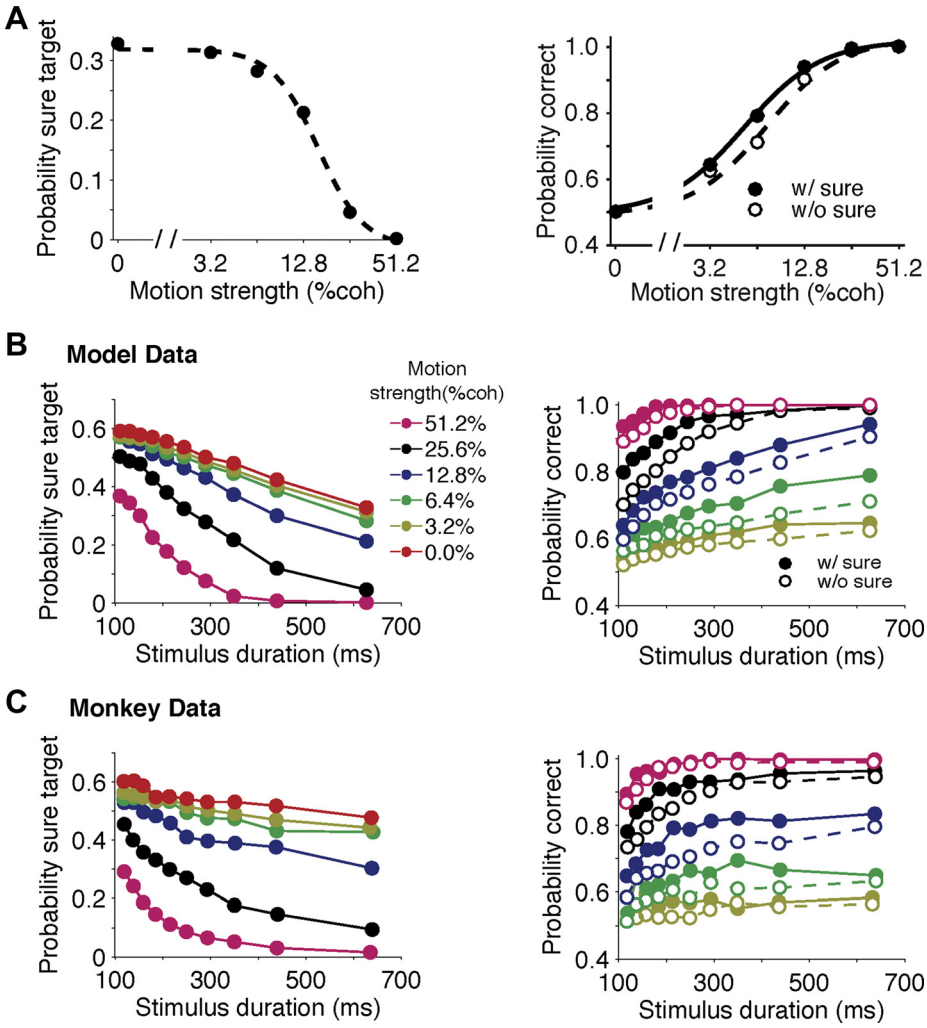

Fig. 3. Behavioral performance. *A*: model performance (at a fixed stimulus duration of 627 ms). *Left*: probability of choosing T_s (P_{sure}) decreases as a sigmoid function of the motion strength. *Right*: accuracy in trials where T_s is not shown ($P_{correct}$) increases as a sigmoid function of the motion strength (dashed black curve), and it is improved in trials when T_s was shown but waived (solid black curve). *B*: at different stimulus durations, P_{sure} decreases with motion strength and stimulus duration; $P_{correct}$ is higher in trials where T_s was shown but waived (solid lines, filled circles) than that where T_s was not shown (dashed lines, open circles). *C*: behavioral data from Kiani and Shadlen (2009) task using awake monkeys. Comparing *B* with *C*, model reproduces salient behavioral observations from the monkey experiment. Experimental data adapted with permission from Kiani and Shadlen (2009).

Fig. 6*A*, black circles). We also assumed that this probability can be predicted as a function of the differential activity $|R_A - R_B|$. We then performed the fit of a logistic function between $|R_A - R_B|$ and $cc = 1 - P_{sure}$. At each differential activity level, i, $cc^i = 1 - P_{sure}$ is computed as the mean of the decision result across the sample trials, k, $<s^i_k>$ for T_s. The decision result, s, is a binary variable, i.e., $s = 1$, if T_s is waived; $s = 0$, if T_s is chosen. To perform the fit, we used the firing activity within a 100-ms time window before T_s onset in the FD task (Fig. 6*A*) for R_A and R_B:

$$cc^i = 1 - P_{sure}(|R^i_A - R^i_B|) = b_1 + a[1 + \exp(k|R^i_A - R^i_B| - b_0)]^{-1}$$

Using all trials in FD task, we obtained $b_0 = 2.22$ Hz; $b_1 = 1.01$; $a = -1.01$; $k = 0.089$ ($R^2 = 0.98$, Fig. 6*A*, red dashed line). Importantly, a real result here is to quantify confidence as a function of the neural activity. Confidence estimation is thus applicable to all trials, even without sure target presentation. We then used these estimated parameters to calculate cc^i for each sample trial in both FD and RT tasks, where R_A and R_B are the average firing rates within a 100-ms time window before T_s onset in FD task (Figs. 6, 7*A*, and 8*C*) and those before one of the bumps reaching a decision threshold in RT task (Fig. 9).

RESULTS

We performed computer simulations of the Kiani-Shadlen task (Fig. 1*A*), using a neuronal circuit model (Furman and Wang 2008). In this model, the pyramidal cells are selective for motion direction as an analog stimulus feature and are uniformly distributed along a ring according to their preferred directions. Pyramidal cells are endowed with strong recurrent excitation, which is balanced by feedback inhibition mediated by interneurons (Fig. 1*B*). We assumed that the neural representation of motion stimulus in MT exhibits normalization (Heeger 1992; Treue et al. 2000) (Fig. 1*C*). The output from MT converges with other visual inputs such as choice targets to the decision circuit. Without loss of generality, we placed the choice targets T_A at 90°, T_B at 270°, and T_s at 180° (Fig. 1*D*). In a short delay after the target onset, a random-dots motion stimulus is presented, and, in our model, the network integrates the motion signal gradually over time. If a categorical choice is formed through attractor dynamics (Furman and Wang 2008; Wang 2002, 2008; Wong and Wang 2006), that choice is maintained in the form of a persistent activity pattern during a delay period. Because the network dynamics is stochastic, a decision may not be reached during stimulus presentation. Network activity continues to evolve during the delay through slow NMDA-mediated reverberation, and this process can be altered in the event that T_s is presented and this third option becomes available.

Network dynamics in an FD task with PDW at zero motion strength. In experiments using single-unit recording, each neuron was recorded at a time, and its selectivity and dynamics were evaluated across trials, whereas, in our model, all neurons are monitored simultaneously in a single trial. At the popula-

tion level, the ramping activity is demonstrated as the gradual development of a bell-shaped activity pattern (bump) around the direction of a selected target. The stimulated neural dynamics in Fig. 2 can be compared directly with single-neuron data from area LIP for a neuron with the preferred direction at T_A; T_A and T_B are equivalent to T_{in} and T_{opp} in the Kiani-Shadlen experiment (2009).

Figure 2*A* shows the spatiotemporal spiking activity of the network model in a trial without T_s presented. Although the input is identical to all the pyramidal cells at zero motion strength, the two activity bumps compete with each other through shared inhibitory feedback and stochastic recurrent dynamics. Eventually one bump ramps up, while the other one decays, leading to a categorical choice (Furman and Wang 2008; Wang 2002; Wong et al. 2007). The ramping-up bump is maintained by the persistent activity in delay (Compte et al. 2000). At the single-unit level, the firing activities of neurons at T_A and T_B (R_A and R_B, respectively) diverge over time after motion onset, leading to the choice of the network (T_A). The winning neural pool persists its firing activity until the end of delay (Fig. 2*B*).

In trials when T_s is presented (Fig. 2, *C* and *D*), again the attractor dynamics dictates the choice behavior of the network, and we identified a choice of T_s by the firing activity of neurons at T_s (R_s); the network selects T_s (Fig. 2*D*, dashed blue line) if R_s persists at a high rate, and it waives T_s (Fig. 2*D*, solid blue line) if R_s decays to a low rate. In the trial where T_s is waived, we observed the same divergence and persistent activity of R_A (choice) and R_B as that without T_s presented (Fig. 2, *C*, *top*, and *D*, solid lines); however, if the network selects T_s, R_A and R_B are indistinguishable at some intermediate rates without significant divergence (Fig. 2*C*, *bottom*, and *D*, dashed lines; $R_A = 32.5$ Hz, $R_B = 33.6$ Hz at T_s input onset), and then both decay to low rates as neurons around T_s win the competition. The average firing activities of R_A, R_B, and R_s across different motion strengths (100 trials for each condition) are shown in Fig. 2, *E* and *F*, which follow the same conventions as those in Fig. 2, *B* and *D*. All these are similar to the observed LIP neuronal data (Kiani and Shadlen 2009) from motion input onset time to the end of trial, which is crucial to capture the neural dynamics of choosing or waiving T_s, although a difference of neuronal activities could exist before motion input onset in our model compared to the experimental observations, which is not important to predict a choice of T_A, T_B, or T_s.

We further studied the neural dynamics that underlies a choice of the network among T_A, T_B, and T_s. Taking trials in Fig. 2*C*, for example, we visualized the neural trajectories in a 3D decision space (R_A, R_B, R_s) following the sequences marked in Fig. 2*D*. In the trial where T_s is waived (Fig. 2*G*, red line), the network first converges to a choice attractor (T_A), then leaves away from and returns to it again after the presentation of T_s; in the trial where T_s is selected (Fig. 2*G*, gray line), the network wanders around the diagonal line ($R_A = R_B$, $R_s \approx 0$) and then converges to the attractor T_s. These trajectories in decision space imply that the presentation of T_s could induce a sure attractor, which behaves similarly to choice attractors; the network could thus act like a three-way competition after the presentation of T_s. To test this, we visualized more sampling trials at the same stimulus condition and explored the basin of attraction for each attractor (Fig. 2*H*). We found that, for trials choosing T_A, T_B, and T_s (Fig. 2*H*, red, green, and gray lines, respectively), the networks converge to the choice attractors T_A (near the R_A axis) and T_B (near the R_B axis) and the sure attractor T_s (near the R_s axis), respectively. The whole decision space is thereby separated into three attractor regions, and, because of this structure of decision space, the location of each neural trajectory at the moment of T_s onset (namely the initial location) can potentially predict the choice of a network (Kiani and Shadlen 2009). Particularly, if the initial location of a network is close to a choice attractor, it would eventually converge back to that choice attractor again after T_s onset, whereas, if its initial location is around the diagonal line, it more likely converges to a sure attractor. Notably, if the initial location of a network is between the diagonal line and a choice attractor, it could either continue converging to that choice attractor or change its mind to T_s (discussed later in Fig. 7). Therefore, our model demonstrates that a categorical choice of a network in this task could be generated by a three-way competition among attractors T_A, T_B, and T_s, which relies internally on the stochastic neural dynamics (Wang 2008).

Behavioral performance. The performance of the model is quantified by the fraction of trials corresponding to a particular behavioral response. Figure 3 exhibits the probability of choosing T_s (P_{sure}) and accuracy ($P_{correct}$) for trials when T_s is not presented or T_s is shown but waived. At a fixed stimulus duration, our model shows that P_{sure} decreases while $P_{correct}$ increases with the motion strength; $P_{correct}$ improves in trials where T_s was shown but waived (Fig. 3*A*). Moreover, Fig. 3*B* shows that P_{sure} decreases with the stimulus duration, wherefore the network selects T_s more often for weaker motion strength or shorter duration; $P_{correct}$ increases monotonically with stimulus duration for trials with or without T_s presented (Fig. 3*B*, *right*, solid and dashed lines, respectively). For a given motion strength and stimulus duration, $P_{correct}$ is higher for trials where T_s is shown but waived than those without T_s presented, implying that P_{sure} is a probe of uncertainty (Kepecs and Mainen 2012; Whiteley and Sahani 2008). In conclusion, the model successfully reproduces the salient behavioral observations in the monkey experiment (Kiani and Shadlen 2009) (Fig. 3*C*).

Choice confidence as a logistic function of the differential activity. Consistent with the neurophysiological observation (Kiani and Shadlen 2009), R_A and R_B undergo brief decreases after T_s onset in our model (Fig. 2*C*). This is because T_s stimulates neurons selective for the sure target, and their increased firing activity recruits more feedback inhibition that reduces R_A and R_B in a three-way competition (Fig. 2*H*). We therefore hypothesized that the network would opt to T_s if it has not converged to a stable attractor for T_A ($R_A >> R_B$) or T_B ($R_A << R_B$); the differential activity $|R_A - R_B|$ at T_s onset determines P_{sure}.

We first observed that R_A and R_B could diverge in the late phase of delay. The long divergent time ($>T_s$ onset) implies that there is a quasistable state at $R_A = R_B$, around which the network could wander, but eventually the network would escape from it and converge to a stable attractor, generating a choice. We next studied whether the network in a state around $R_A = R_B$ would opt to T_s (e.g., Fig. 2*H*, gray lines). For individual sample trials, we visualized R_A and R_B against each other on a decision space. Figure 4, *A* and *B*, shows that the activity of the network falls down along a diagonal line and then separates into three groups for choices of T_A, T_B, and T_s.

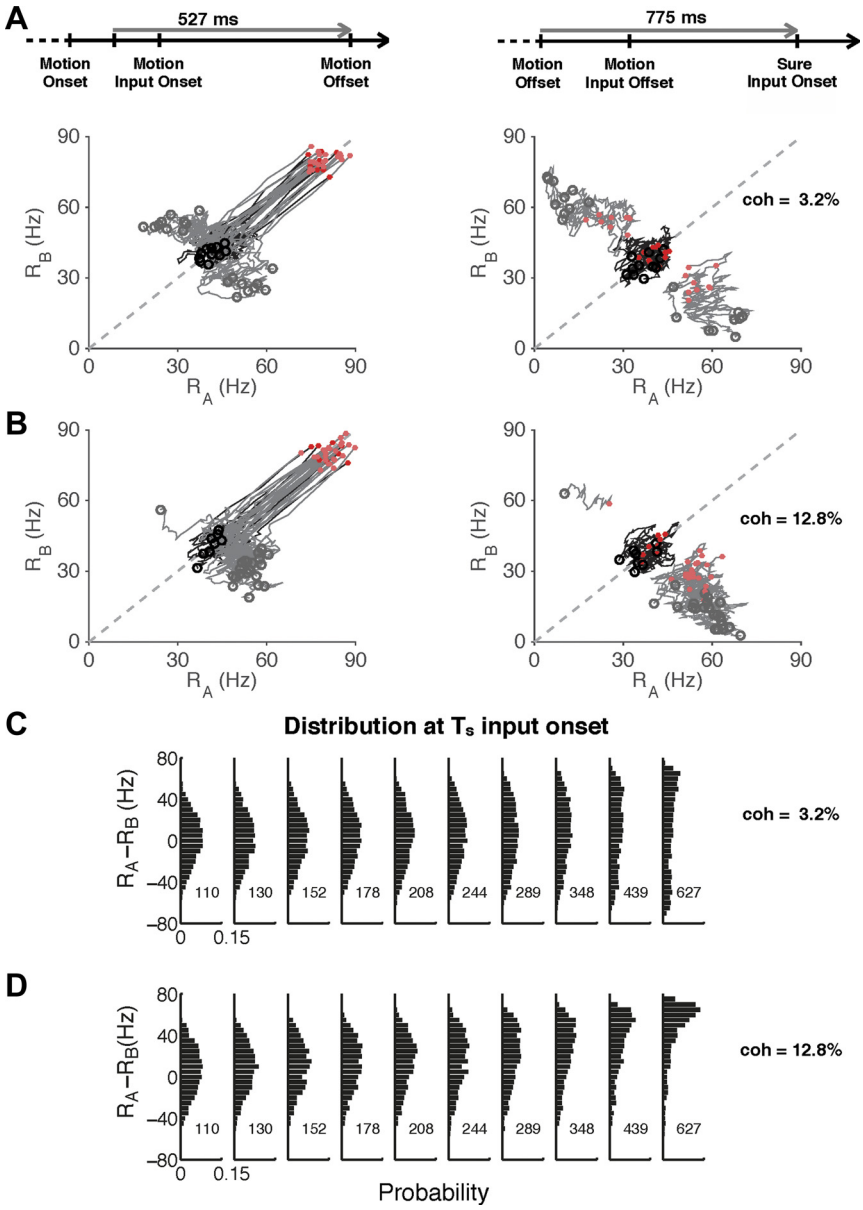

Fig. 4. Differential activity $|R_A - R_B|$ determines whether T_s is waived. *A* and *B*: single-trial dynamics of the network in the decision state space, where the population firing rates R_A and R_B are plotted against each other (the starting point of each network trajectory is marked as a red circle, and ending point is marked as an open circle). Gray: trials when T_s was waived; black: trials when T_s was selected. The dynamical trajectories are shown from 100 ms after motion onset to its offset (*left*), then to T_s input onset (*right*), at different motion strengths (*A*: 3.2%; *B*: 12.8%; stimulus duration: 627 ms). At the onset of motion stimulus, both R_A and R_B are high (~90 Hz), near the diagonal line, because of the presentation of directional targets. The population dynamics first decays along the diagonal line, induced by a suppression of target inputs after motion onset. In trials when T_s was waived, the network trajectory converges to 1 of 2 target attractors (where R_A is high and R_B is low, or vice versa), whereas, in trials when T_s was selected, the population dynamics continues to wander randomly around the diagonal line. The absolute value of differential activity at T_s onset therefore determines whether T_s is waived. *C* and *D*: distribution of $R_A - R_B$ at T_s onset is a function of the motion strength (*C*: 3.2%; *D*: 12.8%) and stimulus duration (presented in each column), where the percentage of trials around 0 decreases with the motion strength and stimulus duration. This explains why P_{sure} decreases with the motion strength and stimulus duration.

In trials where T_s is waived, the network converges to one of two stable attractors, T_A or T_B (Fig. 4, *A* and *B*, gray lines), while, in those where T_s is selected, the network walks randomly around the quasistable state at $R_A = R_B \approx 35$ Hz (Fig. 4, *A* and *B*, black lines). In summary, once the network converges to a stable attractor before T_s onset, T_s is waived; if it wanders around $R_A = R_B$, the network is likely to opt to T_s.

The studies of the similar attractor models (Wang 2002; Wong and Wang 2006; Wong et al. 2007) showed that early divergence of R_A and R_B (bias to one attractor) determines the probability of choosing T_A and T_B as a function of the motion strength and stimulus duration. One can thus expect that early divergence would also result in a decrease of P_{sure} as a function of the motion strength and stimulus duration. To examine this, we investigated the distributions of $R_A - R_B$ with different motion strengths at T_s input onset. Figure 4, *C* and *D*, shows that the percentage of the trials around $R_A = R_B$ decreases with higher motion strength or longer stimulus duration, resulting in a decrease of P_{sure}.

Although the early divergence plays a dominant role in the network dynamics, network continues to evolve via NMDA-mediated reverberatory dynamics; the slow stochastic dynamics could thus drive the network away from $R_A = R_B$ in the later phase of the delay. Consequently, P_{sure} depends on T_s onset time. Figure 5*A* displays the evolution of the distribution of $R_A - R_B$ at different times from motion offset, demonstrating that the slow stochastic dynamics also plays an essential role in the behavior of the network. Across trials, our model predicts that P_{sure} decreases with longer T_s onset times (Fig. 5*B*) because more trials settle down to a stable attractor later in the delay, i.e., the percentage of trials with $R_A = R_B$ decreases with T_s input onset time. Interestingly, we found that $P_{correct}$ was nearly constant with different T_s input onset times (Fig. 5*C*). This happens because $P_{correct} \approx r_{AB}/(1 + r_{AB})$, where r_{AB} is the ratio of the number of trials at attractor T_A (choice) to that at attractor T_B at T_s onset and is saturated at 575 ms after motion offset (Fig. 5*A*). In short, P_{sure} is directly related to the percentage of trials around $R_A = R_B$, whereas $P_{correct}$ is

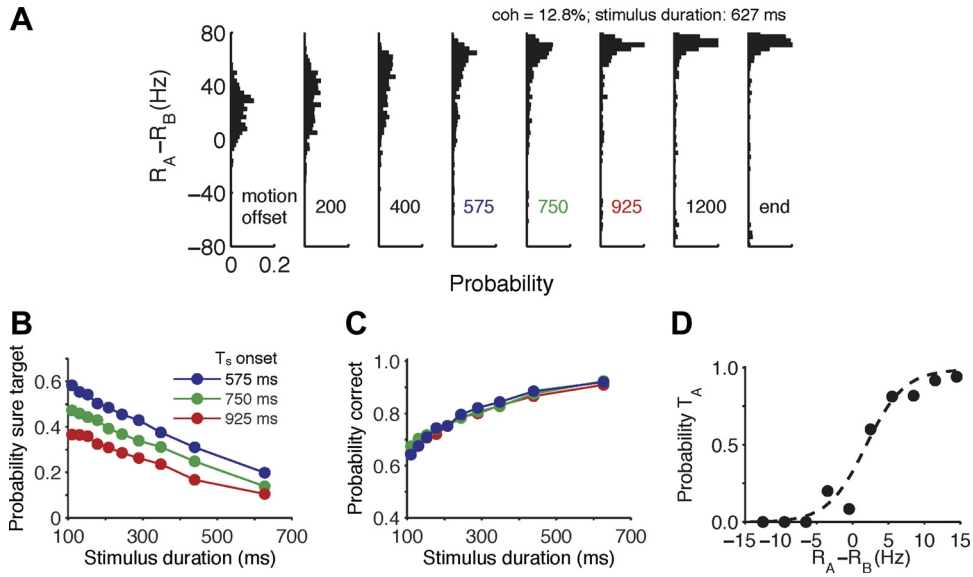

Fig. 5. Onset time of T_s determines the probability of choosing T_s but has little impact on accuracy. *A*: at a fixed motion strength and stimulus duration (12.8% and 627 ms, respectively), $R_A - R_B$ continues to change after motion offset (time presented in each column is relative to motion input offset time) and is settled down only until the late phase of delay (>1,200 ms in simulation). *B* and *C*: P_{sure} decreases as a function of T_s input onset time (575 ms: blue; 750 ms: green; 925 ms: red), while $P_{correct}$ remains unaffected. *D*: probability of choosing T_A (at the end of the delay) depends on the differential activity, $|R_A - R_B|$, at T_s input onset time (filled circles: simulation data from *B* and *C* where $coh = 12.8\%$ and motion direction toward T_A; dashed line: logistic function fit). When $|R_A - R_B|$ is large, the sign of $R_A - R_B$ determines the choice at T_s onset, i.e., positive for T_A and negative for T_B. If $|R_A - R_B|$ is small ($R_A - R_B$ from −5 Hz to 5 Hz), the probability of choosing T_A increases with $R_A - R_B$. Data in this figure are composed of those at all T_s onset times.

associated with the number of trials at T_A and T_B at T_s onset. In this sense, there is a dissociation of confidence estimation from performance (Graziano and Sigman 2009; Graziano et al. 2010). Moreover, without T_s presented, the network continues to converge to T_A or T_B via stochastic dynamics, and the probability to one of them is biased and relies on $R_A - R_B$ in the early phase of delay (Fig. 5*D*). In conclusion, we found that $|R_A - R_B|$ at the moment of T_s input onset determines P_{sure} probabilistically and reflects a degree of the stability of a choice: if a categorical choice is achieved but with small $|R_A - R_B|$, it could be altered to T_s; whereas, if $|R_A - R_B|$ is large, the choice of the network would not be changed by T_s.

Here we define choice confidence as a function of the instantaneous differential activity $|R_A - R_B|$ at sure target input onset time for each trial i, i.e.,

$$cc^i = f(|R^i{}_A - R^i{}_B|).$$

In our model, $|R_A - R_B|$ shows the position of a network in the (R_A,R_B) plane (Fig. 2, *G* and *H*) related to choice attractors in the decision space, i.e., the larger $|R_A - R_B|$ is, the closer a network is to a choice attractor T_A or T_B; $f(\cdot)$ is therefore required to be an increasing function. In the previous studies (Beck et al. 2008; Kepecs and Mainen 2012; Kepecs et al. 2008; Vickers 1979), functions $f(\cdot)$ were given in a variety of ways. One can picture that, as long as $f(\cdot)$ is a monotonic function, we can always equate $f(\cdot)$ from one model to another. Of note, our definition of the confidence stems from the structure of the attractor basin in decision space (Fig. 2, *G* and *H*), i.e., if a choice is confident, then it is easily resistant to the other external input such as a sure target, whereas confidence estimations from models like Beck et al. (2008) and Kepecs and Mainen (2012) are compared directly to log odds of a choice in Bayesian framework. In the studies of Beck et al. (2008), they found from both the experimental data and their model that log odds of choice at A, namely confidence across trials for choice A, is proportionate to $<R_A> - <R_B>$, where $<\cdot>$ is the average across trials. However, such a read-out of confidence would predict a strong correlation between confidence and performance on single trials, which is somewhat inconsistent with experimental observation of a broad performance variation in different confidence categories (Graziano et al. 2010; Juslin and Olsson 1997). Second, this "optimal decoder" $<R_A> - <R_B>$ relies explicitly on the equal variance hypothesis for likelihood (Kepecs and Mainen 2012) or "left-right" symmetry of a linear decoder (Beck et al. 2008). It remains unclear what the biologically plausible mechanisms are to achieve such a fine-tuned neural circuit to compute confidence signals in these models.

In our model, confidence is defined using a monotonic increasing function of $|R_A - R_B|$. Particularly in an opt-out task, confidence can be probed by the probability that a choice stays in its attractor after presenting a sure target. If the choice is confident at $|R_A - R_B|$, then this probability, $1 - P_{sure}$, is low. Using this probe, we found that choice confidence increases as a sigmoid function [i.e., function $f(\cdot)$] of $|R_A - R_B|$ (Fig. 6*A*). Next, we asked whether, across trials, our definition of confidence can also reflect probability of correct at each $|R_A - R_B|$ level as those defined in Bayesian framework (Beck et al. 2008; Kepecs and Mainen 2012; Kepecs et al. 2008). This seems possible as indicated from Fig. 5*D*. Moreover, a detailed analysis should compare probability of correct choice and confidence simultaneously. We performed this analysis using trials without presenting T_s, where we computed both the performance and the confidence at the end of decision averaged across all stimulus conditions (Fig. 6*B*). Figure 6*B* demonstrates that confidence in our model increases monotonically with the performance. Importantly, our model indicates that confidence can be computed as a function of the instantaneous neural activities (like a population code; Beck et al. 2008) at

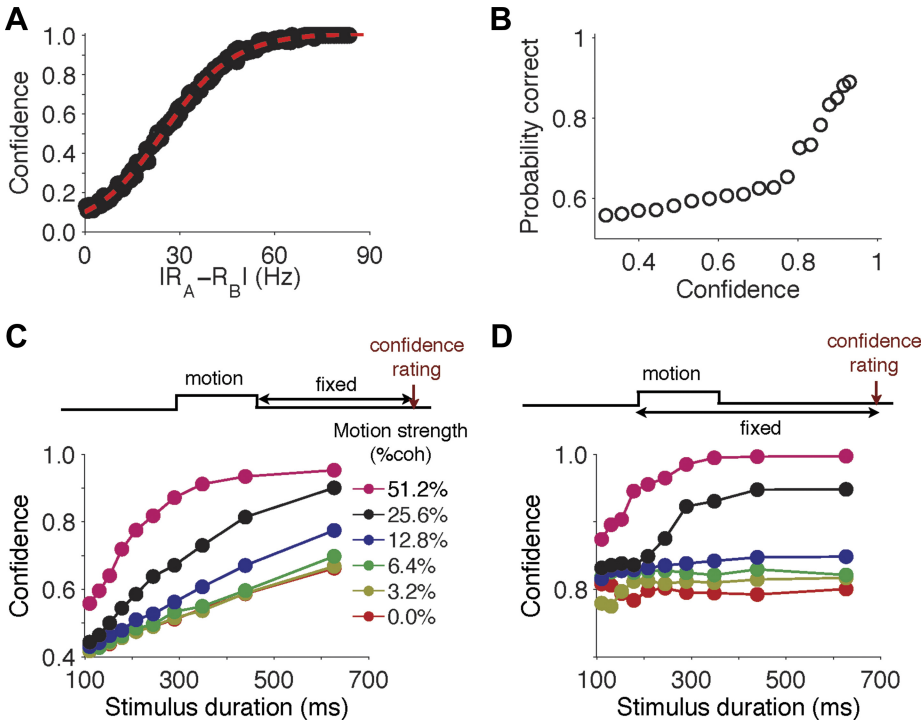

Fig. 6. The probability of waiving T_s reflects choice confidence. *A*: confidence is defined as the probability of waiving T_s at each $|R_A - R_B|$, i.e. the differential activity of 2 competing bumps at the moment of the sure target input onset, in single trials. A logistic function fit (red dash line) is performed on the data from all computer simulations with T_s presented. *B*: comparison of probability of correctness and confidence at each $|R_A - R_B|$ level in single trials. Both confidence and probability of correctness at each $|R_A - R_B|$ level in single trials are computed at decision time of trials without T_s presentation. Probability of correctness increases as a monotonic function of confidence, which implies that confidence in our model would also be a good measurement of the subjective correct rate or log odds of choice. *C*: confidence assessment at T_s input onset (the duration from motion input offset to the time of confidence estimation is fixed, i.e. 575 ms, *top*) increases with the motion strength and stimulus duration. *D*: confidence assessment at an identical time after the motion input onset (the duration from motion onset to confidence estimation is fixed, i.e. 1,550 ms, *top*) saturates after a short period of stimulus duration. In this case, early evidence plays a dominant role in confidence estimation.

any time in a decision circuit without explicit use of elapsed time for integration of the sample (Drugowitsch et al. 2012; Kiani and Shadlen 2009; Kiani et al. 2014; Moreno-Bote 2010). Therefore, although confidence in our model is not defined as a log odds function of the choice (Beck et al. 2008; Drugowitsch et al. 2012; Kepecs and Mainen 2012; Kepecs et al. 2008; Kiani and Shadlen 2009; Kiani et al. 2014), confidence can be a good measurement of the subjective correct rate across trials. Importantly, on single trials, choice confidence in our model is dissociable from performance (Graziano et al. 2010), whereas Bayesian models would predict a strong correlation.

Despite the similarity of $f(\cdot)$, choice confidence in our model is, however, conceptually distinct from those from Bayesian decision models because our definition of confidence fundamentally comes from the structure of the attractor basin in the decision space. Therefore, our model predicts that confidence would be different when estimated at the different times after motion offsets (Fig. 5), and that would be nearly the same in a Bayesian model (Beck et al. 2008; Drugowitsch et al. 2012; Kiani and Shadlen 2009; Moreno-Bote 2010). To test this, we estimated choice confidence using neural activities R_A and R_B in trials without T_s presented at different times after motion offsets. We first estimated the choice confidence at 575 ms after motion input offset (Fig. 6*C*; compared directly with Fig. 3*B*, *left*), where the distribution of $R_A - R_B$ is still evolving, namely the confidence estimation after a short delay (Fig. 5*A*). One can thus expect an increase of choice confidence in trials with longer stimulus durations (Fig. 6*C*), according to the variation of the bimodal distribution of $R_A - R_B$ at different stimulus durations (Fig. 4, *C* and *D*). We next estimated the choice confidence at 1,550 ms after motion input onset, i.e., the same time of a trial, where the internal noise is nearly identical at different stimulus conditions and the strength of the input dominates the choice confidence of the network. In this case, one would expect that confidence should increase as a function with the motion strength and stimulus duration for a noiseless integrator (Beck et al. 2008), unless it is bounded (Drugowitsch et al. 2012; Kiani and Shadlen 2009). While in our model, the attractor dynamics implies that the bimodal distribution of $R_A - R_B$ depends dominantly on the early divergence (Fig. 4, *C* and *D*). As a result, Fig. 6*D* shows that all confidences saturate at stimulus duration >400 ms, suggesting that the early evidence has the greatest effect on confidence estimation. Of note, the saturation time prolongs with lower motion strength.

Low confidence results in changes of mind to T_s. The whole dynamic space, R_A-over-R_B decision space, can be divided into three regions: choice attractor regions ($R_A >> R_B$ or $R_A << R_B$) and an unstable region in between them (Wang 2008). In the previous study, we focused on the trials along the diagonal line ($R_A = R_B$), where a choice of network remains undecided before the presentation of T_s. We then investigated the dynamics of networks that are between the diagonal lines ($R_A = R_B$) and a choice attractor ($R_A >> R_B$ or $R_A << R_B$) on R_A-over-R_B decision space preceding the presentation of T_s, where a trial could be still in the unstable region (and thus could go to the sure attractor after T_s onset) or in a stable attractor region (where the network trajectory would stay in the same choice attractor even after T_s onset). For these trials, we could define an initial choice of the network by its nearby choice attractor T_A (T_B, respectively), where R_A (R_B, respectively) fires above a decision threshold (>50 Hz). Particularly, we explored under which condition the network would more likely continue converging to the attractor of its initial choice, or shift to the sure attractor, when T_s is offered.

Figure 7*A* compares the neural activity in trials with low- and high-confidence initial choices, the analysis of which is on single trials and is missing in Kiani and Shadlen (2009). In low-confidence trials (Fig. 7*A*, *top*), only one of the activity

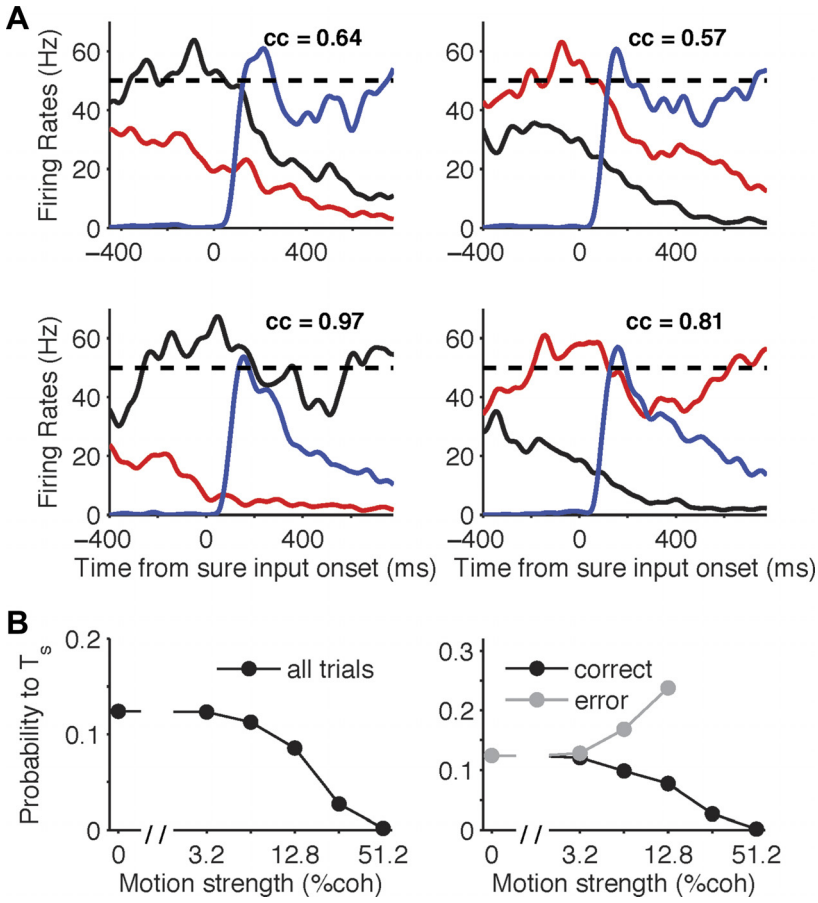

Fig. 7. Low confidence results in changes of mind to T_s. *A*: trials with low confidence exhibit changes of mind to T_s (motion strength: 6.4%; stimulus duration: 243 ms). *Top*: sample trials with low confidence, small $|R_A - R_B|$. Even though the network has reached 1 of the 2 choice attractors [*left*: T_A (black lines); *right*: T_B (red lines)], upon the presentation of T_s, the neural pool selective for T_s takes over (blue lines), so there are changes of mind. *Bottom*: sample trials with high confidence, large $|R_A - R_B|$. No changes of mind take place. Choice confidence, *cc*, for each trial is estimated at the time of T_s input onset and shown at the top of each panel. *B*, *left*: across trials (averaged over different stimulus durations), the probability of shifting to T_s decreases with the motion strength. *Right*: in error (correct) trials, this probability increases (decreases, respectively) with the motion strength.

bumps fires above the decision threshold, but the other remains similar ($|R_A - R_B|$ is small). After T_s onset, both R_A and R_B decay to a low level, while R_s grows to a high level and T_s is chosen. By contrast, in high-confidence trials (Fig. 7*A*, *bottom*), when one of the firing rates reaches a steady state, the other is much smaller ($|R_A - R_B|$ is large). Neurons activated by T_s are suppressed, and T_s is waived. In the latter case, the activity of the winning neural pool exhibits a brief dip upon T_s input onset and then ramps up again to its steady state.

Across trials, the probability of changes of mind to T_s is negatively correlated with choice confidence, i.e., the network exhibits low confidence in trials at low motion strength (Fig. 6*C*) and high probability of changes of mind to T_s (Fig. 7*B*, *left*). To further test whether the network bases the probability of changes of mind to T_s on its performance and confidence, we categorized the trials with initial choices, where either R_A or R_B reaches a decision threshold, 50 Hz (if both of them do not reach the decision threshold, we considered the choice remaining undecided at T_s onset), into correct and error groups and found that network changes its choice to T_s more often in error trials. Furthermore, the probability of choosing T_s in correct (error, respectively) choice decreases (increases, respectively) with the motion strength (Fig. 7*B*, *right*). This finding is reminiscent of the experimental observation that, in a decision making task with a delayed reward, animals moved back to self-restart port more often when the task became more difficult or the confidence was low (Kepecs et al. 2008).

In conclusion, we identified two possibilities for choosing T_s; either an initial choice was not made (along the diagonal line; Fig. 2), or it was made with a low confidence (between the diagonal line and choice attractors; Fig. 7*A*, *top*). For the latter case, $|R_A - R_B|$ reveals the confidence about an initial choice; low confidence of a choice is likely to result in changes of mind to T_s.

A sure target as a probe about the confidence of the system. The introduction of a sure target plays a role of probing the confidence of the system. Specifically, in the monkey experiment, the physical luminance of the sure target was the same as the choice targets. Monkeys were trained to understand what the sure target meant behaviorally, which depended on the amount of reward by choosing it. Therefore, in our model, the amplitude of the sure target input (I_4 in our model) does not correspond to its physical properties but is related to the behavioral significance of the sure target that a monkey learned as the amount of reward he receives by choosing the sure target. One can imagine that, if choosing the sure target yields a negligible (significant, respectively) amount of reward, monkeys would never (always, respectively) have learned to choose it. To test this, we studied the effect of T_s input strength, I_4, on the behavioral performance at a fixed motion strength level (i.e., 12.8%). We found that P_{sure} increases as a function of T_s input strength (Fig. 8*A*). When T_s input strength is low (e.g., I_4 = 192 pA), T_s is always waived; when T_s input strength is high (I_4 = 288 pA), T_s is mostly chosen, as stimulus duration is short. Moreover, in the trials where T_s is shown but waived, our simulation predicts an increase of correct rate at high T_s input strength (Fig. 8*B*). At the network level, these observations in Fig. 8, *A* and *B*, still follow a three-way

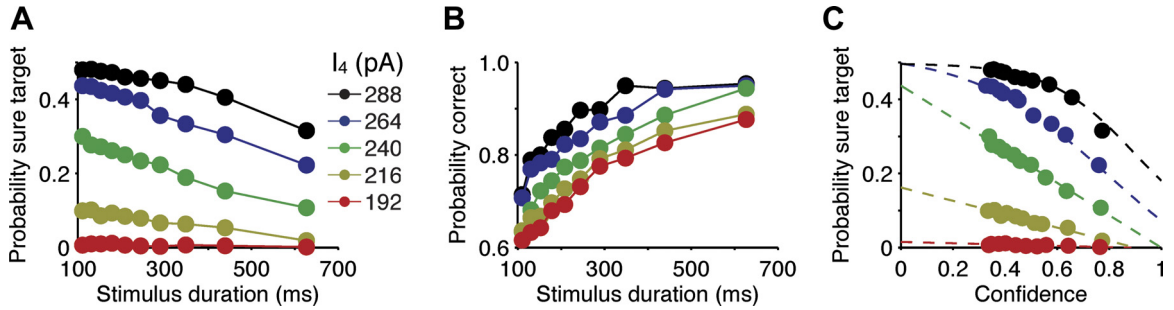

Fig. 8. Effect of T_s input strength on the behavioral performance. In this simulation, the motion strength is fixed at $coh = 12.8\%$, and T_s input strength at $I_4 =$ 240 pA (green circles and line) is the same as those used in Figures 2–7. *A*: P_{sure} increases as a function of T_s input strength. T_s is usually waived (chosen, respectively), when T_s input strength is weak (strong, respectively). *B*: correct rate in the trials, where T_s is waived, increases as a function of T_s input strength. *C*: choice confidence is identical at the moment of T_s input onset (which increases as a function of stimulus duration). For a range of T_s input strength (216 pA $< I_4 <$ 264 pA), P_{sure} decreases as a linear function of choice confidence.

competition among R_A, R_B, and R_S, e.g., when input of the sure target is weak (strong, respectively), it always behaves like a loser (winner, respectively). Last, we examined whether, in a range of T_s input strengths (a selected range of amount of T_s rewards), a sure target can serve as a probe about the confidence of the system, when the network applies the attractor dynamics. We assessed the choice confidence as a function of $|R_A - R_B|$ at the moment of T_s input onset for different choices of T_s input strengths. Figure 8*C* shows that, on average, choice confidence is identical for different T_s input strengths (and increases as a function of stimulus durations), and P_{sure} decreases as a linear function of choice confidence for a broad range of T_s input strengths, i.e., in this range, 216 pA $< I_4 <$ 264 pA, a sure target in our model can be considered as a probe about the confidence of the system. Therefore, a sure target is only a probe, and the confidence measure is valid even without it.

Assessment of choice confidence in an RT task. In our model, confidence can be read out at any time and increases as a function of stimulus duration in an FD task. One may thus argue that the network would exhibit high confidence despite the task difficulty if it freely controls the viewing duration of the stimulus. However, classical literature about confidence in cognitive psychology (Vickers 1979) emphasizes an inverse relationship between confidence and response time, which can be potentially tested in an RT version of discrimination task [developed previously by Furman and Wang (2008)] with direct assessment of choice confidence. This distinguishable difference between confidence estimation in FD vs. RT task in fact comes from two distinct processes; whereas longer viewing time in an FD task enables more integration of evidence (confidence thus increases with motion viewing time), a longer RT means a higher task difficulty in an RT task (confidence thus decreases instead with motion viewing time). We thus want to further test whether our model can nicely explain such a contrasting observation. To do this, we designed an RT version of discrimination task with direct assessment of choice confidence (Fig. 9*A*); the network integrates the motion input until the neurons selective for one of two alternatives fire above a decision threshold and reports the confidence as the function of the instantaneous $|R_A - R_B|$ (Fig. 6*A*) at the moment of choice [a similar human behavioral experiment is performed and reported by Kiani et al. (2014) recently].

Our model exhibits the typical psychometric and chronometric curves of a two-alternative discrimination task (Churchland et al. 2008; Roitman and Shadlen 2002), i.e., $P_{correct}$ increases, whereas RT decreases with the motion strength (Fig. 9, *B* and *C*). Importantly, weaker motion strengths are associated with longer RTs, where $|R_A - R_B|$ will be less at longer RTs. Choice confidence thus increases with the motion strength [Fig. 9*D*; see also Fig. 5 in Beck et al. (2008)] and is positively correlated with the behavioral performance across trials (Barthelmé and Mamassian 2010) (data not explicitly shown). We also found that choice confidence decreases as an inverse function of RT (Fig. 9*E*), which agrees broadly with the human behavioral observations (Vickers 1979). Although an erroneous choice could be associated with high confidence (Graziano and Sigman 2009), the average $|R_A - R_B|$ across trials is higher in correct trials than that in error ones (Wang 2002). Therefore, in our model, confidence increases (decreases, respectively) with motion strength in correct (error, respectively) trials (Fig. 9*F*), consistent with human studies (Pierrel and Murray 1963).

Moreover, we studied correlation between the choice confidence and decision accuracy. Figure 9, *B* and *D*, implies that choice confidence is positively correlated with behavioral performance across trials. Although confidence in our model does not directly represent a subjective estimation of performance (like that in Beck et al. 2008; Drugowitsch et al. 2012; Kiani and Shadlen 2009), one can estimate the subjective performance from choice confidence using a monotonic function, $g(\cdot)$, in Fig. 6*B*. We can thus compare directly our confidence score with performance to study the "hard-easy" effect (Juslin and Olsson 1997). Here we defined underconfidence score as the difference between the choice confidence and accuracy, $cc - P_{correct}$ [one can also use $g(cc) - P_{correct}$], and a hard-easy effect is the observation that the underconfidence score decreases as a function of task difficulties, i.e., in the easy (difficult, respectively) trials, the report is more likely to be overconfident (underconfident, respectively) $cc - P_{correct} > 0$ ($cc - P_{correct} < 0$, respectively). Figure 9, *G* and *H*, shows the variation of underconfidence scores as a function of choice confidence for the FD task with a delay of 627 ms and RT task, respectively; both display the hard-easy effect in the reports. Of note, these results still hold true when comparing estimated subjective performance from choice confidence with the behavioral performance using $g(cc) - P_{correct}$. Such a hard-easy effect in our model mainly stems from sampling of stochastic neural dynamics; sampling duration thus influences the underconfidence score in our model. When the sampling duration is short, the network behaves with more overconfidence. To test

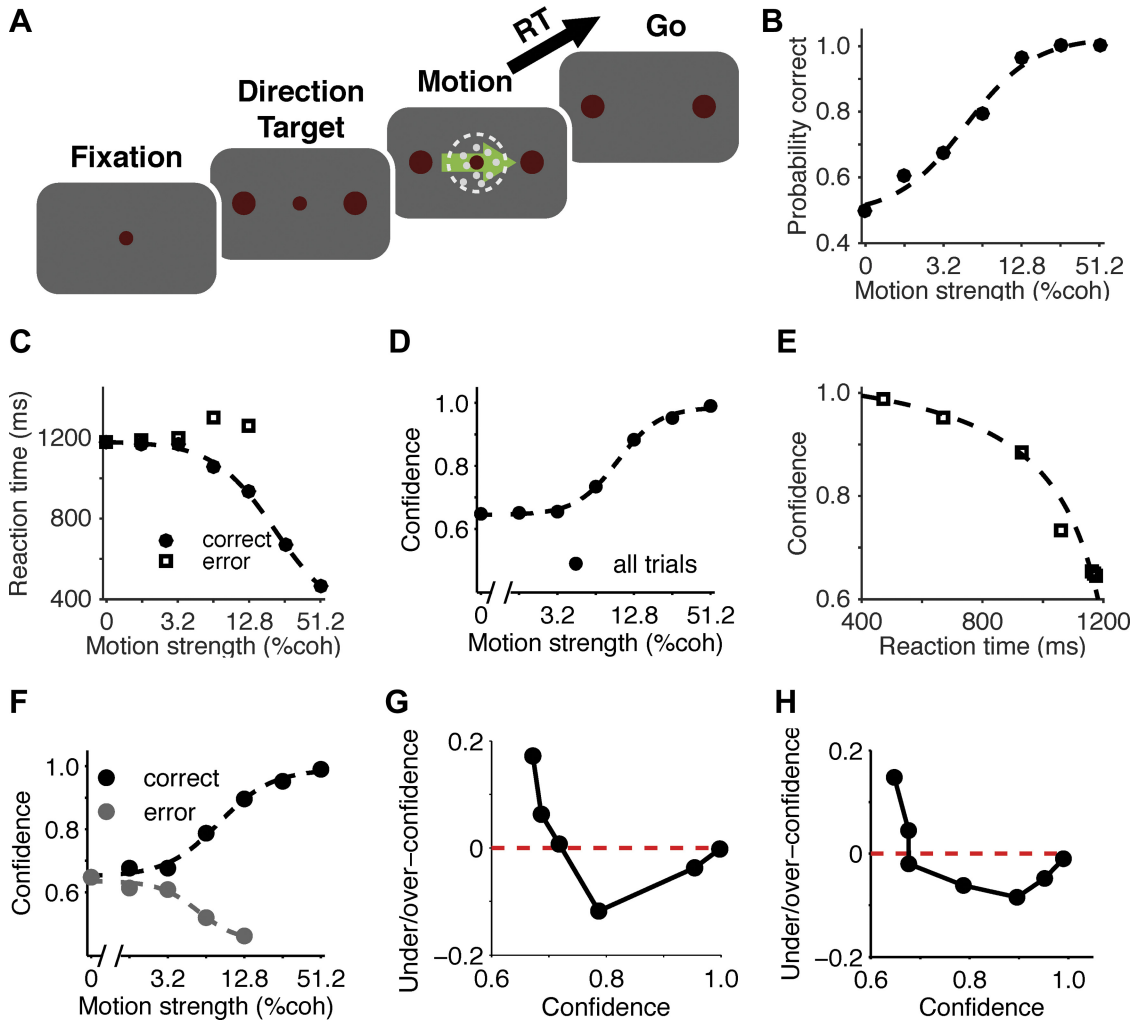

Fig. 9. Choice confidence in a reaction time (RT) task. *A*: RT discrimination task with confidence rating. In task, a subject can indicate its choice at any time after the motion onset simultaneously with a direct report of confidence. *B*: psychometric curves. *C*: chronometric curves. $P_{correct}$ increases while RT decreases with the motion strength. *D–F*: confidence reported as a post hoc feature of decision. *D*: choice confidence increases with motion strength (see also the result in Fig. 5, Beck et al. 2008). *E*: confidence decreases as an inverse function of RT [$cc = a/(t - b) + c$; $a = 91.24$ ms, $b = 1369.35$ ms, $c = 1.089$ are parameters to fit, $R^2 = 0.998$; black line]. *F*: confidence increases (decreases) as a function of the motion strength in correct (error, respectively) trials. *B* and *D* imply that choice confidence increases with $P_{correct}$. We found that, for difficult trials, the simulation exhibits overconfidence (confidence estimation is greater than correct rate), whereas, for easy trials, it exhibits underconfidence (confidence estimation is lower than correct rate). *G* and *H*: variation of under-/overconfidence score with the increase of confidence in the fixed-duration (FD) task with a delay of 627 ms and RT task, respectively. The network behaves with overconfidence (above 0) in very difficult trials (at 0, 3.2% and 6.4% motion strengths for FD task; at 0 and 1.6% motion strengths for RT task), but with underconfidence (below 0) in easy and moderately difficult trials.

this, we compared the scores in the FD and RT tasks, where the average sampling durations in RT tasks are longer than those in FD tasks at low motion strength (from 0 to 6.4%). Consequently, the network exhibits overconfidence more often at low motion strengths in FD task.

DISCUSSION

We have shown that a biologically plausible spiking network model can account for salient physiological and behavioral data from an experiment designed to study confidence (Kiani and Shadlen 2009), and in our model internal stochasticity plays an essential role of choice confidence (see also Whiteley and Sahani 2008). Specifically, at the moment of choice, our model simultaneously generates a neural signal for confidence. Confidence can be estimated as a function of the differential activity of the competing neural populations, $|R_A - R_B|$. Compared with Bayesian inference models, in our model, there is no explicit representation of probabilities such as likelihood or posterior function. Indeed, all computations are carried out by the fluctuating neural network dynamics. Therefore, confidence estimation itself is simply a quantity that stochastically varies over time and from trial to trial under the same stimulus condition.

Our identification of a confidence signal, $|R_A - R_B|$, agrees with the idea that, as a metacognitive process, confidence is estimated directly on a decision process (Graziano and Sigman 2009; Graziano et al. 2010; Middlebrooks and Sommer 2011, 2012). At the same time, choice confidence is also dissociable from whether the decision is correct or wrong in a single trial, as illustrated by high-confidence error trials (Fig. 7*A*, *bottom*, *right*). In line with our model, the EEG data from (Graziano et al. 2010) showed that, at the neural level, choice confidence

could be dissociated from performance. Such dissociation is naturally explained by attractor dynamics, which could yield the same magnitude of the differential activity $|R_A - R_B|$, hence the same confidence rating in correct and error trials. It is worth noting that *1*) R_A and R_B represent the choices of a decision (not necessary to be a directional decision-making process); *2*) confidence estimation does not depend on a specific choice of the decision (i.e., it does not exclusively rely on the activity of the winner bump, or the losing bump) but a differential activity between choices. In this case, one would expect that a sure target is chosen when $|R_A - R_B|$ is small or the downstream neuronal activity is weak and that a nonsure target is chosen when $|R_A - R_B|$ is large or the downstream neuronal activity is strong. This prediction from our model is consistent with the observations in Komura et al. (2013), wherefore the finding of pulvinar neuronal activity (Komura et al. 2013) could be an example of $|R_A - R_B|$ in the downstream read-out circuit of confidence.

In our model, fast early divergence, i.e., the difference of early buildup rates between R_A and R_B, has a predominant effect on the choice and confidence. This is manifested in the dependence of the choice confidence on the stimulus duration, which saturates quickly for sufficiently long stimuli (Fig. 6*D*; Wong et al. 2007). By contrast, in DDM, sensory evidence contributes equally in time to confidence estimation. Future experiments are needed to assess this different characteristic of the attractor network model vs. DDM. Furthermore, the two competing neural pools could also diverge slowly later in a trial. In our model, persistent activity during the delay not only maintains working memory but also continues to slowly integrate signals from memory (Curtis and Lee 2010). This provides a neural mechanism for postdecision sampling (Resulaj et al. 2009). For instance, Fig. 7*B* shows that the probability of switching from an initial decision to T_s is higher in error trials, in agreement with behavioral observation in a rat experiment (Kepecs et al. 2008). This finding also sheds insights into the phenomenon of changes of mind, which may result from the instability (low confidence) of a choice (see also Albantakis and Deco 2011).

Of note, in the monkey experiment, as well as in our model simulations, the introduction of a sure target only serves as a probe about the confidence of the system (Fig. 8). The probability of opting for the sure target is bounded (Figs. 3 and 8*C*), so it represents a good choice for estimating confidence. The real result, we emphasize here, is to quantify confidence as a function of the neural activity $|R_A - R_B|$. Confidence thus quantified should be applicable to all trials, even without sure target presentation. Furthermore, in the analysis of Kiani and Shadlen (2009), they also found that the probability of opting for the sure target can be predicted using either $f(|R_A - \langle R_A \rangle|)$ or $f(|R_B - \langle R_B \rangle|)$. Nevertheless, $f(|R_A - \langle R_A \rangle|)$ or $f(|R_B - \langle R_B \rangle|)$ is not a good measure of confidence for an RT task, for which either R_A or R_B is assumed to reach a fixed threshold at the moment of the choice; therefore, $f(|R_A - \langle R_A \rangle|)$ or $f(|R_B - \langle R_B \rangle|)$ would always be a fixed value [f(|threshold − average|)] rather than a graded quantity that varies from trial to trial (Kiani et al. 2014).

Comparison with existing models. Computational schemes have been proposed for the study of confidence (Kepecs and Mainen 2012; Kepecs et al. 2008; Kiani and Shadlen 2009; Moreno-Bote 2010; Ratcliff and Starns 2009; Rolls et al. 2010a, 2010b; Vickers 1979). These models can be classified into Bayesian inference models and neural network models.

In Bayesian inference models, one can either compute confidence based on a single decision variable (Drugowitsch et al. 2012; Kiani and Shadlen 2009) or an optimal population code (Beck et al. 2008). Kiani and Shadlen (2009) proposed that confidence could be defined in terms of the log posterior ratio for the two choices given the position of a decision variable and elapsed time at decision, using DDM. This looks promising, yet it remains unclear what is a direct representation of a decision variable exclusively for a choice. Moreover, for the RT version of the task, this kind of model implies that the position of a decision variable at decision time would be a deterministic function of RT (either a constant or a time-varying function like that in Drugowitsch et al. 2012); one can thus find that confidence would also decrease deterministically as a monotonic function of RT (Drugowitsch et al. 2012; Kiani and Shadlen 2009; Volkmann 1934) on single trials. This idea, however, failed to explain the widely overlapped RT distributions in different confidence categories (Ratcliff and Starns 2009). Such a strong correlation between confidence and RT or performance can be eliminated through a two-stage DDM (Pleskac and Busemeyer 2010), where additional process for confidence is required. Nevertheless, in our model, performance, RT, and confidence are naturally dissociated with each other on single trials. Importantly, in a classic DDM model, sensory evidence contributes equally in time to confidence estimation, whereas, in our model, confidence estimation would more be dominated by the early sensory evidence. Last, in the Kiani and Shadlen model, choosing T_s uses a hard thresholding process and has little to do with neural activity at T_s response field, whereas, in our model, it generates from the same sampling of stochastic neural dynamics as the other choices (as indicated by data in Fig. 5, Kiani and Shadlen 2009).

On the other hand, the optimal population code model (Beck et al. 2008) posits that confidence could be estimated as the instantaneous differential activity, $|R_A - R_B|$, without explicit use of RT as our model. A notable difference between our model and theirs is that the optimal population code model requires LIP neural circuit to be a fine-tuned noiseless integrator. This can be easily tested experimentally because our model predicts that confidence estimation would differ at different times in the delay, whereas their model would expect that to be constant. Generally, these Bayesian inference models (Beck et al. 2008; Kiani and Shadlen 2009) assert that confidence must be based on explicit neural representation of probability functions, such as likelihood at any moment in time and in single trials. Our model demonstrates an alternative perspective. Whereas probability representations may be a perfectly valid mathematical description of the aggregated statistics across trials, they should not be confused with what actually happens in single trials, which is stochastic neural dynamics.

Insabato et al. (2010) and Rolls et al. (2010a, 2010b) extended the model of Wang (2002) to account for the confidence estimation and its behavioral readout. Insabato et al. (2010) argued that confidence can be read out as a function of $R_A + R_B$, and Rolls et al. (2010a, 2010b) claimed that it can be further approximated as a function of the neural activity of the winning pool. All these models showed some consistencies with the existing data. However, as discussed in our model,

neither of them can exclusively demonstrate the position of a neural trajectory related to the choice attractor and thus the choice confidence at any time during a decision. These results seem only true at the moment when a decision is made exactly around a choice attractor in the decision space, where $|R_A - R_B| \simeq |R_A + R_B| \simeq \max(R_A,R_B)$ because $\min(R_A,R_B) << \max(R_A,R_B)$. Therefore, these models would fail to predict confidence using a fixed decision boundary or capture the relationship between neural activities in LIP with P_{sure} or high-confidence errors in single trials. Alternatively, our model does not require a time-varying decision threshold, estimates confidence simply as a function of instantaneous $|R_A - R_B|$ at the moment of choice on single trials, and can correctly reproduce the salient behavioral relationships between confidence, RT, and performance on single trials and those across trials.

Confidence rating is important for monitoring cognition when there is uncertainty, and two types of uncertainty should be distinguished, namely Brunswikian (external) uncertainty originating from incomplete states of knowledge (noisy or ambiguous sensory data) and Thurstonian (internal) uncertainty attributable to variations intrinsic to the brain (Juslin and Olsson 1997). The noise level in a decision circuit has only recently begun to be examined experimentally (Brunton et al. 2013). Our work provides a computational framework to detect these two effects using a spiking-neuron circuit. Our model can also be extended in several important ways. It still remains an open question how confidence estimation, as a sigmoid function of the differential activity in downstream neural circuits, can be read out for a direct report and to guide future behavior. In fact, confidence is commonly assessed without a verbal report using a two-stage PDW task; subjects perform a first-order discrimination task and then make a high-low bet on the outcome of the decision (Middlebrooks and Sommer 2011, 2012; Smith 2009) (see also Kiani et al. 2011), where the probability for a high bet is considered as a readout of confidence estimation. A plausible neural circuit for explicit representation and memory of a confidence signal is needed for the two-stage PDW (Komura et al. 2013; Middlebrooks and Sommer 2011, 2012) and should be examined in the future. A biologically plausible neural circuit to computing $|R_A - R_B|$ involves neurons in pulvinar (Komura et al. 2013), where the neurons fires highly in nonsure-target trials and low in sure-target trials. Moreover, Kiani et al. (2014) recently found that confidence could also decrease in error trials with the difficulty of task, which poses a challenge to our model prediction (Fig. 9*F*). One possible direction in the future is to understand the mechanism of error trial. Nevertheless, what is robust about the prediction in our model compared with their observation is that the difference of the confidence between correct and error decreases with the difficulty of task (Lak et al. 2014). Second, for the sake of simplicity, we assumed that the amount of T_s reward is encoded as the onset strength of its target input, which mimics the firing activity of midbrain dopamine neurons in response to the targets with different amounts of reward (Tobler et al. 2005). Our model predicts that both P_{sure} and $P_{correct}$ increase with the reward of T_s (data not shown). This then brings up two inquiries for future studies: *1*) what a reasonable amount of T_s reward used to measure confidence in a PDW task would be (Dienes and Seth 2010; Fleming and Dolan 2010; Persaud et al. 2007) and *2*) how the amount of T_s reward used is learned through neural dynamics and applied to the decision circuit (Soltani and Wang 2006). Moreover, one can extend our model to investigate confidence signals for multiple-choice decision tasks and effects of microstimulation on confidence (Fetsch et al. 2014). Specifically, one can incorporate known effects of microstimulation on MT inputs in our model to perform the experiment of Fetsch et al. (2014) using computer simulation and then test its effects on confidence. Finally, confidence may be represented in a distributed network in the brain (Del Cul et al. 2009); the dynamical nature and computational principle remains to be elucidated in future research. In conclusion, we found it remarkable that a previously established model of decision making (Furman and Wang 2008) naturally accounts for all the salient behavioral and neurophysiological observations of the Kiani-Shadlen experiment (2009). Furthermore, it reproduces the observation that confidence decreases with response time in an RT version of the task. The model also offers testable predictions about the changes of mind and, unexpectedly, the hard-easy effect observed in human studies, which naturally emerges from the model. Taken together, our work establishes that a dynamical system of stochastic neural population can underlie even the seemingly abstract metacognitive concept of confidence.

ACKNOWLEDGMENTS

We thank M. Furman for help with the computer code (Furman and Wang 2008) and D. Wang for help with simulations. Model simulations were carried out at Beijing Normal University.

GRANTS

This work was supported by NIH grant R01-MH062349.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Z.W. and X.-J.W. conception and design of research; Z.W. performed experiments; Z.W. analyzed data; Z.W. and X.-J.W. interpreted results of experiments; Z.W. prepared figures; Z.W. and X.-J.W. drafted manuscript; Z.W. and X.-J.W. edited and revised manuscript; Z.W. and X.-J.W. approved final version of manuscript.

REFERENCES

Albantakis L, Deco G. Changes of mind in an attractor network of decision-making. *PLoS Comput Biol* 7: e1002086, 2011.

Barthelmé S, Mamassian P. Flexible mechanisms underlie the evaluation of visual confidence. *Proc Natl Acad Sci USA* 107: 20834–20839, 2010.

Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, Shadlen MN, Latham PE, Pouget A. Probabilistic population codes for Bayesian decision making. *Neuron* 60: 1142–1152, 2008.

Brunton BW, Botvinick MM, Brody CD. Rats and humans can optimally accumulate evidence for decision-making. *Science* 340: 95–98, 2013.

Churchland AK, Kiani R, Shadlen MN. Decision-making with multiple alternatives. *Nat Neurosci* 11: 693–702, 2008.

Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10: 910–923, 2000.

Cook EP, Maunsell JHR. Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5: 985–994, 2002.

Curtis CE, Lee D. Beyond working memory: the role of persistent activity in decision making. *Trends Cogn Sci* 14: 216–222, 2010.

Del Cul A, Dehaene S, Reyes P, Bravo E, Slachevsky A. Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain* 132: 2531–2540, 2009.

Dienes Z, Seth A. Gambling on the unconscious: a comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Conscious Cogn* 19: 674–681, 2010.

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.

Fetsch CR, Kiani R, Newsome WT, Shadlen MN. Effects of cortical microstimulation on confidence in a perceptual decision. *Neuron* 83: 797–804, 2014.

Flavell JH. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *Am J Psychol* 34: 906–911, 1979.

Fleming SM, Dolan RJ. Effects of loss aversion on post-decision wagering: implications for measures of awareness. *Conscious Cogn* 19: 352–363, 2010.

Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G. Relating introspective accuracy to individual differences in brain structure. *Science* 329: 1541–1543, 2010.

Furman M, Wang XJ. Similarity effect and optimal control of multiple-choice decision making. *Neuron* 60: 1153–1168, 2008.

Gold JI, Shadlen MN. The neural basis of decision making. *Annu Rev Neurosci* 30: 535–574, 2007.

Graziano M, Sigman M. The spatial and temporal construction of confidence in the visual scene. *PLoS One* 4: e4909, 2009.

Graziano M, Parra LC, Sigman M. Neurophysiology of perceived confidence. *Conf Proc IEEE Eng Med Biol Soc* 2010: 2818–2821, 2010.

Hanks TD, Ditterich J, Shadlen MN. Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat Neurosci* 9: 682–689, 2006.

Hansel D, Mato G, Meunier C, Neltner L. On numerical simulations of integrate-and-fire neural networks. *Neural Comput* 10: 467–483, 1998.

Heeger DJ. Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9: 181–197, 1992.

Insabato A, Pannunzi M, Rolls ET, Deco G. Confidence-related decision making. *J Neurophysiol* 104: 539–547, 2010.

Juslin P, Olsson H. Thurstonian and Brunswikian origins of uncertainty in judgment: a sampling model of confidence in sensory discrimination. *Psychol Rev* 104: 344–366, 1997.

Kepecs A, Mainen ZF. A computational framework for the study of confidence in humans and animals. *Philos Trans R Soc B Biol Sci* 367: 1322–1337, 2012.

Kepecs A, Uchida N, Zariwala HA, Mainen ZF. Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455: 227–231, 2008.

Kiani R, Shadlen MN. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324: 759–764, 2009.

Kiani R, Shadlen MN, Newsome WT. Belief and time: Response dynamics of lateral intraparietal neurons represent choice certainty. In: *Neuroscience Meeting*. Washington, DC: Society for Neuroscience, 2011, 17.06

Kiani R, Corthell L, Shadlen MN. Choice certainty is informed by both evidence and decision time. *Neuron* 86: 1329–1342, 2014.

Komura Y, Nikkuni A, Hirashima N, Uetake T, Miyamoto A. Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nat Neurosci* 16: 749–755, 2013.

Lak A, Costa GM, Romberg E, Koulakov AA, Mainen ZF, Kepecs A. Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron* 84: 190–201, 2014.

Liu F, Wang XJ. A common cortical circuit mechanism for perceptual categorical discrimination and veridical judgment. *PLoS Comput Biol* 4: e1000253, 2008.

Middlebrooks PG, Sommer MA. Metacognition in monkeys during an oculomotor task. *J Exp Psychol Learn Mem Cogn* 37: 325–337, 2011.

Middlebrooks PG, Sommer MA. Neuronal correlates of metacognition in primate frontal cortex. *Neuron* 75: 517–530, 2012.

Moreno-Bote R. Decision confidence and uncertainty in diffusion models with partially correlated neuronal integrators. *Neural Comput* 22: 1786–1811, 2010.

Persaud N, Mcleod P, Cowey A. Post-decision wagering objectively measures awareness. *Nat Neurosci* 10: 257–261, 2007.

Pierrel R, Murray CS. Some relationships between comparative judgment, confidence, and decision-time in weight-lifting. *Am J Psychol* 76: 28–38, 1963.

Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. *Nature* 400: 233–238, 1999.

Pleskac TJ, Busemeyer JR. Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychol Rev* 117: 864–901, 2010.

Ratcliff R, Smith PL. A comparison of sequential sampling models for two-choice reaction time. *Psychol Rev* 111: 333–367, 2004.

Ratcliff R, Starns JJ. Modeling confidence and response time in recognition memory. *Psychol Rev* 116: 59–83, 2009.

Resulaj A, Kiani R, Wolpert DM, Shadlen MN. Changes of mind in decision-making. *Nature* 461: 263–266, 2009.

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.

Rolls ET, Grabenhorst F, Deco G. Choice, difficulty, and confidence in the brain. *NeuroImage* 53: 694–706, 2010a.

Rolls ET, Grabenhorst F, Deco G. Decision-making, errors, and confidence in the brain. *J Neurophysiol* 104: 2359–2374, 2010b.

Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86: 1916–1936, 2001.

Smith JD. The study of animal metacognition. *Trends Cogn Sci* 13: 389–396, 2009.

Smith JD, Shields WE, Washburn DA. The comparative psychology of uncertainty monitoring and metacognition. *Behav Brain Sci* 26: 317–339, 2003.

Soltani A, Wang XJ. A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J Neurosci* 26: 3731–3744, 2006.

Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science* 304: 1782–1787, 2004.

Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science* 307: 1642–1645, 2005.

Treue S, Hol K, Rauber HJ. Seeing multiple directions of motion-physiology and psychophysics. *Nat Neurosci* 3: 270–276, 2000.

Van Zandt T. ROC curves and confidence judgments in recognition memory. *J Exp Psychol Learn Mem Cogn* 26: 582–600, 2000.

Vickers D. *Decision Processes in Visual Perception*. New York, NY: Academic, 1979.

Volkmann J. The relation of the time of judgment to the certainty of judgment. *Psychol Bull* 31: 672–673, 1934.

Wang XJ. Probabilistic decision making by slow reverberation in neocortical circuits. *Neuron* 36: 955–968, 2002.

Wang XJ. Decision making in recurrent neuronal circuits. *Neuron* 60: 215–234, 2008.

Whiteley L, Sahani M. Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *J Vis* 8: 2.1–2.15, 2008.

Wong KF, Wang XJ. A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci* 26: 1314–1328, 2006.

Wong KF, Huk AC, Shadlen MN, Wang XJ. Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Front Comput Neurosci* 1: 6, 2007.