

 Open access • Posted Content • DOI:10.1101/027102

Medial prefrontal cortex population activity is plastic irrespective of learning

— [Source link](#) 

Abhinav Singh, Adrien Peyrache, Mark D. Humphries

Institutions: University of Manchester, McGill University, University of Nottingham

Published on: 27 Jul 2018 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Prefrontal cortex, Population and Neuroplasticity

Related papers:

- [Dynamics of population code for working memory in the prefrontal cortex.](#)
- [Multimodal Plasticity in Dorsal Striatum While Learning a Lateralized Navigation Task](#)
- [Anterior-posterior gradient of plasticity in primate prefrontal cortex.](#)
- [Rule learning enhances structural plasticity of long-range axons in frontal cortex](#)
- [Learning-Induced Enduring Changes in Functional Connectivity among Prefrontal Cortical Neurons](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/medial-prefrontal-cortex-population-activity-is-plastic-1d00ybi082>

Medial prefrontal cortex population activity is plastic irrespective of learning

Abbreviated title: Population plasticity in prefrontal cortex

Abhinav Singh¹, Adrien Peyrache², Mark D. Humphries^{1,3*},

1 Faculty of Biology, Medicine and Health, University of Manchester, Manchester, M13 9PT, UK

2 Montreal Neurological Institute, McGill University, Montreal, QC H3A 1A1, Canada

3 School of Psychology, University of Nottingham, Nottingham, NG7 2RD, UK

* Corresponding author: mark.humphries@nottingham.ac.uk

Pages: 24

Figures: 8

Tables: 2

Abstract: 236

Introduction: 384

Discussion: 1417

Conflict of interest The authors declare no conflicts of interest.

Acknowledgments We thank Silvia Maggi and Rasmus Petersen for comments on early drafts of this manuscript, and the Humphries lab (Javier Caballero, Mat Evans) for discussions. A.S. and M.D.H were supported by a Medical Research Council Senior non-Clinical Fellowship award MR/J008648/1 to M.D.H, and Medical Research Council Grant MR/P005659/1. A.P. was supported by a Canada Research Chair Tier 2 (154808). The original data were obtained through funding from the EU Framework 6 “ICEA” project.

1 Abstract

2 The prefrontal cortex is thought to learn the relationships between actions and their
3 outcomes. But little is known about what changes to population activity in prefrontal
4 cortex are specific to learning these relationships. Here we characterise the plasticity of
5 population activity in the medial prefrontal cortex of male rats learning rules on a Y-maze.
6 First, we show that the population always changes its patterns of joint activity between
7 the periods of sleep either side of a training session on the maze, irrespective of successful
8 rule learning during training. Next, by comparing the structure of population activity in
9 sleep and training, we show that this population plasticity differs between learning and
10 non-learning sessions. In learning sessions, the changes in population activity in post-
11 training sleep incorporate the changes to the population activity during training on the
12 maze. In non-learning sessions, the changes in sleep and training are unrelated. Finally,
13 we show evidence that the non-learning and learning forms of population plasticity are
14 driven by different neuron-level changes, with the non-learning form entirely accounted
15 for by independent changes to the excitability of individual neurons, and the learning
16 form also including changes to firing rate couplings between neurons. Collectively, our
17 results suggest two different forms of population plasticity in prefrontal cortex during the
18 learning of action-outcome relationships, one a persistent change in population activity
19 structure decoupled from overt rule-learning, the other a directional change driven by
20 feedback during behaviour.

21 Significance statement

22 The prefrontal cortex is thought to represent our knowledge about what action is worth
23 doing in which context. But we do not know how the activity of neurons in prefrontal
24 cortex collectively changes when learning which actions are relevant. Here we show in a
25 trial-and-error task that population activity in prefrontal cortex is persistently changing,
26 irrespective of learning. Only during episodes of clear learning of relevant actions are
27 the accompanying changes to population activity carried forward into sleep, suggesting a
28 long-lasting form of neural plasticity. Our results suggest that representations of relevant
29 actions in prefrontal cortex are acquired by reward imposing a direction onto ongoing
30 population plasticity.

31 Introduction

32 Among the myriad roles assigned to the medial prefrontal cortex a common thread is that
33 it learns a model for the statistics of actions and their expected outcomes, in order to
34 guide or monitor behaviour (Alexander and Brown, 2011; Euston et al., 2012; Holroyd
35 and McClure, 2015; Khamassi et al., 2015; Starkweather et al., 2018; Wang et al., 2018).
36 One way to probe this role is to use rule-switching tasks that depend on trial-and-error to
37 uncover the statistics of each new action-outcome association. Previous work has shown
38 that inactivating medial prefrontal cortex impairs the learning of new rules (Ragozzino et
39 al., 1999,a; Rich and Shapiro, 2007; Floresco et al., 2008), and single pyramidal neurons
40 change their firing times relative to ongoing theta-band oscillations only with successful
41 rule learning (Benchenane et al., 2010). In well-trained animals, a shift in their behavioural
42 strategy in response to a rule change is preceded by a shift in population activity in
43 prefrontal cortex (Durstewitz et al., 2010; Karlsson et al., 2012; Powell and Redish, 2016),
44 consistent with a change to a statistical model of the current action-outcome dependencies.

45 We know little though about how prefrontal cortex population activity changes dur-
46 ing the initial learning of rules (Peyrache et al., 2009; Tavoni et al., 2017; Maggi et al.,
47 2018). The changes to population activity could be continuous or constrained only to
48 periods of overt learning. And these changes could be modulations of firing rates, of firing

49 correlations, or of precise co-spiking between neurons. Knowing the continuity and form
50 of plasticity in population activity would provide strong constraints on theories for how
51 statistical models of the world are acquired and represented by medial prefrontal cortex.

52 To address these questions, here we analyse the continuity and form of population
53 plasticity in the prefrontal cortex of rats learning rules on a Y-maze (Peyrache et al.,
54 2009). We report that the structure of the population's activity markedly changes be-
55 tween the periods of sleep either side of training on the maze. This turnover in neural
56 activity occurs whether or not there is behavioural evidence of learning during training,
57 and can be accounted for entirely by changes to the excitability of individual neurons, with
58 no contribution from changes to correlations. Unique to bouts of learning is that changes
59 to the structure of population activity in training are carried forward into the following
60 periods of sleep. These conserved activity states are created by a combination of changes
61 to individual neurons' excitability and to rate, but not spike, correlations between neu-
62 rons. Thus, prefrontal cortex population activity undergoes constant plasticity, but this
63 plasticity only has a persistent direction during learning.

64 **Materials and Methods**

65 **Task and electrophysiological recordings**

66 Four Long-Evans male rats with implanted tetrodes in prelimbic cortex were trained on a
67 Y-maze task (Figure 1A). Each recording session consisted of a 20-30 minute sleep or rest
68 epoch (pre-training epoch), in which the rat remained undisturbed in a padded flowerpot
69 placed on the central platform of the maze, followed by a training epoch, in which the
70 rat performed for 20-40 minutes, and then by a second 20-30 minute sleep or rest epoch
71 (post-training epoch). Figure 1B shows the structure of these three epochs in the ten
72 identified learning sessions. Every trial in the training epoch started when the rat left
73 the beginning of the departure arm and finished when the rat reached the end of one of
74 the choice arms. Correct choice was rewarded with drops of flavoured milk. Each rat had
75 to learn the current rule by trial-and-error, either: go to the right arm; go to the cued
76 arm; go to the left arm; go to the uncued arm. To maintain consistent context across
77 all sessions, the extra-maze light cues were lit in a pseudo-random sequence across trials,
78 whether they were relevant to the rule or not.

79 The data analysed here were from a total set of 50 experimental sessions taken from
80 the study of (Peyrache et al., 2009), representing training sessions starting from naive
81 until either the final training session, or until choice became habitual across multiple
82 consecutive sessions (consistent selection of one arm that was not the correct arm). The
83 four rats respectively had 13, 13, 10, and 14 sessions. From these we have used here ten
84 learning sessions and up to 17 "stable" sessions (see below).

85 Tetrode recordings were spike-sorted only within each recording session for conservative
86 identification of stable single units. In the sessions we analyse here, the populations ranged
87 in size from 15-55 units. Spikes were recorded with a resolution of 0.1 ms. For full details
88 on training, spike-sorting, sleep identification, and histology see (Peyrache et al., 2009).

89 **Session selection and strategy analysis**

90 We primarily analyse here data from the ten learning sessions in which the previously-
91 defined learning criteria (Peyrache et al., 2009) were met: the first trial of a block of at
92 least three consecutively rewarded trials after which the performance until the end of the
93 session was above 80%. In later sessions the rats reached the criterion for changing the
94 rule: ten consecutive correct trials or one error out of 12 trials. By these criterion, each
95 rat learnt at least two rules.

96 We also sought sessions in which the rats made stable choices of strategy. For each
97 session, we computed $P(rule)$ as the proportion of trials in which the rat's choice of arm

98 corresponded to each of the three rules (left, right, cued-arm). Whereas $P(left)$ and
 99 $P(right)$ are mutually exclusive, $P(cued - arm)$ is not, and has an expected value of 0.5
 100 when it is not being explicitly chosen because of the random switching of the light cue. A
 101 session was deemed to be “stable” if $P(rule)$ was greater than some threshold θ for one of
 102 the rules, and the session contained at least 10 trials (this removed only two sessions from
 103 consideration). Here we tested both $\theta = 0.9$ and $\theta = 0.85$, giving $N = 13$ and $N = 17$
 104 sessions respectively. These also respectively included 2 and 4 of the rule-change sessions.
 105 For the time-series in Figure 1C,E,F we estimated $P(rule)$ in windows of 7 trials, starting
 106 from the first trial, and sliding by one trial.

107 Characterising population activity as a dictionary

108 For a population of size N , we characterised the instantaneous population activity from
 109 time t to $t + \delta$ as an N -length binary vector or *word*. The i th element of the vector was
 110 a 1 if at least one spike was fired by the i th neuron in that time-bin, and 0 otherwise.
 111 Throughout we test bin sizes covering two orders of magnitude, with δ ranging from 1 ms
 112 to 100 ms. For a given bin size, the set of unique words that occurred in an epoch defined
 113 the dictionary of that epoch. The probability distribution for the dictionary was compiled
 114 by counting the frequency of each word’s occurrence in the epoch and normalising by the
 115 total number of time bins in that epoch.

116 For each session we constructed three dictionaries per bin size, and their corresponding
 117 probability distributions $P(Epoch)$: pre-session sleep $P(Pre)$, post-session sleep $P(Post)$,
 118 and trials during training $P(Trials)$. To unambiguously identify sleep periods, and for
 119 comparisons with previous reports of replay in Pfc (Euston et al., 2007; Peyrache et al.,
 120 2009), we used slow-wave sleep bouts for the pre- and post-session sleep dictionaries.

121 We built dictionaries using the number of recorded neurons N , up to a maximum of
 122 35 for computational tractability. The number of neurons used in each analysis is listed in
 123 Tables 1 and 2; where we needed to use less than the total number of recorded neurons, we
 124 ranked them according to the coefficient of variation of their firing rate between the three
 125 epochs, and choose the N least variable; in practice this sampled neurons from across the
 126 full range of firing rates. Only two learning sessions and six stable sessions were capped
 127 in this way.

Session ID	Neurons	Trials		Pre-training SWS		Post-training SWS	
		Duration (ms)	Number	Duration (ms)	Bouts	Duration (ms)	Bouts
201222	31	125.5279	23	724.0082	3	660.9652	3
201227	23	137.8321	18	703.9857	3	829.9588	3
201229	12	153.0175	33	866.0116	3	532.9798	3
181012	35	228.5572	13	481.9801	2	923.9320	5
181020	35	125.8876	29	1117.0111	4	644.9920	3
150628	25	155.9059	29	775.9994	7	1137.0150	4
150630	27	202.6222	15	742.0170	5	907.9818	4
150707	23	217.2740	48	561.9935	4	386.9965	2
190214	20	236.8101	42	130.0125	1	331.0333	2
190228	20	122.9788	26	540.0200	3	198.9732	2

Table 1. Learning session statistics. The Neurons column give the number of neurons used from each of the ten learning sessions to build the words; eight used all recorded neurons, two were capped at 35. For each epoch within a session, we give the total duration of spike-train data used to construct words, and the number of trials or sleep bouts that comprised this total duration. The number of words per epoch at a given bin size b can thus be calculated from this table as: Duration / b .

Session ID	Neurons	Trials		Pre-training SWS		Post-training SWS	
		Duration (ms)	Number	Duration (ms)	Bouts	Duration (ms)	Bouts
150701	21	83.9801	15	866.0116	3	532.9798	3
150706	19	140.8352	20	754.0503	3	937.0184	3
181024	35	152.8336	35	525.9901	4	525.0188	2
181025	35	80.0858	20	682.0686	6	501.0109	4
181026	35	140.2582	34	333.0157	3	779.0119	5
181027	35	132.5743	34	209.9913	2	33.9931	2
181102	35	133.6886	38	572.9771	2	521.0275	7
181103	35	93.0362	22	219.9789	3	418.0025	4
190213	21	142.4687	32	255.9889	2	605.9930	1
190301	19	899.2288	12	693.0521	5	897.0089	5
190302	22	60.0684	14	477.0404	4	279.9953	1
190303	17	132.0855	29	1043.9569	4	661.0032	4
201228	19	217.3680	41	883.9506	2	337.9834	4
201230	21	171.9406	44	180.9926	2	224.9939	3
200102	22	195.2417	42	199.0138	1	162.0023	2
200103	29	289.0712	37	308.9891	3	429.9769	4
200105	12	223.3549	45	215.9840	2	408.0112	4

Table 2. Stable session statistics. Column entries as per Table 1.

128 Comparing dictionaries between epochs

129 We quantified the distance $D(P|Q)$ between two dictionary's probability distributions P
 130 and Q using the Hellinger distance, defined by $D_H(P|Q) = \frac{1}{2} \sum_{i=1}^n (\sqrt{p_i} - \sqrt{q_i})^2$. To a first
 131 approximation, this measures for each pair of probabilities (p_i, q_i) the distance between
 132 their square-roots. In this form, $D_H(P|Q) = 0$ means the distributions are identical, and
 133 $D_H(P|Q) = 1$ means the distributions are mutually singular: all positive probabilities in
 134 P are zero in Q , and vice-versa.

135 To understand if a pair of pre- and post-training sleep dictionaries meaningfully dif-
 136 fered in their structure, we compared the distance between them $D(Pre|Post)$ to the pre-
 137 dicted distance if they had an identical underlying probability distribution (in which case
 138 $D(Pre|Post) > 0$ would be solely due to finite sampling effects). We used a resampling
 139 test to estimate the predicted distance. We first created a single probability distribution
 140 $P(sleep)$ for a session by calculating the probability of each word's appearance in all sleep
 141 bouts across both pre and post-training sleep epochs. We then sampled $P(sleep)$ to create
 142 new time-series of pre- and post-training sleep words, matching the number of emitted
 143 words in each epoch in the original data. By then reconstructing the dictionaries in each
 144 epoch from the resampled data, we obtained a prediction for the distance $D(Pre^*|Post^*)$,
 145 where $*$ denotes the estimate from the resampled data. Repeating the resampling 20 times
 146 gave us a distribution of expected distances assuming an identical underlying probability
 147 distribution for words. The sampling distribution's mean and its 99% confidence interval
 148 are plotted for each session in Figure 3D,E – the intervals are too small to see on this
 149 scale.

150 We quantified the relative convergence of the training dictionary X with the dictionar-
 151 ies in sleep by $[D(Pre|X) - D(Post|X)]/[D(Pre|X) + D(Post|X)]$. Convergence greater
 152 than 0 indicates that the distance between the training epoch $[P(X)]$ and post-training
 153 sleep $[P(Post)]$ distributions was smaller than that between the training and pre-training
 154 sleep $[P(Pre)]$ distributions.

155 Testing hypotheses for changes in dictionary structure

156 To understand what drove the observed changes in the structure of population activity,
157 we tested three hypotheses: independent changes in the excitability of neurons; changes in
158 firing rate co-variations between neuron; and shifts in precise co-spiking between neurons.
159 We tested these hypotheses in two steps:

- 160 1. We tested whether dictionaries constructed from independently firing neurons could
161 account for the observed changes in the structure of population activity, with two
162 possible outcomes:
 - 163 • Yes: then we could conclude that changes in the data were due to independent
164 changes to the excitability of the recorded neurons.
 - 165 • No: this implied that the correlations between neurons were also changed.
- 166 2. To then identify the types of those correlations, we turned to dictionaries constructed
167 from spikes jittered a little in time, and asked if they could account for the observed
168 changes:
 - 169 • No: then we would have evidence that precise co-spiking between neurons con-
170 tributed to the changes in population activity structure.
 - 171 • Yes: then changes to population activity did not depend on precise co-spiking,
172 and could be accounted for by changes to co-variations in rate between neurons.

173 For the independent neuron dictionaries, we shuffled inter-spike intervals for each neu-
174 ron independently, and then constructed words at the same range of bin sizes. As both the
175 training and sleep epochs were broken up into chunks (of trials and slow-wave sleep bouts,
176 respectively), we only shuffled inter-spike intervals within each chunk. This procedure kept
177 the same inter-spike interval distribution for each neuron, but disrupted any correlation
178 between neurons during a trial or during a sleep bout, thus testing for dictionary changes
179 that could be accounted for solely by changes to independent neurons. We repeated the
180 shuffling 20 times.

181 For any given data statistic s_{data} for a single session, we compute the same statistic
182 $s_{shuffle}$ for each shuffled data-set, and plot the difference $\delta = s_{data} - E(s_{shuffle})$ using the
183 mean $E()$ over the shuffled data's statistics. Confidence intervals at 99% for all δ were
184 smaller than the size of the plotted symbol for δ , so are omitted for clarity.

185 For the jittered dictionaries, each spike was jittered in time by a random amount drawn
186 from a Gaussian of mean zero and standard deviation σ . We tested σ from 2 to 50 ms. For
187 each σ we constructed 20 jittered data-sets. Words were constructed from each using 5 ms
188 bins here, both as this time-scale would capture millisecond-precise spike-timing between
189 neurons, and because the biggest effects in the data were most consistently seen at this
190 bin size.

191 We illustrated changes in the rate co-variation between neurons using the coupling
192 between single neuron and ongoing population activity (Okun et al., 2015). Each neuron's
193 firing rate was the spike density function f_i obtained by convolving each spike with a
194 Gaussian of 100 ms standard deviation. Population coupling for the i th neuron is the
195 Pearson's correlation coefficient: $c_i = \text{corr}(f_i, P_{\neq i})$, where $P_{\neq i}$ is the population rate
196 obtained by summing all firing rate functions except that belonging to the i th neuron.

197 Relationship of location and change in word probability

198 To examine the spatial correlates of word occurrence, the maze was linearised, and nor-
199 malised (0: start of departure arm; 1: end of the chosen goal arm). The location of every
200 occurrence of a word during the training epoch's trials ("trial word") was expressed as a
201 normalized position on the linearised maze, from which we computed the word's median

202 location and corresponding interquartile interval. Histograms of median word location
203 were constructed using kernel density, with 100 equally spaced points between 0 and 1.

204 We tested whether the trial words closer in probability to post- than pre-training sleep
205 were from any specific locations, which would suggest a changing representation of a key
206 location. For each word, we computed the difference in its probability between training
207 and pre-training sleep $\delta_{pre} = |p(pre) - p(trial)|$, and the same for post-training sleep $\delta_{post} =$
208 $|p(post) - p(trial)|$, and from these computed a closeness index: $(\delta_{pre} - \delta_{post}) / (\delta_{pre} + \delta_{post})$.
209 Closeness is 0 if the word is equidistant from training to both sleep epochs, 1 if it has an
210 identical probability between training and post-training sleep; and -1 if it has an identical
211 probability between training and pre-training sleep.

212 When assessing identified maze segments, words were divided into terciles by thresholds
213 on the closeness index at $[-0.5, 0.5]$; similar results were obtained if we used percentile
214 bounds of [10, 90]%. We counted the proportion of words in each tercile whose median
215 position fell within specified location bounds on the linearised maze. Confidence intervals
216 on the proportions were computed using 99% Jeffrey's intervals (Brown et al., 2001).

217 Statistics

218 Quoted measurement values are mean \bar{x} and confidence intervals for the mean $[\bar{x} -$
219 $t_{\alpha/2,n}SE, \bar{x} + t_{\alpha/2,n}SE]$, where $t_{\alpha/2,n}$ is the value from the t -distribution at $\alpha = 0.05$ (95%
220 CI) or $\alpha = 0.01$ (99% CI), and given the number n of data-points used to obtain \bar{x} . For
221 testing the changes in convergence, we used the Wilcoxon signed-rank test for a differ-
222 ence from zero; for differences in population-coupling correlations, we used the Wilcoxon
223 signed-rank paired-sample test. Throughout, we have $n = 10$ learning sessions and $n = 17$
224 stable sessions.

225 Data and code availability

226 The spike-train and behavioural data that support the findings of this study are available
227 at CRCNS.org (DOI: 10.6080/K0KH0KH5) (Peyrache et al., 2018). The sessions meeting
228 our learning and stable criteria are listed in Tables 1 and 2.

229 Code to reproduce the main results of the paper is available at:
230 <https://github.com/mdhumphries/PfCDictionary>.

231 Results

232 Signatures of rule-learning on the Y-maze

233 Rats with implanted tetrodes in the prelimbic cortex learnt one of four rules on a Y-maze:
234 go right, go to the randomly-cued arm, go left, or go to the uncued arm (Figure 1A).
235 Rules were changed in this sequence, unsignalled, after the rat did 10 correct trials in
236 a row, or 11 correct trials out of 12. Each rat learnt at least two of the rules, starting
237 from a naive state. Each training session was a single day containing 3 epochs totalling
238 typically 1.5 hours: pre-training sleep/rest, behavioural training on the task, and post-
239 training sleep/rest (Figure 1B). Here we consider bouts of slow-wave sleep throughout,
240 to unambiguously identify periods of sleep. Tetrode recordings were spike-sorted within
241 each session, giving populations of single neuron recordings ranging between 12 and 55
242 per session (see Tables 1 and 2 for details of each session and each epoch within a session).

243
244 In order to test for the effects of learning on the structure of joint population activity,
245 we need to compare sessions of learning with those containing no apparent learning as
246 defined by the rats' behaviour. In the original study containing this data-set, Peyrache et
247 al. (2009) identified 10 learning sessions as those in which three consecutive correct trials
248 were followed by at least 80% correct performance to the end of the session; the first trial

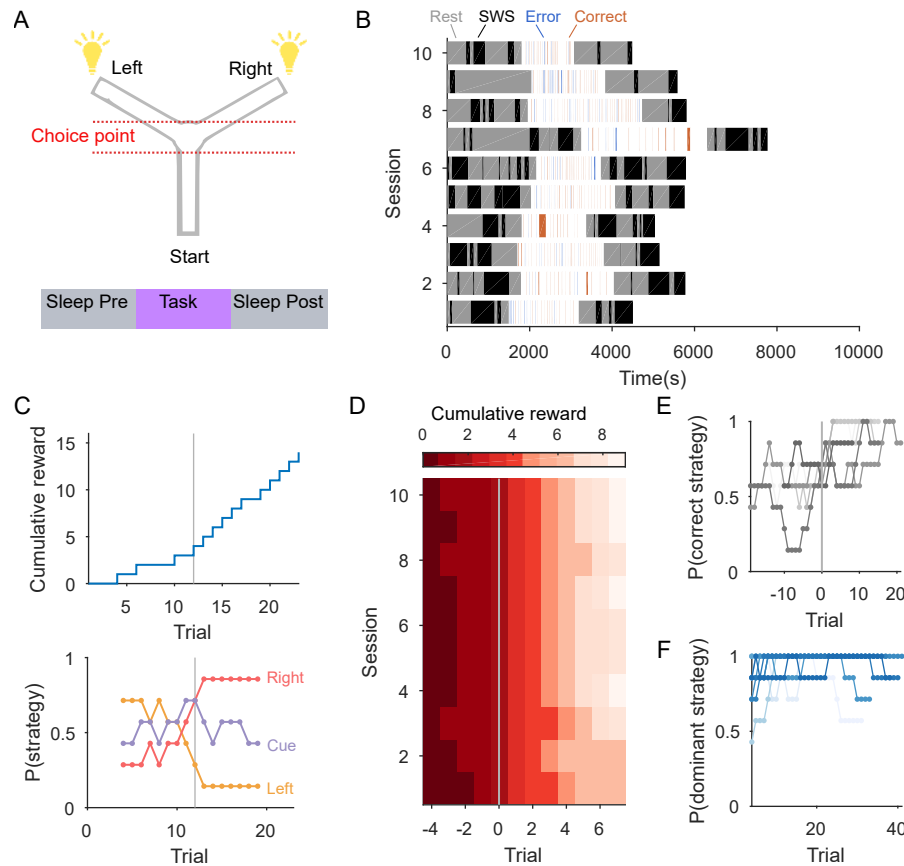


Figure 1. Task and behaviour.

(A) Y-maze task set-up (top); each session included the epochs of pre-training sleep/rest, training trials, and post-training sleep/rest (bottom). One of four target rules for obtaining reward was enforced throughout a session: go right; go to the cued arm; go left; go to the uncued arm. No rat successfully learnt the uncued-arm rule.

(B) Breakdown of each learning session into the duration of its components. The training epoch is divided into correct (red) and error (blue) trials, and inter-trial intervals (white spaces). Trial durations were typically 2-4 seconds, so are thin lines on this scale. The pre- and post-training epochs contained quiet waking and light sleep states (“Rest” period) and identified bouts of slow-wave sleep (“SWS”).

(C) Internally-driven behavioural changes in an example learning session: the identified learning trial (grey line) corresponds to a step increase in accumulated reward and a corresponding shift in the dominant behavioural strategy (bottom). The target rule for this session is ‘go right’. Strategy probability is computed in a 7-trial sliding window; we plot the mid-points of the windows.

(D) Peri-learning cumulative reward for all ten identified learning sessions: in each session, the learning trial (grey line) corresponds to a step increase in accumulated reward.

(E) Peri-learning strategy selection for the correct behavioural strategy. Each line plots the probability of selecting the correct strategy for a learning session, computed in a 7-trial sliding window. The learning trial (grey vertical line) corresponds to the onset of the dominance of the correct behavioural strategy.

(F) Strategy selection during stable behaviour. Each line plots the probability of selecting the overall dominant strategy, computed in a 7-trial sliding window. One line per session.

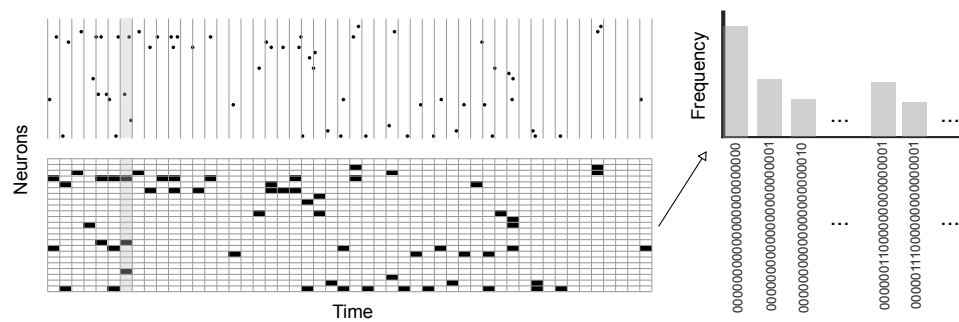


Figure 2. A neural dictionary of population activity in prefrontal cortex.

A snapshot of population activity from $N = 23$ neurons during 500 ms of pre-training sleep, and below is the corresponding binary word structure (black: 1; white: 0) for bins of 10 ms. One bin of the population activity and its corresponding binary word is highlighted in grey. Right: The set of binary words and the frequency of their occurrence over the whole pre-training sleep epoch defines a dictionary of population activity.

249 of the initial three was considered the learning trial. By this criterion, the learning trial
250 occurs before the mid-point of the session (mean 45%; range 28-55%). We first check this
251 criterion corresponds to clear learning: Figure 1C,D shows that each of the ten sessions
252 has an abrupt step change in reward accumulation around the identified learning trial
253 corresponding with a switch to a consistent, correct strategy within that session (Figure
254 1E).

255 We further identify a set of 17 sessions with a stable behavioural strategy throughout,
256 defined as a session with the same strategy choice (left, right, cue) on more than 85% of
257 trials (Figure 1F). This set includes four sessions in which the rule changed. Setting this
258 criterion to a more conservative 90% reduces the number of sessions to 13 (including two
259 rule change sessions), but does not alter the results of any analysis; we thus show the 85%
260 criterion results throughout.

261 Constant plasticity of population activity between sleep epochs

262 We want to describe the joint population activity over all N simultaneously-recorded
263 neurons with minimal assumptions, so that we can track changes in population activity
264 however they manifest. Dividing time into bins small enough that each neuron either spikes
265 ('1') or doesn't ('0') gives us the instantaneous state of the population as the N -element
266 binary vector or *word* in that bin (Figure 2). The dictionary of words appearing in an
267 epoch and their probability distribution together describe the region of joint activity space
268 in which the population is constrained. Comparing dictionaries and their probabilities
269 between epochs will thus reveal if and how learning changes this region of joint activity.

270 If learning during training correlated with changes to the underlying neural circuit in
271 prefrontal cortex then we might reasonably expect population activity in post-training
272 sleep to also be affected by these changes, and so differ from activity in pre-training sleep.
273 We thus compare the dictionaries in pre- and post-training sleep for the learning sessions,
274 and then check if any detected changes also appear during sessions of stable behaviour.

275 A first check is simply if the dictionary content changed during learning and not stable
276 behaviour. We find that the words common to both sleep epochs (Figure 3A) account for
277 almost all of each epoch's activity (Figure 3B) at bin sizes up to 20 ms. Consequently,
278 there are no differences between learning and stable behaviour in the overlap of dictionary
279 contents between sleep epochs (Figure 3A) or in the proportion of activity accounted for
280 by words common to both sleep epochs (Figure 3B). We could thus rule out that learning
281 changes the dictionary content between sleep epochs compared to stable behaviour. Any
282 learning-specific change ought then be found in the structure of the population activity.

283 We capture this structure by the respective distributions $P(Pre)$ and $P(Post)$ for
284 the probability of each word appearing in pre- or post-training sleep. Changes to the
285 detailed structure of the pre- and post-training sleep dictionaries are then quantified by

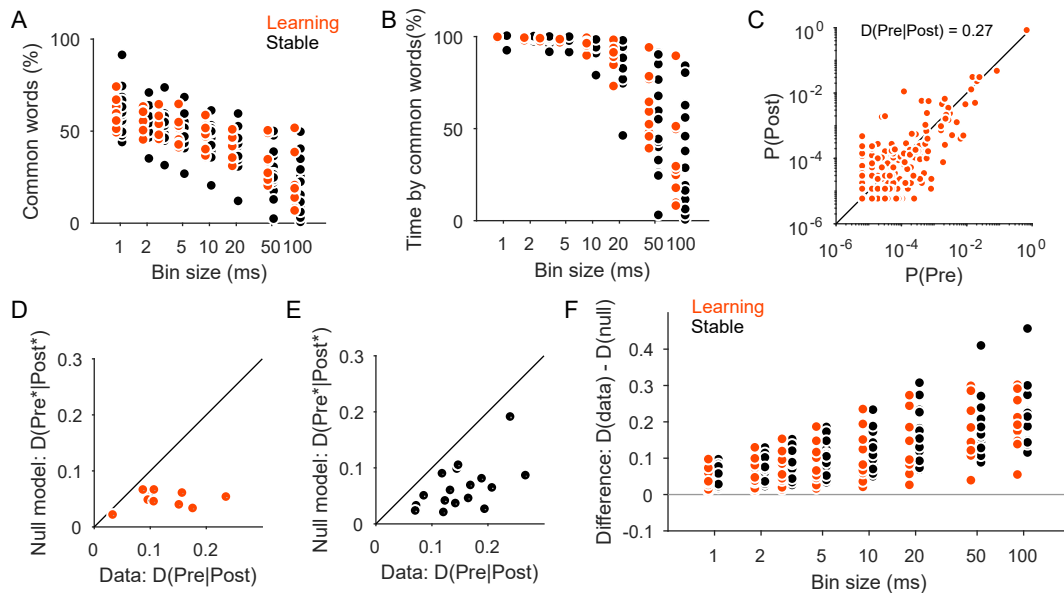


Figure 3. Distributions of word probabilities change between pre- and post-training sleep.

(A) Proportion of words in the pre-training sleep dictionary that are also in the post-training sleep dictionary, per session.

(B) Proportion of the pre-training sleep epoch’s activity that is accounted for by words in common with post-training sleep, per session.

(C) The joint distribution of the probability of every word occurring in pre-training sleep (distribution $P(Pre)$) and post-training sleep (distribution $P(Post)$), for one learning session. $D(Pre|Post)$: the distance between the two probability distributions for words.

(D) Distance between the word probability distributions for pre- and post-training sleep (x-axis) against the expected distance if the sleep activity was drawn from the same distribution in both epochs (y-axis). One symbol per learning session; we plot the mean and 99% confidence interval (too small to see) of the expected distance $D(Pre^*|Post^*)$. Words constructed using 5 ms bins.

(E) Same as (D), for stable sessions.

(F) Bin-size dependence of changes in the dictionary between sleep epochs. Difference between the data and mean null model distance are plotted for each session, at each bin-size used to construct words.

286 the distance between these probability distributions (Figure 3C). These distances will
 287 vary according to both the number of neurons N and the duration of each epoch. So
 288 interpreting them requires a null model for the distances expected if $P(Pre)$ and $P(Post)$
 289 have the same underlying distribution $P(Sleep)$, which we approximate using a resampling
 290 test (see Methods). In this null model any differences between $P(Pre)$ and $P(Post)$ are
 291 due to the finite sampling of $P(Sleep)$ forced by the limited duration of each epoch.

292 In learning sessions the distance between pre- and post-training sleep probability dis-
 293 tributions always exceeds the upper limit of the null model’s prediction (Figure 3D). This
 294 was true at every bin size (Figure 3F), even at small bin sizes where the dictionaries were
 295 nearly identical between the sleep epochs (Figure 3A). Thus, the probability distributions
 296 of words consistently differ between pre and post-training sleep epochs in learning sessions.

297 However, Figure 3E-F shows this consistent difference is also true for the sessions
 298 with stable behaviour. There is quantitative agreement too as the gap between the data
 299 and predicted distances has the same distribution for both learning and stable behaviour
 300 (Figure 3F). We conclude that the probabilities of words do systematically change between
 301 sleep epochs either side of training, but do so whether there is overt learning or not.

302 Learning systematically updates the dictionary

303 This leaves open the question of whether changes in population activity between sleep
 304 epochs are a consequence of changes during training. If the population changes between
 305 sleep epochs are unrelated to population activity in training, then the probability distribu-
 306 tion of words in training will be equidistant on average from that in pre- and post-training
 307 sleep. Alternatively, changes to population activity during training may carry forward into

308 post-session sleep, possibly as a consequence of neural plasticity during the trials changing
309 the region of joint activity space in which the population is constrained. A prediction
310 of this neural-plasticity model is that the directional change would thus occur predomi-
311 nantly during learning sessions, so that only in these sessions is the distribution of word
312 probabilities in training closer to that in post-training sleep than in pre-training sleep.

313 Unpicking the relationship between the sleep changes and training requires that the
314 dictionary in training also appears in the sleep epochs; otherwise changes to word prob-
315 abilities during training could not be tracked in sleep. We find that the structure of
316 population activity in training is highly conserved in the sleep epochs (Figure 4A), both
317 in that the majority of words appearing in trials also appear in the sleep epochs, and that
318 these common words account for almost all of the total duration of the trials. This conser-
319 vation of the training epoch population structure in sleep allows us to test the prediction
320 of a learning-driven directional change in population structure (Figure 4B).

321 To do so, we take the dictionary of words that appear during training, and compute
322 the distance between its probability distribution and the probability distribution of that
323 dictionary in pre-training sleep ($D(Pre|Learn)$), and between training and post-training
324 sleep ($D(Post|Learn)$) (Figure 4C). The prediction of the directional change model is
325 then $D(Pre|Learn) > D(Post|Learn)$. This is exactly what we found: $D(Pre|Learn)$ is
326 consistently larger than $D(Post|Learn)$ at small bin sizes, as illustrated in Figure 4D for
327 5 ms bins.

328 If these directional changes are uniquely wrought by learning, then it follows that we
329 should not see any systematic change to the dictionary in the stable behaviour sessions
330 (Figure 4B). To test this prediction, we similarly compute the distances $D(Pre|Stable)$
331 and $D(Post|Stable)$ using the dictionary of words from the training epoch, and test if
332 $D(Pre|Stable) \approx D(Post|Stable)$. Again, this is exactly what we found: $D(Pre|Stable)$
333 was not consistently different from $D(Post|Stable)$ at any bin size, as illustrated in Figure
334 4E for 5 ms bins.

335 It is also useful to consider not just which sleep distribution of words is closer to
336 the training distribution, but how much closer. We express this as a convergence ratio
337 $C = [D(Pre|X) - D(Post|X)]/[D(Pre|X) + D(Post|X)]$, given the training distribution
338 $X = \{Learn, Stable\}$ in each session. So computed C falls in the range $[-1, 1]$ with a
339 value greater than zero meaning that the training probability distribution is closer to the
340 distribution in post-training sleep than the distribution in pre-training sleep. Figure 4G
341 shows that for learning sessions the word distribution in training is closer to the post-
342 training than the pre-training sleep distribution across an order of magnitude of bin sizes.
343 For stable sessions the absence of relative convergence is consistent across two orders of
344 magnitude of bin size (Figure 4G). Both qualitatively and quantitatively, the structure of
345 prefrontal cortex population activity shows a relative convergence between training and
346 post-training sleep that is unique to learning.

347 **Changes to neuron excitability account for changes between sleep epochs**

348 What then is the main driver of the observed changes in the structure of population
349 activity? These could arise from changes to the excitability of independent neurons, to co-
350 variations in rate over tens to hundreds of milliseconds, or to the millisecond-scale precise
351 timing of co-incident spiking between neurons. We first examine the drivers of the changes
352 between sleep epochs we saw in Figure 3.

353 Individual sessions showed a rich spread of changes to neuron excitability between the
354 sleep epochs (Figure 5A). We thus begin isolating the contribution of these three factors
355 by seeing how much of the change in population structure between sleep epochs can be
356 accounted for by independent changes to neuron excitability. Shuffling inter-spike intervals
357 within each epoch gives us null model dictionaries for independent neurons by removing
358 both rate and spike correlations between them, but retaining their excitability (at least,
359 as captured by their inter-spike interval distribution).

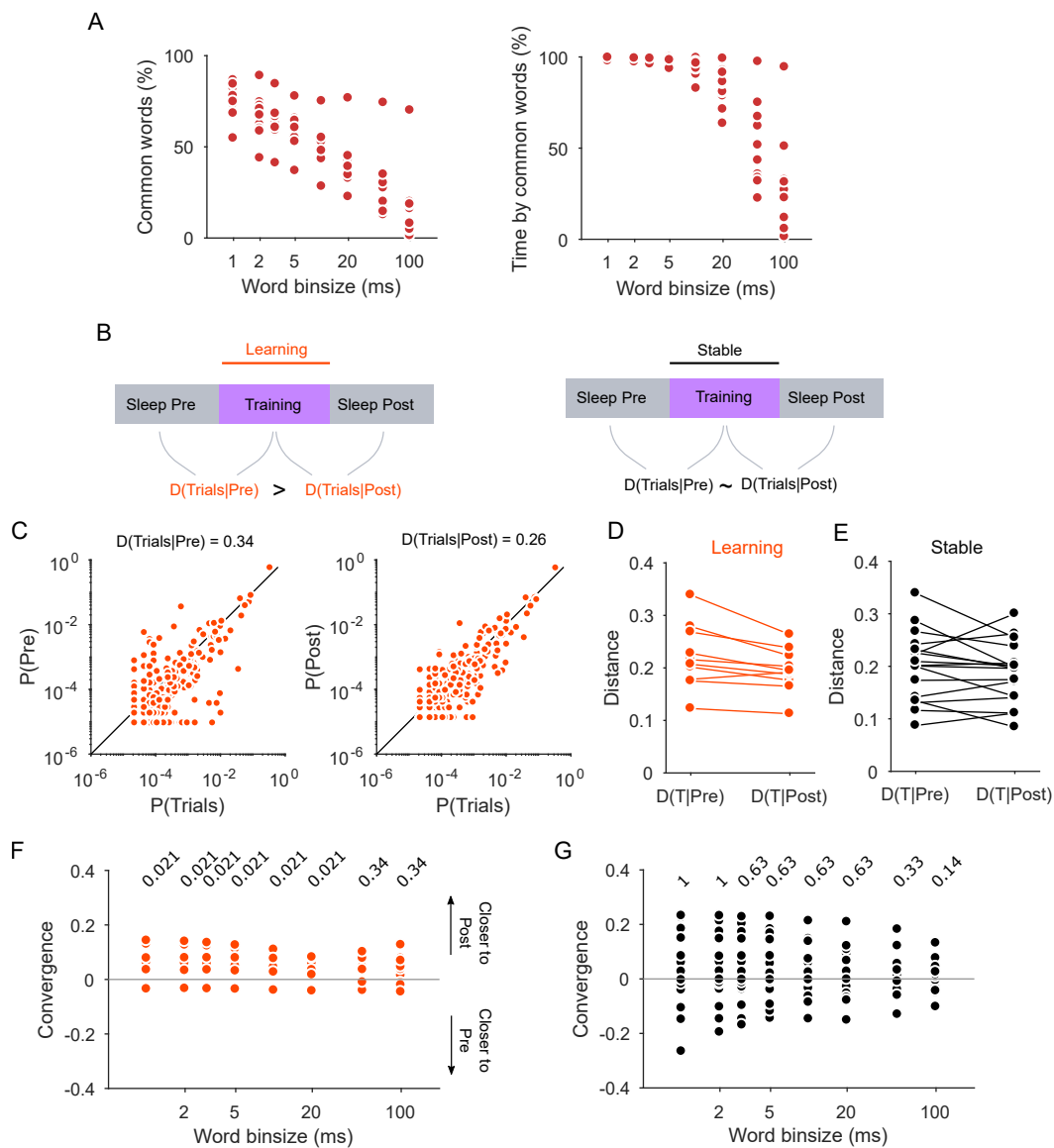


Figure 4. Distributions of word probabilities converge only during learning.

(A) For the training epochs, the proportion of the epoch's dictionary (left) and duration (right) accounted for by words in common with both sleep epochs. One symbol per learning session.

(B) Schematic of comparisons between epochs, and summary of main results. (C) Examples for one learning session of the joint probability distributions for each word in trials and pre-training sleep (left), and trials and post-training sleep (right), using 5 ms bins. $D(\text{Trials}|X)$: the distance between the two probability distributions for words.

(D) Distances for all learning sessions, for words constructed using 5 ms bins. T: Trials.

(E) As for (D), for stable sessions.

(F) Bin-size dependence of the relative convergence between the word distributions in trials and in sleep. Each distance was computed using only the dictionary of words appearing in the trials. Numbers are P -values from two-sided signtests for each median differing from zero.

(G) As for (F), for stable sessions.

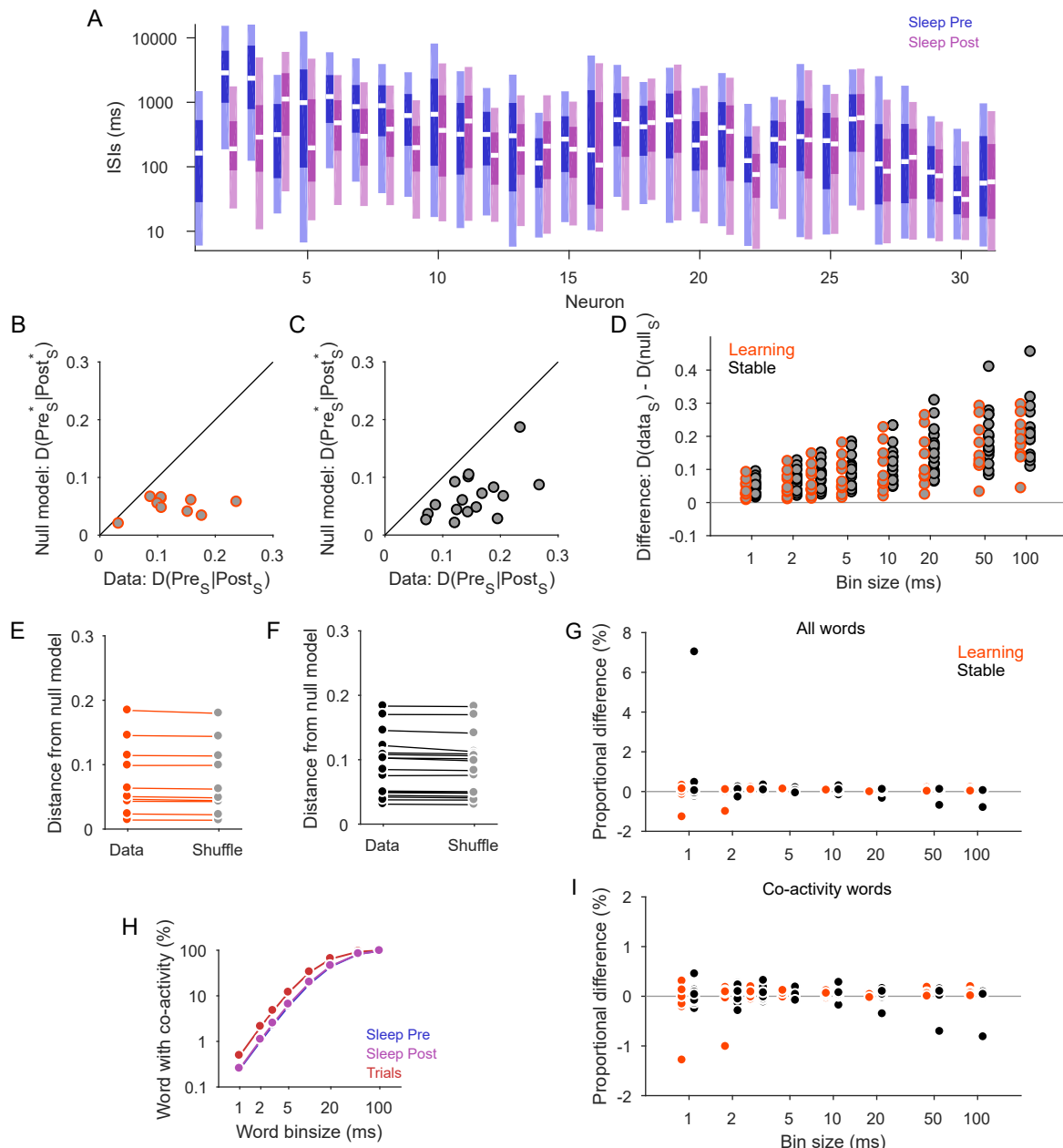


Figure 5. Changes between sleep epochs are accounted for by independently changing neurons.

(A) Example excitability changes between sleep epochs, for one learning session. Each pair of bars plot the distributions of a neuron's inter-spike intervals in the pre- and post-training sleep epochs, each bar showing the median (white line), interquartile range (dark shading) and 95% interval (light shading). Neurons are ranked by the difference in their median interval between sleep epochs. We use a log-scale on the y-axis: some neurons shift their distribution over orders of magnitude between sleep epochs. The first neuron was silent in the post-training sleep epoch.

(B) Distances between sleep epochs for dictionaries of independent neurons (x-axis), and their expected distances from a null model of the same dictionary in both epochs (y-axis). Independent neuron dictionaries are constructed by shuffling inter-spike intervals within trials or sleep bouts. One symbol per learning session; we plot the mean and 99% confidence interval (too small to see) of the expected distance $D(Pre^*|Post^*)$. Words constructed using 5 ms bins. S: shuffled data.

(C) As for (B), for stable sessions.

(D) Independent neuron dictionaries are consistently different between sleep epochs at all bin sizes — compare to results for the data dictionaries in Figure 3F. Each symbol is a mean over 20 shuffled data-sets.

(E) Departure from the expected distance between sleep epochs for each learning session (Data), and the corresponding predicted departure by independent neurons (Shuffle; mean over 20 shuffled data-sets). Words constructed using 5 ms bins.

(F) As for (D), for stable sessions.

(G) Difference between the recorded and shuffled data, as a proportion of the data's departure from the expected distance between sleep epochs. Almost all differences are less than 0.1% of the difference between data and the null model. One symbol per session.

(H) The proportion of words in the dictionary with two or more active neurons, over all learning sessions.

(I) As for panel (G), using dictionaries that contained only words with co-activity. At all bin sizes, there is no systematic difference between recorded and shuffled data.

360 When we analyse the changes between sleep epochs for independent neuron dictionar-
361 ies, the strong similarity with the results from the data dictionaries is compelling. We
362 illustrate this in Figure 5B-D, by repeating the analyses in Figure 3D-F but now using
363 the independent neuron dictionaries – and see the results are essentially the same. The
364 departure from the null model of a single probability distribution in sleep is almost iden-
365 tical between the data and independent neuron dictionaries, illustrated in Figure 5E-F for
366 5 ms bins. And while the data dictionaries tend to depart further from the null model,
367 this excess is negligible, being on the order of 0.1% of the total departure from the null
368 model (Figure 5G).

369 A potential confound in searching for the effects of correlation here are that words
370 coding for two or more active neurons are infrequent at small bin sizes, comprising less
371 than 10% of words at small bins sizes (Figure 5H). As a consequence, any differences
372 between the independent neuron and data dictionaries that depend on correlations between
373 neurons in the data could be obscured. To check for this, we repeat the same analyses of
374 the changes between sleep for both the data and independent neuron dictionaries when
375 they are restricted to include only co-activity words. As Figure 5I shows, this did not
376 uncover any hidden contribution of correlation between neurons in the data; indeed, for co-
377 activity words alone, the difference between the data and the independent model is about
378 zero. Thus, the changes in word probabilities between pre- and post-training sleep can
379 be almost entirely accounted for by independent changes to the excitability of individual
380 neurons (Figure 5A).

381 **Learning-driven changes to the dictionary include rate co-variations**

382 Can independent changes to individual neuron excitability also account for the relative
383 convergence of dictionaries in learning? Repeating the comparisons of training and sleep
384 epoch activity using the independent neuron dictionaries, we observe the same learning-
385 specific convergence of the training and post-training sleep dictionaries, illustrated in
386 Figure 6A for 5 ms bins (compare Figure 4D-E). Figure 6B shows that the difference
387 in convergence score between the data and independent neuron dictionaries is close to
388 zero at most bin sizes. This suggests that the changes in population activity during the
389 trials that are carried forward to the post-training sleep can also be accounted for by the
390 changing excitability of individual neurons.

391 To check this conclusion, we again account for the relative infrequency of co-activity
392 words at small bin sizes by recomputing the distances between sleep and training epochs
393 using dictionaries of only co-activity words. Now we find that, unlike the changes between
394 sleep epochs, the relative convergence between training and post-training sleep for the
395 data dictionaries is greater than for the independent neuron dictionaries (Figure 6C). We
396 conclude that changes to the correlations between neurons during the trials of learning
397 sessions are also detectably carried forward to post-training sleep.

398 These correlations could take the form of co-variations in rate, or precise co-incident
399 spikes on millisecond time-scales. To test for precise co-spiking, we construct new null
400 model dictionaries: we jitter the timing of each spike, and then build dictionaries using 5
401 ms bins to capture spike alignment. If precise co-spiking is contributing to the correlations
402 between neurons, then relative convergence should be smaller for these jittered dictionaries
403 than the data dictionaries. As Figure 6D shows, this is not what we found: across a range
404 of time-scales for jittering the spikes, the difference in relative convergence between the
405 data and jittered dictionaries was about zero. The changed correlations between neurons
406 are then rate co-variations, not precise co-spiking.

407 Figure 6E-H gives some intuition for these changes in rate co-variation. We measure
408 the coupling of each neuron's firing to the ongoing population activity (Figure 6E) as
409 an approximation of each neuron's rate covariation (as population-coupling is fixed to a
410 particular time-scale, so it can only represent part of the co-variation structure captured
411 by the dictionaries of words). The distribution of population coupling across the neurons

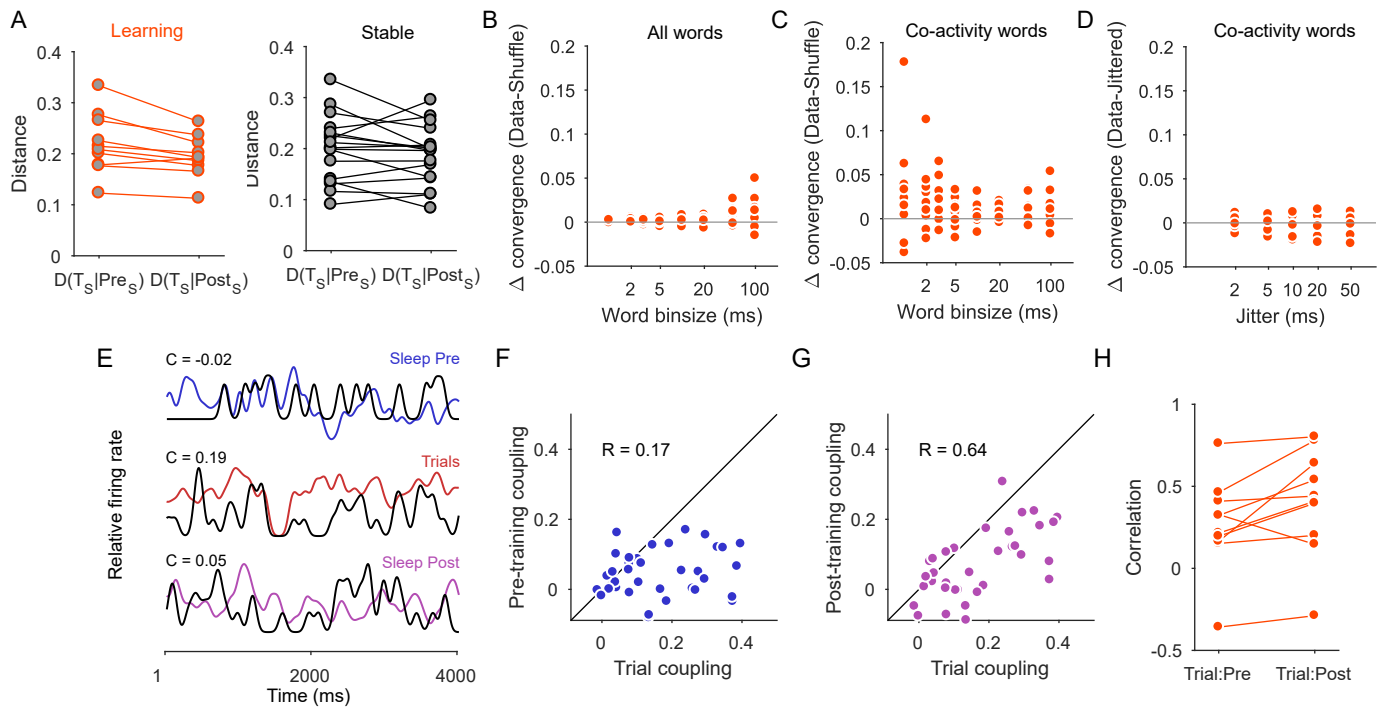


Figure 6. Convergence of dictionaries during learning is partly driven by changes in rate co-variation, but not spike-timing.

(A) Distances between sleep and trial distributions for all learning (left) and stable (right) sessions, in an example shuffled data-set. Words constructed using 5 ms bins. $D(T_S|X_S)$: the distance between the trial probability distribution and the probability distribution of sleep epoch X in the shuffled data.

(B) Difference between the recorded and shuffled data convergence between trial and post-training sleep epochs, in learning sessions.

(C) As for panel B, using distributions containing only words with co-activity.

(D) As for panel C, comparing co-activity word distributions from recorded and jittered data, to test for the contribution of precise spike timing. Spike data were jittered at a range of standard deviations (x-axis), and words constructed using 5 ms bins.

(E) Snapshots of a single neuron's firing rate (black) in comparison to the simultaneous population firing rate (colour) in each epoch. C: population-coupling in each epoch.

(F) Joint distribution of the population coupling for each neuron in the training and pre-training sleep epochs of one learning session. R: Pearson's correlation coefficient between the two distributions of population coupling.

(G) Same as (F), for the training and post-training sleep epochs in the same session.

(H) Correlations between population coupling in training and sleep epochs for all learning sessions.

Population-coupling is more correlated between training and post-training sleep (signed-rank test $P = 0.02$, $rank = 5$).

412 varied between epochs (Figure 6F-G), signalling changes to the co-variations in rate be-
413 tween neurons. Consistent with changes to rate co-variations, the distribution of coupling
414 tended to be more similar between training and post-training sleep than between training
415 and pre-training sleep (Figure 6H).

416 **Locations of dictionary sampling during learning**

417 The changes to population activity in training carried forward to post-training sleep may
418 correspond to learning specific elements of the task. We check for words linked to task
419 elements by first plotting where each word in the training dictionary occurs on the maze
420 during trials. Words cluster at three maze segments, as illustrated in Figure 7A for 3 ms
421 bins: immediately before the choice area, at its centre, and at the end of the chosen arm.
422 This clustering is consistent across all bin sizes (Figure 7B).

423 Repeating this location analysis using the dictionaries of independent words gives the
424 same three clusters (grey lines in Figure 7A-B). This suggests that the clustering of words
425 at particular locations can be largely attributed to the amount of time the animals spent
426 at those locations. The only departures are that the choice region is slightly under-
427 represented in the data dictionaries, and the arm-end slightly over-represented. These
428 departures are potentially interesting, as they correspond to key points in the task: the
429 area of the maze at which the goal arm has to be chosen, and the arrival at the goal arm's
430 reward port.

431 We thus check if words in these three segments are more likely to have their probabilities
432 in training carried forward to post-training sleep. Figure 7C shows that when we plot the
433 closeness of each word's probability in training and sleep, we obtain a roughly symmetrical
434 distribution of locations for words closer to pre-training and post-training sleep. At the
435 three maze segments, we indeed find that a word's probability in training is equally likely
436 to be closer to pre-training sleep, equidistant from both sleep epochs, or closer to post-
437 training sleep (Figure 7D-F). We obtain the same results if we use just co-activity words,
438 or if we divide the closeness distribution into pre/equidistant/post by percentiles rather
439 than the fixed ranges we use in Figure 7D-F (data not shown). There is, then, no evidence
440 in this analysis that words representing specific maze locations, and putatively key task
441 elements, have their changes in training carried forward to post-training sleep. Rather,
442 changes to the structure of population activity during learning are distributed over the
443 entire maze.

444 **Independent neurons capture the majority of structure in prefrontal cor- 445 tex population activity**

446 The above analyses have shown that independently-firing neurons capture much of the
447 changes to and location dependence of population activity in medial prefrontal cortex.
448 This implies that independent neurons can account for much of the population activity
449 structure within each epoch. We take a closer look at this conclusion here.

450 A useful measure of the overall structure of the population spiking activity is the
451 proportion of "1's" that encode two or more spikes. The occurrence rates of these "binary
452 errors" across different bin sizes tell us about the burst structure of the neural activity.
453 Figure 8A shows that increasing the bin size applied to the data interpolates between
454 words of single spikes and words of spike bursts in both training and sleep epochs. At
455 bin sizes less than 10 ms, almost all 1's in each word are single spikes; at bin sizes above
456 50 ms, the majority of 1's in each word are two or more spikes and so encode a burst of
457 spikes from a neuron.

458 Dictionaries of independent neurons largely recapitulate these bin size dependencies
459 for all epochs (Figure 8B-D). Their only departure is about 5% more binary errors than
460 in the data at bin sizes above 20 ms (Figure 8D). As by construction there are the same
461 number of spikes for each neuron in the data and independent neuron dictionaries, this

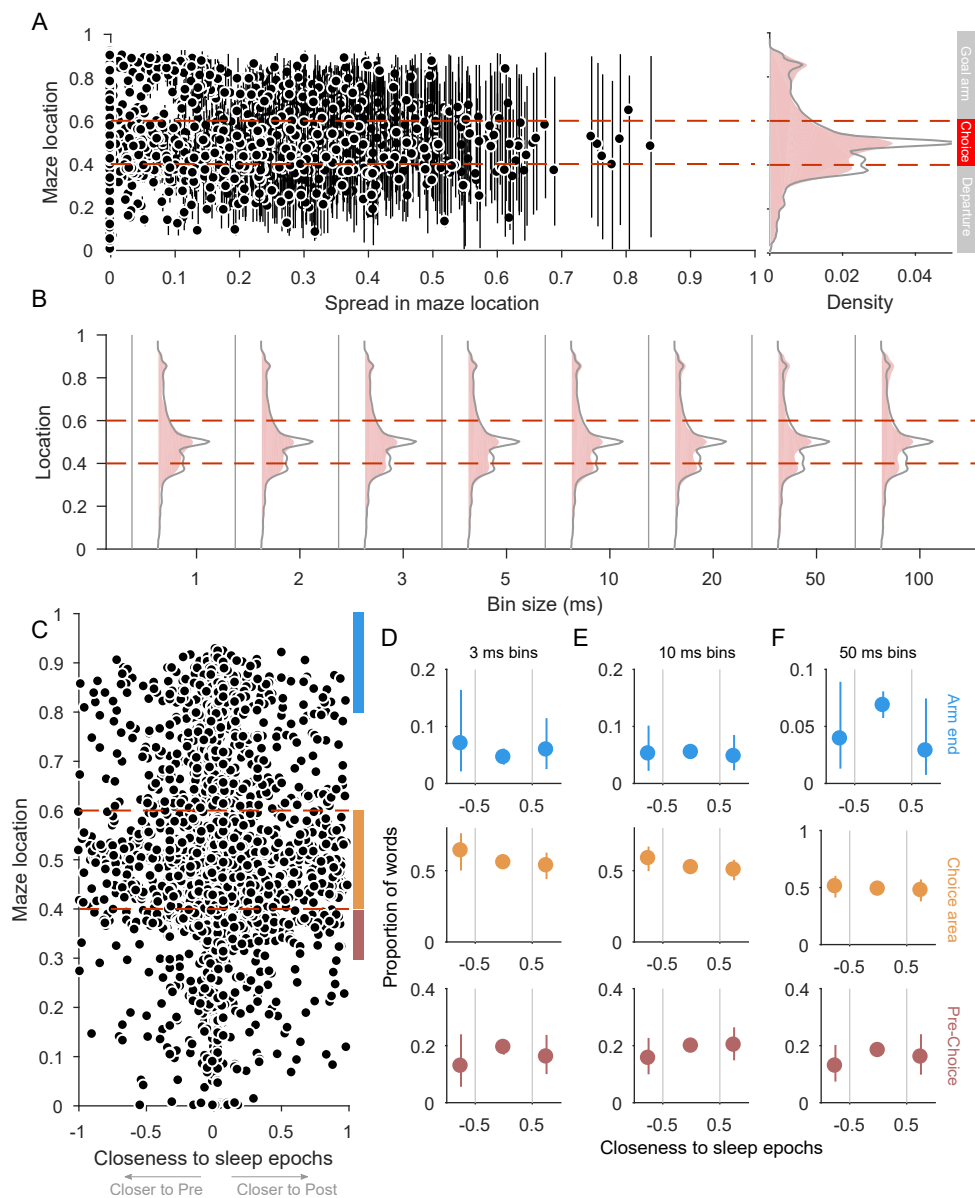


Figure 7. Locations of words during trials of learning sessions.

(A) Scatter of the spread in location against median location for every word in the training epoch dictionaries of the learning sessions, constructed using 3 ms bins. Spread in location is the inter-quartile interval, which we also plot as vertical lines. On the right we plot the density of median locations for the data (red area plot) and independent neuron (grey line) dictionaries.

(B) Density of median locations across all bin sizes, for data (red area plot) and independent neuron (grey line) dictionaries.

(C) For each word in the training epoch dictionaries, we plot its median location against the closeness between its training epoch and sleep epoch probability. Closeness is in the range $[-1, 1]$, where -1 indicates identical probability between training and pre-training sleep, and 1 indicates identical probability between training and post-training sleep. Coloured bars indicate the regions of the maze analysed in panels D-F.

(D) Distributions of word closeness to sleep in specific maze segments, for 3 ms bins. All words with median locations within the specified maze segment are divided into terciles of closeness by thresholds of -0.5 and 0.5 (vertical grey lines). Symbols plot proportions of words falling in each tercile, and error bars plot 99% confidence intervals on those proportions. Blue: arm end; orange: choice point; red: pre-choice segment.

(E) As panel D, for 10 ms bins.

(F) As panel D, for 50 ms bins.

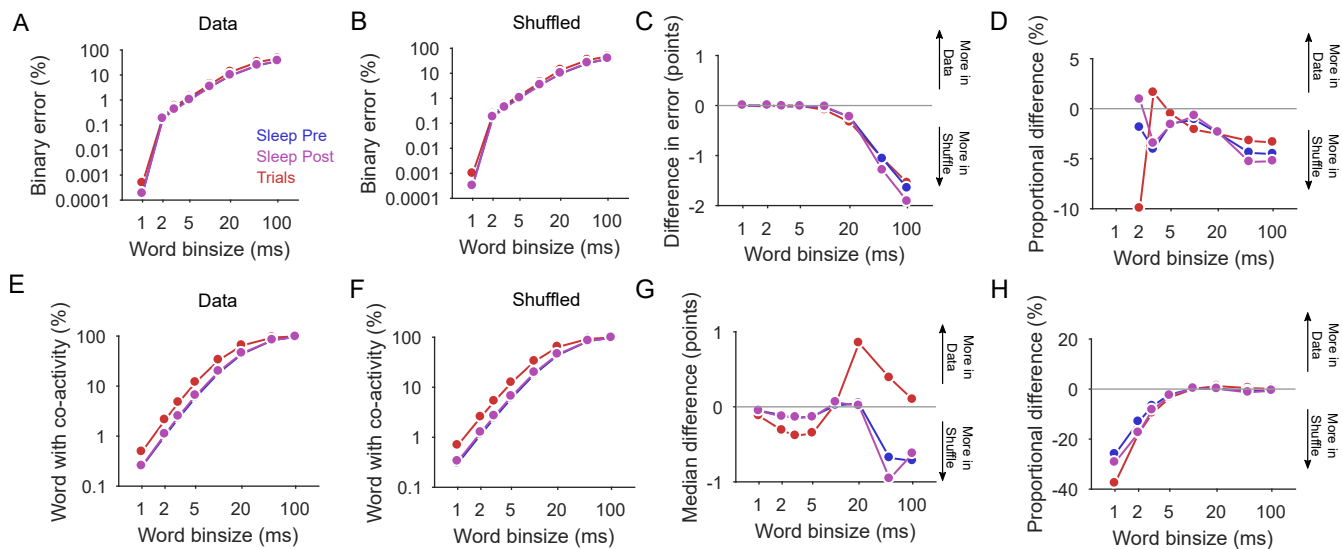


Figure 8. Independent neurons capture a large fraction of population activity structure.

(A) Proportion of “1’s” that encode more than one spike (“binary error”), across all emitted words in all learning sessions. Epoch colours apply to all panels.

(B) As panel (A), for dictionaries of independent neurons derived by shuffling neuron inter-spike intervals to remove correlations. Proportions from 20 shuffled datasets of the learning sessions.

(C) Mean difference between binary error proportions in the data and predicted by independent neurons, in percentile points.

(D) As panel (C), expressed as a proportion of the binary errors in the data.

(E) Proportion of emitted words in each epoch that have more than one active neuron, pooled across all learning sessions (replotted from Figure 5H).

(F) As panel (E), for dictionaries of independent neurons.

(G) Median difference between the proportion of emitted co-activity words in the data and predicted by independent neurons.

(H) As panel (G), expressed as a proportion of the number of co-activity words in the data.

462 implies that the data contain more spikes per burst on 50-100 ms time-scales (so that
463 there are fewer bins with bursts in total).

464 A useful summary of the joint structure of population activity is the fraction of emitted
465 words that code for two or more active neurons. For the data, increasing the bin size
466 increases the fraction of emitted words that contain more than one active neuron (Figure
467 8E), from about 1% of words at 2 ms bins to all words at 50 ms bins and above. There are
468 consistently more of these co-activity words in training epochs than sleep epochs for the
469 same bin size, pointing to more short time-scale synchronous activity during movement
470 along the maze than in sleep.

471 Dictionaries of independent neurons also recapitulate these bin size and epoch de-
472 pendencies of neural co-activity (Figure 8F-H). Figure 8H shows that the independent
473 neuron dictionaries have more co-activity words at small bin sizes. It might be tempt-
474 ing here to conclude that the data dictionaries are constrained to fewer co-activity words
475 than predicted by independent neurons; but these differences are equally consistent with
476 a shadowing effect from spike-sorting, where one or more near-simultaneous spikes from
477 neurons on the same electrode are missed (Harris et al., 2000; Bar-Gad et al., 2001):
478 when the data are shuffled, more near-simultaneous spikes between neurons are possible.
479 Nonetheless, above bins of 5 ms, the disagreement between the data and independent
480 neuron dictionaries is proportionally negligible (Figure 8H). Consequently, much of the
481 population activity in medial prefrontal cortex is well-captured by an independent-neuron
482 model, perhaps pointing to a high-dimensional basis for neural coding.

483 Discussion

484 We studied here how the structure of population activity in medial prefrontal cortex
485 changes during rule-learning. We found the structure of instantaneous population activ-
486 ity in sleep always changes after training, irrespective of any change in overt behaviour
487 during training. This plasticity of population activity could be entirely accounted for by
488 independent changes to the excitability of individual neurons. Unique to learning is that
489 changes to the structure of instantaneous population activity during training are carried
490 forward into the following bouts of sleep. Population plasticity during learning includes
491 both changes to individual neuron excitability and to co-variations of firing rates between
492 neurons. These results suggest two forms of population plasticity in medial prefrontal cor-
493 tex, one a constant form unrelated to learning, and the other correlated with the successful
494 learning of action-outcome associations.

495 To isolate learning and non-learning changes, we found useful the “strong inference”
496 approach of designing analyses to decide between simultaneous hypotheses for the same
497 data. We identified separable sessions of learning and stable behaviour in order to contrast
498 the hypothesis that population structure would only change during overt learning against
499 the hypothesis that population structure is always changing irrespective of behaviour.
500 Similarly, we contrasted three hypotheses for what drove those changes in population
501 structure: changes to excitability of independent neurons; changes in brief co-variations
502 of rates; and changes in precise co-spiking.

503 A dictionary of cortical activity states

504 Characterising the joint activity of cortical neurons is a step towards understanding how
505 the cortex represents coding and computation (deCharms and Zador, 2000; Wöhrer et
506 al., 2013; Yuste, 2015). One clue is that the joint activity of a cortical population seems
507 constrained to visit only a sub-set of all the possible states it could reach (Tsodyks et
508 al., 1999; Luczak et al., 2009; Sadtler et al., 2014; Jazayeri and Afraz, 2017), in part
509 determined by the connections into and within the network of cortical neurons (Galan,
510 2008; Marre et al., 2009; Ringach, 2009; Buesing et al., 2011; Habenschuss et al., 2013;
511 Kappel et al., 2015). This view predicts that changing the network connections through
512 learning would change the set of activity states (Battaglia et al., 2005).

513 We see hints of this prediction in our data. We found changes to the probability of
514 words in training that are detectable in post-training sleep, consistent with the idea that
515 reinforcement-related plasticity of the cortical network has persistently changed the con-
516 strained set of activity states. But changing the network’s connections should change not
517 just the set of activity states, but also their sequences or clustering in time (Tkacik et al.,
518 2014; Ganmor et al., 2015). This suggests that further insights into population plastic-
519 ity with these data could be found by characterising the preservation of word sequences
520 or clusters in time between training and sleep epochs, and comparing those to suitable
521 alternative hypotheses for temporal structure.

522 Excitability drives constant population plasticity

523 A change in the statistics of a population’s neural activity is not in itself evidence of
524 learning (Okun et al., 2012). Indeed, we saw here a constant shifting in statistical structure
525 between sleep epochs, regardless of whether the rats showed any evidence of learning in
526 the interim training epoch. As these shifts between sleep could be seen at all time-scales of
527 words we looked at, and were recapitulated by dictionaries of independent neurons, they
528 are most consistent with a model of independent changes to the excitability of individual
529 neurons.

530 Excitability changes could arise from the spontaneous remodelling of synaptic connec-
531 tions onto a neuron, whether from remodelling of dendritic spines (Fu et al., 2012; Hayashi-

532 Takagi et al., 2015), or changes of receptor and protein expression within a synapse (Wolff
533 et al., 1995; Ziv and Brenner, 2017). Alternatively, these changes could arise from long-
534 lasting effects on neuron excitability of neuromodulators accumulated in medial prefrontal
535 cortex during training (Seamans and Yang, 2004; Tierney et al., 2008; Dembrow et al.,
536 2010; Benchenane et al., 2011). A more detailed picture of this constant population plas-
537 ticity will emerge from stable long-term population recordings at millisecond resolution
538 (Jun et al., 2017) of the same prefrontal cortex neurons throughout rule-learning.

539 **Learning correlates with directional population plasticity**

540 Unique to learning a new rule in the Y-maze was that changes to word probability in
541 training were carried forward to post-training sleep. As this persistence of word probability
542 occurred most clearly for short time-scale words (20 ms or less), and were partly driven
543 by changes in rate co-variations, it is most consistent with a model of synaptic changes
544 to the prefrontal cortex driven by reinforcement. A possible mechanism here is that
545 reinforcement-elicited bursts of dopamine permitted changes of synaptic weights into and
546 between neurons whose co-activity preceded reward (Izhikevich, 2007; Benchenane et al.,
547 2011). Such changes in synaptic weights would also alter the excitability of the neuron
548 itself, accounting for the changes between pre and post-training sleep epochs in learning
549 sessions.

550 A particularly intriguing question is how the constant and learning-specific plasticity
551 of population activity are related. Again, stable long-term recordings of spiking activity
552 in the same population of neurons across learning would allow us to test whether neurons
553 undergoing constant changes in excitability are also those recruited during learning (Lee
554 et al., 2012; Hayashi-Takagi et al., 2015). Another question is how the carrying forward of
555 training changes of population activity into sleep depends on an animal's rate of learning.
556 In each learning session here the identified learning trial was before the half-way mark,
557 meaning that the majority of words contributing to the training dictionary came from trials
558 after the rule was acquired. It is an open question as to whether the same relationship
559 would be seen in sessions of late learning, or in tasks with continual improvement in
560 performance rather than the step changes seen here.

561 **Replay and dictionary sampling**

562 The increased similarity of word probability in training and post-training sleep suggests
563 an alternative interpretation of “replay” phenomena in prefrontal cortex (Euston et al.,
564 2007; Peyrache et al., 2009). Replay of neural activity during waking in a subsequent
565 episode of sleep has been inferred by searching for matches of patterns of awake activity
566 in sleep activity, albeit at much coarser time-scales than used here. The better match of
567 waking activity with subsequent sleep than preceding sleep is taken as evidence that replay
568 is encoding recent experience, perhaps to enable memory consolidation. However, our
569 observation that the probabilities of words in stable sessions' trials are not systematically
570 closer to those in post-training sleep (Figure 4) is incompatible with the simple replay of
571 experience-related activity in sleep. Rather, our results suggest learning correlates with
572 persistent changes to the cortical network, such that words have more similar probabilities
573 of appearing in training and post-training sleep than in training and pre-training sleep. In
574 this view, replay is a signature of activity states that appeared in training being resampled
575 in post-training sleep (Battaglia et al., 2005).

576 **Population coding of statistical models**

577 What constraints do these changes to mPFC population activity place on theories for
578 acquiring and representing statistical models of actions and their outcomes? In this view,
579 the joint activity of the population during the trials represents something like the joint
580 probability $P(a, o|state)$ of action a and outcome o given the current state of the world

581 (Alexander and Brown, 2011); or, perhaps more generally, a model for the transitions
582 in the world caused by actions, $P(\text{state}(t + 1)|a, \text{state}(t))$. Such models could support
583 the proposed roles of medial prefrontal cortex in guiding action selection (by querying
584 the outcomes predicted by the model), or monitoring behaviour (by detecting unexpected
585 deviations from the model). The changes in the structure of population activity during
586 learning are consistent with updating such models based on reinforcement.

587 Our results show these dictionary changes are carried forward to the spontaneous
588 activity of sleep, suggesting that the encoded statistical model is present there too. One
589 explanation for this stems from the sampling hypothesis for probability encoding. In this
590 hypothesis, a population encodes a statistical model in the joint firing rates of its neurons,
591 so that the pattern of activity across the population at each moment in time is a sample
592 from the encoded distribution (Fiser et al., 2010; Berkes et al., 2011). This hypothesis
593 predicts that spontaneous activity of the same neurons must still represent samples from
594 the statistical model: but in the absence of external input, these are then samples from
595 the “prior” probability distribution over the expected properties of the world.

596 According to this hypothesis, our finding that learning-driven changes to population
597 structure are conserved in post-training sleep is consistent with the statistical model now
598 reflecting well-learned expected properties of the world – namely, that a particular set of
599 actions on the maze reliably leads to reward. In other words, the prior distribution for the
600 expected properties of the world has been updated. Further, the sampling hypothesis also
601 proposes a role for the constant changes of excitability without obvious direction – that
602 such spontaneous plasticity explores possible configurations of the network and so acts as
603 a search algorithm to optimise the encoded statistical model (Kappel et al., 2015; Maass,
604 2016). These links, while tentative, suggest the utility of exploring models for probabilistic
605 codes outside of early sensory systems (Fiser et al., 2010; Pouget et al., 2013).

References

- Alexander WH, Brown JW (2011) Medial prefrontal cortex as an action-outcome predictor. *Nature Neurosci* 14:1338–1344.
- Bar-Gad I, Ritov Y, Vaadia E, Bergman H (2001) Failure in identification of overlapping spikes from multiple neuron activity causes artificial correlations. *J Neurosci Meth* 107:1–13.
- Battaglia FP, Sutherland GR, Cowen SL, Mc Naughton BL, Harris KD (2005) Firing rate modulation: a simple statistical view of memory trace reactivation. *Neural Netw* 18:1280–1291.
- Benchenane K, Peyrache A, Khamassi M, Tierney PL, Gioanni Y, Battaglia FP, Wiener SI (2010) Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron* 66:921–936.
- Benchenane K, Tiesinga PH, Battaglia FP (2011) Oscillations in the prefrontal cortex: a gateway to memory and attention. *Curr Opin Neurobiol* 21:475–485.
- Berkes P, Orbán G, Lengyel M, Fiser J (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331:83–87.
- Brown LD, Cai TT, DasGupta A (2001) Interval estimation for a binomial proportion. *Statist Sci* 16:101–133.
- Buesing L, Bill J, Nessler B, Maass W (2011) Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput Biol* 7:e1002211.

- deCharms RC, Zador A (2000) Neural representation and the cortical code. *Annu Rev Neurosci* 23:613–647.
- Dembrow NC, Chitwood RA, Johnston D (2010) Projection-specific neuromodulation of medial prefrontal cortex neurons. *J Neurosci* 30:16922–16937.
- Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* 66:438–448.
- Euston DR, Gruber AJ, McNaughton BL (2012) The role of medial prefrontal cortex in memory and decision making. *Neuron* 76:1057–1070.
- Euston DR, Tatsuno M, McNaughton BL (2007) Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science* 318:1147–1150.
- Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci* 14:119–130.
- Floresco SB, Block AE, Tse MTL (2008) Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behav Brain Res* 190:85–96.
- Fu M, Yu X, Lu J, Zuo Y (2012) Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. *Nature* 483:92–95.
- Galan RF (2008) On how network architecture determines the dominant patterns of spontaneous neural activity. *PloS One* 3:e2148.
- Ganmor E, Segev R, Schneidman E (2015) A thesaurus for a neural population code. *Elife* 4:e06134.
- Habenschuss S, Jonke Z, Maass W (2013) Stochastic computations in cortical microcircuit models. *PLoS Comput Biol* 9:e1003311.
- Harris KD, Henze DA, Csicsvari J, Hirase H, Buzsáki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J Neurophysiol* 84:401–414.
- Hayashi-Takagi A, Yagishita S, Nakamura M, Shirai F, Wu YI, Loshbaugh AL, Kuhlman B, Hahn KM, Kasai H (2015) Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature* 525:333–338.
- Holroyd CB, McClure SM (2015) Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. *Psychological review* 122:54–83.
- Izhikevich EM (2007) Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cereb Cortex* 17:2443–2452.
- Jazayeri M, Afraz A (2017) Navigating the neural space in search of the neural code. *Neuron* 93:1003–1014.
- Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, Lee AK, Anastassiou CA, Andrei A, Aydn C, Barbic M, Blanche TJ, Bonin V, Couto J, Dutta B, Gratiy SL, Gutnisky DA, Hausser M, Karsh B, Ledochowitsch P, Lopez CM, Mitelut C, Musa S, Okun M, Pachitariu M, Putzeys J, Rich PD, Rossant C, Sun WL, Svoboda K, Carandini M, Harris KD, Koch C, O’Keefe J, Harris TD (2017) Fully integrated silicon probes for high-density recording of neural activity. *Nature* 551:232–236.

- Kappel D, Habenschuss S, Legenstein R, Maass W (2015) Network plasticity as Bayesian inference. *PLoS Comput Biol* 11:e1004485.
- Karlsson MP, Tervo DGR, Karpova AY (2012) Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* 338:135–139.
- Khamassi M, Quilodran R, Enel P, Dominey PF, Procyk E (2015) Behavioral regulation and the modulation of information coding in the lateral prefrontal and cingulate cortex. *Cereb Cortex* 25:3197–3218.
- Lee D, Lin BJ, Lee AK (2012) Hippocampal place fields emerge upon single-cell manipulation of excitability during behavior. *Science* 337:849–853.
- Luczak A, Barthó P, Harris KD (2009) Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* 62:413–425.
- Maass W (2016) Searching for principles of brain computation. *Curr Opin Behav Sci* 11:81–92.
- Maggi S, Peyrache A, Humphries MD (2018) An ensemble code in medial prefrontal cortex links prior events to outcomes during learning. *Nature Comms* p. in press.
- Marre O, Yger P, Davison AP, Frégnac Y (2009) Reliable recall of spontaneous activity patterns in cortical networks. *J Neurosci* 29:14596–14606.
- Okun M, Steinmetz NA, Cossell L, Iacaruso MF, Ko H, Bartho P, Moore T, Hofer SB, Mrcic-Flogel TD, Carandini M, Harris KD (2015) Diverse coupling of neurons to populations in sensory cortex. *Nature* 521:511–515.
- Okun M, Yger P, Marguet SL, Gerard-Mercier F, Benucci A, Katzner S, Busse L, Carandini M, Harris KD (2012) Population rate dynamics and multineuron firing patterns in sensory cortex. *J Neurosci* 32:17108–17119.
- Peyrache A, Khamassi M, Benchenane K, Wiener SI, Battaglia F (2018) Activity of neurons in rat medial prefrontal cortex during learning and sleep. doi:10.6080/K0KH0KH5.
- Peyrache A, Khamassi M, Benchenane K, Wiener SI, Battaglia FP (2009) Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat Neurosci* 12:916–926.
- Pouget A, Beck JM, Ma WJ, Latham PE (2013) Probabilistic brains: knowns and unknowns. *Nat Neurosci* 16:1170–1178.
- Powell NJ, Redish AD (2016) Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behaviour. *Nat Commun* 7:12830.
- Ragozzino ME, Detrick S, Kesner RP (1999a) Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *J Neurosci* 19:4585–4594.
- Ragozzino ME, Wilcox C, Raso M, Kesner RP (1999) Involvement of rodent prefrontal cortex subregions in strategy switching. *Behav Neurosci* 113:32–41.
- Rich EL, Shapiro ML (2007) Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *J Neurosci* 27:4747–4755.
- Ringach DL (2009) Spontaneous and driven cortical activity: implications for computation. *Curr Opin Neurobiol* 19:439–444.

- Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, Yu BM, Batista AP (2014) Neural constraints on learning. *Nature* 512:423–426.
- Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74:1–58.
- Starkweather CK, Gershman SJ, Uchida N (2018) The medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. *Neuron* p. in press.
- Tavoni G, Ferrari U, Battaglia FP, Cocco S, Monasson R (2017) Functional coupling networks inferred from prefrontal cortex activity show experience-related effective plasticity. *Network Neurosci* pp. 275–301.
- Tierney PL, Thierry AM, Glowinski J, Deniau JM, Gioanni Y (2008) Dopamine modulates temporal dynamics of feedforward inhibition in rat prefrontal cortex in vivo. *Cereb Cortex* 18:2251–2262.
- Tkacik G, Marre O, Amodei D, Schneidman E, Bialek W, Berry n MJ (2014) Searching for collective behavior in a large network of sensory neurons. *PLoS Comput Biol* 10:e1003408.
- Tsodyks M, Kenet T, Grinvald A, Arieli A (1999) Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science* 286:1943–1946.
- Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, Hassabis D, Botvinick M (2018) Prefrontal cortex as a meta-reinforcement learning system. *Nat Neurosci* 21:860–868.
- Wohrer A, Humphries MD, Machens C (2013) Population-wide distributions of neural activity during perceptual decision-making. *Prog Neurobiol* 103:156–193.
- Wolff JR, Laskawi R, Spatz WB, Missler M (1995) Structural dynamics of synapses and synaptic components. *Behav Brain Res* 66:13–20.
- Yuste R (2015) From the neuron doctrine to neural networks. *Nat Rev Neurosci* 16:487–497.
- Ziv NE, Brenner N (2017) Synaptic tenacity or lack thereof: Spontaneous remodeling of synapses. *Trends Neurosci* 41:89–99.