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*

1

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.

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

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
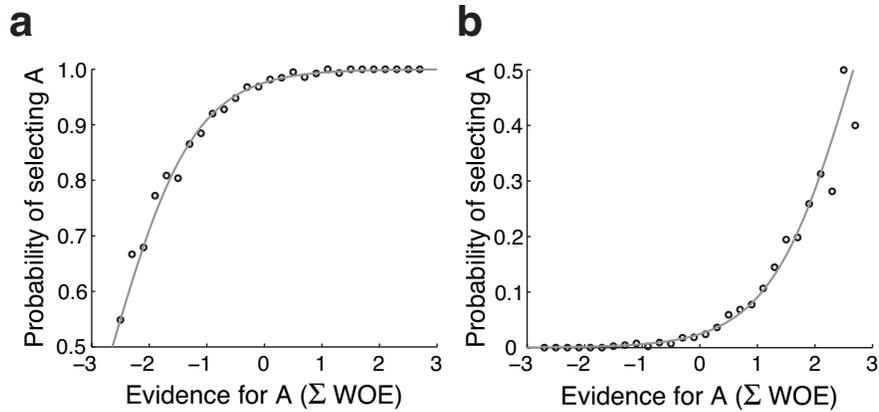
2

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.
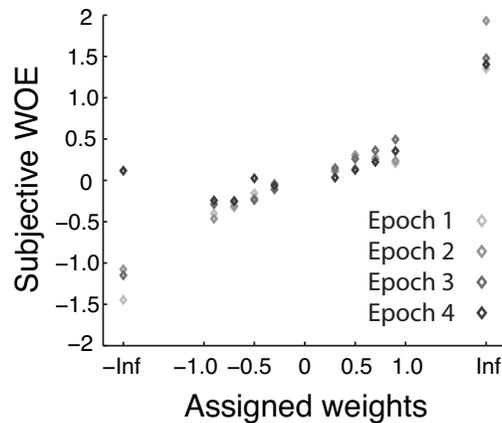
3

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

4

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).
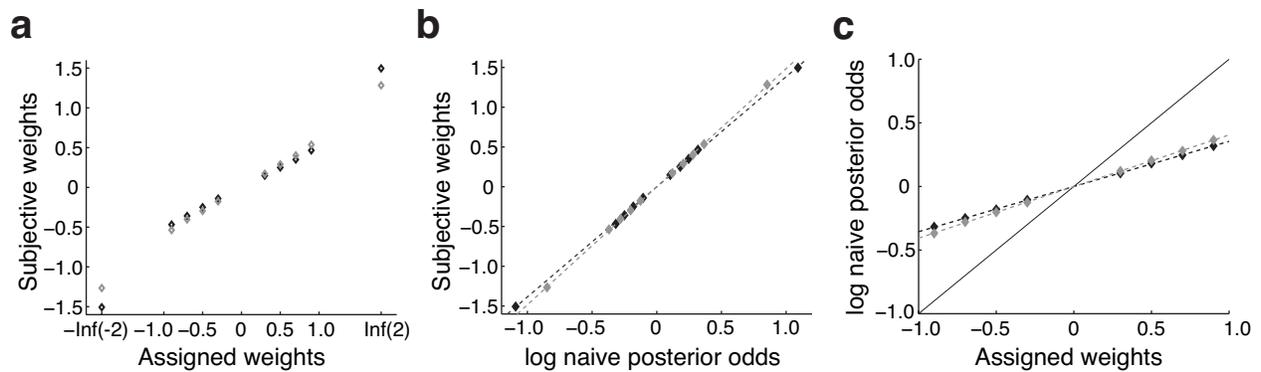
5

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.
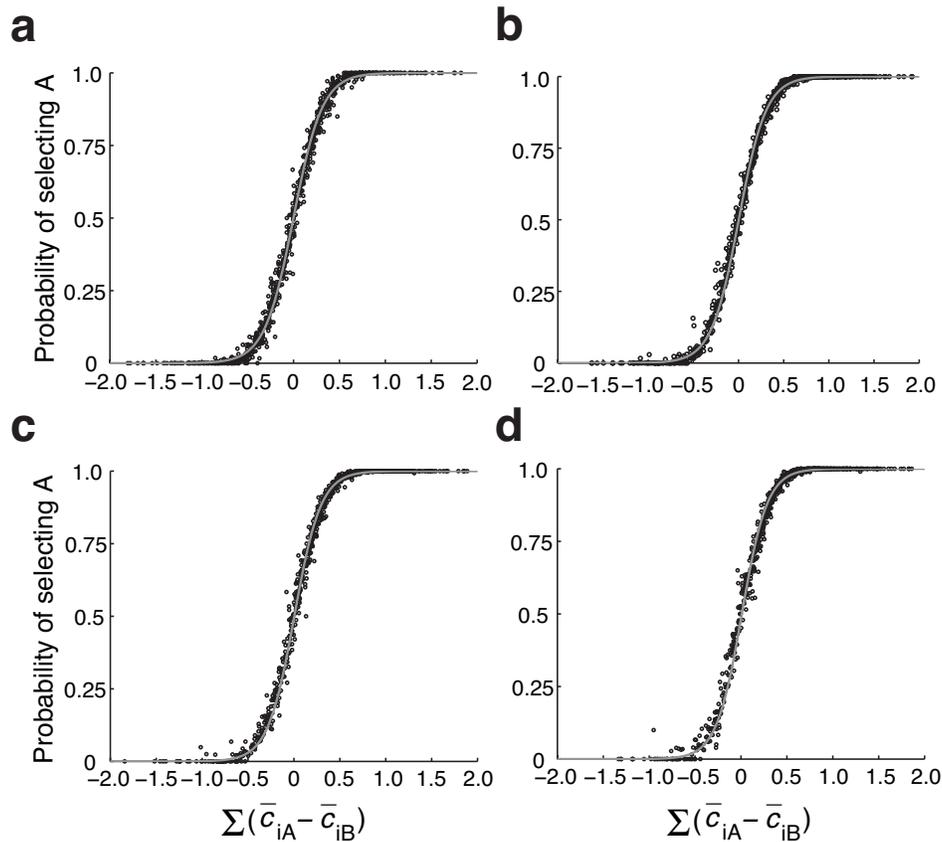
6

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.
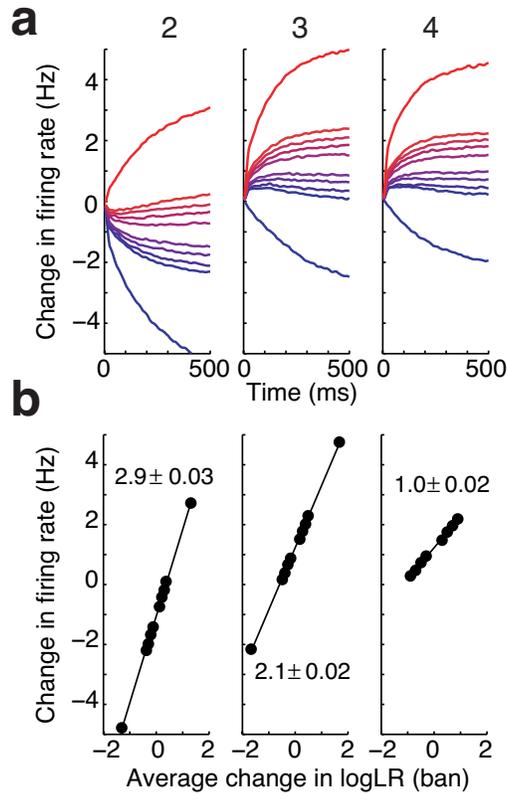
7

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.
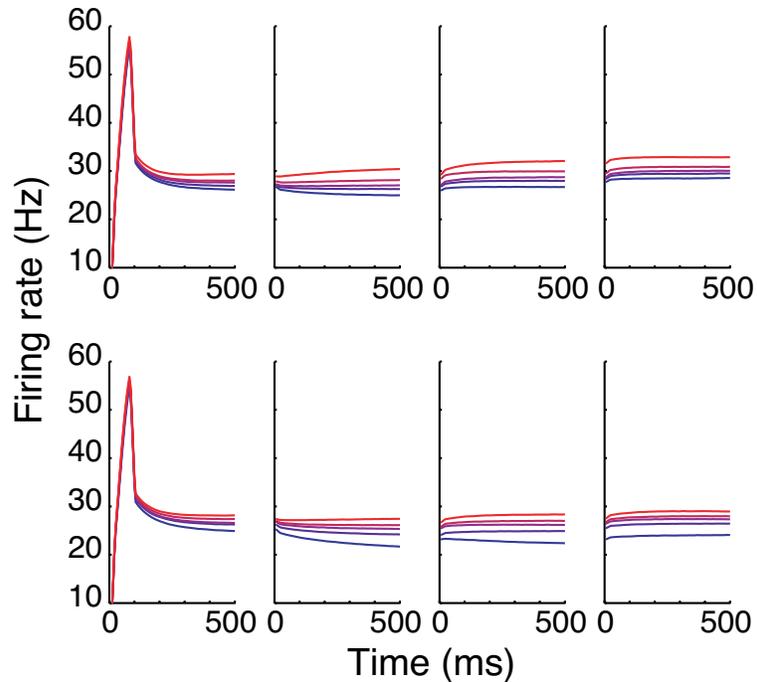
8

Figure S 6: The effect of choice on modulation of neural activity by the logLR. The activity in each epoch is grouped into five quintiles according to the logLR and choice of the model on each trial. Color shadings from blue to red corresponds to larger logLR favoring alternative A. The top (respectively bottom) panels show traces when choice is the preferred (respectively nonpreferred) target of the neural population. On the one hand, if the choice is the preferred target the baseline of activity is higher (specially in later epochs). On the other hand, the modulation of activity by the logLR is larger when the choice is the nonpreferred target (see also **Fig.5c** in the main text).

9

Supplementary Results

Supplementary Note 1:

Steady-state of plastic synapses when multiple cues precede an outcome

In the main text we show that plastic synapses encode posterior probability when a single cue is presented on each trial. Here we show that in the general case in which multiple cues are presented on each trial, plastic synapses encode a quantity proportional to the naive posterior probability (i.e. the probability that a choice alternative is assigned a reward, given a cue is presented in any combination of cues). In order to demonstrate this, we use a weather prediction task in which fours shapes are presented on each trial [1] .

Note that even though there are 10^4 possible patterns (i.e. combination of four shapes) in this task, the reward is assigned to each choice alternative based on the sum of the weight of evidence (WOE) for each pattern, independently of the order in which these shapes are presented on each trial. As a result, there are only 750 unique values for the WOEs of patterns. Lets us refer to all patterns which contain a shape S_i and have the same WOE (equal to t) with C_i^t. Using the same argument as in the case where one shape alone is presented on each trial, we can show that if only this set of patterns is presented, the synaptic strength related to shape S_i and alternative A would be proportional to the posterior probability for this set of patterns, $P(A|C_i^t)$.

However, shape S_i appears in many sets of patterns which have different WOEs. These different sets of patterns drive the synaptic strengths toward different values, and by analogy with

10

equation(1) one can show that the steady-state of the synaptic strength is equal to

$$c_{iA}^{ss} = \frac{\sum_t q_+ P(C_i^t) P_A(C_i^t) P(A|C_i^t)}{\sum_t \left(q_+ P(C_i^t) P_A(C_i^t) P(A|C_i^t) + q_- P(C_i^t) P_A(C_i^t)(1 - P(A|C_i^t)) \right)} \tag{S2}$$

where the sum is over all unique values of the WOE. If the learning rates are equal we get

$$c_{iA}^{ss} = \frac{\sum_t P(C_i^t) P_A(C_i^t) P(A|C_i^t)}{\sum_t P(C_i^t) P_A(C_i^t)} \tag{S3}$$

Note that the denominator in this expression is equal to $\tilde{P}_A(S_i) = \sum_t P(C_i^t) P_A(C_i^t)$, which is the

probability that A is chosen, given the shape S_i is presented in any pattern. From equation(S3) it

is clear that the steady-state of the synaptic strengths depends on the model's choice behavior and

cannot be computed analytically. It is worth noting that in the case in which the choice behavior is

not influenced by plastic synapses (so plastic synapses undergo modification but do not influence

decision making processes), $P_A(C_i^t)$ in equation(S3) cancels out and we get

$$c_{iA}^{ss} = \frac{\sum_t P(C_i^t) P(A|C_i^t)}{\sum_t P(C_i^t)} = \frac{\sum_t P(A, C_i^t)}{\sum_t P(C_i^t)} = \tilde{P}(A|S_i) \tag{S4}$$

where $\tilde{P}(A|S_i)$ is the naive posterior probability (i.e. the probability that A is assigned a reward,

given that shape S_i is presented in any pattern). Therefore, in this case the steady-state of the

synaptic strength is equal to the naive posterior probability.

In general, the steady-state of the synaptic strength for shape S_i and alternative A is a

weighted sum of the posteriors for different patterns which contain shape S_i. This sum is not

exactly equal to the naive posterior probability and moreover, it depends on the prior probabil-

ity. Nevertheless, as we show in the main text, the difference in the steady-state of the synap-

tic strengths is linearly proportional to the log naive posterior odds (**Fig. 2**) due to the fact that

$x - (1 - x) \simeq \log_{10}(x/(1 - x))$, if $0.2 \leq x \leq 0.8$ (**Fig. 1**).

11

Finally, assuming conditional independence between shape presentation in each epoch, one can compute the naive evidence provided by each shape. The naive evidence, $\tilde{P}(S_i|A)$, is equal to the conditional probability that shape S_i is presented in any pattern, given alternative A is assigned a reward. From these likelihoods one can compute the naive weight of evidence (NWOE) which is the log likelihood ratio that a shape is presented in any pattern given an alternative is assigned a reward

$$NWOE_i = \log_{10} \frac{\tilde{P}(S_i|A)}{\tilde{P}(S_i|B)} \tag{S5}$$

This quantity is independent of the prior probability. Because in the original weather prediction task the prior probability for two alternatives are equal [1], the log naive posterior odds is equal to the NWOE

$$\log_{10} \frac{\tilde{P}(A|S_i)}{\tilde{P}(B|S_i)} = \log_{10} \frac{\tilde{P}(S_i|A)}{\tilde{P}(S_i|B)}.$$

Supplementary Note 2:

Relationship between choice behavior and the synaptic strengths

In this section we describe the relationship between the choice behavior and the steady-state of the synaptic strengths. When four shapes are presented on each trial, the choice behavior is determined by the sum of the inputs to decision neurons. We found that although the choice on each trial is stochastic due to neural fluctuations, the probability of selecting A is approximately a sigmoid function of the sum of the difference in synaptic strengths (**Fig.S4**) and can be written as

$$P_A(C^t) = \frac{1}{1 + \exp\left(-\frac{\sum_i (c_{iA}^{ss} - c_{iB}^{ss})}{\sigma}\right)} \tag{S6}$$

12

where $P_A(C^t)$ is the probability of selecting A when a set of patterns C^t (which consists of all patterns with the WOE equal to t) is presented, $1/\sigma$ quantifies the sensitivity of the decision network to the difference in the inputs (with current parameters we found $\sigma \approx 0.15$), and the sum is over all shapes presented in such patterns.

Interestingly, the relationship between the probability of selecting A and the sum of the difference in the synaptic strengths is not influenced by the prior probability that each choice alternative is assigned a reward (**Fig.S4**). In addition, we found that the probability of choosing each alternative is not influenced by the order in which shapes are presented (**Fig.S2**). This is because until the offset of the fixation point the decision network is not in the competition regime and at that point, the activity in two decision pools is determined by the sum of the inputs.

In order to express the choice behavior in terms of behavioral measures such as the log naive posterior odds or log posteriors odds, we need to relate the sum of the difference in the synaptic strengths to these measures. As we show in the main text, the difference in the synaptic strengths related to each shape is linearly proportional to the log naive posterior odds but is also influenced by the prior probability

$$c_{iA}^{ss} - c_{iB}^{ss} = \alpha \log_{10}\left(\frac{\tilde{P}(A|S_i)}{\tilde{P}(B|S_i)}\right) + \beta \log_{10}\left(\frac{P(A)}{P(B)}\right) \tag{S7}$$

where $\tilde{P}(A|S_i)$ is the naive posterior probability and $P(A)$ and $P(B)$ are the prior probabilities that alternatives A and B are assigned a reward, respectively. The values of α and β depend on the network parameters and on the learning rates (see Supplementary Note 4 for details of these dependence) but in general we found the value of β to be negative and $|\beta| < \alpha$ (with current parameters we found $\alpha = 0.48$ and $\beta = -0.31$).

13

We can assess what is learned in each set of synapses by computing the probability of selecting A when one shape alone is presented

$$P_A(S_i) = \frac{1}{1 + \exp\left(-\frac{\alpha \log_{10} \frac{\tilde{P}(A|S_i)}{\tilde{P}(B|S_i)} + \beta \log_{10} \frac{P(A)}{P(B)}}{\sigma}\right)} \qquad \text{(S8)}$$

where $P_A(S_i)$ is the probability of selecting A when shape S_i is presented alone. As can be seen from the last equation, the choice behavior depends on the log naive posteriors but is also biased by the prior information. We define the bias in the choice behavior to be equal to the probability of selecting A when the log naive posterior odds is equal to zero

$$B_s = P_A(S_i)_{|\tilde{P}(A|S_i)=\tilde{P}(B|S_i)} = \frac{1}{1 + \exp\left(-\frac{\beta}{\sigma} \log_{10} \frac{P(A)}{P(B)}\right)} \qquad \text{(S9)}$$

Because β is negative, if $P(A) > P(B)$ we get $B_s < 0.5$ and therefore, the choice behavior is biased toward the less probable choice alternative (i.e. B) if one shape alone is presented.

Note that a bias toward the less probable alternative does not mean that plastic synapses encode prior information in the wrong direction. This can be seen by rewriting equation(S7) in terms of the NWOE (see equation(S5)), a quantity which is independent of the prior

$$c_{iA}^{ss} - c_{iB}^{ss} = \alpha NWOE_i + (\alpha + \beta) \log_{10}\left(\frac{P(A)}{P(B)}\right) \qquad \text{(S10)}$$

The fact that $(\alpha + \beta) > 0$ shows that information encoded by plastic synapses are positively influenced by prior information, but this influence is not strong enough (i.e. $(\alpha + \beta) < \alpha$) and results in a bias toward the less probable alternative when one shape alone is presented.

When four shapes are presented in a given pattern, the sum of the difference in the synaptic

14

strengths is equal to

$$\sum_i (c_{iA}^{ss} - c_{iB}^{ss}) = \alpha \sum_i \log_{10} \left(\frac{\tilde{P}(A|S_i)}{\tilde{P}(B|S_i)} \right) + 4\beta \log_{10} \left(\frac{P(A)}{P(B)} \right) \tag{S11}$$

where the sum is over all shapes in that pattern (note that the factor 4 is due to the fact that there are four shapes in each pattern). Because in this task the WOE for a given pattern is equal to the sum of the WOE of the shapes in that pattern, the sum of the log naive posterior odds of shapes in a set of patterns with a unique sum of the WOEs is linearly proportional to the log posterior odds for that set but is also influenced by log prior odds (**Fig.S7a**)

$$\sum_i \log_{10} \left(\frac{\tilde{P}(A|S_i)}{\tilde{P}(B|S_i)} \right) = \gamma \log_{10} \left(\frac{P(A|C^t)}{P(B|C^t)} \right) + \lambda \log_{10} \left(\frac{P(A)}{P(B)} \right) \tag{S12}$$

where $P(A|C^t)$ is the posterior probability that A is assigned a reward, given that a set of patterns C^t is presented, $\gamma = 0.36$, and $\lambda = 3.64$. Note that the relationship between the sum log naive posterior odds and the log posterior odds is due to task design and it is independent of the model's choice behavior. These linear relationships hold true for all sets of patterns with finite log posterior odds. For sets of patterns with infinite log posterior odds (i.e. those which contain unbalanced number of trump shapes) the choice behavior is not stochastic and one of the two choice alternatives is selected all the time. Therefore, these patterns influence parts of the psychometric function which are close to 0 and 1 and do not contribute to what follows.

Combining the previous equations we can express the choice behavior in terms of the log posterior and the log prior odds, when a pattern of shapes is presented

$$P_A(C^t) = \frac{1}{1 + \exp\left(-\frac{\alpha\gamma \log_{10} \frac{P(A|C^t)}{P(B|C^t)} + (\alpha\lambda + 4\beta) \log_{10} \frac{P(A)}{P(B)}}{\sigma} \right)} \tag{S13}$$
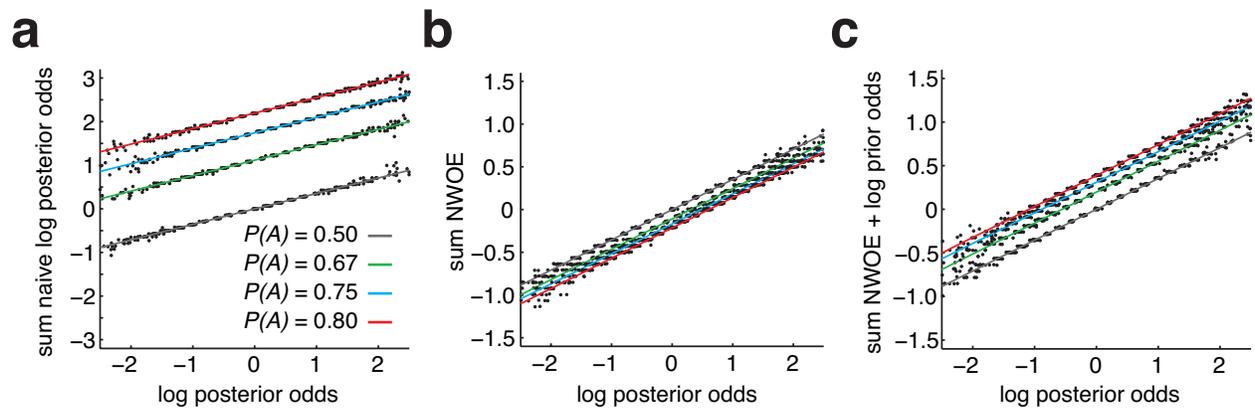
15

Figure S7: Relationship between the log posterior odds, the sum of log naive posterior odds, and the sum of the NWOEs from all shapes in each set of patterns with a finite WOE. (**a**) Log posterior odds for each pattern can be estimated by the sum of the log naive posterior odds of shapes in that pattern. Plotted is the sum of the log naive posterior odds from shapes in a pattern as a function of the log posterior odds for that pattern. The solid lines show the linear fit for each set of points corresponding to different values of prior probability (indicated by the inset). (**b**) Sum of the NWOEs from shapes in a pattern as a function of the log posterior odds for that pattern (the same convention as in **a**). (**c**) Sum of the NWOEs from shapes in a pattern plus the log prior odds as a function of the log posterior odds for that pattern (the same convention as in **a**). Estimating the log posterior odds for each pattern by adding the NWOEs and log prior odds is biased toward the more probable alternative.

16

where $P_A(C^t)$ is the probability that A is selected, given that a set of patterns C^t is presented. The last equation explains why the psychometric function is a sigmoid function of the log posterior odds (and similarly of the sum WOE). It also shows that the choice behavior is biased by the prior probability. We define the bias in the psychometric function to be equal to the probability of selecting A when the log posterior odds is equal to zero

$$B_p = P_A(C^t)_{|P(A|C^t)=P(B|C^t)} = \frac{1}{1 + \exp\left(-\frac{(\alpha\lambda+4\beta)}{\sigma} \log_{10} \frac{P(A)}{P(B)}\right)} \tag{S14}$$

Because $(\alpha\lambda + 4\beta) \approx 0.51$ is positive, we get $B_p > 0.5$ if $P(A) > P(B)$ and therefore, the psychometric function is biased toward the more probable choice alternative (i.e. A). As we show in Supplementary Note 4, for different values of model parameters we consistently found $\beta < 0$ and $(\alpha\lambda + 4\beta) > 0$.

Supplementary Note 3:

Performance of Bayesian observers

In the main text, we compare the choice behavior and performance of our model with those of different ideal Bayesian observers. On each trial of the weather prediction task four shapes (out of 10 possible shapes) are presented and one of the two alternatives is rewarded if it is assigned a reward and is selected. Therefore, it is nontrivial how a Bayesian observer learns the evidence associated with each shape in this task. In order to avoid the issue about learning, we assume that Bayesian observes are able to learn the evidence associated with each shape in two different ways, and we evaluate the performance of these observers after learning.

On the one hand, by assuming that the order of shapes in a pattern is irrelevant, the Bayesian

17

observer can compute the likelihood for a given pattern (the probability that a pattern is presented given alternative A is assigned a reward) and then combines it with prior information to obtain posteriors for that pattern. The posteriors then can be used to make decision according to different decision rules; e.g. strict Bayesian (i.e. selecting the alternative with the larger posterior) or 'probability matching' (i.e. selecting each alternative with a probability equal to the posterior). The performance of these observers is shown in in **Fig.S8** and in **Fig.3a** of the main text. In the case in which priors are equal, the SWOEs of the probabilistic Bayesian observer are equal to the assigned WOEs (**Fig.S8a**). The SWOEs for the strict Bayesian observer are infinite. Therefore, while the choice behavior of the probabilistic Bayesian observer may resemble those of monkeys, the SWOE associated to each shape is different from the experimental data. In the case in which priors are not equal, these Bayesian observers show no bias (**Fig.S8b** and **Fig.6b** in the main text). Moreover, the reward rate (i.e. percent harvested of the assigned rewards) of these Bayesian observers are larger than those of our model and monkeys (**Fig.S8c**).

On the other hand, if the Bayesian observer assumes independence between evidence provided by each shape, then this observer can compute the naive likelihood for each shape, $\tilde{P}(S_i|A)$ (the probability that the shape S_i is presented in any pattern, given alternative A is assigned a reward) as well as the prior probability of each alternative being rewarded. Subsequently, he/she can combine this information to compute the posterior probability that each alternative is assigned a reward given a combination of shapes is presented. For example, he/she can add the log naive likelihood ratio (i.e. NWOE) for all shapes in a given pattern to the log prior odds, in order to estimate the log posterior odds for that pattern.

18

We found that the SWOEs of the "alternative" probabilistic Bayesian observer are equal to the log naive posterior odds, while the SWOEs for the alternative strict Bayesian observer are infinite (**Fig.S9**a). When priors are equal the alternative probabilistic Bayesian observer performs similar to our model. When priors are not equal this observer shows similar bias in the choice behavior while its estimate of the evidence associated to each shape is not biased (**Fig.S9c**). These alternative Bayesian observes show bias in the psychometric function because they combine the NWOEs from all shapes in a given pattern with the log prior odds, in order to estimate the log posterior odds for that pattern. To show this we used equation(S12) to calculate the sum NWOEs

$$\sum_i NWOE_i = \sum_i \log_{10}\left(\frac{\tilde{P}(S_i|A)}{\tilde{P}(S_i|B)}\right) = \gamma \log_{10}\left(\frac{P(A|C^t)}{P(B|C^t)}\right) + (\lambda - 4)\log_{10}\left(\frac{P(A)}{P(B)}\right) \text{ (S15)}$$

The last equation shows that while combining the NWOE of shapes in a given pattern with the log prior odds provides an estimate for the log posterior odds for that pattern, this estimate is biased toward the more probable alternative ($\lambda - 4 > -1$, see **Fig.S7c**). Therefore, the alternative Bayesian observers show bias in their choice behavior while their estimate of evidence associate with each shape is not biased. This is due to the fact that evidence in different epochs are not conditionally independent. Finally, the reward rate of our model is larger than the reward rate for the alternative probabilistic Bayesian and is less than the alternative strict Bayesian (**Fig.S9d**).

Supplementary Note 4:

Dependence of choice behavior on model parameters and optimal strategy

In Supplementary Note 2 we show that the probability of selecting alternative A is approximately a sigmoid function of the sum of the difference in the synaptic strengths (equation(S6)). The sensi-
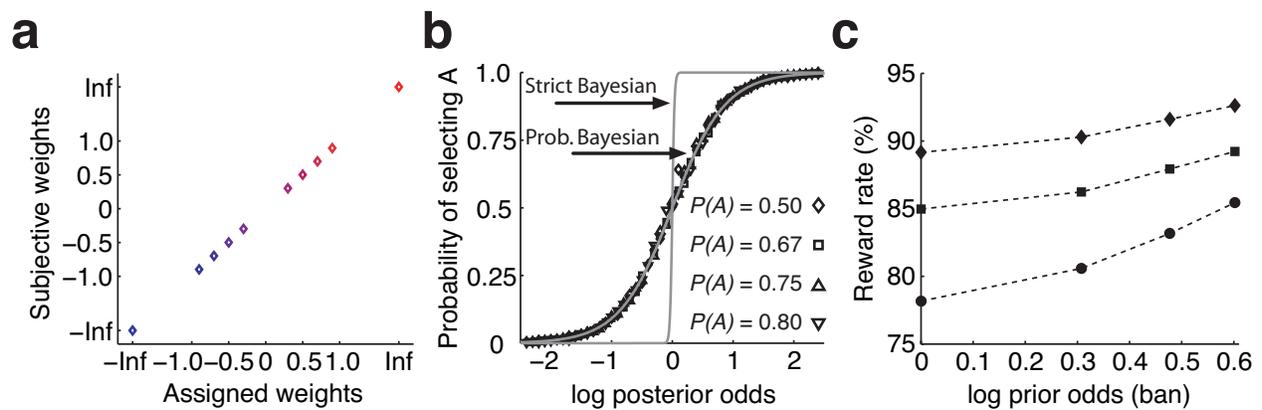
19

Figure S 8: Choice behavior and performance of Bayesian observers. (**a**) Subjective weight of evidence (SWOE) for each shape as a function of the assigned WOE to that shape, for a Bayesian observer who makes decision according to probability matching. For this observer, the SWOEs are equal to the assigned WOEs when the priors are equal. SWOEs for the strict Bayesian observer are plus and minus infinity. (**b**) Psychometric function for patterns with finite log posterior odds is plotted for four values of prior probability (indicated by the inset). The data points show the results for the probabilistic Bayesian observer and gray curves are the logistic function fits for individual observers. (for clarity, data points for the strict Bayesian observer are not shown). The psychometric functions are identical for all values of prior and are not biased toward either alternative. (**c**) The reward rate for the probabilistic (squares) and strict (diamonds) Bayesian observers as well as our model (circles), as a function of the log prior odds. Dashed lines are only to guide the eyes.
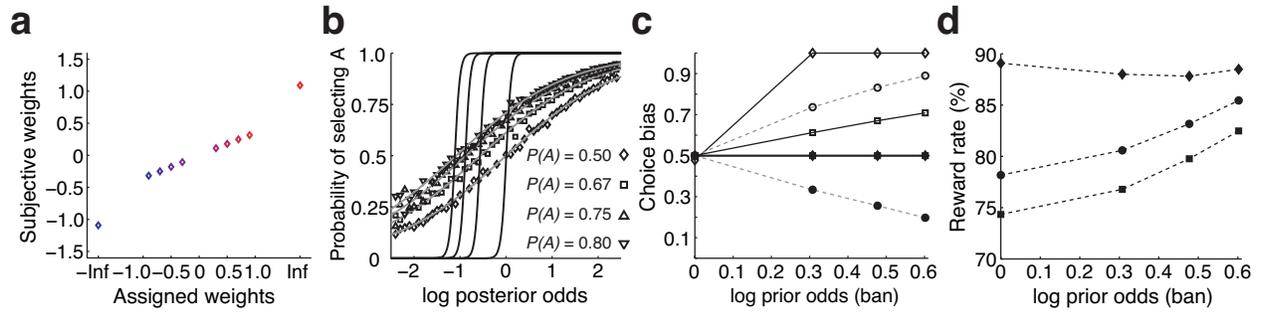
20

Figure S9: Choice behavior and performance of "alternative" Bayesian observers. (**a**) SWOE for each shape as a function of the assigned WOE to that shape, for an alternative Bayesian observer who makes decision according to probability matching. For this observer, the SWOEs are equal to the log naive posterior odds (and NWOEs) when the priors are equal. SWOEs for the alternative strict Bayesian observer are plus and minus infinity. (**b**) Psychometric function of alternative Bayesian observers for four values of prior probability. The data points show the results for the alternative probabilistic Bayesian observer and gray curves are the logistic function fits. The black curves show the fitting results for the alternative strict Bayesian observer (for clarity, data points for this observer are not shown). The psychometric function for both observers are equally biased toward the more probable alternative. (**c**) Choice bias for the alternative strict (diamonds) and probabilistic (squares) Bayesian observers as a function of the log prior odds. Open diamond and squares: bias of the psychometric function. Filled diamonds and squares: bias when one shape is presented alone. For comparison, the bias of the psychometric function (open circles), and of the differential synaptic input (learned evidence) about each shape (filled circles) for our model are also shown. (**d**) The reward rate for the alternative probabilistic (squares) and strict (diamonds) Bayesian observers as well as our model (circles), as a function of the log prior odds. Our model's performance falls between those of Bayesian observers. Dashed lines are only to guide the eyes.

21

tivity of the choice behavior to this quantity is measured by a single parameter $(1/\sigma)$ but it depends on many model parameters such as the input firing rate of cue-selective neurons or the number of plastic synapses. We exploited this result to facilitate exploring the model choice behavior by replacing the decision process on each trial by a simple coin tossing with a probability which is computed according to equation(S6). Using this approach we could explore the dependence of the choice behavior on model parameters only by investigating this dependence on the values of σ and the learning rates, q_+ and q_-.

Firstly, we examined how biases in the choice behavior and the performance (i.e. rate of reward harvest per trial) depend on the learning rates. We found that when the learning rates are small only their ratio strongly influences the choice behavior. **Fig.S10a** shows that both values of α and $|\beta|$, which quantify how evidence about each shape is stored in plastic synapses, increase as the learning rate ratio decreases. These happen because when $q_- > q_+$ plastic synapses transit more often to the depressed than to the potentiated state and as a result, the range of the difference in the synaptic strengths is larger than the case in which $q_+ = q_-$ (the opposite of this happens when $q_+ > q_-$). As a result of increase in the range that plastic synapses represent the log naive posterior odds and log prior odds, both values of α and $|\beta|$ increases. An increase in the value of $|\beta|$ in turn increases the bias in the choice behavior when one shape alone is presented (**Fig.S10b**, note that in equation(S9) the value of B_s is farther from 50% as $-\beta/\sigma$ increases). At the same time, the bias in the psychometric function only decreases slightly due to a slight decrease in $(\alpha\lambda + 4\beta)$ (**Fig.S10b**, note that in equation(S14) the value of B_p is closer to 50% as $(\alpha\lambda + 4\beta)$ decreases). It is worth noting that the bias in the psychometric function goes to zero if $(\alpha\lambda + 4\beta) = 0$. This is

22

not possible in our model because the values of α and β are correlated. Overall, we found that a decrease in the learning rate ratio increases the reward rate (**Fig.S10c**).

Secondly, we examined the dependence of the choice behavior on the sensitivity of the decision network ($1/\sigma$). As the sensitivity of the decision network to the difference in the synaptic strength increases (i.e. smaller value of σ, which can occur by increasing the number of plastic synapses or the input firing rates to these synapses) both values of α and $|\beta|$ decrease (**Fig.S10d**). This happens because for smaller σ values, a small change in the difference in the synaptic strengths can strongly bias the choice behavior. As a result, plastic synapses represent information about the log naive posterior odds and the log prior odds in a smaller range which in turn results in smaller values of α and $|\beta|$. Because the value of $-\beta/\sigma$ determines the bias in the choice behavior when one shape is presented alone (and this value increases as σ decreases), this bias increases with a decrease in the value of σ (**Fig.S10e**). Overall, decreasing σ results in an improvement of performance in terms of reward rate (**Fig.S10f**).

These results point to general biophysical limits for stochastic inference and decision making. On the one hand, there is always a bias in the choice behavior toward the less probable choice alternative if one shape is presented alone (in order to test what is learned about each shape). This bias can be reduced by making decision circuit less sensitive to evidence (information) in the external world, either by increasing the value of σ or by adopting learning rate ratio larger than 1 (i.e. $q_+ > q_-$). Both of these adjustments result in a poorer performance, because the values of α and β are correlated and the choice behavior is stochastic. The values of α and β are correlated in our model because the prior information and evidence about each shape are learned and encoded
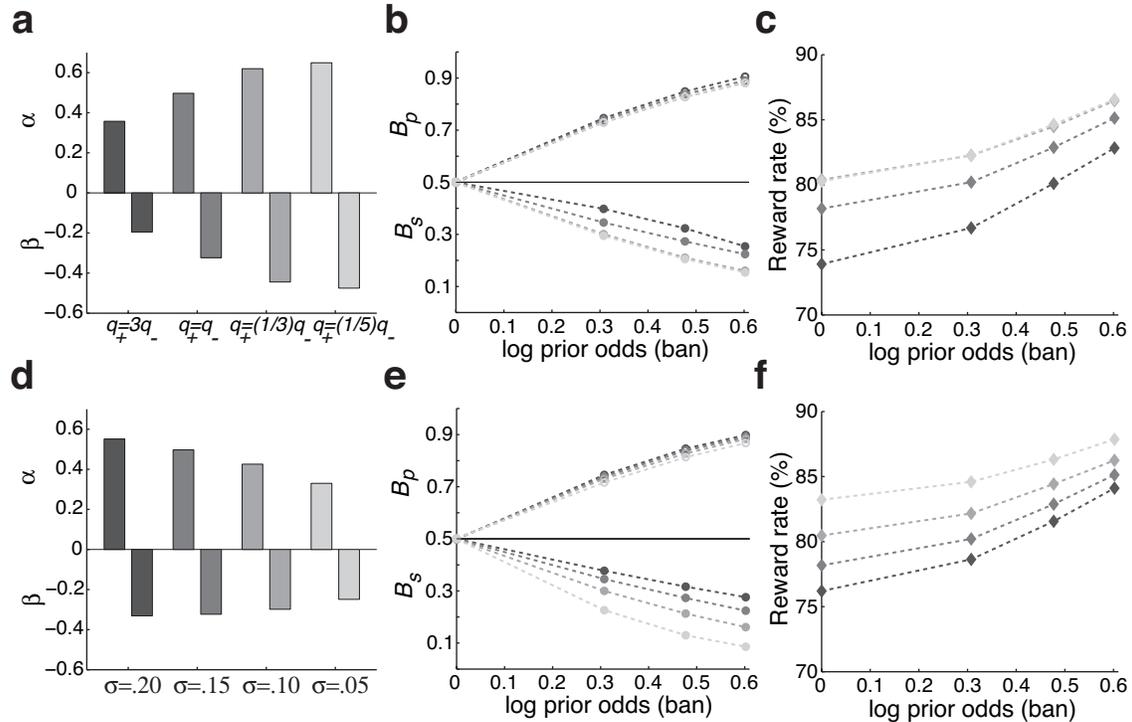
23

Figure S 10: Dependence of the choice bias and performance on the model parameters: learning rates (**a-c**), sensitivity of the decision network (**d-f**) . (**a**) Dependence of the relationship between the difference in the synaptic strengths and the log naive posterior and the log prior odds (quantified with α and β, see equation(S7)) on the learning rate ratio. The learning rates are set to ($q_+ = 0.06, q_- = 0.02$), ($q_+ = 0.02, q_- = 0.02$), ($q_+ = 0.02, q_- = 0.06$), and ($q_+ = 0.01, q_- = 0.05$) in four cases. All bars, markers, and lines in the top (respectively bottom) three panels are coded with different shading corresponding to different values of learning rate ratios (respectively σ) as shown in **a** (respectively **d**). (**b**) The bias in the psychometric function (B_p) and in the choice behavior when one shape is presented alone (B_s), as a function of the log prior odds for different values of the learning rate ratio. (**c**) Reward rate (i.e. percent harvested of the assigned rewards) as a function of the log prior odds. (**d**) to (**f**): similar to **a** to **c** but for different values of σ.

24

in the same set of synapses and there is no separate system to learn prior information. As a result, an increase in the bias of the choice behavior toward the less probable alternative does not reduce the performance but counterintuitively, increases the performance.

On the other hand, the optimal performance in this task could be achieved if the model was extremely sensitive to the difference in the log posterior odds and could always select the choice alternative which has a larger posterior probability. This means that the value of $1/\sigma' = \alpha\gamma/\sigma$ should become very large (see equation(S13)). For this to happen, α should increase and/or σ should decrease (remember that γ is fixed by the task design). But as we showed in **Fig.S10d**, as the value of σ decreases the value of α also decreases (but not as fast as σ) due to the interaction between decision and learning processes. Moreover, a very sensitive decision network can be attracted to a state in which it only selects the alternative which is assigned a reward more often. The same limit also applies to an "alternative" Bayesian observer that employs the naive posterior odds (as an approximation for the posterior odds) to make decision. The difference between such a Bayesian observer and our model is that the former obtains the naive posterior odds for a given pattern by combining the naive likelihood ratios from shapes in that pattern with the prior odds, while the latter directly summates log naive posterior odds from shapes in that pattern. As a result, the choice behavior of an alternative Bayesian observer is not biased toward either alternative when a shape is presented alone, while its psychometric function is biased toward the more probable alternative and its performance is limited by the stochastic nature of the task and decision making processes, similar to our model (**Fig. S9**).

25

Supplementary Note 5:

Model's robustness

In the main text we show that our model is able to perform the weather prediction task. This model is constructed based on certain assumptions and in order to explore the crucial components of the model, we relaxed some of the assumptions and observed the resulting behavior.

One of the assumptions behind the behavioral analysis in the simulated experiment [1] and in our model is that on each trial all presented shapes are registered by the monkeys (i.e. decision maker) and so all shapes influence decision making. More naturally, it is perceivable that on some trials some of the shapes are ignored, especially those which are not very predictive of either outcome. Therefore, an important test for our model is to see if some shapes are not registered randomly on some trials, whether the model is still able to learn and perform the task or not.

In order to test this, we simulated the weather prediction task with the following assumptions. First, a presented shape is registered with a certain probability. Second, only when a shape is registered on a given trial, the corresponding shape-selective population becomes active and influences decision making in that trial. Third, if presentation of a shape is not registered on a given trial, the plastic synapses related to that shape are not updated at the end of that trial (this is a natural outcome of our learning rule).

We found that if the probability of registering a shape is similar for all shapes, the model performs the task as good as before and only the overall sensitivity of the choice behavior to evidence provided by each shape is reduced (i.e. σ' is increased). Moreover, if the probability of registering a shape is proportional to its assigned WOE (so less predictive shapes are registered less often such

26

that the least predictive shapes were registered only 50% of trials) the model is still able to perform the task. In general, the main effect of a shape being registered probabilistically is that the choice behavior is less sensitive to the evidence provided by that shape as expected (proportional to the probability that the shape is missed). Furthermore, we found that the difference in the synaptic strengths in this case is larger than in the case in which all shapes are registered on every trial. This indicates that learning can partially compensate for the effect of shapes not being registered on every trial.

Another assumption of our model was that the model is homogeneous such that the input firing rates from cue-selective to value-coding populations are similar and furthermore, the connections between these populations are identical. More realistically, network can be heterogeneous and these connections can be different. In order to test the robustness against such heterogeneity, we simulated the task by assuming different weights for connections from cue-selective to value-coding populations. We found that if these randomly assigned weights differ by less than 10% for the two value-coding populations, the model performance is not different from the monkey's performance in this task. Overall, the effect of heterogeneity reduces the performance slightly and results in a noisier SWOEs (data not shown).

Finally, we assumed that if a shape is presented more than once on a given trial, the firing rate of the population of neurons selective for that shape is increased by an amount equal to the number of times which that shape is repeated. We examined the importance of this assumption by reducing the effect of shape repetition on the firing rate of cue-selective neurons from 1 to $1/16$ (i.e. repetition factor). That is the repetition of a shape increases the firing rate of the cor-

27

responding cue-selective neurons by 1 to $1/16$ times the amount of activity increase due to the first presentation, respectively. We found that even when the repeated shape evokes $(1/16)$th of the first response, the reward rate is only reduced by less than 5%. Moreover, the model's choice behavior is similar to the monkeys' choice behavior for repetition factors equal to $1/4$ or larger (data not shown). Overall, these results indicated that our model can robustly perform the weather prediction task.

Supplementary Methods

The decision circuit receives three types of inputs. The first are fixed background inputs which mimic the massive projections from other cortical neurons (spiking at spontaneous rate of 3 Hz). These inputs bring all decision neurons close to their firing threshold. The second type are the inputs from value-coding populations that in turn receive their inputs from the sensory neurons through plastic synapses. The sensory neurons become active due to presentation of a new shape and stay active till the shapes disappear. These inputs provide reward-dependent information to all decision populations through plastic synapses (red, green, and blue curves in **Fig.S11**). Moreover, we assume that during synaptic update at the end of each trial, the sensory neurons are reactivated by working memory populations (not modeled) which keep track of the presented shapes in that trial. The third type are purely visual inputs to excitatory populations in the decision circuit which mimic the visual response of neurons in the visual cortex. These inputs also keep the network from entering the competition regime during the presentation of the four shapes and before the extinction of the fixation point. The inputs through value-coding neurons and the purely visual inputs on
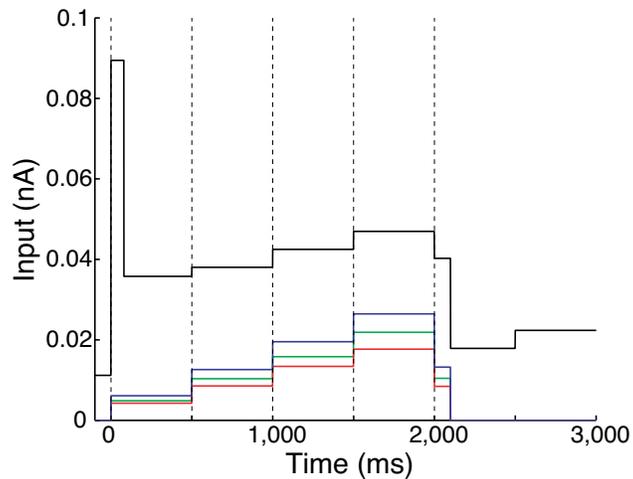
28

Figure S 11: Inputs to decision neurons. The inputs to neurons in the decision-making network consist of fixed background inputs (not shown), a purely visual input (black), and inputs from intermediate value-coding neurons through plastic synapses. Red and green curves show examples of such inputs to two excitatory populations and blue curve shows the input to the inhibitory population.

average constitute about 1.5% and 7% of the total inputs to decision neurons, respectively.

We modeled each of the purely visual inputs as the sum of step functions, starting with an onset of visual response followed by a decrease and then slow increase upon the presentation of each shape (black curve in **Fig.S11**). At the end of the trial, when the fixation point goes off, this visual input is decreased strongly. Consequently, decision network enters the competition regime similar to what has been shown in different network architectures [2–5] . Note that the purely visual inputs are approximately equal to the visual response generated by all visual targets on the screen independently of their reward information.

29

References

1. Yang, T. & Shadlen, M. N. Probabilistic reasoning by neurons. *Nature* **447**, 1075–80 (2007).

2. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci* **26**, 1314–1328 (2006).

3. Wong, K.-F., Huk, A. C., Shadlen, M. N. & Wang, X.-J. Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Front. Comput Neurosci* **1**, 6 (2007).

4. Furman, M. & Wang, X.-J. Similarity effect and optimal control of multiple-choice decision making. *Neuron* **60**, 1153–68 (2008).

5. Liu, F. & Wang, X.-J. A common cortical circuit mechanism for perceptual categorical discrimination and veridical judgment. *PLoS Comput Bio* **4**, e1000253 (2008).