bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Signatures of optimal inference during learning in rat prefrontal cortex

Abhinav Singh[1], Adrien Peyrache[2], Mark D. Humphries[1]

1. Faculty of Life Sciences, University of Manchester, Manchester, M13 9PT, United Kingdom
2. The Neuroscience Institute, School of Medicine, New York University, New York, New York, USA.

Contact: Correspondence should be addressed to M.D.H (mark.humphries@manchester.ac.uk)

Abstract

Cortical population activity may represent sampling from an internal model. Using data from rats learning rules on a maze task, we show that sampling rates of population activity patterns in prefrontal cortex converge between waking and sleep over learning. Sample rate changes were greatest for activity patterns predictive of correct choice and expressed at the choice point. Our results suggest inference-by-sampling is a general computational principle for cortex.

Cortical neurons collectively code and compute information[1,2]. The recent inference-by-sampling hypothesis proposes that cortical population activity at some time t is a sample from an underlying probability distribution[3,4], which can be reconstructed by integrating over samples. A key prediction is that the distribution during "spontaneous" activity (representing the prior) and during evoked activity (representing the posterior) converge over repeated experience. This convergence represents the updating of an internal model to match the relevant statistics of the external world. Just such a convergence has been observed in small populations from ferret V1 over development[4]. Unknown is the extent to which this hypothesis is a general computational principle for cortex[5,6]: whether it can be observed during learning, or in higher-order cortices, or during ongoing behaviour.

To address these issues, we analysed previously-recorded population activity from the medial prefrontal cortex (mPFC) of rats learning rules in a Y-maze task[7] (Fig. 1a). Medial PFC is necessary for learning new rules or strategies[8,9], and change in mPFC neuron firing times correlates with successful rule learning[10], suggesting PFC plays a role in building an internal model of a task. We theorised that mPFC population activity on each trial was sampling from the posterior distribution over (unknown) task parameters; and that "spontaneous" activity in slow-wave sleep (SWS), occurring in the absence of task-related stimuli and behaviour, samples the prior distribution. We focussed on ten sessions where the animal reached the learning criterion for a rule mid-session (Methods), after at least one full prior session with the same rule. In this way, we sought to isolate changes in population activity solely due to rule-learning, assuming that prior experience allowed learning of the basic task parameters, including stimuli and reward locations.

Activity patterns were characterised as a binary vector (or "word") of active and inactive neurons (Fig. 1b). Throughout we consider only "co-activation" patterns with two or more active neurons (see Methods). Consistent with being samples from related

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

posterior (in task) and prior (in sleep) distributions, each recorded population visited the same highly limited sub-space of all possible activity patterns across sleep and behaviour (Supplementary Figure 1)[2,11].

We found that the task-evoked distribution of patterns after learning was more similar to the distribution in post-task than in pre-task SWS (Fig. 1c-f). This increase in similarity only occurred for rewarded trials, not for unrewarded trials (Fig. 1g). The increase was robust to choice of activity pattern binsize (Supplementary Figure 2). Together, these results are consistent with the convergence over learning of the posterior and prior distributions represented by mPFC population activity.

Population firing rate differences between waking and sleep states, and increases in SWS firing after task-learning, could account for the convergence of distributions[12,13]. To control for this, we used the "raster" model[12] to generate surrogate sets of spike-trains that matched both the mean firing rates of each neuron, and matched the moment-to-moment total population activity in each time-bin. These rate effects could not account for the convergence (Fig. 1h).

In the inference-by-sampling hypothesis, the difference between the prior distributions sampled in spontaneous activity before and after learning should reflect the changed coding of the learnt parameter. Here this predicts that the activity patterns differing most in sampling between pre- and post-task SWS should be those related to the learnt rule. Remarkably, this is exactly what we find: the patterns most predictive of outcome (Fig. 2a) are also those that changed their sampling frequency the most between pre- and post-task SWS (Fig. 2b,c).

Consistent with the predictive patterns playing a guiding role in behaviour, we further found that they preferentially occurred around the choice point of the maze (Fig. 3a,b). Particularly striking was that patterns strongly predictive of outcome rarely occurred in the starting arm (Fig. 3a). Together, the selective sampling changes over learning to outcome-specific (Fig. 2) and location-specific (Fig. 3) activity patterns show that the convergence of distributions (Fig. 1) is not a statistical curiousity, but reflects the updating of a behaviourally-relevant probability distribution.

Prefrontal cortex has been implicated in planning and working memory during spatial navigation[14–17]. We find moment-to-moment samples of population activity converge over learning of a spatial navigation task, consistent with an inference-by-sampling computation underpinning task performance. Remarkably we observe this in precise activity patterns down to 2 ms resolution. Previous work observed fine structure in stimulus-evoked population activity patterns[4,18,19]; here in a behavioural task we have identified clues to what such patterns encode – in this case, the decision rule. Our analyses suggest mPFC constructs a probabilistic internal model of task parameters, which is sampled online to guide behaviour.

Methods

Task and electrophysiological recordings The data analysed here were from ten recording sessions in the study of ref.[7]. For full details on training, spike-sorting, and histology please see ref.[7]. Four Long-Evans male rats with implanted tetrodes in prelimbic cortex were trained on the Y-maze task (Fig. 1a). Each recording session consisted of a 20-30 minute sleep or rest epoch (pre-task epoch), in which the rat remained undisturbed in a padded flowerpot placed on the central platform of the maze, followed by a task epoch, in which the rat performed for 20-40 minutes, and then by a second 20-30 minute sleep or rest epoch (post-task epoch). Every trial started when the rat reached the departure

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

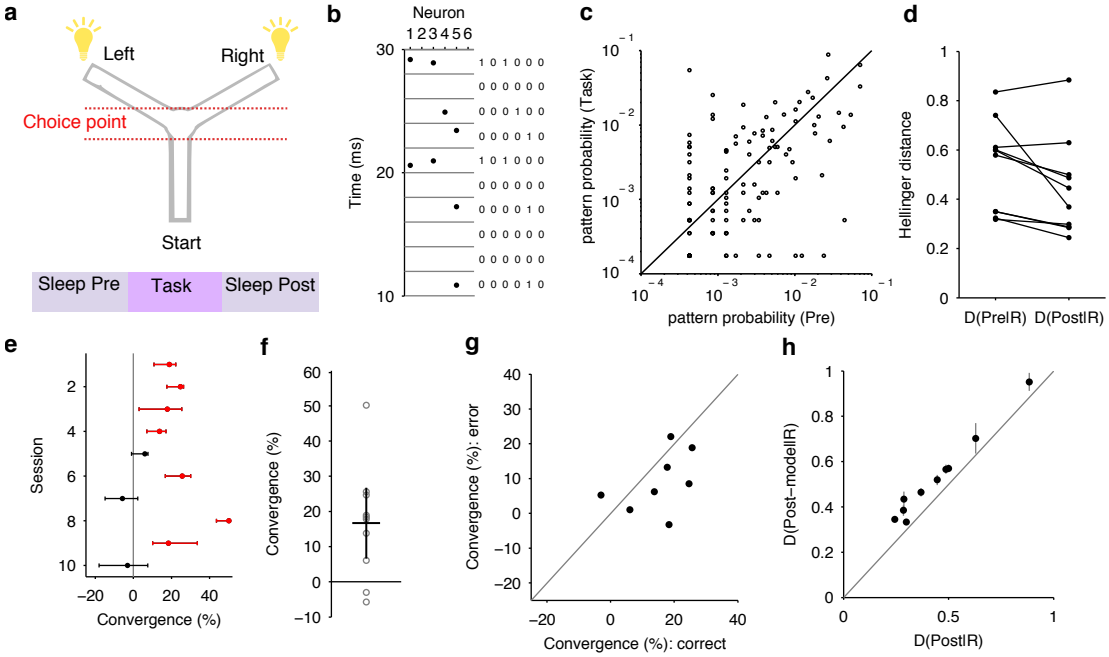

Figure 1: Convergence of activity pattern distributions between the task and post-task sleep. **a** Y-maze task set-up (top); each session included the epochs of pre-task sleep, task trials, and post-task sleep (bottom). One of three target rules for obtaining reward was enforced throughout a session: go right; go left; go to the lit arm. **b** The population activity of simultaneously recorded spike trains was represented as a binary activity pattern in some small time-bin (here 2 ms). **c** Scatter plot of the joint frequency of every occurring pattern in pre-task SWS and task epochs for one session. **d** Distances between the distributions of pattern frequencies in sleep and task epochs; one dot per session. $D(X|Y)$: distance between pattern distributions in epochs X and Y: Pre: pre-task SWS; Post: post-task SWS; R: correct task trials. **e** Median and 95% bootstrapped CI for the relative convergence of task and post-task SWS activity pattern distributions for each session. Convergence greater than zero means that the activity pattern distribution in the task is closer to the distribution in post-task SWS than the distribution in pre-task SWS. Red: sessions with CIs above 0. **f** Scatter of convergence across all sessions (circles). Black lines give mean ± 2 s.e.m (convergence greater than zero at $P = 0.0049$). **g** Scatter of convergence for correct against error trials for all sessions (difference in convergence greater than zero at $P = 0.019$). **h** Distance between task and post-task sleep distributions $D(Post|R)$ is always smaller than predicted by firing rate changes alone $D(Post - model|R)$. For model predictions, dots and bars give means and 95% intervals.

arm and finished when the rat reached the end of one of the choice arms. Correct choice was rewarded with drops of flavoured milk. Each rat had to learn the current rule by trial-and-error, either: go to the left arm; go to the right arm; go to the lit arm. To maintain consistent context across all sessions, the extra-maze light cues were lit in a pseudo-random sequence across trials, whether they were relevant to the rule or not.

We analysed here data from the ten sessions in which the previously-defined learning criterion trial was reached: the first trial of a block of at least three consecutive rewarded trials after which the performance until the end of the session was above 80%. In later sessions (not analysed here) the rats reached the criterion for changing the rule: ten consecutive correct trials or one error out of 12 trials. Thus each rat learnt at least two rules.

Tetrode recordings were spike-sorted only within each recording session for conservative identification of stable single units. In the ten sessions we analyse here, the populations

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

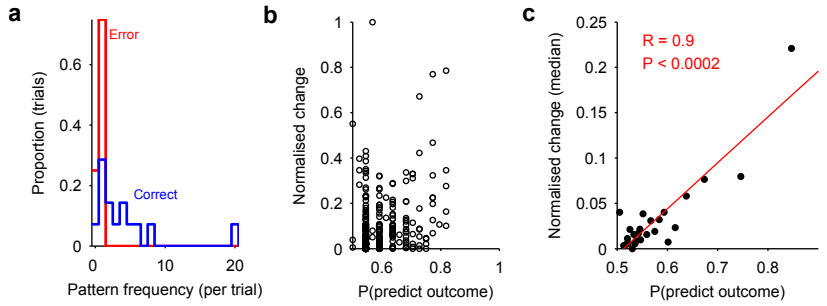

Figure 2: Coding of trial outcome by sampled activity patterns. **a** Example distributions of a pattern's frequency conditioned on trial outcome from one session. **b** For all co-activation patterns in one session, the correlation between outcome prediction and change in pattern frequency between pre- and post-task SWS. **c** Correlation of outcome prediction and change in pattern occurrence between sleep epochs, over all ten sessions. Dots show median change values (for N=90 data-points each; Supplementary Fig. 3 shows that the regression is robust to all choices of N). Red line is the best-fit linear regression ($P < 0.0002$, permutation test).

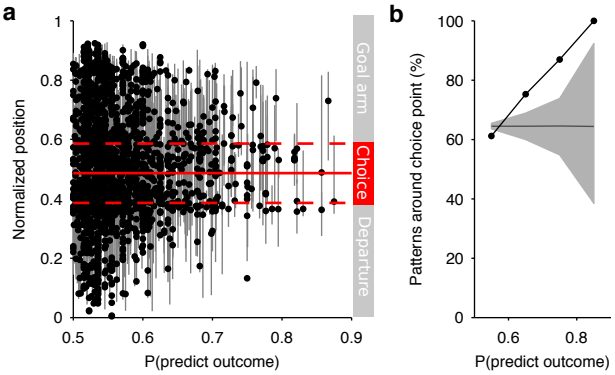

Figure 3: Outcome predicting activity patterns are sampled at the choice area. **a** Scatter plot of each pattern's outcome prediction and sample locations in the maze (dot is median position; grey line is interquartile range); all positions given as a proportion of the linearised maze from start of departure arm. Red lines indicate the approximate centre (solid) and boundaries (dashed) of the maze's choice area. **b** Proportion of activity patterns whose interquartile range of sample locations crosses the choice area (black dots and line). Grey region shows mean (line) and 95% range (shading) of proportions from permutation test.

ranged in size from 15-55 units.

Activity pattern distributions For a population of size N, we characterised population activity from time t to $t+\delta$ as an N-length binary vector with each element being 1 if at least one spike was fired by that neuron in that time-bin, and 0 otherwise. In the main text we use a binsize of $\delta = 2$ ms throughout, and report the results of using larger binsizes in Supplementary Figure 2. We build patterns using the number of recorded neurons N, up to a maximum of 35 for computational tractability.

The probability distribution for these activity patterns was compiled by counting the frequency of each pattern's occurrence and normalising by the total number of pattern occurrences.

Throughout the main text we consider only co-activation patterns consisting of two or more active neurons. We do this to focus on genuine changes in population activity patterns distinct from pure firing rate changes: patterns with only one active neuron could

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

change in occurrence either because of changes in relative spike-timing of that (or other) neurons, or because that neuron changed its firing rate. Moreover, when considering the predictive coding (Fig. 2) and position (Fig. 3) properties of individual patterns, this choice allows us to rule out pure firing rate codes. Our main results of convergence were obtained despite this restriction: if we do include all patterns in our distributions, then we find that the convergence between task and post-task sleep distributions is both stronger and consistent across every session (Supplementary Figure 4).

Comparing distributions We quantified the distance $D(P|Q)$ between probability distributions P and Q using the standard Hellinger distance. Given we have two discrete probability distributions $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$, the Hellinger distance is $D(P|Q) = \sqrt{\sum_{i=1}^{n}(\sqrt{p_i} - \sqrt{q_i})^2}$. To a first approximation, this measures for each pair of probabilities (p_i, q_i) the Euclidean distance between their square-roots. In this form, $D(P|Q) = 0$ means the distributions are identical, and $D(P|Q) = \sqrt{2}$ means the distributions are mutually singular: all positive probabilities in P are zero in Q, and vice-versa.

To compare the distances between pairs of distributions, we compute convergence as $100 \times (D(Pre|R) - D(Post|R))\,/D(Pre|R)$. Convergence greater than 0% indicates that the distance between the task (R: correct trials) and post-task SWS ($Post$) distributions is smaller than that between the task and pre-task SWS (Pre) distributions.

Our main goal was to compare the distribution of activity patterns occurring during task trials (while the animal was putatively learning) with the distributions in pre- and post-task SWS. Under the inference-by-sampling hypothesis, we expect the distribution in the task is closer to that in post-task SWS, reflecting the updated prior distribution from which the population is sampling during spontaneous activity. If mPFC is playing a role in rule-learning, then we would expect that patterns who change sampling frequency due to learning would only unambiguously appear on correct trials. Consequently, the convergence between distributions could be confounded by correct arm choices made by chance without learning the rule: in these trials, sampled activity patterns are not causal for the correct choice, and their sampling rate would not be matched in the "spontaneous" distribution after successful learning. To control for this, we constructed distributions for correct trials from all correct trials after the learning criterion trial, which attempts to identify the time-point at which the animal has clearly learnt the correct rule. Similarly, only error trials before the learning criterion trial were included in the error-trial distribution. Consistent with this confound, when including all correct trials in our task distribution, incorporating likely success by random choice, we see a weakening of convergence between task and post-task SWS distributions (Supplementary Fig. 5).

Statistics Bootstrapped confidence intervals (in Fig. 1e) for each session were constructed using 1000 bootstraps of each epoch's distribution. Each bootstrap was a sample-with-replacement of activity patterns from the data distribution X to get a sample distribution X^*. For a given pair of bootstrapped distributions X^*, Y^* we then compute their distance $D^*(X^*|Y^*)$. Given both bootstrapped distances $D^*(Pre|R)$ and $D^*(Post|R)$, we then compute the bootstrapped convergence $(D^*(Pre^*|R^*) - D^*(Post^*|R^*))\,/D^*(Pre^*|R^*)$.

All hypothesis tests used the non-parametric Wilcoxon ranksum test for either: a one-sample test that the sample median is greater than zero (Fig. 1f; Supplementary Figures 4 and 5); or a paired-sample test that the difference in sample medians is greater than zero (Fig. 1g). Equivalent results were obtained using t-tests despite low power.

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Raster model To control for the possibility that changes in activity pattern occurrence were due solely to changes in the firing rates of individual neurons and the total population, we used the raster model - see ref.[12] for details. For a given data-set of spike-trains N and binsize δ, the raster model constructs a synthetic set of spikes such that each synthetic spike-train has the same mean rate as its counterpart in the data, and the distribution of the total number of spikes per time-bin matches the data. In this way, it predicts the frequency of activity patterns that should occur given solely changes in individual and population rates. For Fig 3g, we generated 1000 raster models per session using the spike-trains from the post-task SWS in that session. For each generated raster model, we computed the distance between its distribution of activity patterns and the data distribution for correct trials in the task. This comparison gives the expected distance between task and post-task SWS distributions due to firing rate changes alone. We plot the mean and 95% interval of that distance over the 1000 raster models (Fig. 1h).

Outcome prediction To check whether individual activity patterns coded for the outcome on each trial, we used standard receiver-operating characteristic (ROC) analysis. For each pattern, we computed the distribution of its occurrence frequencies separately for correct and error trials (as in the example of Fig. 2a). We then used a linear classifier to compute coding of outcome: for a given frequency threshold T, we find the fraction of correctly classified correct trials (true positive rate) and the fraction of error trials incorrectly classified as correct trials (false positive rate). Plotting the false positive rates against the true positive rates for all values of T gives the ROC curve. The area under the ROC curve gives the probability that a randomly chosen pattern frequency will be correctly classified as from a correct trial; we report this as P(predict outcome).

Relationship of sampling change and outcome prediction Within each session, we computed the change in each pattern's occurrence between pre- and post-task SWS. These were normalised by the maximum change within each session. Maximally changing patterns were candidates for those updated by learning during the task. Correlation between change in pattern sampling and outcome prediction was done on normalised changes pooled over all sessions. Change scores were binned using variable-width bins of P(predict outcome), each containing the same number of data-points to rule out power issues affecting the correlation. We regress P(predict outcome) against median change in each bin, using the mid-point of each bin as the value for P(predict outcome). Our main claim is that prediction and change are dependent variables (Fig. 2c). To test this claim, we compared the data correlation against the null model of independent variables, by permuting the assignment of change scores to the activity patterns. For each permutation, we repeat the binning and regression. We permuted 5000 times to get the sampling distribution of the correlation coefficient R^* predicted by the null model of independent variables. To check robustness, all analyses were repeated for a range of fixed number of data-points per bin between 20 and 100 (Supplementary Fig. 3).

Relationship of location and outcome prediction The location of every occurrence of an activity pattern was expressed as normalized position on the linearised maze (0: start of departure arm; 1: end of the chosen goal arm). Our main claim is that activity patterns strongly predictive of outcome occur predominantly around the choice point of the maze, and so prediction and overlap of the choice area are dependent variables (Fig. 3b). To test this claim, we compared this relationship against the null model of independent variables, by permuting the assignment of average location (median and interquartile range) to the

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

activity patterns. For each permutation, we compute the proportion of patterns whose interquartile range overlaps the choice area, and bin as per the data. We permuted 5000 times to get the sampling distribution of the proportions predicted by the null model of independent variables: we plot the mean and 95% range of this sampling distribution as the grey region in Figure 3b.

Author Contributions M.D.H and A.S. designed the analyses; A.S. and M.D.H. analysed the data; all authors discussed the results; M.D.H wrote the paper with contributions from A.S. and A.P.

Acknowledgements We thank the Humphries lab for discussions, and M. Okun for making his raster model code publicly available. A.S. and M.D.H were supported by a Medical Research Council Senior non-Clinical Fellowship award to M.D.H. A.P. was supported by Human Frontier Science Program Fellowship LT000160/2011-l and National Institute of Health Award K99 NS086915-01.

References

1. Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E. Probabilistic brains: knowns and unknowns. *Nat Neurosci* **16**, 1170–1178 (2013).

2. Wohrer, A., Humphries, M. D. & Machens, C. Population-wide distributions of neural activity during perceptual decision-making. *Prog Neurobiol* **103**, 156–193 (2013).

3. Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci* **14**, 119–130 (2010).

4. Berkes, P., Orbán, G., Lengyel, M. & Fiser, J. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* **331**, 83–87 (2011).

5. Buesing, L., Bill, J., Nessler, B. & Maass, W. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput Biol* **7**, e1002211 (2011).

6. Habenschuss, S., Jonke, Z. & Maass, W. Stochastic computations in cortical microcircuit models. *PLoS Comput Biol* **9**, e1003311 (2013).

7. Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I. & Battaglia, F. P. Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat Neurosci* **12**, 916–926 (2009).

8. Ragozzino, M. E., Detrick, S. & Kesner, R. P. Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *J Neurosci* **19**, 4585–4594 (1999).

9. Rich, E. L. & Shapiro, M. L. Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *J Neurosci* **27**, 4747–4755 (2007).

10. Benchenane, K. *et al.* Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron* **66**, 921–936 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

11. Luczak, A., Barth, P. & Harris, K. D. Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* **62**, 413–425 (2009).

12. Okun, M. *et al.* Population rate dynamics and multineuron firing patterns in sensory cortex. *J Neurosci* **32**, 17108–17119 (2012).

13. Fiser, J., Lengyel, M., Savin, C., Orbán, G. & Berkes, P. How (not) to assess the importance of correlations for the matching of spontaneous and evoked activity. arXiv:1301.6554 (2013).

14. Baeg, E. H. *et al.* Dynamics of population code for working memory in the prefrontal cortex. *Neuron* **40**, 177–188 (2003).

15. Fujisawa, S., Amarasingham, A., Harrison, M. T. & Buzsáki, G. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat Neurosci* **11**, 823–833 (2008).

16. Ito, H. T., Zhang, S.-J., Witter, M. P., Moser, E. I. & Moser, M.-B. A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation. *Nature* **522**, 50–55 (2015).

17. Spellman, T. *et al.* Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature* **522**, 309–314 (2015).

18. Schneidman, E., Berry, M. J., Segev, R. & Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012 (2006).

19. Tkacik, G. *et al.* Searching for collective behavior in a large network of sensory neurons. *PLoS Comput Biol* **10**, e1003408 (2014).

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

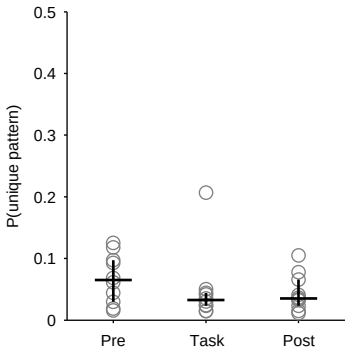

Figure S1: Consistent sampling of activity patterns across session epochs. Each circle is the proportion of co-activation patterns that appeared only in that epoch of the session. Black bar and line give the median and interquartile range across the 10 sessions.

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

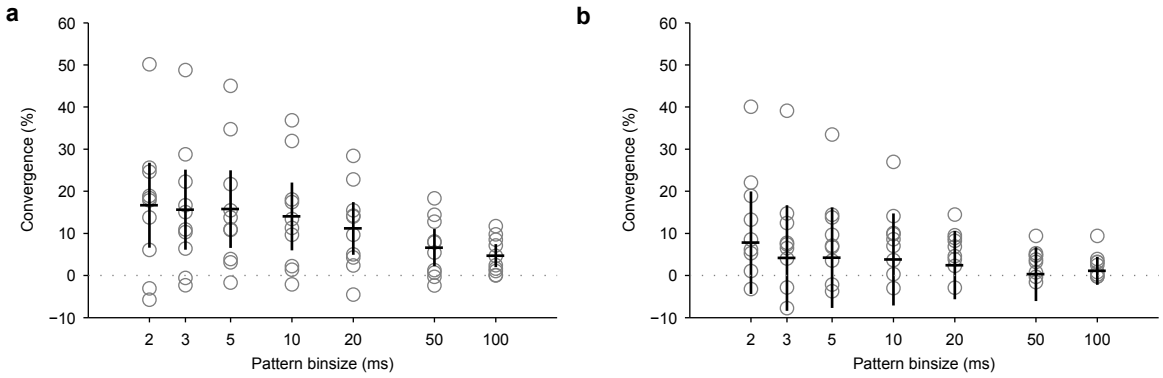

Figure S2: Consistent convergence of distributions over choice of pattern binsize. We plot here the dependence of the convergence of task and post-task SWS distributions on the binsize used for constructing the activity patterns. We see that the convergence of the distribution of patterns on correct task trials (**a**) is robust to order of magnitude changes in binsize (the distribution at the binsize of 2ms is plotted in Fig. 1f). Similarly, we see that the non-convergence of the distribution of patterns on incorrect task trials (**b**) is also robust to order of magnitude changes in binsize. Circles are convergence in individual sessions. Black lines give mean ± 2 s.e.m.

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

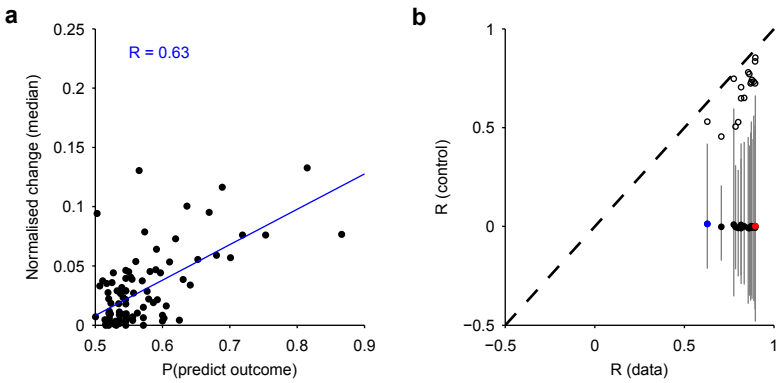

Figure S3: Consistent correlation between choice prediction and change in sampling rate. a Correlation of outcome prediction and change in pattern occurrence between sleep epochs, over all ten sessions. Here we plot the worst-case correlation amongst all tested. Dots show median change values over N=25 data-points each. Blue line is the best-fit linear regression ($P < 0.0002$, permutation test). **b** Scatter of correlation in the data against predicted correlation by the null model of independent relationship between outcome prediction and sampling change. The data axis plots the correlation value (R) found for every tested data regression using between N=20 and N=100 data-points per bin in steps of 5. The blue dot is the worst-case fit in panel a; the red dot is the fit shown in Figure 2c, main text. The control axis plots the median (dot) and 95% range (line) of correlation coefficient (R) for regressions between the outcome prediction scores and permuted change scores (for 5000 permutations). All data points fall beyond the maximum correlation reported by all 5000 permutations (black circles), giving $P < 0.0002$ for every data correlation.

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

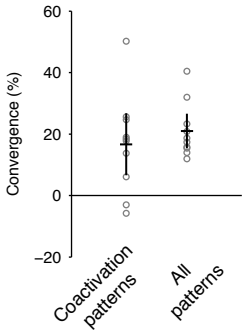

Figure S4: Using all activity patterns increases convergence.. In the main text, we conversatively base our analyses on co-activation patterns of two or more active neurons, to help rule out firing rate effects (see Methods). We replot here the scatter of convergence scores for co-activation patterns from Fig. 1f. If we include all activity patterns in our analysis of distances between distributions, we find clear convergence between task and post-task SWS distributions for every session (convergence greater than zero at $P = 0.001$); if anything, this convergence is more consistent than when using only co-activation patterns. Circle: convergence for a session. Black lines give mean $\pm$ 2 s.e.m.

bioRxiv preprint doi: https://doi.org/10.1101/027102; this version posted September 21, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

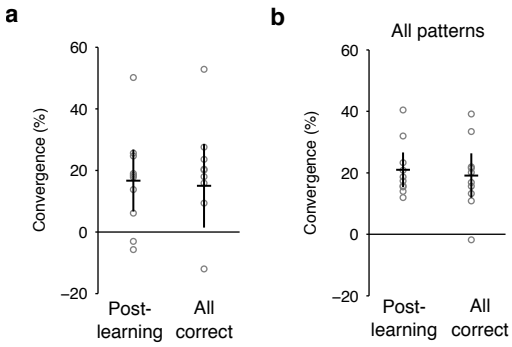

Figure S5: Effect of including all correct trials. In the main text, we construct all task-related distributions by considering only correct trials after the learning criterion trial, to isolate the population activity changes due to learning: correct trials before obtaining the rule may not contain causal samples from the population (see Methods). **a** If we include all correct trials of a session, we find that convergence between task and post-task SWS distributions is still present (convergence greater than zero at $P = 0.042$). However it is clearly weaker than when considering only trials post-learning. This indicates that the learning criterion trial does usefully separate periods of random choice from rule-learning, and the consequent difference in patterns in the mPFC population activity. Note that the convergence between task and post-task SWS distributions is greater for post-learning trials, even though that task distribution is built from fewer samples, and so might be expected to be noisier. **b** If we include all correct trials in the task distribution and include all activity patterns, then convergence between task and post-task SWS distributions is still strongly present (convergence greater than zero at $P = 0.002$).