Choice-correlated activity fluctuations underlie learning of neuronal category representation

Tatiana A. Engel[1,2,*], Warasinee Chaisangmongkon[1,*], David J. Freedman[3] & Xiao-Jing Wang[1,4,5]

The ability to categorize stimuli into discrete behaviourally relevant groups is an essential cognitive function. To elucidate the neural mechanisms underlying categorization, we constructed a cortical circuit model that is capable of learning a motion categorization task through reward-dependent plasticity. Here we show that stable category representations develop in neurons intermediate to sensory and decision layers if they exhibit choice-correlated activity fluctuations (choice probability). In the model, choice probability and task-specific interneuronal correlations emerge from plasticity of top-down projections from decision neurons. Specific model predictions are confirmed by analysis of single-neuron activity from the monkey parietal cortex, which reveals a mixture of directional and categorical tuning, and a positive correlation between category selectivity and choice probability. Beyond demonstrating a circuit mechanism for categorization, the present work suggests a key role of plastic top-down feedback in simultaneously shaping both neural tuning and correlated neural variability.

[1] Department of Neurobiology, Yale University School of Medicine, Kavli Institute for Neuroscience, 333 Cedar Street, New Haven, Connecticut 06510, USA. [2] Department of Bioengineering, Stanford University, 318 Campus Drive, Stanford, California 94305, USA. [3] Department of Neurobiology, The University of Chicago, 5812 S. Ellis Ave., Chicago, Illinois 60637, USA. [4] Center for Neural Science, New York University, 4 Washington Place, New York, New York 10003, USA. [5] NYU-ECNU Joint Institute of Brain and Cognitive Science, NYU-Shanghai, Shanghai 200122, China. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to X.-J.W. (email: xjwang@nyu.edu).

Through experience we can learn to classify a continuum of sensory stimuli into discrete meaningful categories, which are critical for guiding behaviour[1,2]. Training improves our ability to discriminate stimuli belonging to different categories and to group together perceptually dissimilar items within the same category. Such learning and refinement of categorical discriminations occur continuously in everyday life; however, their neural basis is poorly understood.

During training on visual tasks, perceptual improvements are accompanied by only moderate tuning changes in the early visual cortex[3,4], whereas more dramatic changes occur in inferior temporal and posterior parietal cortices. In monkeys trained to classify directions of random dot motion into two arbitrary categories, neurons in the lateral intraparietal (LIP) area encoded learned motion categories in an almost binary manner[5], whereas in naive animals LIP neurons represent directions uniformly with bell-shaped tuning functions[6]. In contrast, categorization training did not induce any apparent change in motion tuning of neurons in the middle temporal (MT) area. Similarly, changes in responses of LIP but not MT neurons were associated with improved behavioural sensitivity on visual discrimination tasks[7–9], which had been attributed to refinements of functional connectivity between MT and LIP through reinforcement learning[10,11]; however, the underlying circuit mechanism remains unknown.

We examined whether changes in tuning of LIP neurons induced by training on a motion categorization task can emerge in a neural circuit model through biophysically plausible Hebbian synaptic plasticity modulated by reward prediction error (RPE) signals[12–15]. Unlike the classical two-layer categorization model[16], our model incorporated a layer of neurons intermediate to sensory and decision layers. We found that neurons in the intermediate layer develop stable category representation if fluctuations of their firing rates are correlated with behavioural choices. In contrast, behavioural performance and neuronal tuning deteriorate with training in networks where activity fluctuations are not correlated with choices. Weak but systematic correlations between neural fluctuations and choices, termed choice probability (CP), have been found in many cortical areas[17,18]. Here we show that CP is critical for successful learning through reward-dependent Hebbian plasticity, which generally holds across different network architectures and behavioural tasks.

Our model predicts that a mixture of directional and categorical tuning and bimodal distribution of preferred directions emerge in the intermediate-layer neurons through learning. This prediction was confirmed by analysis of LIP responses recorded in monkeys trained on the motion categorization task. Moreover, the model predicts that neurons with larger CP exhibit a larger increase in their category sensitivity (CS), leading to a positive correlation between these measures, which was also found in the LIP data. Finally, the model suggests that task-specific noise correlations arise from the plasticity of top-down connections and makes testable predictions about changes of noise correlations throughout learning.

Results

A neural circuit model of category learning. We trained a neural circuit model to perform a motion categorization task[5]. Twelve motion directions were assigned to two categories, C1 and C2, defined by an arbitrary category boundary (Fig. 1a), and the model learned through trial and error to decide on the category membership of these stimuli.

Our model is a recurrent neural network comprising three interconnected circuits (Fig. 1b). Sensory neurons (MT) encode motion directions with bell-shaped tuning functions (Fig. 1c), arising from direction-selective bottom-up inputs and structured recurrent excitation[19]. Association neurons (LIP) are also tuned to motion directions initially (Fig. 1c)—just like LIP neurons in naive monkeys[6]—because synaptic weights are initialized to be stronger between sensory and association neurons with similar preferred directions. Over the course of learning, tuning of association neurons changes through synaptic plasticity. The activity of association neurons is pooled by the decision network, which consists of two competing populations (C_1 and C_2, Fig. 1b,c) firing at higher rates for the two respective category decisions[20,21]. These neurons encode the model's choice and represent a subpopulation of neurons within LIP or in the prefrontal cortex. Synaptic connections between association and decision neurons are initialized at random values; therefore, the model's categorization decisions are completely random initially.

Our model has plastic feedforward connections from sensory to association ($c^{S \to A}$) and from association to decision ($c^{A \to D}$) circuits, and plastic feedback connections from decision to association circuit ($c^{D \to A}$, Fig. 1b). At the end of each trial, the strength c of each plastic synapse is updated according to a reward-dependent Hebbian plasticity rule:

$$c \to c + q(R - \langle R | \theta \rangle) r_{\text{pre}} \ r_{\text{post}}, \qquad (1)$$

where r_{pre} and r_{post} are the trial-average firing rates of pre- and postsynaptic neurons, q is the learning rate parameter, R is the reward received on each trial (1 or 0 for correct and incorrect decisions, respectively), θ stands for a motion direction stimulus and $\langle R | \theta \rangle$ is a stimulus-specific reward expectation, which may be encoded in the orbitofrontal cortex or basal ganglia. For simplicity, we computed $\langle R | \theta \rangle$ as a running average of reward history[14]. Phasic activity of dopamine neurons encodes the difference $R - \langle R | \theta \rangle$, called the RPE signal[12,22,23], and dopamine concentration modulates long-term plasticity[24,25]. In our model, positive RPE signals lead to potentiation, while negative RPE signals lead to depression. Finally, the synaptic strengths c are bounded between 0 and 1.

Model learning performance. We compared the learning performance of our model with that of two control networks: a network without feedback, which had only feedforward connections between the local circuits, and a network with fixed tuning of association neurons, which had only feedforward connections and no plasticity of synapses between sensory and association neurons (effectively, a classical two-layer categorization model[16]). Initially, performance of all models rapidly improved from the chance level to ~80% correct responses over several thousand trials (Fig. 2a). During this short period of associative learning, the models learn to associate motion directions and categories, driven by plasticity of the synapses from association to decision neurons. Plasticity transforms the profile of these synapses from random to nearly binary: association neurons with preferred directions in category C1 have strong weights to C_1 and nearly zero weights to C_2 decision neurons, and *vice versa* (Supplementary Fig. 1b). As a result, motion directions from category C1 generate stronger input into the C_1 decision population, which makes C_1 choices more likely, because the probability of choice in our model is determined by the difference in input currents to two competing populations[21]. At this stage of learning, the performance is less accurate for stimuli closest to (15°) the category boundary (Fig. 2d). Near-boundary stimuli activate a subpopulation of association neurons with preferred directions in both categories (Fig. 1c), resulting in comparable inputs to both decision populations and less reliable categorization behaviour.
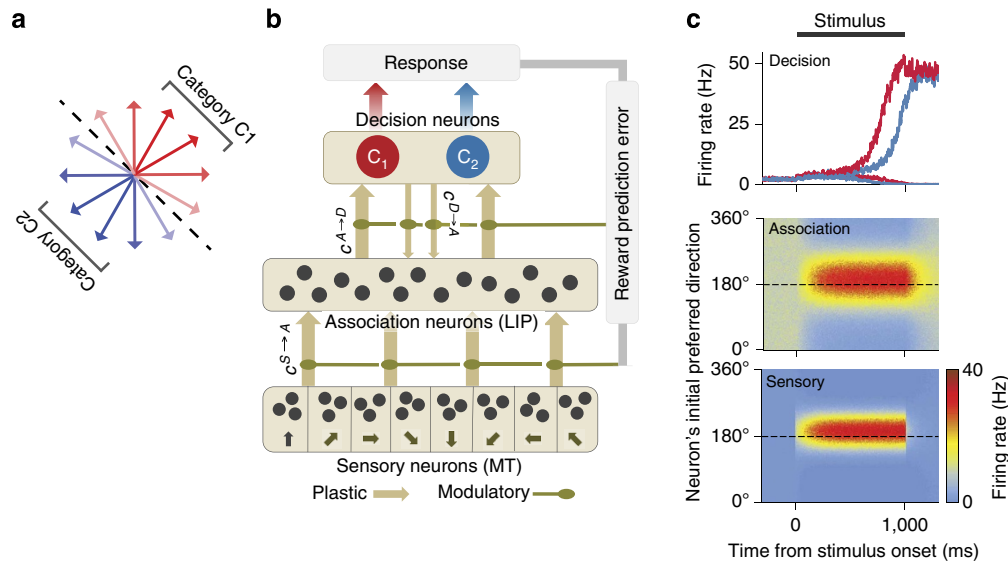
Figure 1 | Categorization task and the neural circuit model. (**a**) A set of 12 motion direction stimuli is divided into two categories, C1 and C2 (red and blue arrows), separated by a category boundary (black dashed line). On each trial, one randomly chosen motion stimulus is presented, and the model learned through trial and error to indicate its category membership. (**b**) Schematic of the circuit model. The network comprises a sensory (MT), an association (LIP) and a decision neural circuits. Neurons in the sensory circuit are tuned to motion directions (indicated by arrows). They receive directional bottom-up inputs and provide inputs to the association neurons through feedforward synapses ($c^{S \to A}$). The decision circuit (C_1 and C_2 populations) pools activity of association neurons through feedforward synapses ($c^{A \to D}$) and generates a category decision through competitive attractor dynamics. The model has feedback connections from the decision to association neurons ($c^{D \to A}$). All synaptic connections between the local circuits undergo Hebbian plasticity modulated by a reward prediction error signal. (**c**) An example network activity before categorization training. A motion direction stimulus (195°) is presented for 1 s (grey bar). The sensory and association neurons show direction-tuned responses in their spatiotemporal activity patterns (lower and middle panels, respectively). x axis, time; y axis, neurons labelled by the preferred direction; firing rate is colour-coded. The decision circuit generates categorical choice through a winner-take-all competition between the C_1 and C_2 populations (upper panel). Firing rates of the C_1 and C_2 populations are shown for two trials, where C_1 (red line) and C_2 (blue line) choice was made for the same stimulus.
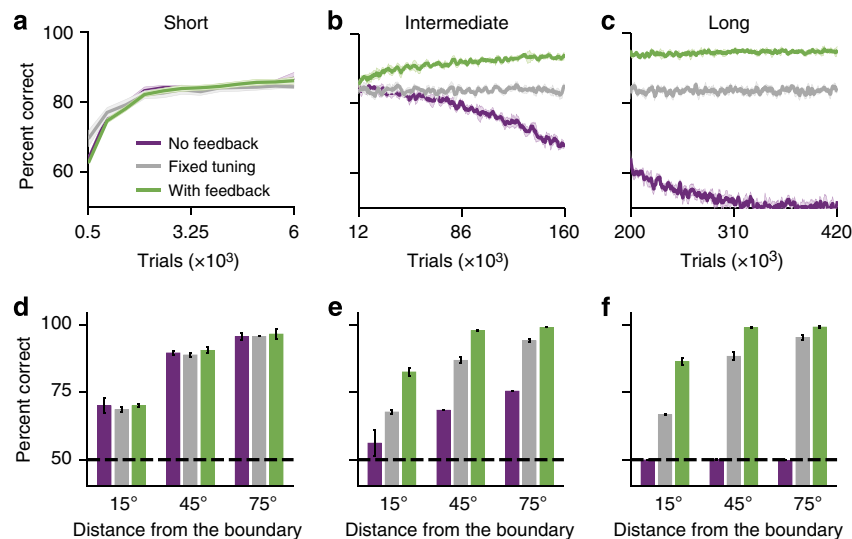

Figure 2 | Behavioural performance of the network models during training on the motion categorization task. Performance of the networks with feedback (green), without feedback (purple) and with fixed tuning of association neurons (grey) is shown at three stages of learning: short (**a,d**), intermediate (**b,e**) and long (**c,f**). (**a–c**) Overall percent correct responses as a function of the number of trials performed. (**d–f**) Psychometric functions evaluated at the end of each training epoch: percent correct responses for stimuli close to (15°) and farther from (45° and 75°) category boundary. At the short training stage, performance improved equally in all three models. As the training progressed, performance of the network with feedback steadily improved especially for the near-boundary stimuli, while performance of the network without feedback gradually deteriorated and eventually dropped to the chance level. Performance of the network with fixed tuning of association neurons remained at the level attained by the end of the short training stage. Shaded area in **a–c** and error bars in **d–f** indicate s.d. across five independent realizations of each network type.

As training progressed, the three models began to exhibit markedly different performance trends (Fig. 2b,e). The network with feedback steadily improved performance over a hundred thousand trials (several months of training for monkeys), mainly due to increasing accuracy for the near-boundary stimuli (Fig. 2e). In contrast, the performance of network without

feedback gradually deteriorated, whereby accuracy decreased for all motion directions. The network with fixed tuning of association neurons maintained the same performance level as attained by the end of the associative learning period. These performance trends were preserved throughout extensively long training (Fig. 2c,f), by the end of which the performance of the network without feedback dropped to the chance level.

Transformation of tuning in association neurons. The striking differences in learning performance of the three models cannot be explained by the synaptic connections from the association to decision neurons, as they are shaped equally in all networks during associative learning and remain virtually unchanged later on (Supplementary Fig. 1b). The reason for the observed performance differences is the change in tuning of association neurons, driven by the plasticity of synapses between sensory and association neurons (Supplementary Figs 1 and 2). In the networks with and without feedback, association neurons have initially the same uniform direction tuning, which is only slightly altered after a short period of learning (6,000 trials, Fig. 3a, upper row), but becomes dramatically different in the two models after extensively long training (420,000 trials, Fig. 3a, lower row). In the network without feedback, the direction tuning deteriorates: the association neurons fire at the same rate for all motion directions. Consequently, the decision circuit receives non-selective inputs and the performance is at the chance level. In contrast, tuning transforms from directional to categorical in the network with feedback: two nonoverlapping subpopulations emerge in the association circuit that respond selectively to stimuli from their preferred categories. As a result, category decisions are very accurate even for near-boundary stimuli.

To quantify the development of category selectivity throughout learning, we computed the average category-tuning index[5] (CTI) of association neurons in the model with feedback. Categorical tuning entails that neurons respond differently to stimuli in different categories and do not differentiate between stimuli in the same category. Accordingly, the CTI varies from -1.0 to 1.0, where positive values indicate larger response differences for stimuli in different categories and negative values indicate larger differences within each category (see Methods). Before learning, the average CTI of association neurons was zero, indicating uniform direction tuning (Fig. 3b), and then CTI gradually increased. At the intermediate learning stage corresponding to the amount of categorization training received by monkeys (65,000 trials or $\sim$ 10–12 weeks), the average CTI was 0.18, comparable to the CTI value 0.125 previously reported for LIP neurons[5].

The gradual increase in the CTI was accompanied by changes in the tuning curves of individual association neurons, which followed two systematic trends. In neurons that initially preferred directions near category centres, tuning curves broadened (Fig. 3c, right), while in neurons that initially preferred directions near category boundaries, tuning curves shifted so that their preferred directions moved towards centres of the respective categories (Fig. 3c, left). Broadening and shifting of tuning curves led to mixed tuning, whereby direction and category signals were combined on the single-cell level. To quantify this mixture, we fitted the tuning curve of each association neuron with a generalized linear model (GLM)[26], which contained a linear combination of two regressor functions: a direction (bell-shaped, equation (12)) and a category (binary step-like, equation (13)) tuning profiles (see Methods). The tuning was classified as pure directional, pure categorical or mixed, according to GLM coefficients that were significantly different from 0. At the intermediate learning stage (65,000 trials), 15.6% of association neurons exhibited a significant influence of category on their tuning curves, while 84.4% remained purely direction-tuned. We examined the distribution of preferred directions in direction-tuned neurons, and found that more neurons were tuned to category centres than to category boundaries (Fig. 3d, the result did not change if all neurons were included).
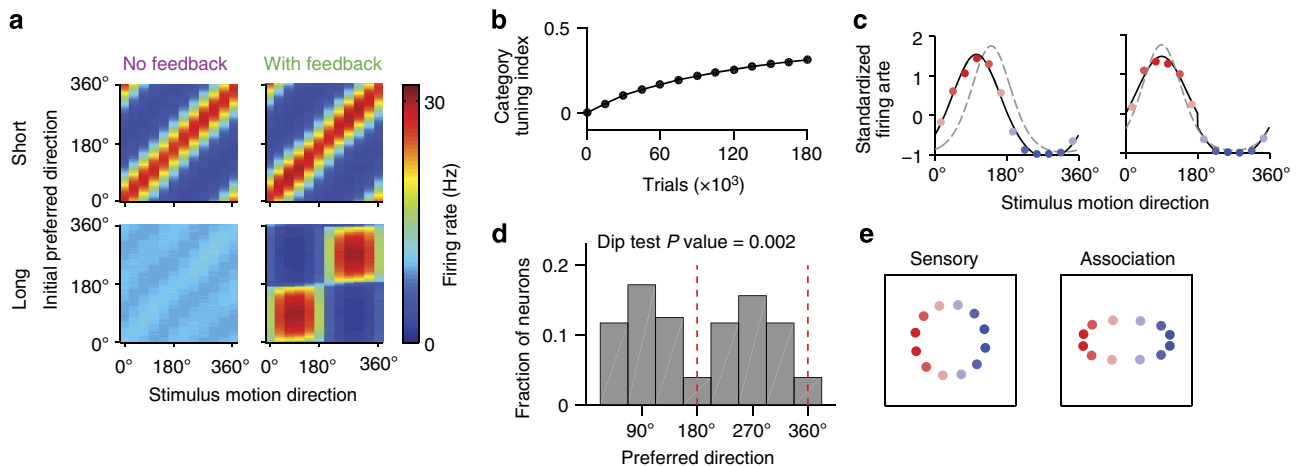

Figure 3 | Mixed direction and category tuning emerges in association neurons through learning. (**a**) Tuning profiles of association neurons after short (6×10^3 trials, upper row) and long (420×10^3 trials, lower row) periods of training for the network without feedback (left column) and with feedback (right column). *x* axis, stimulus motion direction; *y* axis, neurons arranged and labelled by their preferred direction before learning; firing rate is colour-coded. After extensive categorization training, motion tuning deteriorates in the network without feedback, whereas categorical tuning develops in the network with feedback. (**b**) The average category-tuning index of association neurons steadily increases over the course of training. (**c**–**e**) Mixed direction and category tuning of the association neurons at the intermediate stage of training (65×10^3 trials). (**c**) Tuning profiles of two example association neurons before (grey dashed line) and after learning (coloured dots—firing rates, black solid line—best-fitted tuning function). Tuning curves broaden (right panel) and shift towards category centres (left panel) in neurons with initial preferred directions near category centres and category boundaries, respectively. (**d**) Bimodal distribution of preferred directions in direction-tuned association neurons (Hartigan's dip test, $P = 0.002$). Majority of neurons are tuned away from the category boundary (indicated by red dashed lines). (**e**) Multidimensional scaling analysis reveals a circular configuration of motion directions in the representation of sensory neurons (left panel), and an elliptical configuration elongated along the axis perpendicular to the category boundary in the representation of association neurons (right panel).

Broadening and shifting of tuning curves alter the representation of motion directions in a way that facilitates the discrimination of categories. We visualized the ensuing representation on the population level using classical multidimensional scaling[27] (MDS). In this framework, stimuli are represented as vectors in a high-dimensional space of neural firing rates, where each dimension corresponds to a neuron in the population. The MDS algorithm finds a two-dimensional configuration of the stimuli that preserves the distances between them as much as possible. In the sensory circuit, the MDS algorithm yields a circular configuration (Fig. 3e, left) that faithfully reproduces the arrangement of directions in the physical space. In the association circuit, the configuration is elongated along the axis perpendicular to the category boundary (Fig. 3e, right), which increases the distances between near-boundary stimuli in different categories making them more easily discriminable and decreases distances between stimuli within the same category making them less discriminable.

Mixed direction and category tuning in LIP neurons. We compared tuning changes in our model to the tuning (during the period of stimulus presentation) of MT and LIP neurons recorded in monkeys trained to categorize motion directions[5]. Such a comparison is meaningful, if the model and monkeys experienced similar amount of categorization training and reached similar behavioural performance. In the model, the time course of learning depends on the learning rate q and the maximal strength of feedback connections $g_{max}^{D \to A}$ (Supplementary Fig. 3). We simulated the model for a range of q and $g_{max}^{D \to A}$ and used the parameters that provided good match to experimental data for the similar number of training trials (that is, 65,000 trials, see Supplementary Fig. 4).

We fitted the tuning curve of each neuron in our database (67 MT and 156 LIP neurons) with direction and category-tuning functions and then classified tuning as directional, categorical or mixed following the same procedure that was used for model neurons. The majority of MT (91.0%) and LIP neurons (69.9%) exhibited pure direction tuning (Fig. 4a, upper panels, Fig. 4b). In agreement with our model prediction, the distribution

of preferred directions was significantly bimodal among direction-tuned LIP neurons (Hartigan's dip test $P = 0.003$, Fig. 4c), but not among MT neurons (Hartigan's dip test $P = 0.08$). A considerable fraction of LIP neurons (18.0%) showed a mixture of directional and categorical tuning (Fig. 4a, lower panels). The distribution of preferred directions remained significantly bimodal when the mixed-tuned LIP neurons were included in the analysis (Hartigan's dip test $P < 10^{-7}$). A small fraction of LIP neurons (3.9%) exhibited pure category tuning (Fig. 4a, middle panels), and the rest (8.3%) were not stimulus-selective. As a control, we repeated the analyses in different time epochs during the trial (Supplementary Table 1) and using a smoothed category-tuning function (Supplementary Table 2), and obtained similar results.

The representation of motion directions at the population level was consistent with the model prediction as well: the MDS algorithm revealed a nearly circular configuration of motion directions in MT (Fig. 4d, upper panel), whereas in LIP motion directions were arranged on an elongated ellipse with the major axis perpendicular to the category boundary (Fig. 4d, lower panel, see Supplementary Note 1 for statistical significance test). Similarly, CTI was significantly higher in LIP than in MT as has been previously reported for the same dataset[5]. Although the LIP population demonstrated high heterogeneity, the main tuning features in LIP bear a remarkable resemblance to the tuning transformation induced by learning in our model.

Reward-driven learning depends on choice probability. To understand effects of learning on tuning of association neurons, we need to examine the reward-dependent Hebbian plasticity rule (equation (1)). The plasticity rule entails that the expected weight change for each stimulus $\langle \Delta c | \theta \rangle$ is proportional to the covariance between the reward R and neural activity $N = r_{pre} r_{post}$ (ref. 15) (see Supplementary Note 2):

$$\langle \Delta c | \theta \rangle = q \, \mathrm{Cov}[R, N \,|\, \theta]. \tag{2}$$

This means that average synaptic weight changes across many trials are driven by covariation between trial-to-trial fluctuations of the firing rates and reward. Thereby, synapses change to
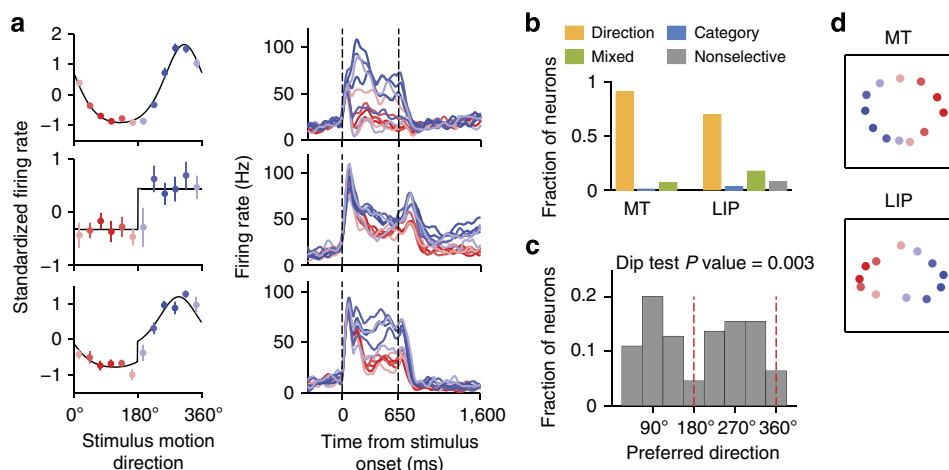

Figure 4 | Mixed direction and category tuning in LIP neurons. (**a**) Average response to the 12 motion directions for three example LIP neurons, classified as direction-tuned (upper row), category-tuned (middle row) and mixed direction and category-tuned (lower row). Left column: tuning curves, the average standardized firing rates during stimulus presentation (coloured dots) overlaid by the best-fitted tuning function (black line). Error bars indicate s.e.m. Right column: average firing rate traces in response to stimuli from category C1 (red lines) and C2 (blue lines). (**b**) The fraction of MT ($N = 67$) and LIP ($N = 156$) neurons that exhibited directional tuning, category tuning, mixed tuning or were not stimulus-selective. (**c**) Bimodal distribution of preferred directions in direction-tuned LIP neurons (Hartigan's dip test, $P = 0.003$, $N = 109$), in support of the model prediction (Fig. 3d). (**d**) Multidimensional scaling analysis reveals a nearly uniform representation of motion directions in the MT population (upper panel), and an elongated representation in the LIP population (lower panel), where the distances are larger between the stimuli from different categories, c.f. the model prediction in Fig. 3e.

increase the expected reward. If for a particular synapse the neural activity is systematically higher on trials when the reward is above its mean, then the covariance is positive, the synapse is potentiated, and hence the mean neural activity and the expected reward increase (and analogously for negative covariance). Fluctuations of both reward and neural activity are critical for learning: if either R or N is deterministic, the covariance equals zero and learning does not increase expected reward.

Covariation between neural activity and reward entails covariation between neural activity and choices, if reward is assigned on the basis of behavioural responses. This simple intuition can be formalized mathematically, if we express the covariance $\mathrm{Cov}[R,N|\theta]$ in terms of expectations conditioned on choices. For tasks with only two possible choices, we obtain a simple expression (see Methods for derivation and generalization to arbitrary number of choices):

$$\mathrm{Cov}[R, N \mid \theta] = \left(N_{1,\theta} - N_{2,\theta}\right)\left(R_{1,\theta} - R_{2,\theta}\right)P_{1,\theta}P_{2,\theta}. \quad (3)$$

Here $P_{i,\theta}$ is the probability that C_i choice is made for the stimulus θ; $R_{i,\theta} = \langle R|\theta,C_i\rangle$ is the reward expected for choosing C_i for stimulus θ; and $N_{i,\theta} = \langle N|\theta,C_i\rangle$ is the expected neural activity conditioned on the stimulus θ and choice C_i.

The term $(N_{1,\theta} - N_{2,\theta})$ represents the difference between the means of two neural activity distributions obtained on trials when different choices are made for the same stimulus θ, and is monotonically related to a measure called choice probability[17,28] (CP, Supplementary Fig. 5a). CP quantifies the accuracy with which an ideal observer could predict choices given neuronal firing rates on a trial-by-trial basis. A CP of 0.5 indicates no correlation between neural fluctuations and choices ($N_{1,\theta} \approx N_{2,\theta}$, Fig. 5b), whereas a CP of 1 (or 0) indicates that the neuron's firing rate is always higher (or lower) on trials when C_1 is chosen than on trials when C_2 is chosen for the same stimulus θ ($N_{1,\theta} > N_{2,\theta}$ in Fig. 5c; our convention of computing CP differs from refs 17,29, see Supplementary Note 3).

Equation (3) demonstrates that synaptic updates lead to increase in expected reward if CP ≠ 0.5 for pre- or postsynaptic neurons; however, if CP ≈ 0.5 for both pre- and postsynaptic neurons, the covariance $\mathrm{Cov}[R,N|\theta]$ vanishes irrespective of the reward expectation. This result is a general property of reward-modulated Hebbian plasticity and holds across different tasks and network architectures. It can be illustrated using a single toy-model neuron (Fig. 5a), whose firing rates for C_1 and C_2 choices are sampled from two Gaussian distributions with different means, without specifying mechanisms generating CP. We assumed that C_1 choices are rewarded, leaving other task details unspecified. The synapse of this toy-model neuron is updated according to the reward-modulated Hebbian plasticity rule. As predicted by equation (3), CP determines the direction and magnitude of synaptic changes in the toy model. If CP > 0.5, the covariance $\mathrm{Cov}[R,N]$ is positive and the synapse is potentiated (red traces in Fig. 5d), and if CP < 0.5 the synapse is depressed (blue traces in Fig. 5d). The covariance magnitude is larger for larger |CP − 0.5|, resulting in faster synaptic changes. If CP ≈ 0.5, the covariance vanishes; hence, synaptic modifications are driven by noise similar to a random walk (yellow traces in Fig. 5d) and over a long period of learning any synaptic weight becomes equally likely (Fig. 5e).

This general principle explains both the fast associative learning and slower behavioural improvements in our model. Since activities of decision neurons directly represent the model's choices, the magnitude of their CP is large; hence, the synapses of decision neurons change rapidly towards increasing expected reward, underpinning fast associative learning. In the network with feedback, CP arises via feedback from the decision circuit, which produces multiplicative rate modulations in association
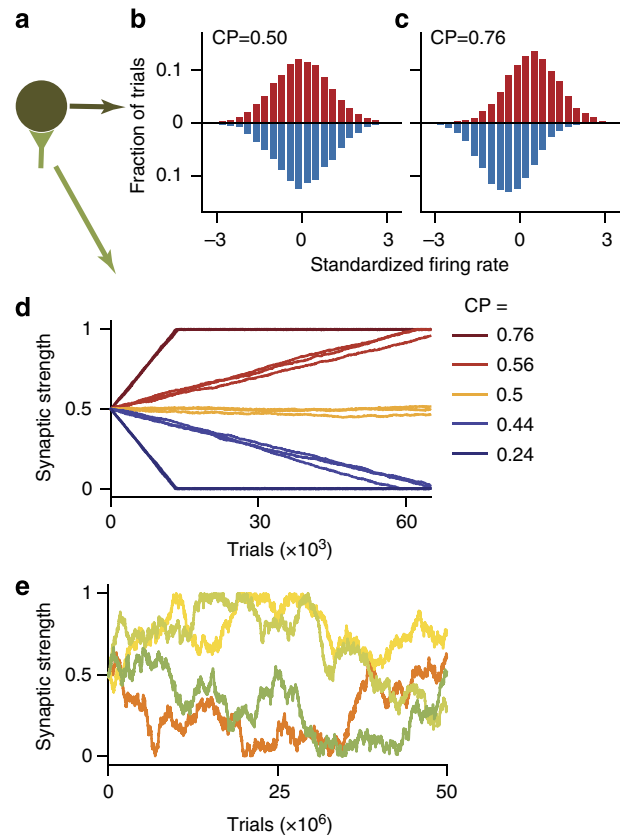

Figure 5 | Choice probability determines the direction and magnitude of synaptic changes. (**a**) Schematic of a toy model: firing rate of a neuron is correlated with choices in a behavioural task, where choice C_1 is rewarded and C_2 is not. Synapse of this neuron is updated according to reward-dependent Hebbian plasticity. (**b,c**) Firing rate distributions of a toy-model neuron on trials when different choices are made (C_1—red bars, C_2—blue bars). CP is close to 0.5, if the distributions largely overlap, that is, rate fluctuations are not correlated with choices (**b**). CP deviates from 0.5 towards 1 (or 0) when the distributions are well separated, that is, rate fluctuations are correlated with choices (**c**). (**d**) The sign of CP − 0.5 determines the direction of synaptic changes: CP > 0.5 leads to potentiation and CP < 0.5 leads to depression (CP is computed relative to the rewarded choice C_1). The larger is the deviation of CP from 0.5, the faster is the rate of synaptic changes. Synaptic strengths are shown across learning trials for different CP values over an intermediate period of learning (three independent realizations for each CP value). (**e**) Synaptic strengths are shown across many learning trials for CP ≈ 0.5 (corresponds to yellow traces in **d**, note the difference in scale of the x axis between **d** and **e**). Synaptic changes are random, and any weight becomes equally likely over a long period of learning. Four independent realizations are shown.

neurons[30,31] (Supplementary Fig. 2). Initially, CP is scattered around 0.5; however, when feedback connections become structured (∼500 trials), neurons receiving stronger input from the C_1 (C_2) decision population fire at higher rates when C_1 (C_2) choices are made and exhibit CP > 0.5 (CP < 0.5, Fig. 6b). The magnitude of CP is smaller in association than in decision neurons; therefore, the tuning changes of association neurons and ensuing behavioural improvements happen more slowly than associative learning. In the network without feedback, CP ≈ 0.5 in all association neurons and at all learning stages (Fig. 6a), because local noise in the decision circuit—required to attain realistic behavioural performance in the categorization task—diminishes the influence of association neurons' rate fluctuations on choices (see Supplementary Note 4 for details). Resulting unstructured
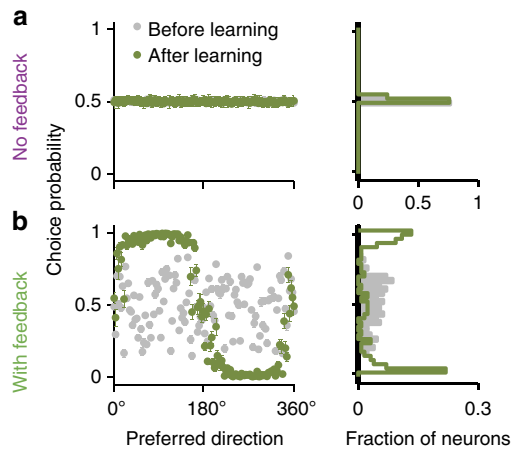
Figure 6 | Association neurons in the network with feedback, but not in the network without feedback, exhibit choice-correlated fluctuations. (**a**) In the network without feedback, CP is close to 0.5 in all association neurons and does not change throughout learning. (**b**) In the network with feedback, CP is randomly scattered around 0.5 before learning (grey dots), but a bimodal profile of CP develops after a short period of learning (500 trials, green dots), such that CP > 0.5 in neurons with preferred directions in category C1, and CP < 0.5 in neurons with preferred directions in category C2. CP (*y* axis) is plotted for all association neurons labeled by their initial preferred direction (*x* axis) before (grey dots) and after (olive dots) learning. Histograms to the right show the corresponding CP distributions.

synaptic changes lead to deterioration of tuning and behavioural performance. Regardless of which mechanism—feedforward or feedback—is more plausible for generating CP in real neurons, our results demonstrate the significance of CP for reward-dependent learning.

Choice-correlated fluctuations shape neural tuning changes. Over many trials, synaptic weight changes Δc_{ij} between the association neuron i and sensory neurons $j = 1 \ldots N$ follow the same two trends as observed in tuning functions (Fig. 3c). For neurons tuned to category centres, the initial bell-shaped profile widens on both sides until it transforms into a step-like profile aligned with the category boundary (Fig. 7e); hence, the tuning curves broaden. For neurons tuned to directions near category boundaries, synapses are strengthened on one side and weakened on the other side of the initial bell-shaped profile (Fig. 7f); hence, the tuning curves shift towards the category centre. Using equation (2), the expected weight change for stimulus θ can be expressed as $\langle \Delta c_{ij} | \theta \rangle = q \ \mathrm{Cov}[R, r_i r_j | \theta] \approx q \ \langle r_j | \theta \rangle \mathrm{Cov}[R, r_i | \theta]$ (see Supplementary Note 2). The overall expected weight change is then the average of $\langle \Delta c_{ij} | \theta \rangle$ across all stimuli. Thus, synaptic changes are determined by the covariance $\mathrm{Cov}[R, r_i | \theta]$ weighted by the rates of sensory neurons.

For neurons initially tuned to directions in category C1, CP > 0.5 and the covariance $\mathrm{Cov}[R, r_i | \theta]$ is positive for stimuli $\theta \in \mathrm{C1}$ and negative for $\theta \in \mathrm{C2}$ (Fig. 7a,b), since the term $(R_{1,\theta} - R_{2,\theta})$ in equation (3) changes sign for θ in different categories. The covariance magnitude is proportional to the product of probabilities of the correct response and error, $P_{1,\theta}(1 - P_{1,\theta})$, which is largest for near-boundary stimuli ($P_{1,\theta} \sim 0.5$). When this covariance is combined with the firing rates of sensory neurons, the overall synaptic weight change is step-like for neurons tuned to category centres (Fig. 7c), and skewed towards the category centre for neurons tuned near category boundaries (Fig. 7d). For neurons initially tuned to directions in category C2, CP < 0.5; hence, the covariance has just the opposite sign leading to the preference for

category C2. Such tuning changes lead to behavioural improvements because the feedforward $c_i^{A \to D}$ and feedback $c_i^{D \to A}$ connections become aligned through learning.

Plastic top-down feedback induces task-specific correlations. In our model, category-tuning and neural fluctuations are simultaneously shaped through plasticity of feedforward and feedback connections to association neurons, giving rise to testable model predictions.

First, our model predicts that association neurons with larger CP exhibit greater sensitivity of their tuning curve to the stimulus category (Fig. 8a). The latter is quantified by category sensitivity (CS), which is the accuracy with which an ideal observer could discriminate between stimuli from categories C1 and C2 given neuron's firing rates on correct trials. A positive correlation between CP and CS arises because of reciprocal interaction of plasticity on the feedforward $c^{S \to A}$ and feedback $c^{D \to A}$ connections to association neurons. On one hand, plasticity of feedforward connections from sensory neurons leads to a greater increase in CS for neurons with larger CP (Fig. 5). On the other hand, plasticity of feedback connections from decision neurons generates a greater difference in top-down inputs from two decision populations, hence larger CP, for neurons with larger CS. The correlation between CP and CS is not an *a priori* given, because these measures quantify independent aspects of neuronal response. CS measures the difference in response to stimuli from different categories on correct trials, whereas CP measures the difference in response to the same stimulus on correct versus error trials. The correlation between CP and CS is abolished if the learned profile of feedback connections is randomized (Supplementary Fig. 6c).

We tested whether the predicted correlation between CP and CS exists in MT and LIP neurons. The overall magnitude of CP was significantly greater in LIP than in MT population (Wilcoxon rank-sum test comparing distributions of $|\mathrm{CP} - 0.5|$, $P = 0.0006$, Fig. 8c). Ten LIP neurons (11.4%, $N = 88$) and none of MT neurons (0%, $N = 31$) showed individually significant CP (shuffle test with 1,000 shuffles and two-sample *t*-test, $P < 0.05$, see Methods). In agreement with the model prediction, CP and CS were significantly correlated in the LIP (Fig. 8b, Pearson correlation, $r = 0.494$, $N = 88$, $P = 10^{-6}$), but not in MT population ($r = -0.181$, $N = 31$, $P = 0.33$, Supplementary Fig. 6d). We also repeated the analyses using CP computed relative to the preferred category of each neuron[17,29] and obtained similar results (Supplementary Note 3 and Supplementary Fig. 6e–h). Although CP magnitude is slightly lower in LIP data than in the model, smaller CP magnitudes can be obtained in the model with weaker top-down connections (Supplementary Fig. 3). In addition, since recorded LIP neurons were sampled randomly, some of them might not be engaged in the categorization task and some were not visually responsive. This sampling heterogeneity may reduce the average effect size in the data and it is not incorporated in our model.

Second, our model predicts that interneuronal correlations depend on CS. In association neurons, correlations between their trial-to-trial rate fluctuations, termed noise correlations r_{ij}^{noise} (a Pearson correlation coefficient between rate fluctuations of neurons i and j), arise from shared recurrent and feedforward inputs. In the network without feedback, noise correlations simply decrease with the difference in neurons' preferred directions reflecting the bell-shaped profile of their recurrent and feedforward connections (Supplementary Fig. 5b). In the network with feedback, association neurons with the same category preference also share top-down input from decision neurons, consequently noise correlations are stronger among neurons that contribute to the same category decision (Fig. 8d),
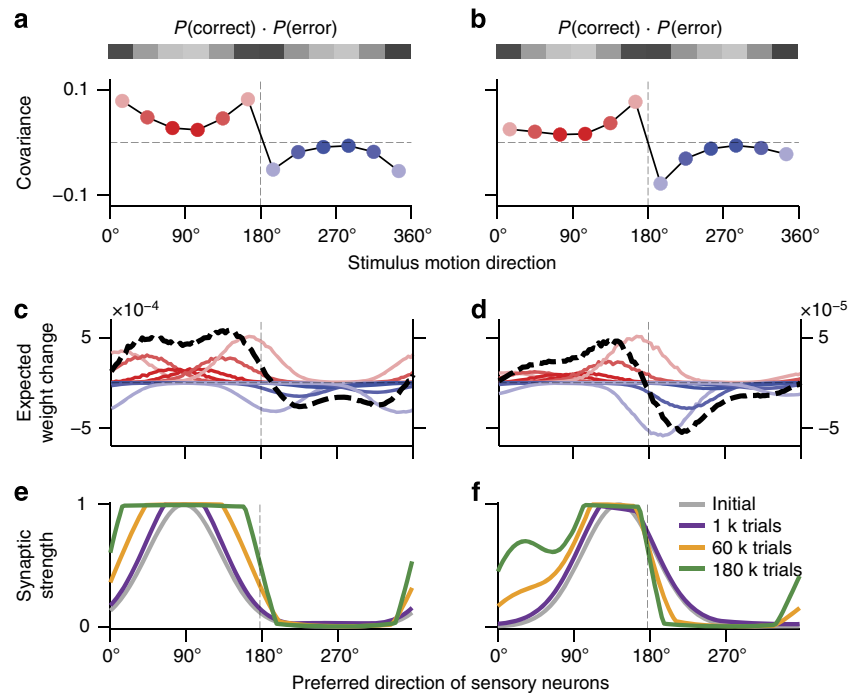
Figure 7 | The covariance between reward and neural activity drives tuning changes in association neurons. (**a,b**) Covariance Cov[$R, r_i|\theta$] between the firing rate and reward (*y* axis) measured for each of the 12 motion direction stimuli (*x* axis) after 1,000 trials of learning. The covariance profile is shown for two association neurons with preferred directions near category centre (**a**) and near category boundary (**b**), the same sample neurons as in Fig. 3c. The covariance has opposing sign for stimuli in different categories (see equation 3). The magnitude of the covariance is proportional to the product of probability to make a correct response and the probability to make an error (grey-scale code bar above the graph), which is the largest for near-boundary stimuli. (**c,d**) The expected weight change for the synapses from all sensory neurons (*x* axis) to a single association neuron; **c,d** correspond to the association neurons shown in **a,b**, respectively. The weight changes expected for each stimulus (red lines—stimuli from category C1, blue lines—stimuli from C2, left *y* axis) add up to produce the total expected weight change (black dashed line, right *y* axis). (**e,f**) The weight changes from **c,d** accumulate over multiple trials. As a result, the synaptic profile from sensory neurons is gradually broadened for the association neuron tuned to the category centre, and shifted towards the category centre for the association neuron tuned near the category boundary. These changes of the synaptic profiles underlie neural tuning changes shown in Fig. 3c.

similar to previous experimental reports[32]. Moreover, noise correlations are larger in neural pairs with smaller absolute difference in their category sensitivities ($\Delta CS_{ij} = |CS_i - CS_j|$): r_{ij}^{noise} is positive in pairs with similar CS ($\Delta CS_{ij} \sim 0$) and negative in pairs with opposite category preference ($\Delta CS_{ij} \sim 1$, Fig. 8e). In addition, the magnitude of noise correlations is larger in neural pairs with higher CS strength, defined as $(|CS_i - 0.5| + |CS_j - 0.5|)/2$ (Fig. 8e). Such structured noise correlations—that remained static throughout learning—were required in a feedforward model[10,29] to capture the correlation between CP and task sensitivity observed in several experimental studies[7,8,29,33]. However, the *a priori* assumption that noise correlations depend on CS is not realistic, since categories are assigned arbitrarily. Alternatively, our model suggests that plasticity of feedback connections represents a common mechanism by which the structure of noise correlations, CP and CS all develop dynamically through learning.

Discussion

Here we proposed a neural circuit mechanism for visual category learning. Our findings represent two major advances going beyond a model for categorization. First, we demonstrated that choice-correlated activity fluctuations, ubiquitous across cortical areas[7–9,17,34], are critical for learning through reward-dependent Hebbian plasticity, which generally holds across different network architectures and behavioural tasks. Second, we showed how behavioural improvements, neuronal tuning changes, CP and noise correlations can be all simultaneously shaped by a common

plasticity mechanism in a network incorporating top-down feedback. Several model predictions about ensuing interdependences between these measures were confirmed by the analysis of LIP recordings.

The reward-dependent Hebbian plasticity in our model belongs to the family of covariance-based learning rules[15] using a stimulus-specific RPE signal, which is critical for successful learning[14] (Supplementary Fig. 7). The idea to harness local fluctuations for reward-dependent learning has been first proposed for connectionist networks[35], and later instantiated in networks of spiking neurons by exploiting either randomness of Poisson spiking[36,37] or stochasticity of synaptic transmission[38]. Such plasticity rules can successfully learn precise spike patterns in networks of just a few neurons, but fail in larger networks and when behavioural outcomes are determined by population firing rates rather than by spike times of individual neurons[39,40]. The reason for their failure in these situations is precisely the lack of correlation between population-level choices and local activity fluctuations. To overcome this problem, plasticity rules have been employed incorporating behavioural choice explicitly as a multiplicative factor[10,41,42]. In contrast, our solution does not require any special plasticity rule, but instead utilizes network architecture where feedback from decision neurons generates choice-correlated variability.

Task-specific neural representations develop in many training paradigms across different cortical areas[43–55]. Our model demonstrates how such task-specific representations can emerge through reward-dependent plasticity. Although task-specific
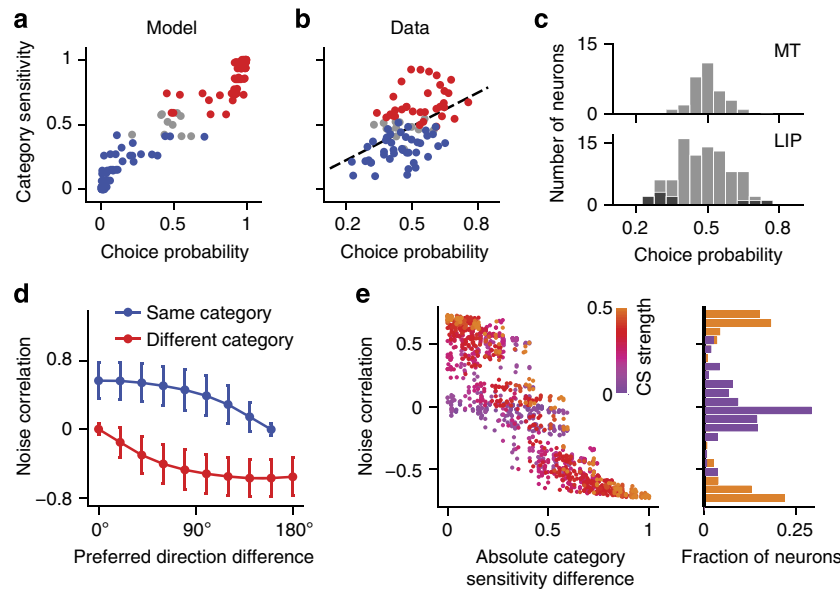
Figure 8 | Model predicts interdependence between the CP, CS and noise correlations. (**a**) In the association neurons, a positive correlation between the CP and CS arises through learning. In the scatter plot, each dot represents a CP–CS pair for an association neuron (measured after 500 trials), colour-coded according to its category preference (C1—red, C2—blue, nonselective—grey). (**b**) A significant positive correlation between the CP and CS was found in LIP neurons recorded from behaving monkeys, similar to the model prediction in **a**. (**c**) Histograms of CP in MT and LIP populations. Black shading indicates neurons with individually significant CP. Overall magnitude of CP was significantly greater in LIP than in MT. (**d,e**) Plasticity of the feedback connections from decision neurons gives rise to task-specific noise correlations. (**d**) After learning, noise correlations decrease with the difference in preferred directions of two neurons, but are stronger for neurons preferring the same category, than for neurons preferring different categories. Error bars indicate s.d. across neurons. (**e**) Noise correlations are stronger in pairs of association neurons with more similar and individually larger CS. Scatter plot of noise correlation (*y* axis) versus the absolute difference between the category sensitivities (*x* axis). Each dot is coloured according to the average CS strength in the pair (purple-to-orange colour code corresponds to low-to-high CS strength). Histograms show the distributions of noise correlations for neural pairs with the high (>0.4) and low (<0.1) average CS strength (orange and purple, respectively).

selectivity could arise through activity modulation via plastic feedback connections[56], in our model, top-down modulation has a negligible effect on selectivity of association neurons (Supplementary Fig. 2), yet it is critical to guide learning of task-relevant features[57].

Tuning changes of association neurons in our model allow for more accurate categorization of near-boundary stimuli than in the classical categorization model with fixed tuning[16]. In our model, tuning changes arise from plasticity of feedforward synapses from sensory (MT) to association (LIP) neurons; however, similar results are obtained if plasticity acts only on the recurrent synapses within the association circuit, or on both the feedforward and recurrent synapses (Supplementary Fig. 8). In our model, the initial direction tuning of association neurons sets the profile of choice-correlated fluctuations, which in turn governs tuning changes. However, initial tuning is not required for successful learning: a population of nonselective neurons carrying choice-correlated fluctuations develops categorical tuning just as well. In this case, neurons develop purely binary category selectivity with the category preference determined solely by their CP (Supplementary Fig. 9). Last, retraining on a categorization task with a new category boundary results in readjustment of neural tuning (Supplementary Fig. 10) similar to experimental observations[5].

It has been speculated that category signals in LIP represent abstract perceptual decisions: category C1 versus C2 (ref. 58). In the motion categorization task, but not in classic motion discrimination work in LIP[7], abstract decisions were dissociated from the actions signalling those decisions by using a two-interval match-to-category design, where the required motor response was unknown at the time of the first stimulus presentation. Moreover, receptive fields of LIP neurons in the motion

discrimination task were aligned with the saccadic choice targets and not with the motion stimulus as in our case; hence, that design was better suited to examine response-related rather than perceptual signals in LIP. Accordingly, these data were interpreted using a feedforward model, where LIP neurons represent a decision-variable pooling activity of MT neurons with weights adjusted by a reinforcement learning rule[10], and behavioural improvements were ascribed to selective strengthening of connections from the most sensitive sensory to decision neurons[8,10]. In contrast, we find that during motion categorization the representation of motion stimuli in LIP constitutes a mixture of directional and categorical tuning that facilitates discrimination of learned categories. Therefore, both mechanisms—that co-exist in our model—may be concurrently employed in the brain: refinements of sensory representations and of their readout by decision neurons.

In our model, mixed selectivity is robustly observed over a period from a few thousand to several hundred thousand trials, accompanied by increasing category tuning. Consistent with high CTI values reported previously in LIP[5], we find that two factors contribute to the increasing population CTI: shift of preferred directions and emergence of mixed and pure category tuning. Some LIP neurons carried category selectivity throughout the delay period of the match-to-category task, which indicates that category encoding may not be a purely feedforward effect.

Our work demonstrates the significance of CP for reward-dependent learning regardless of its origin. The origin of CP has been recently debated, with accumulating evidence for top-down contributions[18,34,59]. Notably, CP signals we observe in LIP are distinct from signals related to reward, attention and upcoming movements[5,54,60,61]. Although origins of CP may differ between earlier sensory areas such as MT and more cognitive

areas such as LIP, our model provides a common framework for understanding the impact of CP on plasticity of neuronal representation.

We proposed a novel model for how CP influences plasticity in LIP, although CP effects in sensory areas (for example, MT) have been modelled previously[18,29,62]. Our model demonstrates how a task-specific structure of CP, noise correlations and CS can arise dynamically through reward-dependent plasticity of top-down connections and predicts that neurons with larger CP develop larger CS. Thus, learning-induced tuning changes may be more pronounced in cortical areas that exhibit greater CP (Supplementary Fig. 3). Interestingly, both CP and CS were found to be significantly larger in LIP than in the prefrontal cortex[54]. Similarly, low CP of MT neurons might explain the absence of obvious tuning changes in this area through categorization training. Small but significant learning-related tuning changes have been observed in other sensory areas[4] that also exhibit CP[63]. Therefore, our findings may generalize across sensory areas not limited to LIP.

Methods

Neural circuit model. *Network architecture.* The network model comprises three interconnected local circuits: sensory, association and decision. All three are strongly recurrent networks with dynamics governed by local excitation and feedback inhibition[19,20,64]. In simulations, we used a reduced mean-field model that has been shown to reproduce neural activity of a full spiking neural network[21]. The dynamics of each excitatory neural population is described by a single variable s representing the fraction of activated N-methyl-D-aspartate receptor conductance, governed by

$$\frac{ds}{dt} = -s/\tau_s + (1-s)\gamma r,$$ (4)

with $\gamma = 0.641$ and $\tau_s = 60$ ms. The firing rate r is a function of the total synaptic current I (refs 21,65):

$$r = f(I) = \frac{aI - b}{1 - \exp[-d(aI - b)]},$$ (5)

with $a = 270$ Hz nA^{-1}, $b = 108$ Hz and $d = 0.154$ s.

The total synaptic current I consists of recurrent and noisy components, $I = I_r + I_n$. Recurrent input to a neuron i in the population A originating from the population B reads:

$$I_{r,i}^{B\rightarrow A} = \frac{1}{N_B}\sum_{j\in B} g_{ij}^{B\rightarrow A}s_j^B,$$ (6)

where $g_{ij}^{B\rightarrow A}$ is the synaptic coupling between the neuron j in the population B and the neuron i in the population A. The current is normalized by the number of presynaptic neurons N_B. Noisy current replicates background synaptic inputs and obeys: $\tau_n dI_n/dt = -(I_n - I_0) + \sqrt{\tau_n}\sigma_n\eta(t)$, where $\eta(t)$ is a white Gaussian noise, $I_0^S = I_0^D = 0.3297$ nA, $I_0^A = 3.1$ nA, $\tau_n = 2$ ms and $\sigma_n = 0.009$ nA.

The sensory and association circuits were each simulated by 128 discrete units with equally spaced preferred directions from 0° to 360°. Within each circuit, the synaptic couplings g_{ij} between neurons with preferred directions θ_i and θ_j have a periodic Gaussian profile:

$$g(\theta_i - \theta_j) = J_- + J_+ \exp\left(-(\theta_i - \theta_j)^2/2\sigma^2\right),$$ (7)

with $\sigma = 43.2°$. Parameters J_- and J_+ determine the amount of recurrent excitation and inhibition. In sensory and association networks, the recurrent inhibition is stronger than recurrent excitation, $J_-^{S\rightarrow S} = -0.5$, $J_+^{S\rightarrow S} = 1.43$, $J_-^{A\rightarrow A} = -10.0$ and $J_+^{A\rightarrow A} = -0.4$ nA. The particularly strong recurrent inhibition in the association circuit sets this module in the normalization regime[66], where the total population activity remains approximately constant for different stimuli[19].

The decision circuit consists of two populations (C_1 and C_2) representing categorical choice, which pool activity of the association neurons. When stimulated, activities of the C_1 and C_2 populations diverge according to winner-take-all dynamics. This behaviour is attained through global inhibition and structured recurrent excitation within the decision circuit[21]: $g_{ij}^{D\rightarrow D} = J_{ij}$ with $J_{C1,C1} = J_{C2,C2} = 0.3725$ nA, $J_{C1,C2} = J_{C2,C1} = -0.1137$ nA.

Plastic synapses. All synapses connecting three local circuits (from sensory to association, and between association and decision neurons) are plastic and excitatory. Synaptic strengths of plastic connections are expressed as $g_{ij} = g_{max}c_{ij}$, where g_{max} is the maximal connection strength and c_{ij} is bounded between 0 and 1, and represents the fraction of potentiated synapses between neurons i and j. At the end of each trial, all c_{ij} are updated according to the Hebbain plasticity rule modulated by the RPE as specified in equation (1), where the learning rate $q = 0.00003$, and r_{pre} and r_{post} are average firing rates during the stimulus period. The stimulus-specific predicted reward $\langle R|\theta\rangle$ was estimated by a running trial

average[14]: $\bar{R}_{n+1}^\theta = \bar{R}_n^\theta + (R_n - \bar{R}_n^\theta)/\tau_R$, where $\tau_R = 5$, and n enumerates trials with stimulus θ.

Plastic synapses between sensory and association neurons $\left(c_{ij}^{S\rightarrow A}\right)$ were initialized with the periodic Gaussian profile as in equation (7) with $J_-^{S\rightarrow A} = 0$, $J_+^{S\rightarrow A} = 1$. Plastic synapses between association and decision neurons ($c_{ij}^{A\rightarrow D}$ and $c_{ij}^{D\rightarrow A}$) were initialized randomly from a uniform distribution on [0.25, 0.75]. The maximal connection strengths of plastic synapses were $g_{max}^{S\rightarrow A} = 1$, $g_{max}^{A\rightarrow D} = 0.03$ and $g_{max}^{D\rightarrow A} = 0.01$ nA.

Simulation protocol and external inputs. Each simulation trial starts with a 200-ms pre-stimulus period (no external inputs), followed by a 1-s presentation of a motion direction stimulus and then by a 500-ms intertrial interval. When a motion direction stimulus θ_s is presented, neurons in the sensory network receive additional input current I_s that depends on the neuron's preferred direction θ:

$$I_s = g_s\exp\left(-(\theta_s - \theta)^2/2\sigma_s^2\right),$$ (8)

where $\sigma_s = 43.2°$ and $g_s = 0.1$ nA. Neurons in the decision circuit receive a non-selective gating current of 0.01 nA during the stimulus period, which sets the circuit in the decision-making regime, and a brief -0.08 nA reset current during the first 300 ms of the intertrial interval, which represents the corollary discharge[67] and resets activity to the spontaneous level.

The model's response on each trial was determined by comparing firing rates of two decision populations with a 20-Hz threshold during the last 25-ms of the stimulus period. The response is considered invalid if both or neither population reach threshold, or either population reaches threshold before the stimulus onset. Across trials, choices of the decision network are stochastic and are characterized by a sigmoidal dependence of the probability of choice C_1 on the difference ΔI in synaptic input currents to two competing populations[68]. Reward equals $R = 1$ on valid correct trials, $R = 0$ on valid incorrect trials and no plasticity is triggered on invalid trials.

Noise correlation r_{ij}^{noise}, CP and CS for the model neurons were estimated from 10,000 simulated trials with synapses 'frozen' (that is, no plasticity) at values attained after specified number of learning trials. Noise correlation r_{ij}^{noise} was computed as the Pearson correlation coefficient between the firing rates of neurons i and j across trials for the same stimulus, and then averaged across stimuli. CP and CS were computed as described in the Data analysis section, except for the CP estimation the model's choice was known explicitly and did not have to be inferred.

Simulations were performed using a custom code written in Matlab implementing Heun integration with a time step of 1 ms. Code implementing the model is available upon request via email.

Derivation of equation (3). The covariance in equation (2), $\mathrm{Cov}[R,N|\theta] = \langle RN|\theta\rangle - \langle R|\theta\rangle\langle N|\theta\rangle$, can be expressed in terms of expectations conditioned on the choice:

$$\mathrm{Cov}[R, N \mid \theta] = \sum_{i=1}^{n} P_{i,\theta}\langle RN \mid \theta, C_i\rangle - \sum_{i=1}^{n} P_{i,\theta}\langle R \mid \theta, C_i\rangle \sum_{i=1}^{n} P_{i,\theta}\langle N \mid \theta, C_i\rangle,$$ (9)

for a task with n possible choices C_i, which are selected with probabilities $P_{i,\theta}$ for stimulus θ. In tasks where reward is delivered on the basis of behavioural response, the reward is independent of neural activity when conditioned on the choice; therefore,

$$\langle RN \mid \theta, C_i\rangle = \langle R \mid \theta, C_i\rangle\langle N \mid \theta, C_i\rangle, = R_{i,\theta}N_{i,\theta},$$ (10)

where $R_{i,\theta}$ and $N_{i,\theta}$ denote the conditional expectations of reward and neural activity, respectively, for choice C_i and stimulus θ. In these terms, equation (9) can be rewritten as

$$\mathrm{Cov}[R, N \mid \theta] = \sum_{i=1}^{n}\sum_{j=1}^{n}(N_{i,\theta} - N_{j,\theta})R_{i,\theta}\,P_{i,\theta}P_{j,\theta}.$$ (11)

For tasks with only two possible choices, equation (11) simplifies to equation (3). In the categorization task reward is a deterministic function of choice (1 and 0 for correct and error choice, respectively); hence, the term $(R_{1,\theta} - R_{2,\theta})$ in equation (3) becomes $+1$ or -1 for stimuli $\theta \in C1$ or $\theta \in C2$, respectively.

Toy-model neuron. We simulated a toy-model neuron (Fig. 5) to illustrate that CP drives synaptic changes independently of a particular network architecture and behavioural task. On each trial, a choice C_1 or C_2 was selected with probability 0.5. The firing rate of the toy-model neuron was then sampled from a Gaussian distribution with the mean N_i for choice C_i and variance 5 Hz. To generate different CP values, the following (N_1,N_2) pairs were used: (55, 50), (51, 50), (50, 50), (50, 51) and (50, 55) Hz. Synaptic changes were simulated with the plasticity rule in equation (1). For simplicity, the firing rate of neuron on the other synaptic side was assumed to be static through learning and set to 1. The mean firing rate and CP of the toy-model neuron were also assumed not to change through learning for simplicity. As in the circuit model, the predicted reward $\langle R\rangle$ was estimated by the running average with $\tau_R = 5$, the learning rate was $q = 0.00003$ and the synapse was initialized at 0.5.

Behavioural task and neurophysiological recordings. All monkey data are from ref. 5, where experimental protocol and recording procedures were described in detail[69]. Two rhesus monkeys (*Macaca mulatta*, weighing about 14 kg) were trained to classify random-dot motion stimuli according to an arbitrary category boundary, which divided 360° of motion directions into two 180°-wide categories. Stimuli were circular patches (9° in diameter) of high-contrast square dots that moved with 100% motion coherence and at a speed of $12° s^{-1}$. Stimuli were always centred in the response field (RF) of the neuron under study. To dissociate categorical decisions from motor or premotor signals, the animals indicated category membership of the first stimulus (sample) by reporting (with a hand movement) whether it matched the category of the second stimulus (test). We focused on the categorization process of the sample stimulus and studied neural activity during the sample period (150–750 ms after stimulus onset, stimulus duration was 650 ms). To combine data from the two monkeys, all stimulus directions were rotated so that the category boundary was aligned with a 0°–180° axis.

The monkeys were implanted with a head post, scleral search coil and recording chamber. Recording chambers were implanted in accordance with coordinates (approximate centres at P3, L10) determined by magnetic resonance imaging, and allowed access to both the intraparietal sulcus (IPS) and the superior temporal sulcus by means of a dorsal approach. All surgical and experimental procedures followed the Harvard Medical School and National Institutes of Health guidelines. During LIP recordings, electrode penetrations sequentially encountered both the medial and lateral banks of the IPS. Most IPS neurons were tested with a memory-saccade task and a passive viewing flash-mapping task to generate detailed spatial maps of neuronal RFs. Neurons were considered to be in LIP if they showed spatially selective delay activity during the memory-saccade task or were located between such neurons in that electrode penetration. LIP neurons were not prescreened for direction selectivity. Area MT neurons were distinguished by direction-selective responses to moving spots and bars, and RF sizes that were roughly proportional to their eccentricity.

Data analysis. *Tuning curve characterization.* The firing rates of MT and LIP neurons were transformed to standard z-scores. Tuning curves $r(\theta)$ were then constructed by computing average standardized firing rates in response to 12 motion direction stimuli θ. Tuning curves $r(\theta)$ of MT and LIP neurons, as well as those of association neurons in the circuit model were fitted by directional and categorical tuning profiles (least squares fit). The directional tuning profile was modelled by an exponential cosine function:

$$r(\theta) = r_0 + r_{max}\exp[w(\cos(\theta - \theta_0) - 1)], \qquad (12)$$

where r_0 is the baseline firing rate, r_{max} is the peak amplitude, w is the tuning width parameter and θ_0 is the preferred direction. First, we obtained the median tuning width w for each population from the unconstrained fit, and then refitted tuning curves with w constrained within the 10 percentile range around the median (93.4°–126.1° for MT and 101.4°–142.7° for LIP and association neurons) to avoid very broad low amplitude (that is, nearly flat) directional fits. The resulting median tuning width was 120.9° for LIP and 104.9° for MT neurons, similar to previous reports[6]. The categorical tuning profile was modelled by a step function:

$$r(\theta) = \begin{cases} \langle r \mid \theta \rangle_{\theta \in C1}, & \text{if } \theta \in C1, \\ \langle r \mid \theta \rangle_{\theta \in C2}, & \text{if } \theta \in C2, \end{cases} \qquad (13)$$

where $\langle r \mid \theta \rangle_{\theta \in C_i}$ is the average firing rate across stimuli in category C_i. We repeated the analysis with more complex categorical tuning profiles (a periodic sigmoid function and a step-like function with a smoother firing rate change near category boundaries), and it did not change the conclusions of our study.

We then used a regularized GLM[26] to determine the relative contribution of fitted directional and categorical tuning profiles to neural firing rates. Regularized GLM provides a principled way to assess the relative strength of direction and category tuning in each neuron, without overfitting and avoiding confounds because of correlation between direction and category-tuning profiles for neurons tuned to category centres. The regression algorithm solves the matrix equation $\beta = (X^T X + \lambda I)^{-1} X^T r$, where X is the matrix of three factors: fitted directional tuning profile, categorical tuning profile and a baseline, r is the vector of neuron's firing rates across trials, β are the regression coefficients for each of the factors in X and λ is a ridge regression coefficient. The value of λ was chosen on the basis of a leave-one-trial-out cross-validation procedure, such that λ minimized the mean squared difference between predicted and actual firing rates[70]. To determine whether the resulting β coefficients were significantly different from zero, we used a standard t-test to compare β against the distribution of shuffled β values, which was obtained by randomizing the trial order and then refitting the linear regression model (1,000 reshuffles). Each neuron was then classified as direction-tuned or category-tuned if the corresponding β was significantly different from zero ($P < 0.05$), mixed direction- and category-tuned if both β's were significantly different from zero and nonselective if neither β was significantly different from zero.

CTI and CS. The CTI measured the difference in firing rate (averaged across all trials for each direction) for each neuron between pairs of directions in different categories (a between-category difference) and the difference in activity between pairs of directions in the same category (a within-category difference). The CTI was defined as the difference between the within-category and between-category

differences divided by their sum. Values of the index could vary from 1 (strong differences in activity to directions in the two categories) to −1 (large activity differences between directions in the same category, no difference between categories). A CTI value of 0 indicates the same difference in the firing rate between and within categories.

CS was estimated using a receiver-operating characteristic (ROC) analysis[17] applied to the distributions of firing rates on correct trials with stimuli from categories C1 and C2. CS is the area under the ROC curve, which ranges between 0 and 1, and indicates the accuracy with which an ideal observer can assign category membership of a stimulus on the basis of the neuron's trial-by-trial firing rate. Values of 1 and 0 correspond to strong preference for categories C1 and C2, respectively. Values of 0.5 indicate complete overlap of the firing rate distributions for the two categories, that is, no category selectivity.

Estimation of CP in MT and LIP neurons. CP was estimated on trials for which the test stimulus was far from (45° or 75°) the category boundary. The monkeys were proficient in categorizing such stimuli (97% correct when both sample and test were far from the boundary); therefore, we assumed that on these trials the test stimulus was categorized correctly and inferred the monkey's decision about the sample category to be the same as the test category if the monkey responded match, and different category if the monkey responded nonmatch[54]. For each stimulus, CP was estimated using an ROC analysis applied to the distributions of firing rates on trials with different category decisions for the same stimulus (that is, correct versus error trials). CP is the area under the ROC curve that ranges between 0 and 1 and indicates the accuracy with which an ideal observer can predict the monkey's category decision on a trial-by-trial basis given neuron's firing rate. Values of 1 and 0 correspond to strong preference (higher firing rate) for C_1 and C_2 category decisions, respectively. Values of 0.5 indicate complete overlap of the firing rate distributions for two decisions. To reliably estimate CP, only stimuli with at least three trials for each category choice were included in the analysis, and only those neurons were included that had a valid CP estimate for at least one stimulus in each category, which resulted in 88 LIP and 31 MT neurons left for the analysis. The CP reported for each neuron was the average CP across all stimuli that passed the inclusion criteria. Significance of CP values for individual neurons was assessed with a shuffle test. To this end, choices of the monkey were randomly assigned to the firing rate data (separately for each stimulus), and then CP was recomputed (1,000 reshuffles). The actual CP was compared with the shuffled distribution with a two-sample t-test.

References

1. Rosch, E. Principles of categorization. in: *Concepts: Core Readings* (eds Margolis, E. & Laurence, S.) 189–206 (MIT Press, Cambridge, Massachusetts, 1999).
2. Ashby, F. G. & Maddox, W. T. Human category learning. *Annu. Rev. Psychol.* **56**, 149–178 (2005).
3. Ghose, G. M., Yang, T. & Maunsell, J. H. R. Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophys.* **87**, 1867–1888 (2002).
4. Yang, T. & Maunsell, J. H. R. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* **24**, 1617–1626 (2004).
5. Freedman, D. J. & Assad, J. A. Experience-dependent representation of visual categories in parietal cortex. *Nature* **443**, 85–88 (2006).
6. Fanini, A. & Assad, J. A. Direction selectivity of neurons in the macaque lateral intraparietal area. *J. Neurophys.* **101**, 289–305 (2009).
7. Law, C.-T. & Gold, J. I. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* **11**, 505–513 (2008).
8. Purushothaman, G. & Bradley, D. C. Neural population code for fine perceptual decisions in area MT. *Nat. Neurosci.* **8**, 99–106 (2005).
9. Uka, T., Sasaki, R. & Kumano, H. Change in choice-related response modulation in area MT during learning of a depth-discrimination task is consistent with task learning. *J. Neurosci.* **32**, 13689–13700 (2012).
10. Law, C.-T. & Gold, J. I. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat. Neurosci.* **12**, 655–663 (2009).
11. Rombouts, J., Bohte, S. & Roelfsema, P. Neurally plausible reinforcement learning of working memory tasks. *Adv. Neural Inf. Process. Syst.* **25**, 1880–1888 (2012).
12. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
13. Schultz, W. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).
14. Fremaux, N., Sprekeler, H. & Gerstner, W. Functional requirements for reward-modulated spike-timing-dependent plasticity. *J. Neurosci.* **30**, 13326–13337 (2010).
15. Loewenstein, Y. & Seung, H. S. Operant matching is a generic outcome of synaptic plasticity based on the covariance between reward and neural activity. *Proc. Natl Acad. Sci. USA* **103**, 15224–15229 (2006).
16. Bishop, C. M. *Neural Networks for Pattern Recognition* (Oxford University Press, 1995).
17. Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
18. Nienborg, H., Cohen, M. R. & Cumming, B. G. Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annu. Rev. Neurosci.* **35**, 463–483 (2012).

19. Engel, T. A. & Wang, X.-J. Same or different? A neural circuit mechanism of similarity-based pattern match decision making. *J. Neurosci.* **31,** 6982–6996 (2011).

20. Wang, X. J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36,** 955–968 (2002).

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

22. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299,** 1898–1902 (2003).

23. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47,** 129–141 (2005).

24. Reynolds, J. N., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. *Nature* **413,** 67–70 (2001).

25. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321,** 848–851 (2008).

26. Dobson, A. *An Introduction to Generalized Linear Models* (Chapman & Hall/CRC, 2001).

27. Borg, I. & Groenen, P. J. F. *Modern Multidimensional Scaling: Theory and Applications* (Springer, 1997).

28. Marzban, C. The ROC curve and the area under it as performance measures. *Wea. Forecasting* **19,** 1106–1114 (2004).

29. Shadlen, M. N., Britten, K. H., Newsome, W. T. & Movshon, J. A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16,** 1486–1510 (1996).

30. Martinez-Trujillo, J. C. & Treue, S. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr. Biol.* **14,** 744–751 (2004).

31. Ardid, S., Wang, X.-J. & Compte, A. An integrated microcircuit model of attentional processing in the neocortex. *J. Neurosci.* **27,** 8486–8495 (2007).

32. Cohen, M. R. & Newsome, W. T. Context-dependent changes in functional circuitry in visual area MT. *Neuron* **60,** 162–173 (2008).

33. Gu, Y., Angelaki, D. E. & DeAngelis, G. C. Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* **11,** 1201–1210 (2008).

34. Nienborg, H. & Cumming, B. G. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* **459,** 89–92 (2009).

35. Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **8,** 229–256 (1992).

36. Xie, X. & Seung, H. S. Learning in neural networks by reinforcement of irregular spiking. *Phys. Rev. E* **69,** 1–10 (2004).

37. Pfister, J. P., Toyoizumi, T., Barber, D. & Gerstner, W. Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning. *Neural Comput.* **18,** 1318–1348 (2006).

38. Seung, H. S. Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* **40,** 1063–1073 (2003).

39. Vasilaki, E., Fremaux, N., Urbanczik, R., Senn, W. & Gerstner, W. Spike-based reinforcement learning in continuous state and action space: when policy gradient methods fail. *PLoS Comp. Biol.* **5,** e1000586 (2009).

40. Urbanczik, R. & Senn, W. Reinforcement learning in populations of spiking neurons. *Nat. Neurosci.* **12,** 250–252 (2009).

41. Roelfsema, P. R. & van Ooyen, A. Attention-gated reinforcement learning of internal representations for classification. *Neural Comput.* **17,** 2176–2214 (2005).

42. Soltani, A. & Wang, X.-J. Synaptic computation underlying probabilistic inference. *Nat. Neurosci.* **13,** 112–119 (2010).

43. Fitzgerald, J. K., Freedman, D. J. & Assad, J. A. Generalized associative representations in parietal cortex. *Nat. Neurosci.* **14,** 1075–1079 (2011).

44. Goodwin, S. J., Blackman, R. K., Sakellaridi, S. & Chafee, M. V. Executive control over cognition: stronger and earlier rule-based modulation of spatial category signals in prefrontal cortex relative to parietal cortex. *J. Neurosci.* **32,** 3499–3515 (2012).

45. Toth, L. J. & Assad, J. A. Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature* **415,** 165–168 (2002).

46. Stoet, G. & Snyder, L. H. Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron* **42,** 1003–1012 (2004).

47. Ferrera, V. P., Yanike, M. & Cassanello, C. Frontal eye field neurons signal changes in decision criteria. *Nat. Neurosci.* **12,** 1458–1462 (2009).

48. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503,** 78–84 (2013).

49. Swaminathan, S. K., Masse, N. Y. & Freedman, D. J. A comparison of lateral and medial intraparietal areas during a visual categorization task. *J. Neurosci.* **33,** 13157–13170 (2013).

50. Sigala, N. & Logothetis, N. K. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* **415,** 318–320 (2002).

51. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci.* **23,** 5235–5246 (2003).

52. Wallis, J. & Miller, E. From rule to response: neuronal processes in the premotor and prefrontal cortex. *J. Neurophys.* **90,** 1790–1806 (2003).

53. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291,** 312–316 (2001).

54. Swaminathan, S. K. & Freedman, D. J. Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. *Nat. Neurosci.* **15,** 315–320 (2012).

55. Cromer, J. A., Roy, J. E. & Miller, E. K. Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron* **66,** 796–807 (2010).

56. Szabo, M. *et al.* Learning to attend: modeling the shaping of selectivity in infero-temporal cortex in a categorization task. *Biol. Cyber* **94,** 351–365 (2006).

57. Roelfsema, P. R., van Ooyen, A. & Watanabe, T. Perceptual learning rules based on reinforcers and attention. *Trends Cogn. Sci.* **14,** 64–71 (2010).

58. Freedman, D. J. & Assad, J. A. A proposed common neural mechanism for categorization and perceptual decisions. *Nat. Neurosci.* **14,** 143–146 (2011).

59. DeLa Rocha, J., Wimmer, K., Renart, A., Roxin, A. & Compte, A. in *Society for Neuroscience Annual Meeting* (New Orleans, LA, USA (2012).

60. Freedman, D. J. & Assad, J. A. Distinct encoding of spatial and nonspatial visual information in parietal cortex. *J. Neurosci.* **29,** 5671–5680 (2009).

61. Rishel, C. A., Huang, G. & Freedman, D. J. Independent category and spatial encoding in parietal cortex. *Neuron* **77,** 969–979 (2013).

62. Haefner, R. M., Gerwinn, S., Macke, J. H. & Bethge, M. Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat. Neurosci.* **16,** 235–242 (2013).

63. Shiozaki, H. M., Tanabe, S., Doi, T. & Fujita, I. Neural activity in cortical area V4 underlies fine disparity discrimination. *J. Neurosci.* **32,** 3830–3841 (2012).

64. Compte, A., Brunel, N., Goldman-Rakic, P. S. & Wang, X. J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex* **10,** 910–923 (2000).

65. Abbott, L. F. & Chance, F. S. In *Cortical Function: a View from the Thalamus, vol. 149 of Progress in Brain Research* (eds Guillery, V. C. R. & Sherman, S.) 147–155 (Elsevier, 2005).

66. Carandini, M. & Heeger, D. J. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13,** 51–62 (2011).

67. Crapse, T. B. & Sommer, M. A. Corollary discharge circuits in the primate brain. *Curr. Opin. Neurobiol.* **18,** 552–557 (2008).

68. Soltani, A. & Wang, X.-J. A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* **26,** 3731–3744 (2006).

69. Freedman, D. & Assad, J. Distinct encoding of spatial and nonspatial visual information in parietal cortex. *J. Neurosci.* **29,** 5671–5680 (2009).

70. Hastie, T., Tibshirani, R. & Friedman, J. H. *Data Mining, Inference, and Prediction* (Springer, 2009).

Acknowledgements

Author contributions

T.A.E., W.C. and X.-J.W. designed research. T.A.E. and W.C. performed model simulations and analysed data. D.J.F. designed and performed experiments. T.A.E., W.C., D.J.F. and X.-J.W. discussed results and wrote the paper.

Additional information