SUPPLEMENTARY MATERIALS
Synaptic Computation Underlying Probabilistic Inference

Alireza Soltani[1,2] and Xiao-Jing Wang[1]

[1]*Department of Neurobiology and Kavli Institute of Neuroscience, Yale University School of Medicine, New Haven, Connecticut, 06510*

[2]*Division of Biology, California Institute of Technology, Pasadena, California, 91125*

Nature Neuroscience: doi:10.1038/nn.2450

In the following sections we provide some additional details about the model as well as some results which are of interest to the reader.

Supplementary Figures

Two of the shapes used in the weather prediction task are assigned a infinite WOE. That is the presence of any of these shapes alone (which are called trump shapes) is fully predictive of the reward outcome on one of the two choice alternatives. Nevertheless, it has been shown that the presence of these shapes does not completely determine the monkeys choice or in other words, trump shapes exert finite weights on decision-making processes [1]. In order to show that this is also the case in our model, we computed the probability that alternative A is selected, for all patterns which contain one of the two trump shapes. We found that the probability of selecting A depends on the evidence provided by the non-trump shapes in these patterns (**Fig.S1**). Therefore, even in the presence of a trump shape in a given pattern the evidence provided by other shapes in that pattern contributes to decision making.

We also show in the main text that the influence of each shape on decision making can be extracted from the average choice behavior using a logistic regression fit. In order to test whether this influence depends on the epoch in which a shape is presented, we performed the same analysis using a logistic regression fit with a term for each of the 10 shapes in each of four epochs.

$$P_A = \frac{10^Q}{1+10^Q} \quad \text{where} \quad Q = \sum_{i=1}^{10}\sum_{j=1}^{4} q_{ij}N_{ij} \tag{S1}$$

where P_A is the probability of choosing alternative A, N_{ij} is equal to 1 if shape S_i is presented in epoch j (and zero otherwise) for a given pattern, and regression coefficient q_{ij} is the subjective

Nature Neuroscience: doi:10.1038/nn.2450

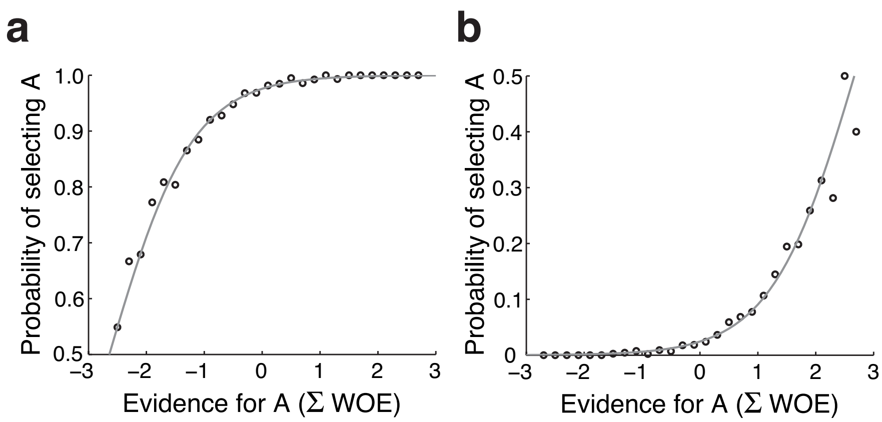

Figure S 1: Influence of trump shapes on choice behavior is finite. Probability of selecting alternative A is computed for all patterns which contain one repetition of the trump shape which is predictive of alternative A (**a**) or B (**b**), as a function of the evidence provided by the rest of the shapes in those patterns (i.e. the sum of the WOE of the non-trump shapes). Although the presence of a trump shape strongly biases the choice behavior toward one of the two alternatives, the evidence provided by the rest of shapes influences the choice selection.

Nature Neuroscience: doi:10.1038/nn.2450

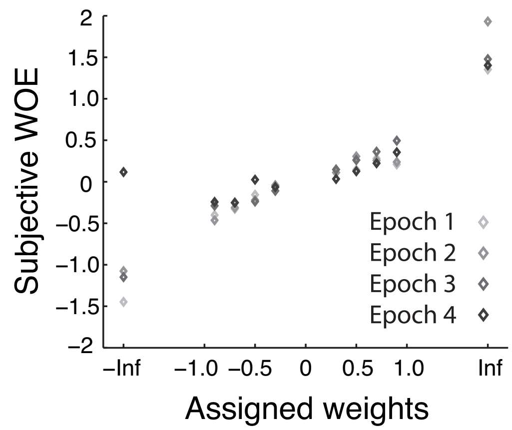

Figure S2: Effect of shapes on decision making is independent of the epoch. The SWOEs are the coefficients of logistic regression fit of probability of choosing A (equation(S1)). There are four diamonds for each shape which show the SWOEs for four different epochs.

weight of evidence (SWOE) for shape S_i in epoch j. We found that regression with 40 parameters does not provide a much better fit than regression with only 10 parameters. Moreover, the result of this fit indicated that the influence of each shape on decision processes is independent of the epoch in which the shape is presented (**Fig.S2**).

Also, we argue that the SWOEs are less than the assigned WOEs because of the concurrence of different shapes and especially the presence of trump shapes on many trials. In order to show this, we reduced the assigned WOEs for the trump shapes (from ∞ and $-\infty$ to 2 and -2, respectively) while leaving the assigned WOE for the non-trump shapes intact. We found that the SWOE as well as the log naive posterior odds of the non-trump shapes are increased due to this alteration (**Fig.S3**). This happens because in the case in which the assigned WOEs for the trump shapes are reduced these shapes have less influence on decision processes and therefore on the

Nature Neuroscience: doi:10.1038/nn.2450

predictive power of the non-trump shapes through learning process. Note that in both cases, the SWOE is linearly proportional to the log naive posterior odds.

As we argue in the main text, the stochastic choice behavior of the model is determined by the overall difference in the synaptic strengths. This is shown in **Fig.S4** where we plot the probability of choosing A as a function the overall difference in the synaptic strengths. We found that the choice probability can be fit as a sigmoid function of the overall difference in the synaptic strengths and this relationship is not influenced by the prior probability that each alternative is assigned a reward (**Fig.S4a-d**).

In the main text we show that the activity of neurons in the decision circuit is modulated by the logLR provided by presented shapes. Another way to show the influence of the logLR on the neural activity is to examine the change of the population activity due to the presentation of a new shape. Here we computed the incremental change of the population firing rate across successive epochs (for each shape presentation) by subtracting the average activity during the last 200 msec of the previous epoch from the activity in the current epoch (**Fig.S5a**). We also computed the average of this change and we found that the average change of activity is proportional to the average change of the logLR (ΔlogLR) caused by the presentation of a new shape (**Fig.S5b**).

In addition to the effect of the logLR on the neural activity, we also examined how the choice on a given trial affects the modulation by the logLR. We found that the average population activity is influenced by both choice and the accumulated logLR over time (**Fig.S6**, see also **Fig.5c** in the main text).

Nature Neuroscience: doi:10.1038/nn.2450

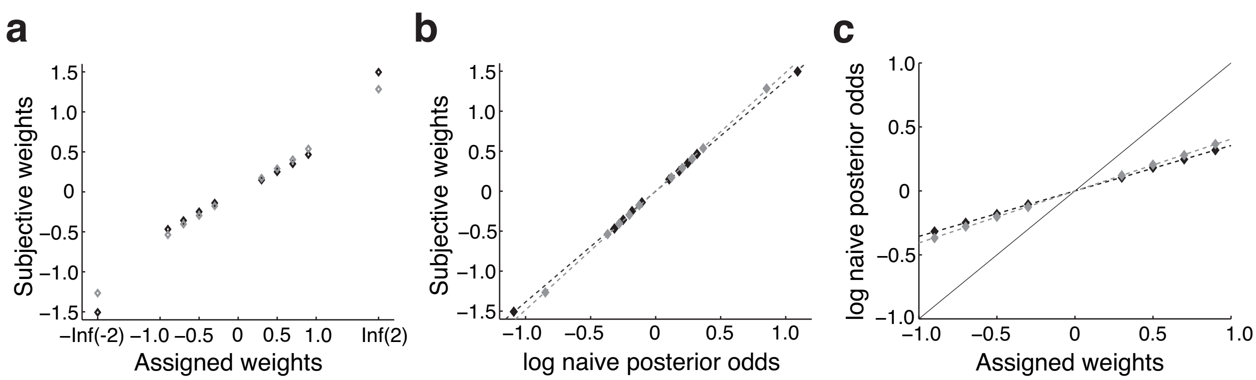

Figure S3: Relationship between the subjective weight of evidence (SWOE), the assigned WOE, and the log naive posterior odds for two cases: normal case (black), the case with a finite WOE for the trump shapes (gray). (**a**) The SWOE extracted from the model's choice behavior as a function of the assigned WOE in each case. The SWOEs for the trump shapes are reduced while the SWOEs for nun-trump shapes are increased in the latter case compared to the normal case. (**b**) SWOE as a function of the log naive posterior odds (which is equal to the NWOE (equation(S5)) when priors are equal). In both cases the SWOE is a linear function of the log naive posterior odds. (**c**) Log naive posterior odds is a linear function of the assigned WOE for non-trump shapes. When the assigned WOE for the trump shapes are reduced, the log naive posterior odds for the non-trump shapes are increased. Note that the relationship between the log naive posterior odds and the assigned WOEs is due to task design and is independent of the model's choice behavior. The dashed lines show the linear fits and the black solid line is the diagonal line.

Nature Neuroscience: doi:10.1038/nn.2450

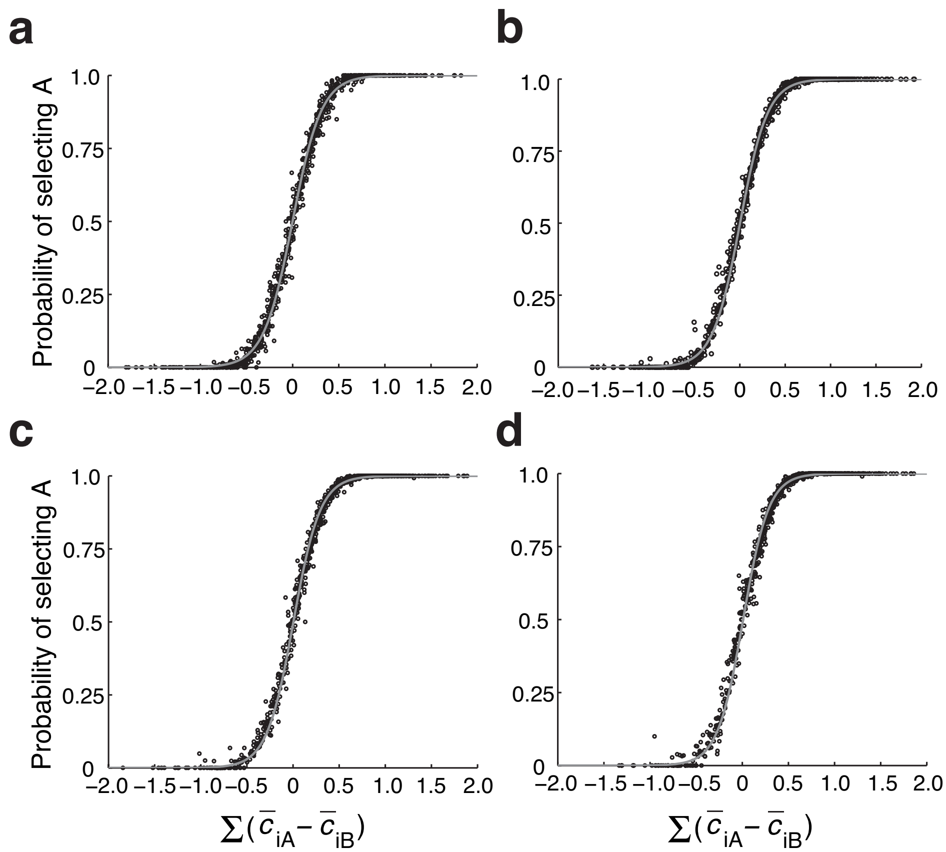

Figure S 4: The probability of choosing A is a sigmoid function of the overall difference in the synaptic strengths and this relationship is not influenced by prior probability. For each set of patterns with a unique WOE, the probability of selecting A is plotted versus the sum of the average difference in the synaptic strengths for that set of patterns. Panels **a** to **d** correspond to the cases in which the prior probability that alternative A is assigned with reward is equal to 0.5, 0.67, 0.75, and 0.8, respectively. The gray curves are the results of the logistic regression fit (equation(S6)). The values of σ obtained from fitting for **a** to **d** are equal to 0.16, 0.15, 0.15, 0.14, respectively.

Nature Neuroscience: doi:10.1038/nn.2450

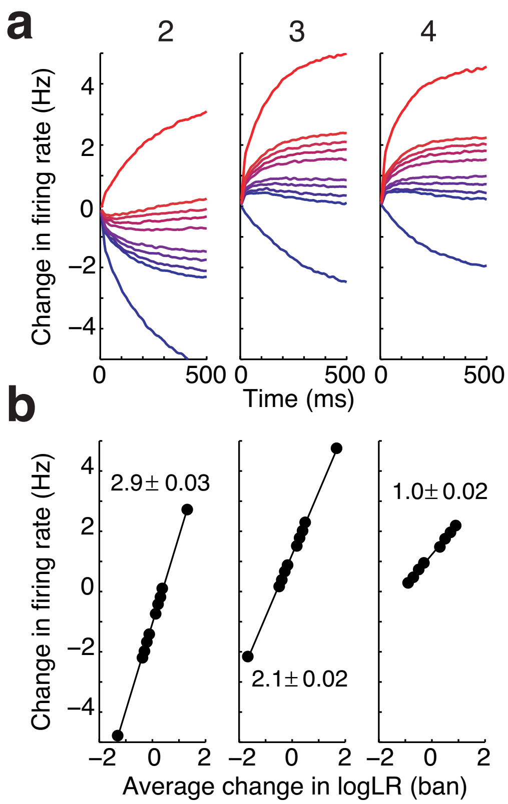

Figure S5: Change of neural activity as a result of each shape presentation. (**a**) The average change in the population activity due to presentation of a new shape is plotted for epoch two, three, and four (indicated at top of the panels). Different color shades from blue to red correspond to shape S_1 to S_{10}. (**b**) Average change of the activity in the last 250 msec of each epoch is plotted as a function of the logLR related to that shape. In this analysis we did not include data related to the presentation of trump shapes in the fourth epoch. The insets show the slope (with estimated s.e.m.) of the linear fit of points in each epoch. Note that part of the decrease in slope for later epochs is because of the fact that change of the logLR due to presentation of any shapes is larger in later epochs. A better approach would be to fit the data according to the NWOEs (equation(S5)) or the log naive posterior odds, the only parameters which possibly can be computed by the subject or a model and do not depend on the epoch.

Nature Neuroscience: doi:10.1038/nn.2450