



# Prediction of decisions from noise in the brain before the evidence is provided

Edmund T. Rolls<sup>1\*</sup> and Gustavo Deco<sup>2</sup>

<sup>1</sup> Oxford Centre for Computational Neuroscience, Oxford, UK

<sup>2</sup> Institució Catalana de Recerca i Estudis Avançats, Computational Neuroscience, Universitat Pompeu Fabra, Barcelona, Spain

## Edited by:

Shu-Chen Li, Max Planck Institute for Human Development, Germany

## Reviewed by:

Hauke R. Heekeren, Max Planck Institute for Human Development, Germany

Rafal Bogacz, University of Bristol, UK

## \*Correspondence:

Edmund T. Rolls, Oxford Centre for Computational Neuroscience, Oxford, UK.

e-mail: edmund.rolls@oxcns.org

URL: www.oxcns.org

Can decisions be predicted from brain activity? It is frequently difficult in neuroimaging studies to determine this, because it is not easy to establish when the decision has been taken. In a rigorous approach to this issue, we show that in a neurally plausible integrate-and-fire attractor-based model of decision-making, the noise generated by the randomness in the spiking times of neurons can be used to predict a decision for 0.5 s or more before the decision cues are applied. The ongoing noise at the time the decision cues are applied influences which decision will be taken. It is possible to predict on a single trial to more than 68% correct which of two decisions will be taken. The prediction is made from the spontaneous firing before the decision cues are applied in the two populations of neurons that represent the decisions. Thus decisions can be partly predicted even before the decision cues are applied, due to noise in the decision-making process. This analysis has interesting implications for decision-making and free will, for it shows that random neuronal firing times can influence a decision before the evidence for the decision has been provided.

**Keywords:** prediction, decision-making, decision prediction, attractor network, noise in the brain, fMRI, computational neuroscience, free will

## INTRODUCTION

There are fMRI analyses of how early one can predict from neural activity what decision will be taken (Haynes and Rees, 2005a,b, 2006; Pessoa and Padmala, 2005; Lau et al., 2006; Hampton and O'Doherty, 2007; Haynes et al., 2007; Rolls et al., 2009). For example, in one investigation subjects held in mind which of two tasks, addition or subtraction, they intended to perform. It was possible, while they held it in mind in a delay period, to decode or predict with fMRI (functional magnetic resonance neuroimaging) from medial prefrontal cortex activations whether addition or subtraction would later be performed, with accuracies in the order of 70% (where chance was 50%; Haynes et al., 2007). There is also evidence that the ongoing variations in neural activity measured for example with fMRI may be related to whether a signal is detected and to perceptual decisions (Ress et al., 2000; Boly et al., 2007; Hesselmann et al., 2008, 2010; Sadaghiani et al., 2010).

A problem with such studies is that it is often not possible to know exactly when the decision was taken at the mental level, or when preparation for the decision actually started, so it is difficult to know whether neural activity that precedes an action or report in any way predicts the actual decision that will be taken (Rolls and Deco, 2010). In fMRI studies, the temporal precision is also poor. In these circumstances, is there anything rigorous that our understanding of the neural mechanisms involved in the decision-making can provide? It turns out that there is, as we show here using an integrate-and-fire (IF) attractor network model of decision-making.

We simulated an attractor network with two possible decision states, D1 and D2. After 2 s of spontaneous firing, decision cues for D1 and D2 were applied to the network. The decision cues for these simulations had equal magnitude, and each decision state

was chosen on approximately 50% of the trials, which is the chance performance that was expected. We however looked backward in time to the period before the decision cues were applied, to investigate whether the noisy firing (i.e., variable because each neuron emitted close to Poisson spike trains) before the decision cues were applied in any way was related to which attractor, D1 or D2, won on a particular trial.

## MATERIALS AND METHODS

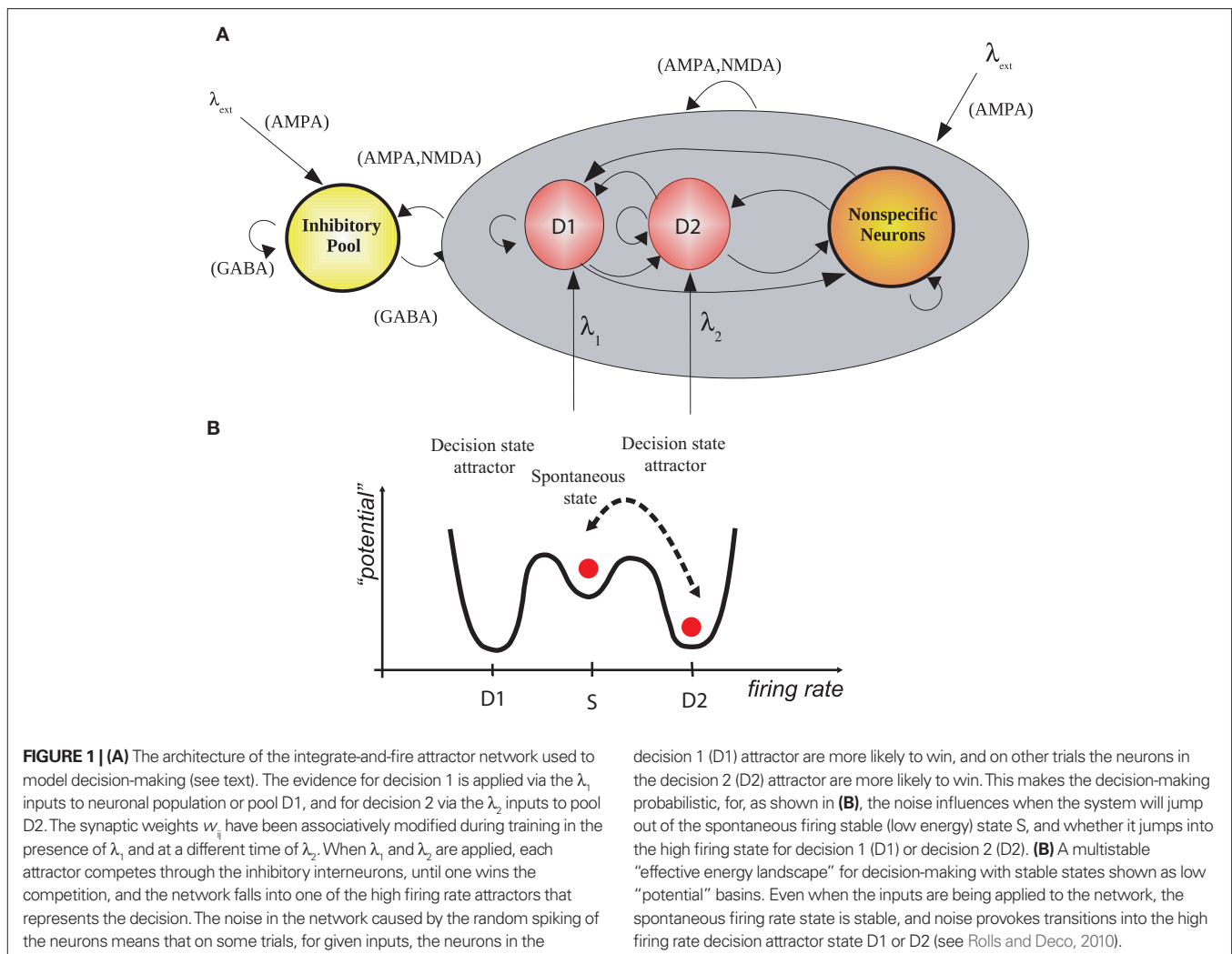
We capitalize on recent advances in theoretical understanding of how choice decisions are made using an IF attractor network that makes probabilistic decisions from the spontaneous low firing state into one of two or more high firing rate stable attractor states each implemented by a set of coupled neurons that receives the inputs for one of the decisions (Wang, 2002, 2008; Deco and Rolls, 2006; Rolls, 2008; Deco et al., 2009; Rolls and Deco, 2010). The choice made is probabilistic because of the noise contributed to by the almost random spiking times of the neurons for a given firing rate, which have close to a Poisson distribution (Rolls, 2008; Rolls and Deco, 2010). On a particular trial, the attractor that happens to have more spikes in its neurons by chance is more likely to win the competition. Such attractor networks are implemented by excitatory connections between cortical pyramidal cells (Rolls, 2008, 2010a), and provide a neural architecture not only for decision-making but also for short-term memory (Goldman-Rakic, 1995; Amit and Brunel, 1997; Rolls, 2008), and memory recall (Rolls, 2008; Rolls and Deco, 2010).

The theoretical framework of the model used here was introduced by Wang (2002) and developed further (Deco and Rolls, 2006; Deco et al., 2007, 2009; Marti et al., 2008; Wang, 2008; Rolls

and Deco, 2010), and the results described here apply generically to IF attractor network models of decision-making. In this framework, we model probabilistic decision-making by a network of interacting neurons organized into a discrete set of populations, as depicted in **Figure 1**. Populations or pools of neurons are defined as groups of excitatory or inhibitory neurons sharing the same inputs and connectivities. The network contains  $N_E$  (excitatory) pyramidal cells and  $N_I$  inhibitory interneurons. In our simulations, we use  $N_E = 400$  and  $N_I = 100$ , and a larger network with  $N_E = 3200$  and  $N_I = 800$ , consistent with the neurophysiologically observed proportion of 80% pyramidal cells vs 20% interneurons (Abeles, 1991; Rolls and Deco, 2002). The neurons are fully connected (with synaptic strengths as specified below). In the model, the specific populations D1 (for decision 1) and D2 encode the categorical result of the choice between the two stimuli that activate each of these populations. Each specific population of excitatory cells contains  $rN_E$  neurons (in our simulations  $r = 0.1$ ). In addition there is one non-specific population, named “Non-specific,” which groups all other excitatory neurons in the modeled brain area not involved in the present task, and one inhibitory population, named “Inhibitory,” grouping the local inhibitory neurons in

the modeled brain area. The latter population regulates the overall activity and implements competition in the network by spreading a global inhibition signal.

Because we are mainly interested in the non-stationary probabilistic behavior of the network, the proper level of description at the microscopic level is captured by the spiking and synaptic dynamics of one-compartment IF neuron models (Deco and Rolls, 2005; Rolls and Deco, 2010). An IF neuron integrates the afferent current generated by the incoming spikes, and fires when the depolarization of the cell membrane crosses a threshold. At this level of detail the model allows the use of realistic biophysical time constants, latencies, and conductances to model the synaptic current, which in turn allows a thorough study of the realistic time scales and firing rates involved in the time evolution of the neural activity. Consequently, the simulated neuronal dynamics, that putatively underlie cognitive processes, can be quantitatively compared with experimental data. For this reason, it is very useful to include a thorough description of the different time constants of the synaptic activity. The IF neurons are modeled as having three types of receptor mediating the synaptic currents flowing into them: AMPA, NMDA (both activated by glutamate), and GABA



receptors. The excitatory recurrent post-synaptic currents (EPSCs) are considered to be mediated by AMPA (fast) and NMDA (slow) receptors; external EPSCs imposed onto the network from outside are modeled as being driven only by AMPA receptors. Inhibitory post-synaptic currents (IPSCs) to both excitatory and inhibitory neurons are mediated by GABA receptors. The details of the mathematical formulation are summarized in previous papers (Brunel and Wang, 2001; Deco and Rolls, 2006), and are provided in the Appendix. The simulations of the spiking dynamics of the network were integrated numerically using the second order Runge–Kutta method with a step size 0.05 ms.

We set the conductance values for the synapses between pairs of neurons by connection weights, which can deviate from their default value 1. The structure and function of the network are achieved by differentially setting the weights within and between populations of neurons. We assume that the connections are already formed, by for example earlier self-organization mechanisms, as if they were established by Hebbian learning, i.e., the coupling will be strong if the pair of neurons have correlated activity (i.e., covarying firing rates), and weak if they are activated in an uncorrelated way. As a consequence of this, neurons within a specific excitatory population (D1 and D2) are mutually coupled with a strong synaptic weight  $w_+$ , set to 2.1 for the simulations described here. This value was chosen based on a consistent mean-field analysis so that when the decision cues were being applied, the spontaneous firing state as well as the two high firing rate decision states were all stable (Deco and Rolls, 2006; Loh et al., 2007; Rolls and Deco, 2010). We refer to this as multistability, and in this regime it is noise generated by the neuronal firing that provokes the transition from the spontaneous state into a high firing rate decision state (Deco and Rolls, 2006; Loh et al., 2007; Rolls and Deco, 2010). Furthermore, the populations encoding these two decisions are likely to have anti-correlated activity in this behavioral context, resulting in weaker than average connections between the two different populations. Consequently, for these weights we choose a weaker value  $w_- = 1 - r(w_+ - 1)/(1 - r)$ , so that the overall recurrent excitatory synaptic drive in the spontaneous state remains constant as  $w_+$  is varied (Brunel and Wang, 2001). Neurons in the inhibitory population are mutually connected with an intermediate weight  $w = 1$ . They are also connected with all excitatory neurons in the same layer with the same intermediate weight, which for excitatory-to-inhibitory connections is  $w = 1$ , and for inhibitory-to-excitatory connections is denoted by a weight  $w_i$ . Neurons in a specific excitatory population are connected to neurons in the non-selective population in the same layer with a feedforward synaptic weight  $w = 1$  and a feedback synaptic connection of weight  $w_-$ .

Each individual population is driven by two different kinds of input. First, all neurons in the model network receive spontaneous background activity from outside the module through  $N_{\text{ext}} = 800$  external excitatory connections. Each synaptic connection carries a Poisson spike train at a spontaneous rate of 3 spikes/s, which is a typical spontaneous firing rate value observed in the cerebral cortex. This results in a background external input summed over all 800 synapses of 2400 spikes/s for each neuron. Second, the neurons in the two specific populations additionally receive added firing to the external inputs that encode the evidence for the decision to be made. When stimulating, the rate of the Poisson train to the neurons of the

specific population D1 is increased by an extra value of  $\lambda_1$ , and to population D2 by  $\lambda_2$ , as these encode the two stimuli to be compared. The absolute difference between the two inputs measures the magnitude of the evidence for the decision, and is defined as  $\Delta I = |\lambda_1 - \lambda_2|$ .

The simulations were run for 2 s of spontaneous activity, and then for a further 2 s while the stimuli were being applied. During the spontaneous period, the stimuli applied to each synapse of each neuron in D1 and D2 (and to all the other neurons in the network) had a value of 3 Hz. During the decision period, the mean input to the synapses for  $\lambda_1$  and  $\lambda_2$  on the D1 and D2 neurons was increased to 3.04 Hz per synapse (an extra 32 Hz per neuron). For  $\Delta I = 0$ , we added 32 extra Hertz to the spontaneous and applied this to each  $\lambda_1$  and  $\lambda_2$  synapse in each neuron of both D1 and D2. For  $\Delta I = 16$ ,  $32 + 8$  Hz was the extra applied to D1 and corresponds to  $\lambda_1$  in **Figure 1**, and  $32 - 8$  Hz was the extra applied to the synapses for  $\lambda_2$  inputs to D2, etc. The mean firing rates of each of the four populations of neurons, and the absolute value of the sum of the synaptic currents (AMPA, NMDA, and GABA, defined in the Appendix), for 50 ms time windows were saved every 50 ms for later analysis. The criterion for which population won, that is for which decision was taken, was a mean rate for the last second of the simulation that was 10 Hz greater than that of the other population. (This is in the context that the spontaneous rate was 3 spikes/s, and that the winning population typically had a mean firing rate of 35–40 spikes/s, as will be shown).

The parameters for the synaptic weights and input currents were chosen using the mean-field equivalent of this network (Brunel and Wang, 2001; Deco and Rolls, 2006; Rolls and Deco, 2010) so that in the absence of noise when the input stimuli are being applied there were three possible stable states, the spontaneous firing rate state (with a mean firing for the pyramidal cells of approximately 3 spikes/s), and one of two possible high firing rate attractor states (with a mean firing for the pyramidal cells of approximately 40 spikes/s), with one neuronal population (D1) representing decision 1, and the other population (D2) decision 2. In particular,  $w_+$  was set to 2.1.

The particular model chosen using the mean-field analysis (Brunel and Wang, 2001; Deco and Rolls, 2006; Rolls and Deco, 2010) had three stable states in the absence of the spiking noise, a spontaneous state and one for each of two decisions, as this model has many interesting decision-making properties, including implementing Weber's law ( $\Delta I/I = k$ , i.e., the difference of intensity  $I$  that can be reliably detected divided by the intensity value has a linear component; Deco and Rolls, 2006; Deco et al., 2009). With operation in this multistable regime, it is the approximately random spiking times of the neurons (i.e., approximately Poisson firing at a given mean rate) that causes statistical fluctuations that makes the network jump from the spontaneous firing state into one of the high firing rate attractor (decision) states. The randomness of the firing dynamically and probabilistically provokes transitions from the spontaneous firing state to one of the high firing rate attractor basins that represent a decision when the decision cues are applied (see **Figure 1B**). The parameters of the network made it operate in the same place in the state space with multiple stability determined by the mean-field analysis as in previous investigations (Deco and Rolls, 2006; Loh et al., 2007; Rolls et al., 2008; Rolls and Deco, 2010). With these parameters, on a proportion of the trials

(typically 290/1000 in the small network specified above) the noise from the neuronal spiking provokes a transition from the spontaneous firing rate before the decision cues are applied into a high firing rate state. We were careful to exclude any such trial from each set of 1000 trials, using a criterion for stability of a mean firing rate of less than 5 spikes/s in each of the pools in the last 0.2 s before the decision cues were applied. This is in the context that the mean spontaneous firing rate determined using the mean-field analysis and verified in the spiking simulations was 3 spikes/s, and that the mean firing rate when in a decision attractor was 40 spikes/s. Further investigations to show that instability of the spontaneous state of firing does not account for the results are shown in **Figure 3**.

The stability of an attractor is characterized by the average time in which the system stays in the basin of attraction under the influence of noise. The noise provokes transitions to other attractor states, as shown schematically by the “potential” landscape in **Figure 1B**. (Further details are provided by Rolls and Deco (2010), and at the bifurcation point the system is one-dimensional and implements a non-linear diffusion process; Roxin and Ledberg, 2008.) The behavior results from the interplay between the Poissonian character of the spike timing and the finite size effect due to the limited numbers of neurons in the network (Rolls and Deco, 2010).

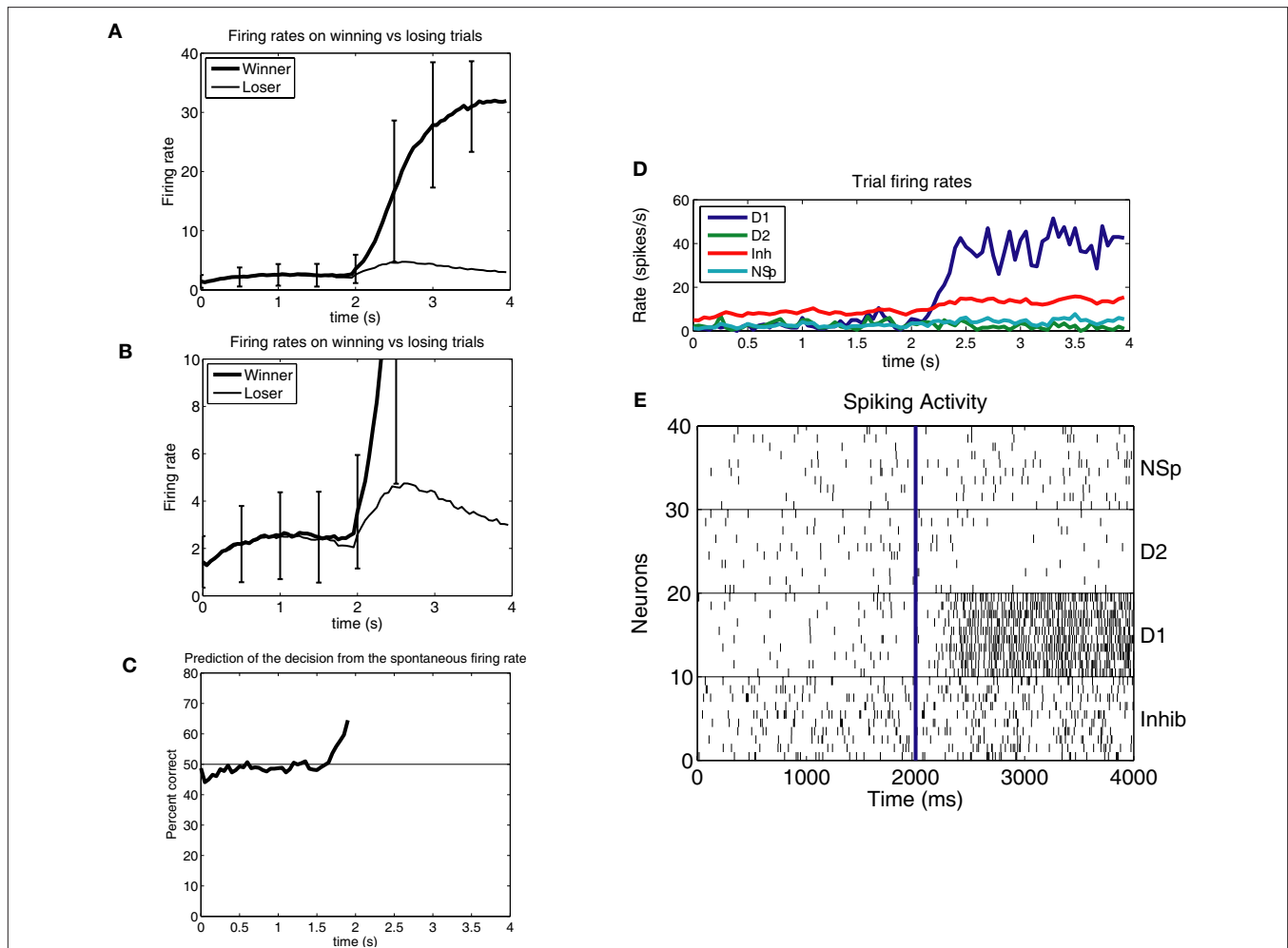
## RESULTS

The results found in simulations in which the firing rate in the spontaneous firing period is measured before a particular attractor population won or lost the competition or even received any evidence relevant to the decision are illustrated in **Figures 2A,B** with  $\Delta I = 0$  so that the chance level for each decision pool is 50% correct. The firing rate averaged over approximately 650 winning (correct) and losing (error) trials for the attractor shows that the firing rate when the attractor will win is on average higher than that for when the attractor will lose at a time that starts in this case approximately 300 ms before the decision cues are applied. Statistical analyses show that the firing rates were significantly different ( $p < 0.005$ ,  $t$ -test, 2.51 vs 2.22 spikes/s) by  $t = 1.7$  s, that is at 300 ms before the decision cues were applied. An example of the activity of the network on a single trial to illustrate the operation of the network is shown in **Figures 2D,E**. It is possible to predict the decision from the firing rates before the decision cues are applied, as shown in **Figure 2C**. The prediction became better (53.6% correct from a single 50 ms bin) than the chance level of 50% correct 300 ms before the decision cues were applied, was significant ( $p < 0.02$ ) at 55.9% correct at 250 ms before the decision cues were applied, and reached a level of 64.5% correct just before the decision cues were applied at  $t = 2.0$  s. The predictions plotted in **Figure 2C** were made from the firing rates of each selective population in 100 ms periods.

The network used for the simulations shown in **Figure 2** contained 500 neurons, with 40 neurons in each of the two decision pools. Because this is quite a small network, the statistical fluctuations caused by the close to Poisson spiking times of the neurons in the spontaneous period of 0–2 s caused the network to sometimes jump out of the spontaneous state into a high firing rate state. (The parameters of the network were selected with a mean-field analysis so that the spontaneous state was stable, and stability would also be the state in an infinite size spiking network. The noise just described in the small network is what causes the

network to on some trials jump out of the spontaneous state.) A jump from the spontaneous state occurred on approximately 290/1000 trials with this size network. We carefully excluded any such unstable trials from the analysis of the results in **Figure 2** by excluding all trials on which the firing rate was  $>5$  spikes/s in the 0.2-s period before the decision cues were applied. This was in a situation in which the parameters had been selected with the mean-field analysis to produce an average spontaneous firing rate of 3 spikes/s.

Further investigations as follows were performed to confirm that the firing rate in each decision pool of neurons had not moved from its stable spontaneous firing before the decision cues were applied. These further investigations involved greatly increasing the size of the network, by eight times so that it contained 4000 neurons, with therefore 3200 excitatory neurons each with 3200 recurrent collateral synapses from the other excitatory neurons. There were now 320 excitatory neurons in each of the two decision pools. As the attractor network becomes larger, the finite size fluctuation-related noise effect due to the neuronal spiking becomes smaller (and the noise is zero when the network has infinite size), as described and analyzed for this network by Deco and Rolls (2006), and Rolls and Deco (2010). With again the strict criterion for stability of the spontaneous state of  $<5$  Hz for the mean firing rate across neurons of each of the decision pools in the last 200 ms before the decision cues were applied, the number of trials with stable spontaneous activity was now 998/1000. Using these 998 trials, so that instability could not be an issue, we found the same effects as those described in **Figure 2**. These further results are illustrated in **Figure 3**, which is for  $\Delta I = 0$ . Of 887 trials with a clear winner using the criteria described above, it was possible to predict which pool would win the competition on 63.7% of the trials (and with chance = 50%, Fisher  $p \ll 0.0001$ ) from the firing rates in a 100 ms period 200 ms before the decision cues were applied. The mean rate was significantly higher in this time window of the spontaneous firing for the winning pool (2.78 spikes/s) than for the losing pool (2.44 spikes/s; paired  $t$ -test,  $p < 10^{-14}$ ). The prediction from a 50-ms bin just before the decision cues were applied was 68% correct. **Figure 3B** shows the percentage correct predictions from 100 ms of firing taken at different pre-stimulus times. The \* indicates the start of the first 100 ms period in which within a 100-ms period a statistically significant prediction could be made ( $p < 0.03$ , Fisher test). This 100 ms period was 750 ms before the decision cues were applied. Further evidence that instability of the spontaneous state was not contributing to these results is that the decision times on all trials were 300 ms or more after the decision cues were applied. Thus with 4000 neurons in the network and with  $\Delta I = 0$  it is still possible to predict which pool will make the decision with approximately 68% correct from the firing rates of the neurons in each of the attractors in the period of spontaneous activity before the decision cues are applied. In this large network, the effects of the spiking-related noise are smaller in that the spontaneous attractor is almost always stable, and with less noise it is more easily possible from the spontaneous firing rates in the predecision period to predict which decision will be taken. This result from a scaled-up network indicates that the processes described here are likely to apply in the neocortex in which the number of recurrent collateral connections onto each neuron is in the order of thousands.



**FIGURE 2 | (A)** Prediction of a decision before the evidence is applied. In this integrate-and-fire simulation of decision-making, the decision cues were turned on at  $t = 2$  s, with  $\Delta I = 0$ . The firing rate averaged over approximately 650 winning vs losing trials for the attractor shows that the firing rate when the attractor will win is on average higher than that for when the attractor will lose at a time that starts in this case at 300 ms before the decision cues are applied. (At  $t = 2$  s with  $\Delta I = 0$  the input firing rate on each of the 800 external input synapses onto every neuron of both of the selective attractor populations is increased from 3.00 to 3.04 spikes/s, as described in the text.) The error bars in this and subsequent Figures show the standard deviation of the firing rate calculated across trials for the 50-ms bins to indicate the noisy operation of this decision-making system. The large standard deviations in the period after the decision cues are applied at 2 s reflect the different decision times of the network on different trials. **(B)** As **(A)**, but with an expanded firing rate axis so

that the difference in the firing rates of the pool that will win and of the pool that will lose can be illustrated. **(C)** The prediction of which pool will win from 100 ms periods of the firing of the two pools at different times before the decision cues are applied at  $t = 2$  s. The network size was 500 neurons, with 400 excitatory neurons, 400 excitatory recurrent collateral synaptic connections on each neuron, and 40 neurons in each of the two decision pools. **(D)** Example from a single trial of the firing rates of the four populations of neurons for a correct decision (for which  $\Delta I = 16$ ). From the top right the plot order is: D1 is the firing rate of the correct and winning attractor D1. Inh is the inhibitory population that uses GABA as a transmitter. NSp is the non-specific population of neurons (see **Figure 1**). D2 is the firing rate of the correctly losing attractor D2. **(E)** Rastergrams for the same trial shown in **D** to illustrate the probabilistic spiking of each neuron. Ten neurons from each of the four pools of neurons are shown. Each vertical line is the spike from a neuron.

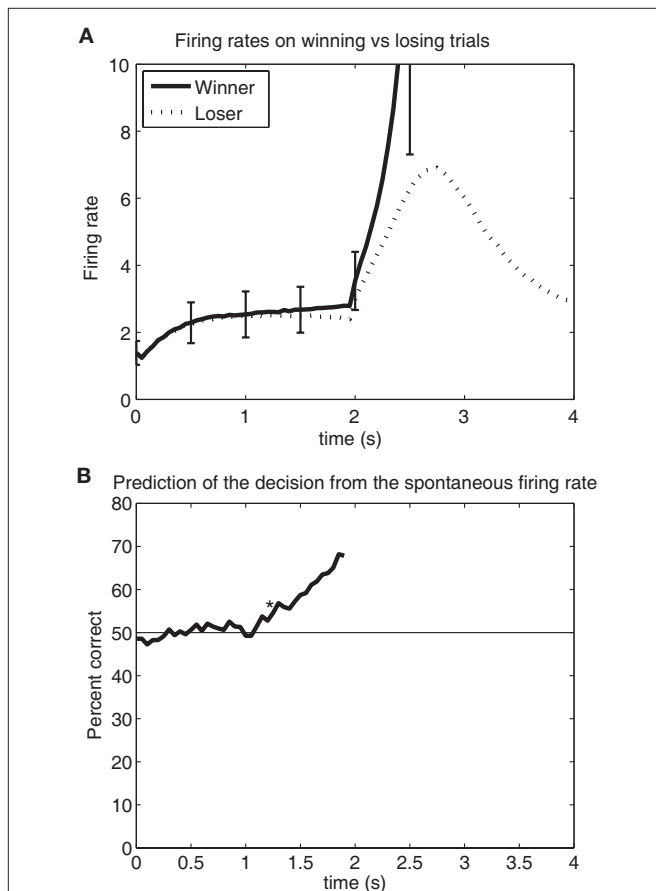
The time scale over which a decision can be partly predicted is of the order of 750 ms in the large simulations described here and illustrated in **Figure 3**. This is a surprisingly long time, given that the neuronal and synaptic time constants of the network are in the order of tens of millisecond or less (with the NMDA receptor time constant somewhat longer at 100 ms). We suggest that the time constant of the system is effectively longer than this, because of the positive feedback in the recurrent collateral synaptic connections, which may encourage statistical fluctuations in the spontaneous activity to feed back and build up over longer time periods. To

obtain evidence on this, we computed the autocorrelation function for the spontaneous firing rate of the D1 population. The function was calculated from the mean firing rates of the population of D1 neurons calculated and sampled every 50 ms. To obtain 4.5 s of data, the decision cues were omitted from these 6 s simulation trials. Data were used from 930 trials in which the spontaneous firing rate remained stable and below 5 spikes/s at the end of the 6-s trial (and this was all but seven of the trials as the network contained 4000 neurons and the spontaneous state was quite stable). The data were from 50 ms firing rate bins starting 1000 ms after the start

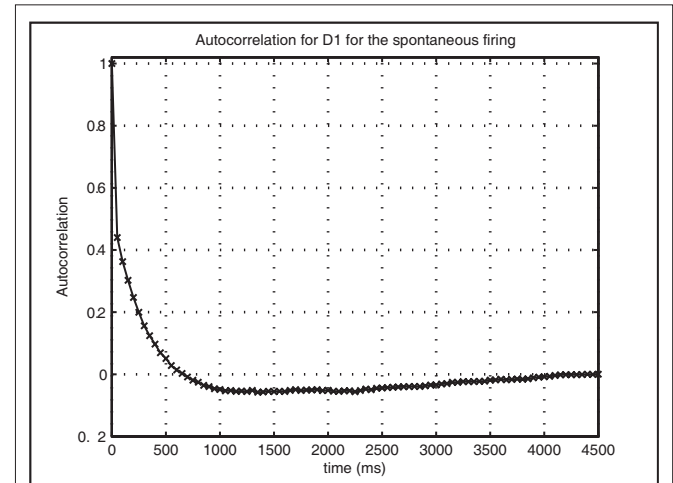
of the 6-s simulation, and the average autocorrelation function computed over the 930 trials is shown in **Figure 4**. It is clear that the autocorrelation function remains positive for lags up to 700 ms, which indicates that fluctuations in the firing rates are positively correlated with the rates up to 700 ms later. This indicates that fluctuations in the firing rates may have effects 700 ms later, and the implication in the present context is that statistical fluctuations in the spontaneous firing rate due to the spiking-related noise may well be expected to have effects on decision-making taking place

700 ms later, which is what has been shown in the results described here. The positive period in the autocorrelation function of 700 ms matches quite closely that of the period in which predictions that are better than chance can be made with this size of network with 320 neurons in each of pools D1 and D2, which become significant 750 ms before the decision cues are applied, as shown in **Figure 3B**.

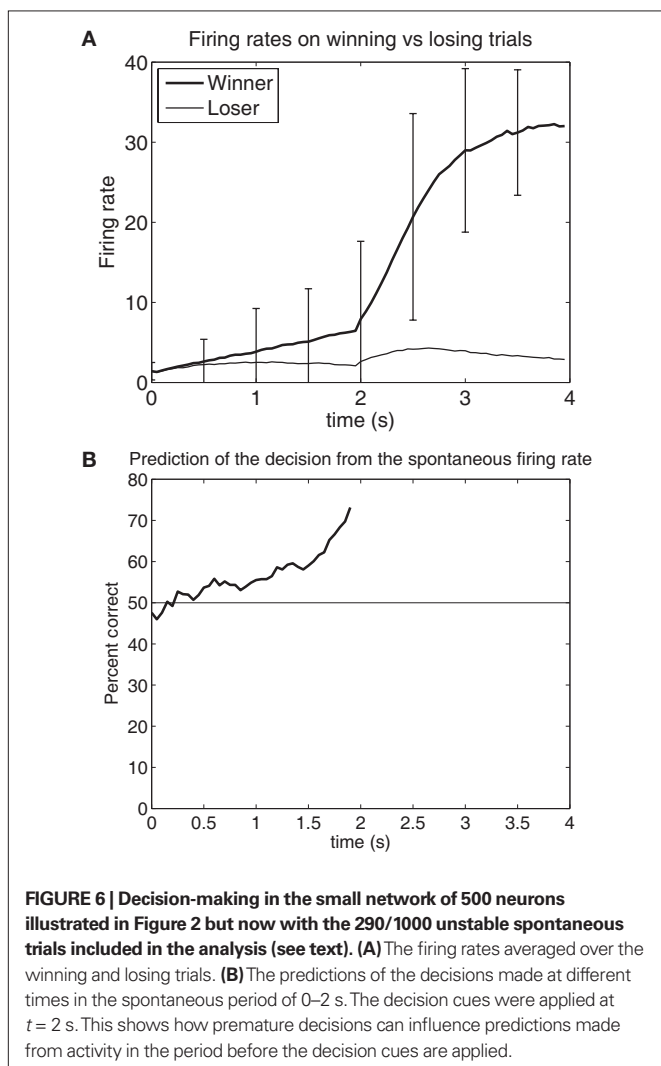
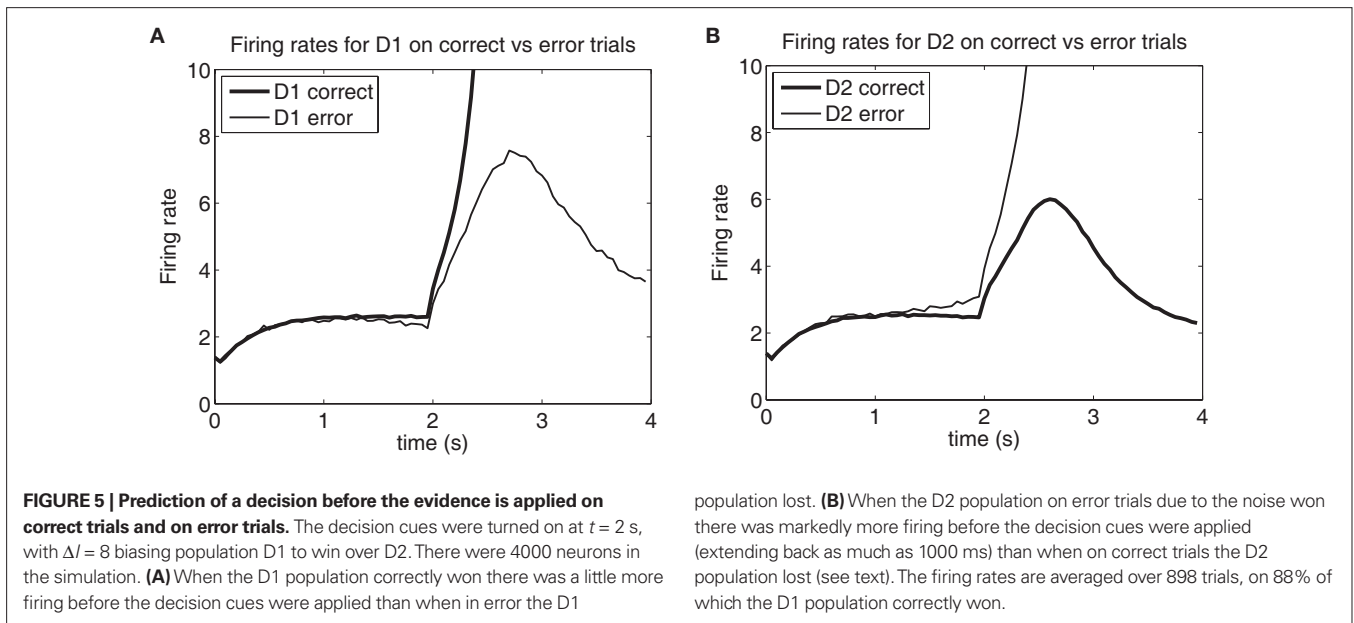
The results shown in **Figures 2 and 3** are for  $\Delta I = 0$ , that is when the input to the D1 attractor was the same as that to the D2 attractor, and each attractor won on approximately 50% of the trials. Analogous effects were found when  $\Delta I$  was greater than 0, and the D1 attractor was biased to win more by the evidence for decision 1. But very interestingly, the influence of the noise of the firing rate in the spontaneous firing period before the decision cues were applied had to be greater on error trials than on correct trials, in order to overcome the bias provided by the external decision cues pushing the network toward a correct decision. This is illustrated in **Figure 5**, which again is for a network with 4000 neurons so that the spontaneous firing state before the decision cues were applied was stable. (There were only two unstable trials out of 1000 analyzed, and these two were excluded.) The decision cues were turned on at  $t = 2$  s, with  $\Delta I = 8$  biasing population D1 via  $\lambda_1$  to win over D2. **Figure 5A** shows that when the D1 population correctly won there was a little more firing (2.58 spikes/s for each neuron in the 150-ms before the decision cues were applied) than when in error the D1 population lost (2.39 spikes/s,  $p < 0.003$ ; a difference of 0.19 spikes/s). **Figure 5B** shows that when the D2 population on error trials due to the noise won there was markedly more firing (2.96 spikes/s) before



**FIGURE 3 | (A)** Prediction of a decision before the evidence is applied. In this integrate-and-fire simulation of decision-making, the decision cues were turned on at  $t = 2$  s, with  $\Delta I = 0$  (as in **Figure 2**), but the network size was now 4000 neurons, with 3200 excitatory neurons, 3200 excitatory recurrent collateral synaptic connections on each neuron, and 320 neurons in each of the two decision pools. The network remained in the state of stable firing at a spontaneous firing rate for 998 of the 1000 trials. The firing rate averaged over 887 of these 998 trials for which there was a clear winner (see Materials and Methods) in the 100-ms pre-cue period starting 200 ms before the decision cues were applied was 2.78 spikes/s for the pool that would win, and 2.44 spikes/s for the pool that would lose ( $p < 10^{-14}$ ). The firing rate for the winning population asymptotated at approximately 36 spikes/s as in the 500-neuron simulation shown in **Figure 2**. **(B)** The percentage correct predictions from 100 ms of firing taken at different pre-stimulus times. The \* indicates the start of the first 100 ms period in which within a 100-ms period a statistically significant prediction could be made ( $p < 0.03$ , Fisher test). This 100 ms period was 750 ms before the decision cues were applied. The predictions were 68% correct in the last 100 ms before the decision cues were applied, and 73% in the last 150 ms. The horizontal line at 50% shows what would be expected by chance.



**FIGURE 4 | Autocorrelation function for the spontaneous firing rate of the D1 population.** The function was calculated from the mean firing rates of the population of neurons D1 calculated and sampled every 50 ms. To obtain 4.5 s of data, the decision cues were omitted from these simulations. Data were used from 930 trials in which the spontaneous firing rate remained stable and below 5 spikes/s at the end of the 6-s trial (and this was all but seven of the trials as the network contained 4000 neurons and the spontaneous state was quite stable). The data were from bins each 50 ms in duration starting 1000 ms after the start of the 6-s simulation, and the average autocorrelation function computed over the 930 trials is shown. The negative values of the autocorrelation function at lags of greater than 700 ms just reflects the setting of the average value of the time series to a mean of zero, with the actual autocorrelation being zero at lags of greater than 700 ms.



the decision cues were applied (extending back as much as 700 ms) than when on correct trials the D2 population lost (2.49 spikes/s,  $p < 10^{-11}$ ; a difference of 0.47 spikes/s). (The firing rates are averaged over 794 trials on which D1 correctly won, and 104 trials on which D2 in error won, =88% correct with  $\Delta I = 8$ .) The cause of the greater effect apparent in the spontaneous, predecision, period on error trials for the losing population than on correct trials with  $\Delta I > 0$  is described in the Section “Discussion.”

We now suggest that a second mechanism may contribute to what are sometimes described as effects of ongoing neural activity that can influence decisions. Given that the spontaneous state of networks in the brain may sometimes be unstable, with the spiking-related (and possibly other) noise sometimes provoking transitions to a high firing rate state in the absence of stimuli, it could be that some instances of what appear to be effects of ongoing neural activity on decisions are in fact just premature decisions being made before the decision cues are applied. The decision in this second case would reflect a transition across a bifurcation in the state space toward a high firing rate decision attractor state during the spontaneous period before the decision cues are applied. To investigate this, we repeated the simulation shown in Figure 2, but did not exclude the 290/1000 trials in which the spontaneous state was unstable. The results are shown in Figure 6A, in which, due to the transitions to high activity that occur at random times in the spontaneous period, the mean spontaneous rate across trials increases throughout the spontaneous period for the pool that wins. In a sense, on these unstable trials, the winning pool has taken a premature decision before the decision cues are applied. This is reflected in the ability to predict the decision far back in time from when the decision cues are applied, as shown in Figure 6B.

## DISCUSSION

What is the mechanism for the prediction of decisions from the spontaneous neural activity before the decision cues are applied, which in our simulations reached 68% correct? We suggest that

the mechanism is as follows. There will be noise, i.e., statistical fluctuation, in the approximately Poisson neuronal firing that will lead to low, but different, firing rates at different times in the period before the decision cues are applied of the two selective populations of neurons that represent the different decisions. If the firing rate of say population D1 (representing decision 1) is higher than that of the D2 population at a time just as the decision cues are being applied, this firing will add to the effect of the decision cues, and make it more likely that the D1 population will win. These statistical fluctuations in the spontaneous firing rate will have a characteristic time course that will be influenced by the time constants of the synapses etc., in the system, so that if a population has somewhat higher firing at say 500 ms before the cues are applied, it will be a little more likely to also have higher firing some time later. By looking backward in time one can see how long the effects of such statistical fluctuations can influence the decision that will be reached, and this is shown in **Figure 2** to be approximately 300 ms in the small network we studied with 500 neurons, and approximately 750 ms in the larger network we studied with 4000 neurons. We emphasize that this gradual increase of firing rate for the attractor that will win before the decision cues are applied is an effect found by averaging over very many trials, and that the fluctuations found on an individual trial (illustrated in **Figures 2D,E**) do not reveal obvious changes of the type illustrated in **Figure 2A**.

What could be the cause of the larger difference in the firing rate on error than on correct trials with  $\Delta I > 0$  illustrated in **Figure 5**? It is suggested that the reason is that on the small proportion of trials on which there are errors, the noise on error trials to force the incorrect, D2, attractor into winning (**Figure 5B**) will need to be especially large to overcome the bias in the network against D2 winning. Thus a neurophysiological prediction is that on error trials for the incorrectly winning attractor (D2) neurons, there will on average be a larger noise contribution to the decision than on the many correct trials for the correct, D1, neurons. In the latter case, small amounts of noise in either direction will not force the D1 attractor away from the correct decision, and will average out to small mean effects when many trials are averaged together (**Figure 5A**).

Thus we have a rigorous and definite answer and understanding of one way in which decisions that will be taken later are influenced by and can be probabilistically predicted from the prior state of the network. It is possible to make a probabilistic prediction of which decision will be taken from the prior activity of the system, before the decision cues are applied. This conclusion emerges from a fundamental understanding of how noise in the brain produces statistical fluctuations that can influence neural processes (Rolls and Deco, 2010). The noise in the simulations arises from the randomness in the firing times of the neurons (for a given mean firing rate), which approximate a Poisson process (Rolls and Deco, 2010).

It would be difficult to obtain evidence for this from fMRI studies, because of the long time course of the hemodynamic response. But we make the prediction that if neuronal activity is recorded in a decision-making task from the neurons involved in the choice process, there will be neuronal firing rate changes that influence the decision and that can be measured before the decision cues are applied. Part of the value of the present approach is that we have a well-defined system, and know when the decision cues are applied, and when the decision is taken. The present finding makes

an important advance because it is difficult to make conclusions about whether neural activity precedes decisions in studies in which there is no precise knowledge of the time at which the decision is taken, or in which the time resolution is poor (Haynes et al., 2007; Soon et al., 2008). However, we note that when multiple attractor networks perhaps in different brain areas interact, the time scale of the fluctuations could be longer, and so correspondingly could the time period over which predictions of a decision could be made from the statistical fluctuation in the preceding period. This is an interesting topic for future research.

The time scale over which a decision can be predicted is of the order of 700 ms in the simulations described here. This is a surprisingly long time, given that the neuronal and synaptic time constants of the network are in the order of tens of millisecond or less (with the NMDA receptor time constant somewhat longer at 100 ms). We suggest that the time constant of the system is effectively longer than this, because of the positive feedback in the recurrent collateral excitatory connections, which may encourage statistical fluctuations in the spontaneous activity to feed back and build up over longer time periods, and evidence for this is provided in **Figure 4** which shows the autocorrelation function for the spontaneous activity of one of the decision pools does not cross 0 until a time lag of 700 ms. The underlying mechanisms for the effects over quite long periods include the properties of the noise (Marti et al., 2008), effects introduced by the positive feedback in the recurrent collateral synaptic connections (Mattia and Del Giudice, 2004), and effects that can be introduced if some of the inhibitory neurons in the network have time constants of longer than 10 ms (Smerieri et al., 2010).

The instability of the spontaneous state provides a second mechanism through which it may seem that decisions can be predicted even before the evidence is provided. This mechanism could contribute to the apparent ability to predict decisions from the fMRI BOLD signal before the decision is apparently made, or reported. The same mechanism may also contribute to many real-life situations in which for example a runner starts a race before the gun sounds. The process is illustrated in **Figure 6**, and is a result of the spiking-related noise. Instability of the spontaneous state has been noted also by others (Salinas, 2003). It could thus be that some instances of what appear to be effects of ongoing activity on decisions are in fact just premature decisions being made before the decision cues are applied. The decision in this second case would reflect a transition across a bifurcation in the state space toward a high firing rate decision attractor state, and would be clearly different from the influence of noise in the pre-cue period that does not actually provoke a transition to a high firing rate attractor decision state, but influences the decision taken after the decision cues are applied. This latter is the first mechanism described in this paper.

A strength of the current approach is that the model incorporates effects from many levels, including the synaptic and biophysical levels, and the abstract level of statistical fluctuations in dynamical systems, and is able to make predictions all the way from synaptic currents, to neuronal activity, to fMRI signals, and to behavioral choice and subjective confidence, and offers a quantitative approach to our understanding at all these levels (Rolls, 2008; Rolls and Deco, 2010). It is in this sense a mechanistic approach, which takes into account details of the underlying biophysical and network



mechanisms in the brain, and then can account for behavioral, as well as neurophysiological and fMRI data. Moreover, the approach generalizes to other processes implemented by attractor networks in the brain including memory recall (Rolls, 2008; Rolls and Deco, 2010). Indeed, the model described here is different in a number of ways (Wang, 2002; Deco et al., 2009; Rolls and Deco, 2010) from accumulator, counter, or race models of decision-making in which the evidence for different choices accumulates, which may include a noise term, and which undergo a random walk in real time, which is a diffusion process (Vickers, 1979; Vickers and Packer, 1982; Ratcliff et al., 1999; Smith and Ratcliff, 2004). First, in most accumulator models, a mechanism for computing the difference between the stimuli is not described, whereas in the current attractor model this is achieved by the feedback inhibition. (Some more recent accumulator models do suggest a process for calculating the difference; Usher and McClelland, 2001; Ditterich, 2006.) Second, in the attractor network model the decision corresponds to high firing rates in one of the attractors, and there is no arbitrary threshold that must be reached, or arbitrary stopping time. In the network there are of course neuronal thresholds, but these are biologically realistic with the operating points of the network (in particular  $w_+$  and  $w_{inh}$ ) selected using the mean-field analysis to produce when the decision cues are applied a stable low firing rate state of 3 spikes/s, and a high firing rate attractor state for one of the attractors with a rate of approximately 40 spikes/s. Third, the noise in the attractor model is not arbitrary, but is accounted for by finite size noise effects of the spiking dynamics of the individual neurons with their Poisson-like spike trains in a system of limited size (Deco et al., 2009; Rolls and Deco, 2010). (Simulated Poisson noise has been used in some more recent accumulator models; Mazurek et al., 2003; Beck et al., 2008; Zhang and Bogacz, 2010.) Fourth, the attractor model we describe is biologically plausible as a mechanism used in the brain consistent with neuronal firing during decision-making (Deco et al., 2009; Rolls and Deco, 2010), and therefore makes testable predictions about the firing rates of neurons, the synaptic currents, and the fMRI BOLD signals as described elsewhere (Rolls and Deco, 2010). (The predictions of the BOLD signals are made by convolving the firing rates or the synaptic currents in the simulated network with the hemodynamic response function; Rolls et al., 2010a,b.) In contrast, accumulator and related models are not neurobiological models, though they have been used to make neurobiological predictions (Hanks et al., 2006; Yang and Shadlen, 2007; Kiani et al., 2008); and of course the starting point from which the accumulation starts (which might be thought of as an initial noise state) can in this system influence the decision (Laming, 1968; Ratcliff and Rouder, 1998). In fact, if one assumed that the accumulation process generated some biological signal, one might postulate that there would be less such signal in accumulator models on easy trials, as the process would stop sooner on easier trials. On the other hand, the attractor model predicts larger BOLD signals on easy vs difficult trials (Rolls et al., 2010a), and on correct than error trials (Rolls et al., 2010b). Moreover, as the process is implemented in an attractor network which also implements short-term memory, the attractor network mechanism also allows the decision choice to be maintained until it is needed, which may be after a delay (Rolls and Deco, 2010). In these senses, the model described in this paper incorporates realistically modeled

neurons, synapses, and cortical architecture, and captures some of the brain processes involved in decision-making in a biologically plausible way. In particular, our aim has been to show that with the natural source of internal noise in the brain, the approximately Poisson spike timing of the neurons, and the way in which decisions are taken in biologically realistically large attractor networks with several thousand neurons and with IF neurons with biologically realistic parameters for the synaptic and neuronal time constants, decisions can be predicted well above chance (68% correct) on a single trial in advance of the evidence being provided to the neuronal network. (It is important to model a system with large numbers of neurons because the noise arises from statistical fluctuations due to the spike timing in a system of finite size; Wang, 2002; Deco and Rolls, 2006; Rolls and Deco, 2010, and therefore systems with biologically realistic numbers of synapses onto each neuron and numbers of neurons in the system were modeled here).

These findings raise interesting issues about free will, a topic of interest in modern neuroscience (Koch, 2009), and the influence of spike timing-related noise on decision-making (Rolls, 2010b; Rolls and Deco, 2010). First, we can note that in so far as the brain operates with some degree of randomness due to the statistical fluctuations produced by the random spiking times of neurons, brain function is to some extent non-deterministic, as defined in terms of these statistical fluctuations. That is, the behavior of the system, and of the individual, can vary from trial to trial based on these statistical fluctuations (Rolls and Deco, 2010). Indeed, given that each neuron has this randomness; that there are sufficiently small numbers of synapses on the neurons in each network (between a few thousand and 20,000) that these statistical fluctuations are not smoothed out; that there are a number of different networks involved in typical thoughts and actions each one of which may behave probabilistically; and with  $10^{11}$  neurons in the brain each within the order of  $10^4$  synapses, the system has so many degrees of freedom that it operates effectively as a non-deterministic system. (Philosophers may wish to argue about different senses of the term deterministic, but it is being used here in a precise and quantitative way, which has been defined within the framework of stochastic neurodynamics; Rolls and Deco, 2010.) These effects of internal noise are especially evident when the evidence for a particular decision is weak, that is when  $\Delta I$  is close to zero, and as  $\Delta I$  increases, the choice increases toward 100% for one of the stimuli, and the choice variability and the influence of the noise decreases (Rolls and Deco, 2010). Second, the findings described here indicate that internal noise in the brain, due for example to random neuronal firing times for a given mean firing rate, can place the system into a state that influences the decision that will be taken before the evidence is even made available, and this may appear surprising in relation to feelings we have about free will and how we take decisions (Rolls, 2010b).

## ACKNOWLEDGMENTS

Gustavo Deco received support from the McDonnell Centre for Cognitive Neuroscience at Oxford University, from European Union grant EC005-024, from the Spanish Research Project SAF2010-16085, and from the CONSOLIDER-INGENIO 2010 Programme CSD2007-00012. The research was supported by the Oxford Centre for Computational Neuroscience.

## REFERENCES

- Abeles, M. (1991). *Corticomics – Neural Circuits of the Cerebral Cortex*. New York: Cambridge University Press.
- Amit, D. J., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252.
- Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., and Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron* 60, 1142–1152.
- Boly, M., Balteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., Phillips, C., Peigneux, P., Maquet, P., and Laureys, S. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12187–12192.
- Brunel, N., and Wang, X. J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63–85.
- Deco, G., and Rolls, E. T. (2005). Attention, short-term memory, and action selection: a unifying theory. *Prog. Neurobiol.* 76, 236–256.
- Deco, G., and Rolls, E. T. (2006). Decision-making and Weber's Law: a neurophysiological model. *Eur. J. Neurosci.* 24, 901–916.
- Deco, G., Rolls, E. T., and Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Prog. Neurobiol.* 88, 1–16.
- Deco, G., Scarano, L., and Soto-Faraco, S. (2007). Weber's law in decision making: integrating behavioral data in humans with a neurophysiological model. *J. Neurosci.* 27, 11192–11200.
- Ditterich, J. (2006). Stochastic models of decisions about motion direction: behavior and physiology. *Neural Netw.* 19, 981–1012.
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485.
- Hampton, A. N., and O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1377–1482.
- Hanks, T. D., Ditterich, J., and Shadlen, M. N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* 9, 682–689.
- Haynes, J. D., and Rees, G. (2005a). Predicting the stream of consciousness from activity in human visual cortex. *Curr. Biol.* 15, 1301–1317.
- Haynes, J. D., and Rees, G. (2005b). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat. Neurosci.* 8, 686–691.
- Haynes, J. D., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534.
- Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., and Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Curr. Biol.* 17, 323–328.
- Hesselmann, G., Kell, C. A., Eger, E., and Kleinschmidt, A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10984–10989.
- Hesselmann, G., Sadaghiani, S., Friston, K. J., and Kleinschmidt, A. (2010). Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PLoS ONE* 5, e9926. doi: 10.1371/journal.pone.0009926
- Kiani, R., Hanks, T. D., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.
- Koch, C. (2009). "Free will, physics, biology, and the brain," in *Downward Causation and the Neurobiology of Free Will*, eds N. Murphy, G. F. R. Ellis, and T. O'Connor (Berlin: Springer), 31–52.
- Laming, D. (1968). *Information Theory of Choice Reaction Time*. Wiley, Chichester.
- Lau, H. C., Rogers, R. D., and Passingham, R. E. (2006). On measuring the perceived onsets of spontaneous actions. *J. Neurosci.* 26, 7265–7271.
- Loh, M., Rolls, E. T., and Deco, G. (2007). A dynamical systems hypothesis of schizophrenia. *PLoS Comput. Biol.* 3, e228. doi: 10.1371/journal.pcbi.0030228
- Marti, D., Deco, G., Mattia, M., Gigante, G., and Del Giudice, P. (2008). A fluctuation-driven mechanism for slow decision processes in reverberant networks. *PLoS ONE* 3, e2534. doi: 10.1371/journal.pone.0002534
- Mattia, M., and Del Giudice, P. (2004). Finite-size dynamics of inhibitory and excitatory interacting spiking neurons. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 70(5 Pt 1), 052903.
- Mazurek, M. E., Roitman, J. D., Ditterich, J., and Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex* 13, 1257–1269.
- Pessoa, L., and Padmala, S. (2005). Quantitative prediction of perceptual decisions during near-threshold fear detection. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5612–5617.
- Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.
- Ratcliff, R., Van Zandt, T., and McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychol. Rev.* 106, 261–300.
- Ress, D., Backus, B. T., and Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nat. Neurosci.* 3, 940–945.
- Rolls, E. T. (2008). *Memory, Attention, and Decision-Making: A Unifying Computational Neuroscience Approach*. Oxford: Oxford University Press.
- Rolls, E. T. (2010a). Attractor networks. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 119–134.
- Rolls, E. T. (2010b). "Noise in the brain, decision-making, determinism, free will, and consciousness," in *New Horizons in the Neuroscience of Consciousness*, eds E. Perry, D. Collerton, F. Lebeau, and H. Ashton (Amsterdam: John Benjamins), 113–120.
- Rolls, E. T., and Deco, G. (2002). *Computational Neuroscience of Vision*. Oxford: Oxford University Press.
- Rolls, E. T., and Deco, G. (2010). *The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function*. Oxford: Oxford University Press.
- Rolls, E. T., Grabenhorst, F., and Deco, G. (2010a). Choice, difficulty, and confidence in the brain. *Neuroimage* 53, 694–706.
- Rolls, E. T., Grabenhorst, F., and Deco, G. (2010b). Decision-making, errors, and confidence in the brain. *J. Neurophysiol.* 104, 2359–2374.
- Rolls, E. T., Grabenhorst, F., and Franco, L. (2009). Prediction of subjective affective state from brain activations. *J. Neurophysiol.* 101, 1294–1308.
- Rolls, E. T., Loh, M., and Deco, G. (2008). An attractor hypothesis of obsessive-compulsive disorder. *Eur. J. Neurosci.* 28, 782–793.
- Roxin, A., and Ledberg, A. (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4, e1000046. doi: 10.1371/journal.pcbi.1000046
- Sadaghiani, S., Hesselmann, G., Friston, K. J., and Kleinschmidt, A. (2010). The relation of ongoing brain activity, evoked neural responses, and cognition. *Front. Syst. Neurosci.* 4:20. doi: 10.3389/fnsys.2010.00020
- Salinas, E. (2003). Background synaptic activity as a switch between dynamical states in a network. *Neural Comput.* 15, 1439–1475.
- Smerieri, A., Rolls, E. T., and Feng, J. (2010). Decision reaction time, slow inhibition, and theta rhythm. *J. Neurosci.* 30, 14173–14181.
- Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.
- Soon, C. S., Brass, M., Heinze, H. J., and Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543–545.
- Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592.
- Vickers, D. (1979). *Decision Processes in Visual Perception*. New York: Academic Press.
- Vickers, D., and Packer, J. (1982). Effects of alternating set for speed or accuracy on response time, accuracy and confidence in a unidimensional discrimination task. *Acta Psychol.* 50, 179–197.
- Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.
- Wang, X. J. (2008). Decision making in recurrent neuronal circuits. *Neuron* 60, 215–234.
- Yang, T., and Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature* 447, 1075–1080.
- Zhang, J., and Bogacz, R. (2010). Optimal decision making on the basis of evidence represented in spike trains. *Neural Comput.* 22, 1113–1148.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 September 2010; accepted: 02 March 2011; published online: 17 March 2011.

Citation: Rolls ET and Deco G (2011) Prediction of decisions from noise in the brain before the evidence is provided. *Front. Neurosci.* 5:33. doi: 10.3389/fnrs.2011.00033

This article was submitted to *Frontiers in Decision Neuroscience*, a specialty of *Frontiers in Neuroscience*.

Copyright © 2011 Rolls and Deco. This is an open-access article subject to an exclusive license agreement between the authors and Frontiers Media SA, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

## APPENDIX

### THE NETWORK MODEL

We used the mathematical formulation of integrate-and-fire (IF) neurons and synaptic currents described by Brunel and Wang (2001). Here we provide a brief summary of this framework, which we have extended to multiple interacting networks. The dynamics of the sub-threshold membrane potential  $V$  of a neuron are given by the equation:

$$C_m \frac{dV(t)}{dt} = -g_m(V(t) - V_L) - I_{\text{syn}}(t)$$

where  $C_m$  is the membrane capacitance taken to be 0.5 nF for excitatory neurons and 0.2 nF for inhibitory neurons;  $g_m$  is the membrane leak conductance taken to be 25 nS for excitatory neurons and 20 nS for inhibitory neurons;  $V_L$  is the resting potential of  $-70$  mV and  $I_{\text{syn}}$  is the synaptic current. The firing threshold is taken to be  $V_{\text{thr}} = -50$  mV and the reset potential  $V_{\text{reset}} = -55$  mV. (see McCormick et al., 1985).

The synaptic current is given by a sum of glutamatergic, AMPA ( $I_{\text{AMPA,rec}}$ ), and NMDA ( $I_{\text{NMDA,rec}}$ ) mediated, recurrent excitatory currents, one AMPA ( $I_{\text{AMPA,ext}}$ ) mediated external excitatory current and one inhibitory GABAergic current ( $I_{\text{GABA}}$ ):

$$I_{\text{syn}}(t) = I_{\text{AMPA,ext}}(t) + I_{\text{AMPA,rec}}(t) + I_{\text{NMDA,rec}}(t) + I_{\text{GABA}}(t).$$

The synaptic currents are defined by:

$$I_{\text{AMPA,ext}}(t) = g_{\text{AMPA,ext}}(V(t) - V_E) \sum_{j=1}^{N_{\text{ext}}} s_j^{\text{AMPA,ext}}(t)$$

$$I_{\text{AMPA,rec}}(t) = g_{\text{AMPA,rec}}(V(t) - V_E) \sum_{j=1}^{N_E} w_j s_j^{\text{AMPA,rec}}(t)$$

$$I_{\text{NMDA,rec}}(t) = \frac{g_{\text{NMDA,rec}}(V(t) - V_E)}{1 + [\text{Mg}^{++}] \exp(-0.062V(t))/3.57} \times \sum_{j=1}^{N_E} w_j s_j^{\text{NMDA}}(t)$$

$$I_{\text{GABA}}(t) = g_{\text{GABA}}(V(t) - V_I) \sum_{j=1}^{N_I} s_j^{\text{GABA}}(t)$$

where  $V_E = 0$  mV,  $V_I = -70$  mV,  $w_j$  are the synaptic weights, each receptor has its own fraction  $s_j$  of open channels, and its own synaptic conductance  $g$ . The NMDA synaptic current is dependent on the potential and controlled by the extracellular concentration of magnesium [ $\text{Mg}^{++} = 1$  mM] (Jahr and Stevens 1990). The values for the synaptic conductances for excitatory neurons are  $g_{\text{AMPA,ext}} = 2.08$  nS,  $g_{\text{AMPA,rec}} = 0.208$  nS,  $g_{\text{NMDA}} = 0.654$  nS, and  $g_{\text{GABA}} = 2.5$  nS; and for inhibitory neurons  $g_{\text{AMPA,ext}} = 1.62$  nS,  $g_{\text{AMPA,rec}} = 0.162$  nS,  $g_{\text{NMDA}} = 0.516$  nS, and  $g_{\text{GABA}} = 1.946$  nS. These

values are obtained from the ones used in Brunel and Wang (2001) by multiplication by a factor which corrects for the difference in number of neurons used in our and Brunel and Wang's model. In their work the conductances were calculated so that in an unstructured network the excitatory neurons have a spontaneous spiking rate of 3 Hz and the inhibitory neuron 0073 a spontaneous rate of 9 Hz.

The fractions of open channels are described by:

$$\frac{ds_j^{\text{AMPA,ext}}(t)}{dt} = -\frac{s_j^{\text{AMPA,ext}}(t)}{\tau_{\text{AMPA}}} + \sum_k \delta(t - t_j^k)$$

$$\frac{ds_j^{\text{AMPA,rec}}(t)}{dt} = -\frac{s_j^{\text{AMPA,rec}}(t)}{\tau_{\text{AMPA}}} + \sum_k \delta(t - t_j^k)$$

$$\frac{ds_j^{\text{NMDA}}(t)}{dt} = -\frac{s_j^{\text{NMDA}}(t)}{\tau_{\text{NMDA,decay}}} + \alpha x_j(t)(1 - s_j^{\text{NMDA}}(t))$$

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_{\text{NMDA,rise}}} + \sum_k \delta(t - t_j^k)$$

$$\frac{ds_j^{\text{GABA}}(t)}{dt} = -\frac{s_j^{\text{GABA}}(t)}{\tau_{\text{GABA}}} + \sum_k \delta(t - t_j^k),$$

where the rise time constant for NMDA synapses is  $\tau_{\text{NMDA,rise}} = 2$  ms (Hestrin et al., 1990; Spruston et al., 1995), the rise time constants for AMPA and GABA are neglected because they are smaller than 1 ms, and  $\alpha = 0.5 \text{ ms}^{-1}$ . All synapses have a delay of 0.5 ms. The decay time constant for the AMPA synapses is  $\tau_{\text{AMPA}} = 2$  ms (Hestrin et al., 1990; Spruston et al., 1995), for NMDA synapses is  $\tau_{\text{NMDA,decay}} = 100$  ms (Hestrin et al., 1990; Spruston et al., 1995), and for GABA synapses  $\tau_{\text{GABA}} = 10$  ms (Salin and Prince 1996; Xiang et al., 1998). The sums over  $k$  represent a sum over spikes formulated as  $\delta$ -Peaks [ $\delta(t)$ ] emitted by presynaptic neuron  $j$  at time  $t_j^k$ .

## REFERENCES

- Hestrin, S., Sah, P., and Nicoll, R. A. (1990). Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron* 5, 247–53.
- Jahr, C. E., and Stevens, C. F. (1990). Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *J. Neurosci.* 10, 3178–3182.
- McCormick, D. A., Connors, B. W., Lighthall, J. W., and Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J. Neurophysiol.* 54, 782–806.
- Salin, P. A., and Prince, D. A. (1996). Spontaneous GABA-A receptor-mediated inhibitory currents in adult rat somatosensory cortex. *J. Neurophysiol.* 75, 1573–1588.
- Spruston, N., Jonas, P., and Sakmann, B. (1995). Dendritic glutamate receptor channels in rat hippocampal CA3 and CA1 pyramidal neurons. *J. Physiol.* 482, 325–352.
- Xiang, Z., Huguenard, J. R., and Prince, D. A. (1998). GABA-A receptor-mediated currents in interneurons and pyramidal cells of rat visual cortex. *J. Physiol.* 506, 715–730.