bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

Independent population coding of the past and the present in prefrontal cortex during learning

Silvia Maggi[1] and Mark D. Humphries[1,2*]

1. School of Psychology, University of Nottingham, Nottingham, UK.
2. Faculty of Biology, Medicine, and Health, University of Manchester, Manchester, UK.

* Contact: mark.humphries@nottingham.ac.uk

Abstract

Medial prefrontal cortex (mPfC) plays a role in present behaviour and short-term memory. Unknown is whether the present and the past are represented in the same mPfC neural population and, if so, how the two representations do not interfere. Analysing mPfC population activity of rats learning rules in a Y-maze, we find population activity switches from encoding the present to encoding the past of the same events after reaching the arm-end. We show the switch is driven by population activity rotating to orthogonal axes, and the population code of the present and not the past reactivates in subsequent sleep, confirming these axes were independently accessible. Our results suggest mPfC solves the interference problem by encoding the past and present on independent axes of activity in the same population, and support a model of the past and present encoding having independent functional roles, respectively contributing to on-line learning and off-line consolidation.

Keywords: decision making, mPFC, learning, neural ensembles, sleep, replay

Introduction

The medial prefrontal cortex (mPfC) plays key roles in adaptive behaviour. It is involved in reshaping behaviour in response to changes in a dynamic environment (Euston et al., 2012) and in response to errors in performance (Narayanan and Laubach, 2008; Laubach et al., 2015). Damage to mPfC prevents shifting behavioural strategies when the environment changes (Laskowski et al., 2016; Guise and Shapiro, 2017). Single neurons in mPfC shift the timing of spikes relative to hippocampal theta rhythms just before acquiring a new action-outcome rule (Benchenane et al., 2010). And multiple labs have reported that global shifts in mPfC population activity precede switching between behavioural strategies (Rich and Shapiro, 2009; Durstewitz et al., 2010; Karlsson et al., 2012; Schuck et al., 2015; Powell and Redish, 2016) and the extinction of learnt associations (Russo et al., 2020).

Adapting behaviour depends on knowledge of both the past and the present. Deep lines of research have established that mPfC activity represents information about both. The memory of the immediate past is maintained in mPfC activity, both in tasks requiring explicit use of working memory (Miller and Cohen, 2001; Baeg et al., 2003; Averbeck and Lee, 2007; Miller et al., 2005; Fujisawa et al., 2008; Machens et al., 2010; Spellman et al.,

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

2015) and those that do not (Jun et al., 2010; Maggi et al., 2018). The use of such memory is seen in both the impairment arising from mPfC lesions (Rich and Shapiro, 2007; Young and Shapiro, 2009; Laskowski et al., 2016), and the role of mPfC in error monitoring (Laubach et al., 2015). Representations of stimuli and events happening in the present have been reported in a variety of decision-making tasks (Averbeck et al., 2006; Rigotti et al., 2013; Erlich et al., 2015; Hanks et al., 2015; Ito et al., 2015; Siegel et al., 2015; Guise and Shapiro, 2017).

Little is known though about the relationship between representations of the past and present in mPfC activity. Prior studies have shown that past and upcoming choices can both modulate activity of neurons in the same mPfC population (for example Baeg et al., 2003; Barraclough et al., 2004; Ito et al., 2015), but none have compared the encodings of the past and present, nor determined how the encoding of the present becomes the encoding of the past. Thus important questions remain: how the past and present are encoded in a mPfC population so they do not interfere with each other, and how the encoding of features in the present transforms into the encoding of the past.

To address these questions, we reanalyse here mPfC population activity from rats learning new rules on a Y-maze (Peyrache et al., 2009). Crucially, this task had distinct trial and inter-trial interval phases, in which we could respectively examine the population encoding of the present (in trials) and the past (in the intervals) of the same task features or events. We first established that small mPfC populations did indeed encode both the present and past of the same features of the task, respectively in the trial and in the inter-trial interval. We found that these encodings were orthogonal, so that the present and the past were encoded by activity evolving along independent coding axes. This independent encoding was functional: population activity encoding the present was reactivated in post-training sleep, but activity encoding the same features in the past was not reactivated. Finally, we show that the pattern of synchrony across the population is also independent between trials and inter-trial intervals. By encoding the past and present of the same events on orthogonal axes, a single mPfC population prevents interference between them, and allows their independent recall.

Results

To address how the mPfC encodes the past and the present, we analyse here data from rats learning rules in a Y maze, which had tetrodes implanted in mPfC before the first session of training (Figure 1a). The animal self-initiated each trial by running along the central stem of the Y maze and choosing one of the arms. The trial finished at the arm's end, and a reward delivered if the chosen arm matched the current rule being acquired. During the following inter-trial interval, the rat made a self-paced return to the start of the central arm to initiate the next trial. This task thus allowed us to study the representation of choice and its environmental context in both the present (the trial) and the immediate past (the inter-trial interval).

Population activity encodes multiple features of the present

In order to compare representations of the same choice and features in the past and present, we first had to establish that these were indeed represented in mPfC population activity. Previously we had shown that these mPfC populations encode the past choice of arm, the past outcome, and the past light position during the inter-trial interval (Maggi et al., 2018) – below we develop these findings further. We thus report here our decoding of the

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

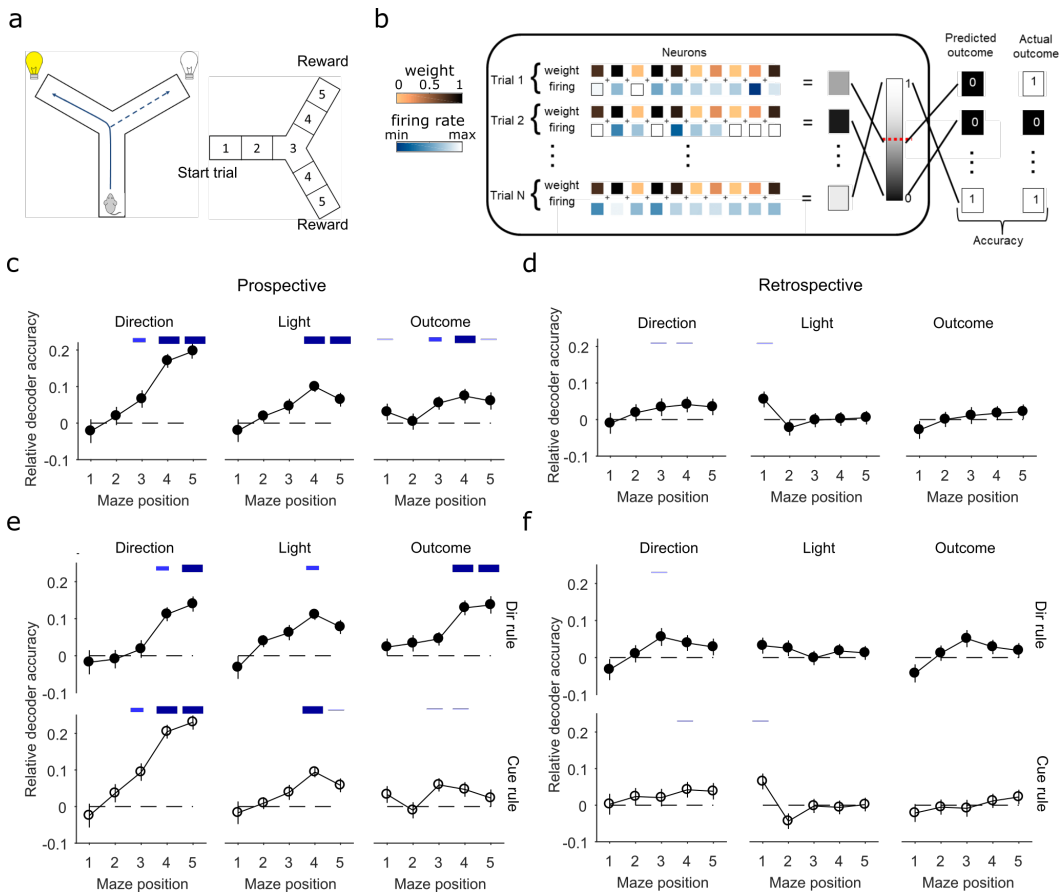

Figure 1: Population encoding of task features and events during trials.
(**a**) The Y maze task. Left: task, schematic with rat at the start position. A trial runs from the start position to the end of the chosen arm; the inter-trial interval is the return from the arm end to the start position. Across sessions, animals were asked to learn one of 4 rules in sequence (go to the right arm, go to the lit arm, go to the left arm, go to the dark arm). Rules were switched after 10 correct choices (or 11 out of 12). Right: division of the maze into 5 positions for analysis of position-dependent encoding.
(**b**) Graphical representation of a classifier. The population vector firing rate for each trial (shade of blue squares) is used as input for a linear classifier that fits the weight (shade of yellow squares) for each neuron across trials. A linear combination of weight and population firing rate for each trial is compared to a threshold (red dashed line) to predict the category to which the trial belongs. Classifier accuracy is the proportion of correctly predicted held-out trials.
(**c**) Decoding the animal's choice (of direction), the cue position, and the outcome of each trial as a function of maze position. We plot the relative decoder accuracy, the difference between the absolute decoder accuracy and the shuffled control for each feature (we plot example absolute accuracy results in Supplementary Figure 1b). Decoding accuracy above chance appear from the choice point (position 3) to the end of the maze. Here and in all panels: symbols are mean ± SEM over all 49 sessions; blue bars indicate positions with decoding performance significantly better than shuffled controls (Wilcoxon sign rank test, $p < 0.05$ thin light blue; $p < 0.01$ middle thickness blue; $p < 0.005$ thick dark blue line).
(**d**) Decoding of the previous trial's choice, cue position, and outcome as a function of maze position on the present trial ($N = 49$).
(**e**) Breakdown of the prospective decoding results in panel (**c**) by rule type (16 direction rule sessions; 34 cue rule sessions).
(**f**) Breakdown of the retrospective decoding results in panel (**d**) by rule type.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

present choice, light position, and imminent outcome from the trials' mPfC population activity.

To examine the evolution of encoding over the trial, we divided the maze into five equally sized sections, and constructed population firing rate vectors for each position in each trial (Figure 1a,b). The set of population vectors across all trials of a session were used to train and test a cross-validated classifier that predicted features – direction or outcome or light position. Only neurons active in every trial could thus contribute to the population vectors; these active populations ranged from 4-22 neurons across 49 sessions, of between 7-51 trials each (Supplementary Figure 1a). We trained the same classifiers using the same population vectors but with features shuffled across trials (see Methods and Supplementary Figure 1b), to define appropriate chance levels for each classifier given the unbalanced distribution of some task features (such as outcome).

We found that the current trial's direction choice, light position, and outcome were significantly encoded in mPfC population activity, and typically from when the animal was at the choice point of the maze until it reached the chosen arm's end (Figure 1c). Moreover, this decoding was a property of the population, not of a small number of highly tuned neurons within the population: when we restricted the analysis to sessions with no significantly tuned neurons ($n = 27$ for direction choice, $n = 39$ for the light position and $n = 25$ for outcome) we could still decode all these task features from the choice point onwards (Supplementary Figure 1c).

These multiple feature encodings were present whether the rule being learnt was a direction (go-left, go-right) or cued (go-to-the-lit-arm) association (Figure 1e). While the choice of arm and the outcome are confounded for direction rules, Figure 1e shows that for cued rules we could still decode the trial's outcome above chance around the choice point. These decodings of the present direction choice, light position, and anticipated outcome were strongly robust, being recapitulated by decoders applied to the whole maze rather than separate sections (Supplementary Figure 1d), and by a range of different classifiers (Supplementary Figure 2).

In order to determine if learning itself affected any mPfC representations of the present, we separated the sessions into two behavioural groups: putative learning sessions ($n = 10$), identified by a step-change in task performance (Supplementary Figure 3a-c), and the remaining sessions, called here "Other" ($n = 39$). We found population coding of task features was similar when comparing learning sessions and all Other sessions (Supplementary Figure 3d-g). However, there were differences in representations between the learning of direction and cue rules. Direction choice could be decoded from the very first session of each animal, suggesting that this representation is not learnt (Supplementary Figure 4). By contrast, the encoding of light position was not present during the initial sessions that enforced a direction rule (Supplementary Figure 4), but most clearly emerged during cue rules when the light was relevant to solve the task, suggesting the representation of light position was learnt (Supplementary Figure 4).

In contrast to the robust encoding of the present, we found scant evidence that mPfC activity during a trial encoded the events of the previous trial (Figure 1d,f; whole maze encoding is in Supplementary Figure 1e; breakdown by session type in Supplementary Figure 3e-g). Previous studies have reported encoding of past choices in mPfC population activity during trials (Baeg et al., 2003; Sul et al., 2010). We found some evidence of the population coding of past arm choice (Figure 1d - 'Direction') but the encoding was weak, and differed between rule types (Figure 1f - 'Direction'). Notably, there was no evidence of any encoding of the previous trial's outcome. Thus, during trials population activity in the prefrontal cortex had robust, sustained encoding of multiple events of the present,

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

but at best weakly and transiently encoded events of the past.

Orthogonal encoding of the past and the present

Having established that population activity in a trial encodes multiple features of the present, we then asked if this encoding is carried forward into the following inter-trial interval. Previously we reported that during the inter-trial interval these populations robustly encode the features - choice, light, and outcome - of the preceding trial (Maggi et al., 2018). Indeed, as we show in Figure 2a, we could increasingly decode the rats' present direction choice as they ran to the arm end to complete the trial, and then continue to decode this direction choice, now in the past, as the rats returned to the start position. Supplementary Figure 5a shows the encoding of light and outcome were similarly contiguous.

One hypothesis is that sustained activity in mPfC continues from the trial into the inter-trial interval, creating a memory trace of the encoding during the trial. Another plausible hypothesis is that the population activity in the trial reactivates during the inter-trial interval, in some form of replay of waking activity. Both hypotheses predict that the population encoding of a feature in the trial and in the following inter-trial interval should be the same. We show here it is not.

One simple way to rule out the memory trace hypothesis is if the active neurons during the trial and inter-trial interval were different. However, the active neurons during the trials were also active during the inter-trial interval (Supplementary Figure 5b), and we could decode features well from activity during the whole trial (Supplementary Figure 1d,e) and whole inter-trial interval (Figure 2c), so this shared common population could, in principle, carry on encoding the same task features.

We used this common population to test the memory trace hypothesis: if the encoding is broadly the same, then the activity in the trial and following inter-trial interval should be interchangeable when predicting the same feature, such as the chosen direction. In this cross-decoding test, we first trained a classifier for features of the present using the common population's activity during the trials, and then tested the accuracy of the classifier when using the common population's activity during the inter-trial interval. If the encoding in the trials really is carried forward into the inter-trial interval, then this cross-decoding should be accurate.

We found that cross-decoding of features was consistently poor in both directions, whether we trained on trial activity and tested on inter-trial intervals, or vice-versa (Figure 2b). Decoding of all features was at or close to chance, strikingly at odds with the within-trial (Supplementary Figure 1d and 3f) or within-interval (Figure 2a,c) decoding. As an additional control, we used leave-one-out cross-decoding instead (Methods) and this still resulted in decoding at or about chance levels (Supplementary Figure 5c,d). These results suggest that population encoding of prior events in the inter-trial interval is not simply a memory trace or reactivation of similar activity in the trial. Instead, they show that the same mPfC population is separately and independently encoding the present and past of the same features.

To quantify this independence, we turned to the vector of decoding weights for the trials and the equivalent vector for the inter-trial intervals of the same session. These weights, obtained from the decoder trained once on all trials and once on all inter-trial intervals, give the relative contribution of each neuron to the encoding of task features. We found that the trial and inter-trial interval weight vectors were approximately orthogonal for all three features: the angles cluster at or close to $\pi/2$ (or, equivalently, their dot-product clusters at or around zero) (Fig 2d; see Supplementary Figure 5e for a breakdown

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

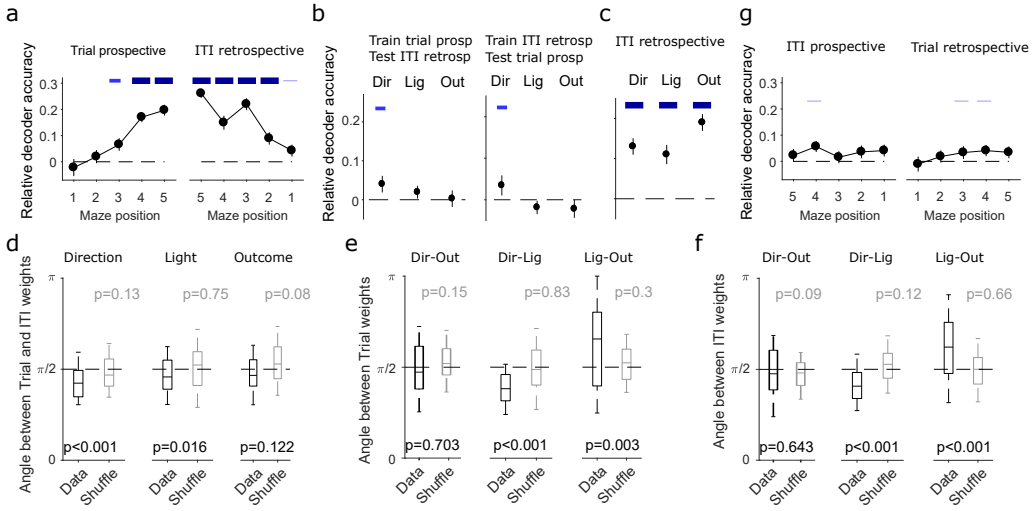

Figure 2: Orthogonal population encoding of past and present task events.
(**a**) Decoding of features is contiguous between the trials and following inter-trial intervals. We plot here the decoding accuracy in the trials for the animal's current choice of direction (left), and the decoding accuracy in the inter-trial intervals for the animal's past choice of direction (right). The left panel is redrawn from Figure 1c. In panels (a-c) and (g): symbols show means ± SEM; blue bars indicate decoding performance significantly better than chance (Wilcoxon sign rank test, $p < 0.05$ thin light blue; $p < 0.01$ middle thickness blue; $p < 0.005$ thick dark blue line; $N = 49$ for all).
(**b**) Cross-decoding performance. Left: relative decoding accuracy of a logistic regression classifier trained on all trials prospectively and tested on all inter-trial intervals retrospectively. Black dashed line shows the chance levels obtained training the classifier on shuffled labels for the trials and testing on inter-trial intervals given the same shuffled labels. Right: relative decoding accuracy for a classifier trained on all inter-trial intervals retrospectively and tested on all trials prospectively.
(**c**) Decoding accuracy for features of the preceding trial from inter-trial interval population activity over the whole maze.
(**d**) Comparison of the decoding vector weights between trials and inter-trial intervals. For each session we plot the angle between its trial and inter-trial interval decoding weight vectors, obtained from the trained classifiers in panel (b). For reference, we also compute the angle between trial and inter-trial interval decoding vectors obtained by training on shuffled label data (grey). Boxplots show median (line), inter-quartile range (box), and 95% interval (tails). P-values are from Wilcoxon ranksum tests for the difference from $\pi/2$.
(**e**) As for (d), but comparing the decoding weight vectors between features, within trials.
(**f**) As for (e), for within inter-trial intervals.
(**g**) Decoding of features is not contiguous between the inter-trial intervals and following trials. We plot here the decoding accuracy in the inter-trial interval of the animal's choice in the following trial (left), and the decoding accuracy in the following trial of the animal's choice in the previous trial (right) – i.e. the choice that was encoded during the preceding inter-trial interval.

by session types). Median angles for direction and light position were significantly less than $\pi/2$ (ranksum test), but the difference was small: 0.067π for direction and 0.045π for light position. Thus, the population encoding in the inter-trial interval was not a memory trace: to a good approximation, the past and present are orthogonally encoded in the same mPfC population.

We considered a range of alternative explanations for these results. One is that the orthogonality arises from the curse of dimensionality: the distance between two i.i.d random vectors with a mean of zero tends to grow with their increasing dimension. If the decoding

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

weights were random vectors, then the apparent orthogonality could be driven by just the largest populations. However, the decoding weights for the whole trial (present) or whole inter-trial interval (past) are not random vectors, for if they were then decoding performance would be at chance, whereas we find clear decoding of all features (Supplementary Figure 1d and Figure 2c). Another explanation is that the independent encoding axes between the trials and inter-trial intervals is somehow driven by differing properties of the trials and inter-trial intervals. For example, they differ in duration (mean 6.5 ± 0.01 seconds for trials, 55.7 ± 0.03 seconds for inter-trial intervals), and hence also in average movement speed. If switching between trials and inter-trial intervals could account for encoding differences, then these differences should be symmetric: we should see encodings change whether the transition was from the trial to inter-trial interval, or from the inter-trial interval back to a trial. However, the encodings were asymmetric: we saw strong encoding during the transition from trial to inter-trial interval (Figure 2a and Supplementary Figure 5a), but no encoding during the transition from inter-trial interval back to the trial (Figure 2g and Supplementary Figure 5f; see Maggi et al. (2018) for full details of the absence of prospective encoding in the inter-trial intervals). In the absence of any encoding, there cannot be an orthogonal shift in encoding.

We also examined the relationship between the features' encoding vectors during the trial and during the inter-trial interval. The encoding axes within an epoch were less independent than between epochs: angles between the encoding vectors for light and direction and for light and outcome were significantly different from $\pi/2$ (Figure 2e,f). But the distributions of angles between the encoding vectors were preserved between the trials and the inter-trial intervals, with outcome-direction around $\pi/2$, light-direction centered below $\pi/2$, and light-outcome centred above $\pi/2$. Thus, while each encoding axis rotated to an orthogonal direction between the trial and inter-trial interval, the internal relationships between the feature encodings was preserved.

Population activity rotates between trials and inter-trial intervals

That all three feature encodings were independent between the trials and inter-trial intervals of a session predicts that the population activity itself should be independent. If true, then trial and inter-trial interval population activity vectors should be easily separable. To test this prediction, we projected all population activity vectors of a session (Fig 3a) into a low dimensional space (Fig 3b), and then quantified how easily we could separate them into trials and inter-trial intervals. Using just one dimension was sufficient for near-perfect separation in many sessions; using two was sufficient for above-chance performance in all sessions (Fig 3c; and see Supplementary Figure 6 for a breakdown of each session's dependence on the number dimensions). Population activity was thus about as independent between the trials and inter-trial intervals as it possibly could be.

The independence in the population activity might arise from the continuous evolution of mPfC population activity across the contiguous trial and inter-trial interval period, such as the sequential activation of PfC neurons observed in previous studies (e.g. Fujisawa et al., 2008). If sequential activation was ongoing, then we should also observe consistently independent population activity between consecutive sections of the maze during trials and during inter-trial intervals (Fig 3d-g). Instead, we found population activity was not independent between contiguous maze sections within trials or within inter-trial intervals. For example, Figure 3d-f shows that population vectors of maze sections 4 and 5 in the inter-trial interval were not easily separable in four or fewer dimensions. Across the whole maze, adjacent sections within trials and inter-trial intervals had classification errors consistently greater than any found between trials and inter-trial intervals (Figure 3g), even

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

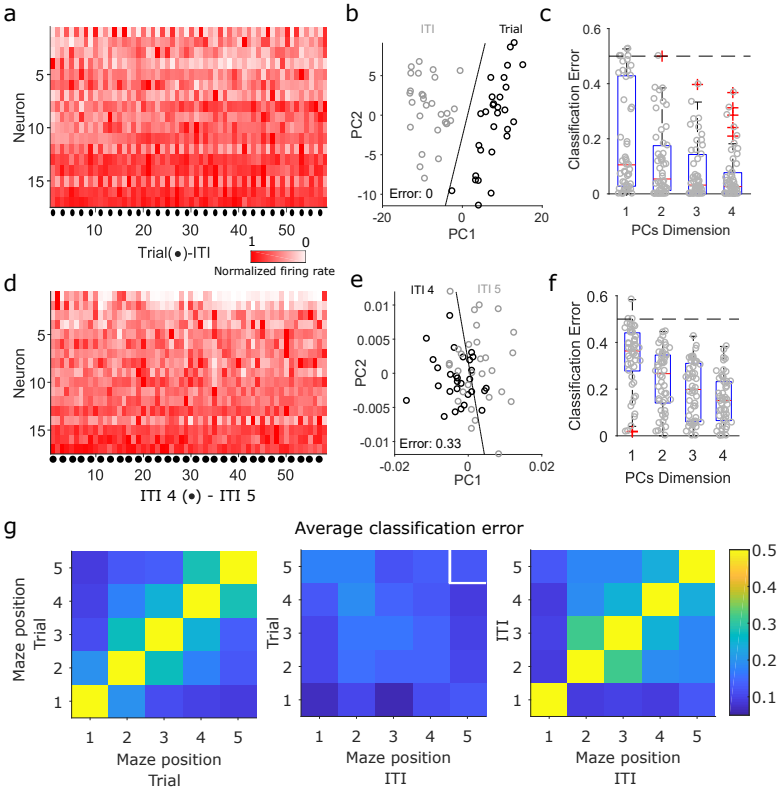

Figure 3: Population activity is independent between trials and inter-trial intervals
(**a**) Population activity vectors for the trials (•) and following inter-trial intervals of one session. The heat-map shows the normalized firing rate for each neuron.
(**b**) Projection of that session's population activity vectors on to two dimensions shows a complete separation of trial and inter-trial interval activity. The black line is the linear separation found by the classifier. PC: principal component.
(**c**) Summary of classification error over all sessions, as a function of the number of dimensions. Each grey dot is the error for one session at that number of projecting dimensions. Dashed line gives chance performance. Boxplots show medians (red line), interquartile ranges (blue box), and outliers (red pluses).
(**d-f**) Same as panels (a-c), but comparing population activity vectors for maze sections 4 and 5 in the inter-trial interval.
(**g**) Heat maps for the average classification error in the separation of population activity vectors between each pair of maze sections within trials (left panel), between trials and inter-trial intervals (middle panel), and within inter-trial intervals (right panel). White square in the middle panel indicates the arm-end position, where the transition from trial to inter-trial interval occurs. Classification error is for projections of the population activity vectors in a two dimensional space.

when the animal was in the same maze position (white square, middle panel). These results confirm that the marked changes in decoding accuracy across the maze (Figure 2a) corresponded to evolving population activity within trials and inter-trial intervals; but they also show that this evolving activity happened along independent directions in the trials and inter-trial intervals.

Population representations of trial features re-activate in sleep

Encoding the past and present of the same features in the same population on independent axes means that, in principle, the representations of past and present can be addressed or recalled without interference. We thus sought further evidence of this independent

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

encoding by asking if either representation could be recalled independently of the other.

Prior reports showed that patterns of mPfC population activity during training are preferentially repeated in post-training slow-wave sleep (Euston et al., 2007; Peyrache et al., 2009; Singh et al., 2019), consistent with a role in memory consolidation. However, it is unknown what features these repeated patterns encode, and whether they encode the past or the present or both. Thus, we asked which, if any, of the trial and inter-trial interval codes are reactivated in sleep, and thus whether they were recalled independently of each other.

We first tested whether population activity representations in trials reactivated more in post-training than pre-training sleep. For each feature of the task happening in the present (e.g choosing the left arm), we followed the decoding results by creating a population vector of the activity specific to that feature during a session's trials. To seek their appearance in slow-wave sleep, we computed population firing rate vectors in pre- and post-training slow-wave sleep in time bins of 1 second duration, and correlated each sleep vector with the feature-specific trial vector (Figure 4a). We thus obtained a distribution of correlations between the trial-vector and all pre-training sleep vectors, and a similar distribution between the trial-vector and all post-training sleep vectors. Greater correlation with post-training sleep activity would then be evidence of preferential reactivation of feature-specific activity in post-training sleep.

We examined reactivation separately between learning and Other sessions, seeking consistency with previous reports that reactivation of waking population activity in mPfC most clearly occurs immediately after rule acquisition (Peyrache et al., 2009; Singh et al., 2019). Figure 4b (upper panels) shows a clear example of a learning session with preferential reactivation. For all trial features, the distribution of correlations between the trial and post-training sleep population activity is right-shifted from the distribution for pre-training sleep. For example, the population activity vector for choosing the right arm is more correlated with activity vectors in post-training (Post-R) than pre-training (Pre-R) sleep.

Such post-training reactivation was not inevitable. In Figure 4b (lower panels), we plot another example in which the trial-activity vector equally correlates with population activity in pre- and post-training sleep. Even though specific pairs of features (such as the left and right light positions) differed in their overall correlation between sleep and trial activity, no feature shows preferential reactivation in post-training sleep.

These examples were recapitulated across the data (Figure 4c). In learning sessions, feature-specific activity vectors were consistently more correlated with activity in post- than pre-training sleep. By contrast, the other sessions showed no consistent preferential reactivation of any feature vector in post-training sleep. As a control for statistical artefacts in our reactivation analysis, we looked for differences in reactivation between paired features (e.g. left versus right arm choice) within the same sleep epoch and found these all centre on zero (Figure 4d). Thus, population representations of task features in the present were reactivated in sleep, and this consistently occurred after a learning session.

To check whether reactivation was unique to step-like learning, we turned to the Other sessions: there we found a wide distribution of preferential reactivation, from many about zero to a few reactivated nearly as strongly as in the learning sessions (Figure 4c, blue symbols). Indeed, when pooled with the learning sessions, we found reactivation of a feature vector in post-training sleep was correlated with the increase in accumulated reward during the session's trials (Fig 4e). Consequently, reactivation of population encoding during sleep may be directly linked to the preceding improvement in performance.

Prior reports suggest that the reactivation of activity patterns in sleep can be faster

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

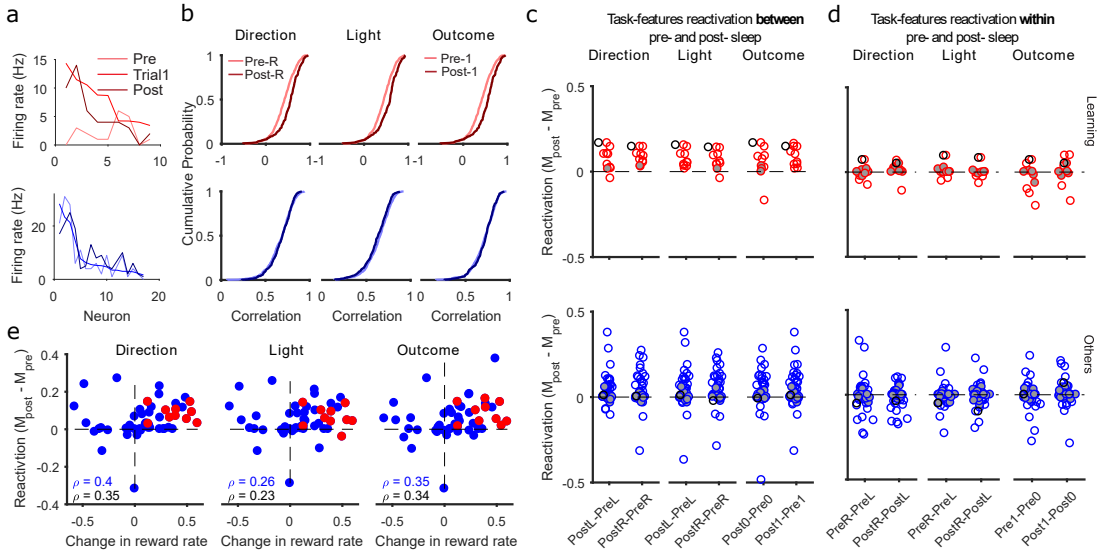

Figure 4: Reactivation of trial population coding in post-training sleep.
(**a**) Example population activity vectors. Upper panel: from one learning session, we plot the average firing rate vector for correct trials (Trial1). For comparison, we also plot examples of firing rate vectors from pre- and post-training slow-wave sleep (1s bins). Neurons are ranked in order of their firing rates in the trial vector. Lower panel: as for the upper panel, for an example session not classified as learning.
(**b**) Example distributions of Spearman's rank correlations between trial and sleep population activity. Upper panels: for the same learning session as panel (a), we plot the distributions of correlations between each vector of feature-specific trial activity and the population activity vectors in pre- and post-training slow-wave sleep. Lower panels: as for the upper panels, for the example non-learning session in panel (a). R: right arm; 1: rewarded trial.
(**c**) Summary of reactivations across all sessions. For each feature, we plot the difference between the medians of the pre- and post-training correlation distributions. A difference greater than zero indicates greater correlation between trials and post-training sleep. Each symbol is a session. Empty symbols are sessions with significantly different correlation distributions at $p < 0.05$ (Kolmogorov-Smirnov test). Grey filled symbols are not significantly different. One black circle for learning and one for non-learning sessions identify the two example sessions in panels (a) and (b).
(**d**) As for panel c, but plotting the median differences between distributions for paired features within the same sleep epoch. For example, in the left-most column, we plot the difference between the correlations with pre-session sleep activity for right-choice and left-choice specific trial vectors (PreR - PreL).
(**e**) Reactivation as a function of the change in reward rate in a session. One symbol per session: learning (red); Other (blue). ρ: Spearman's correlation coefficient. Black ρ is for all 49 sessions; blue ρ, using only sessions with any incremental improvement in performance ($N = 33$ in total, 10 learning and 23 Other sessions; see Methods). We plot here reactivation of vectors corresponding to left (direction and light) or correct; correlations for other vectors are similar in magnitude: 0.37 (choose right), 0.35 (cue on right), 0.2 (error trials) for all 49 sessions; 0.37 (choose left), 0.33 (cue on right) and 0.26 (error trials) for sessions with incremental improvement in performance.

or slower during sleep than they were during waking activity. We tested the time-scale dependence of feature-vector reactivation by varying the size of the bins used to create population vectors in sleep, with larger bins corresponding to slower reactivation. We found that preferential reactivation in post-training sleep in learning and (some) other sessions was robust over orders of magnitude of vector widths (Figure 5a). Notably, in the learning sessions only the vectors for rewarded outcome were significantly reactivated.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

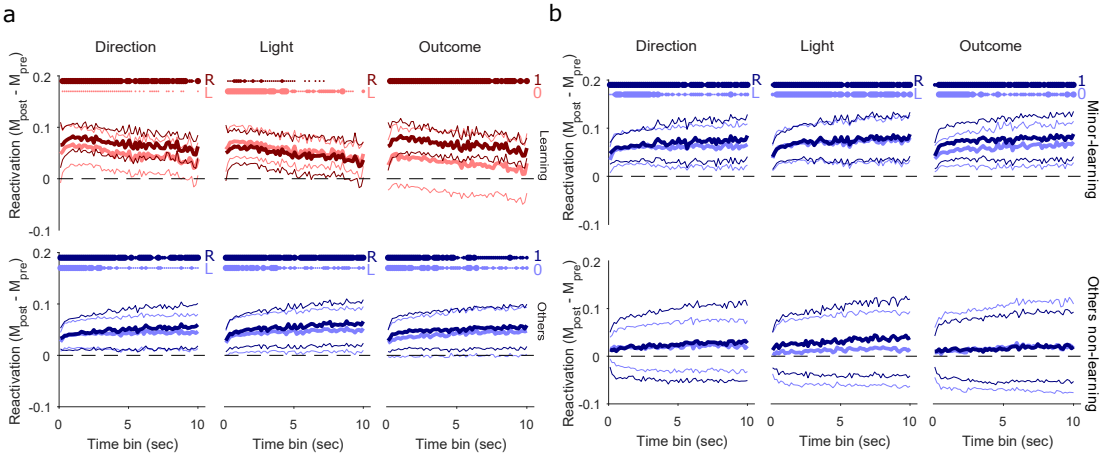

Figure 5: Robust reactivation of trial population coding across time-scales of sleep activity.
(**a**) At each time bin used to construct population activity vectors in sleep, we plot the distribution over sessions of the median differences between pre- and post-training correlation distributions, for learning (top), and other (bottom) sessions. Distributions are plotted as the mean (thick lines) $\pm$ 2 SEM (thin lines); at the 1s bin, these summarise the distributions shown in full in Figure 4c. Each panel plots two distributions, one per pair of features: lighter colours indicate left or error trials (L or 0); while darker colours indicate right or correct trials (R or 1). Time bins range from 100 ms to 10 s, tested every 150 ms. Dotted lines at the top of each panel indicate bins with reactivation significantly above zero (Wilcoxon sign rank test, $p < 0.05$ thin dot; $p < 0.01$ middle size dot; $p < 0.005$ thicker dots; $N = 10$ learning, $N = 39$ Other sessions).
(**b**) Here we divide the Other sessions from panel (a) into those showing any increment in performance from the animal ($N = 23$, "Minor-learning", see Methods) and those that did not ($N = 16$).

Moreover, among Other sessions, the reactivation in post-training sleep was significant only for those sessions in which the animal's performance improved (however slightly) within the session (Figure 5b). This consistency across broad time-scales suggests that it is the changes during trials to the relative excitability of neurons within the mPfC population that are carried forward into sleep (Singh et al., 2019).

No re-activation in sleep of inter-trial interval feature representations

To ask if this reactivation was unique to encoding of the present, we repeated the same reactivation analysis for population vectors from the inter-trial interval. Again, following our decoding results, each population feature vector was created from the average activity during inter-trial intervals after that feature (e.g. choose left) had occurred. We then checked for reactivation of this feature vector in pre- and post-training slow-wave sleep.

We found absent or weak preferential reactivation of population encoding in post-training sleep, for any feature in any type of session (Figure 6a). Consistent with this, we found no correlation between the change in performance during a session and the reactivation of feature vectors after a session (Figure 6b). The orthogonal population encoding during sessions (Figure 2) thus appears functional: population encoding of features in the present was reactivated in sleep, but encoding of the same features until the start of the next trial was not.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

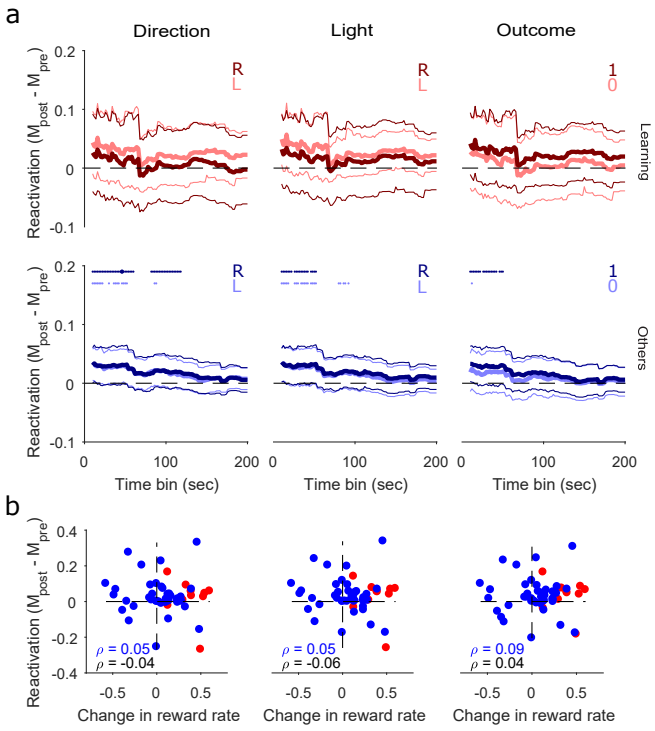

Figure 6: No preferential reactivation of population encoding of the past.
(**a**) Similar to Figure 5, for reactivation of population feature-vectors constructed from inter-trial interval activity. We plot the distribution over sessions of the median differences between pre- and post-training correlation distributions, for learning (top), and other (bottom) sessions. Note that the range of sleep vector time-bins is an order of magnitude larger than for trials, as the inter-trial intervals themselves are an order of magnitude longer than trials (Maggi et al., 2018). Dotted lines at the top indicate significant reactivation (Wilcoxon sign rank test, $p < 0.05$ thin dot; $p < 0.01$ middle size dot; $p < 0.005$ thicker dots). Lighter colours indicate left or error trials (L or 0); while darker colours indicate right or correct trials (R or 1)
(**b**) Similar to Figure 4e, reactivation of the inter-trial interval population vector as a function of the change in reward rate in a session. Reactivation is computed for 22 s bins. One symbol per session: learning (red); Other (blue). ρ: Spearman's correlation coefficient; black, all sessions; blue, only sessions with any incremental improvement in performance. We plot here reactivation of vectors corresponding to left (direction and light) or correct trials; correlations for other vectors are similar in magnitude: -0.004 (choose right), 0.02 (cue on right), -0.08 (error trials) for all sessions; -0.005 (choose right), 0.01 (cue on right) and -0.1 (error trials) for sessions with incremental improvement in performance.

Population synchrony is also independent between past and present

Our analyses thus far had shown evidence of independent coding of the past and the present in the same mPfC population. We then asked whether this independence of activity also extended to the independence of synchrony across the population. To do so, we characterised population synchrony as the matrix of pairwise similarities between neurons on each trial (Appendix - Figure A.1). By comparing between trials within a session, we found that the population's synchrony was consistently more similar between trials with upcoming rewards than errors during learning sessions (Appendix - Figure A.2). We previously showed that a population's synchrony was also consistently more similar between inter-trial intervals after a reward (Maggi et al., 2018). A natural question was then whether this repeated pattern of synchrony was carried forward from the trials to the inter-trial intervals during learning. However, we found that the patterns of population

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

synchrony in a trial and its following inter-trial interval were as different from each other as shuffled data (Appendix - Figure A.4a). Moreover, we found that the pattern of population synchrony in the trials was recapitulated more in post than pre-session sleep, but the population synchrony in the inter-trial intervals was not (Appendix - Figure A.4c-f). The synchrony between neurons in the population thus seems as independent between the task periods as the population encoding of past and present.

Discussion

We have shown that medial PfC population activity independently represents the past and present of the same task features. Three lines of evidence support this conclusion. First, the same task feature, such as the choice of arm, is independently encoded in the trials and the inter-trial intervals. Second, these independent encodings are also functionally independent: population activity representations of features during the trials are re-activated in post-training sleep, but inter-trial interval representations are not. Third, the reward-evoked pattern of population synchrony is also independent between trials and inter-trial intervals, and only the trials' synchrony pattern is recapitulated in post-training sleep.

Mixed population coding in mPfC

Consistent with prior reports of mixed or multiplexed coding by single mPfC neurons (Jung et al., 1998; Horst and Laubach, 2012; Rigotti et al., 2013; Fusi et al., 2016), we found that small mPfC populations can sustain mixed encoding of two or more of the current direction choice, light position, and anticipated outcome. These encodings were also position-dependent. Encoding of direction choice reliably occured from the maze's choice point onwards, but it is unclear whether this represents a causal role in the choice itself, or an ongoing representation of a choice being made.

We also report that these mixed encodings of the present within each population reactivate in post-training sleep. This finding goes beyond prior reports that specific patterns of trial activity reactivate in sleep (Euston et al., 2007; Peyrache et al., 2009; Singh et al., 2019) to show what those patterns were encoding – multiple features of the present, but not the past. It seems mixed encoding is a feature of sleep too.

As we showed in (Maggi et al., 2018) and extended here, population activity during the inter-trial interval also has mixed encoding of features of the past. This encoding decays in strength as the rat gets closer to the start of the maze: we showed here this mixed encoding is still evident even if we disregard maze positions. Collectively, our results show that population activity in mPfC can switch from mixed encoding of the present in a trial to mixed encoding of the past in the following inter-trial interval.

Independent population codes solve interference of past and present

There are multiple hypotheses for how this transition from coding the present to the past could happen. One hypothesis is that there are groups of neurons separately dedicated to encoding the past and present. We ruled out this idea by only decoding from neurons active in every trial and inter-trial interval, so showing that the transition from present to past happened within the same group. Moreover, by showing that we could decode each task feature from population activity even in the absence of any single neuron tuned to that feature, we ruled out that there were distinct groups dedicated to each feature. Thus, this mixed encoding and its transition from present to past were sustained across

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

small mPfC populations, at most 22 neurons, rather than by distinct groups within each population.

Another hypothesis, as we noted in the Results, is that the switching from a population encoding of the present to encoding of the past is explained by population activity in the trials being carried forward into the inter-trial interval, whether by persistent activity acting as a memory trace, or by the recall of patterns of trial activity during the inter-trial interval. But our demonstration of independent encoding in the population between trials and the following intervals rules out this hypothesis.

Our results support dynamic coding in mPfC: population encoding evolved within both the trials and the inter-trial intervals, consistent with the underlying changes we observed in the population activity. The evolution of population dynamics over the inter-trial interval is consistent with reports of dynamic changes of PfC activity during the delay period of working memory tasks (Murray et al., 2017; Spaak et al., 2017; Cavanagh et al., 2018; Wasmuht et al., 2018). They thus support the hypothesis that working memory is sustained by population activity rather than the persistent activity of single neurons (Constantinidis et al., 2018; Lundqvist et al., 2018). Crucially, the evolution of activity within trials and inter-trial intervals was continuous, with adjacent maze sections containing more similar population activity, yet the transition from the trial to the inter-trial interval was discontinuous, with both encoding and population activity becoming orthogonal. Our results thus show that evolution of encoding of the present and past were along two independent axes.

Any neural population encoding both the past and the present in its activity faces problems of interference: of how to prevent the addition of new information in the present from overwriting the encoded information of the short-term past (Libby and Buschman, 2019); of how inputs to the population can selectively recall only the past or the present, but not both; and of how downstream populations can access or distinguish the encodings of the past and the present. Representing the present and past on independent axes solves these problems. It means that the encoding of the present can be updated without altering the encoding of the past, that inputs to the population can activate either the past or the present representations independently, and that downstream populations can distinguish the two by being tuned to read-out from one axis or the other. Indeed, we showed that in post-session sleep the encoding of the present can be accessed independently of the encoding of the past.

An open question is how much the clean independence between the encoding of the past and present depends on the behavioural task. In the Y-maze task design, there is a qualitative distinction between trials (with a forced choice) and inter-trial intervals (with a self-paced return to the start arm), which we used to clearly distinguish encoding of the present and the past. Such independent coding may be harder to uncover in tasks without a distinct separation of decision and non-decision phases. For example, tasks where the future choice of arm depends on recent history, such as double-ended T-mazes (Jones and Wilson, 2005), multi-arm sequence mazes (Poucet et al., 1991), or delayed non-match to place (Spellman et al., 2015), blur the separation of the present and the past. Comparing population-level decoding of the past and present in such tasks would give useful insights into when the two are, and are not, independently coded.

Mechanisms for rapid switching of population codes

The independent encoding and independent population activity between the trial and immediately following inter-trial interval implies a rapid rotation of population activity. How might such a rapid switch of network-wide activity be achieved?

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

Such rapid switching in the state of a network suggests a switch in the driver inputs to the network. In this model, drive from one source input creates the network states for population encoding A; a change of drive – from another source, or a qualitative change from the same source -— creates the network states for population encoding B (either set of states may of course arise solely from internal dynamics). One option for a switching drive is the hippocampal-prefrontal pathway.

In several spatial tasks (Jones and Wilson, 2005; Jadhav et al., 2012; Peyrache et al., 2009; Benchenane et al., 2010) cortico-hippocapal regions show coherent theta oscillation. This coherence increases during learning at the choice point of the Y maze task (Benchenane et al., 2010; Peyrache et al., 2009) described here, and recurs during slow-wave ripples in post-training sleep. These data and our analyses here are consistent with the population encoding of the trials being (partly) driven by hippocampal input, and with the re-activation of only the trial representations in sleep being the recruitment of those states by hippocampal input during slow-wave sleep. The increased coherence between hippocampus and mPfC activity may act as a window for synaptic plasticity of that pathway (Benchenane et al., 2010, 2011). Consistent with this, we saw a correlation between performance improvement in trials and reactivation in sleep (see also Maingret et al., 2016).

All of which suggests the encoding of the past during the inter-trial interval is not driven by the hippocampal input to mPfC, as its representation is not re-activated in sleep. (Spellman et al. 2015 report hippocampal input to mPfC is necessary for the maintenance of a cue location; though unlike in our task, actively maintaining the location of this cue was necessary for a later direction decision). Rather, the population coding during the inter-trial interval could reflect the internal dynamics of the mPfC circuit. Indeed, network models of working memory in the prefrontal cortex focus on attractor states created by its local network (Durstewitz et al., 2000; Miller et al., 2005; A. Compte, 2000; Wimmer et al., 2014). If somewhere close to the truth, this account of rapid switching suggests that the hippocampal input to mPfC drives population activity in the trial, and a change or reduction in that input allows the mPfC local circuits to create a different internal state during the inter-trial interval. A prediction of this account is that perturbation of the hippocampal input to the mPfC could disrupt its encoding of the past and present in different ways.

Reconciling mPfC roles in memory and choice

We propose that our combined results here and previously (Maggi et al., 2018) support a dual-function model of mPfC population coding, where the independent coding of the past and present respectively support on-line learning and consolidation. This model is somewhat counter-intuitive: our data suggest the representation of the present in mPfC is used for offline learning, whereas the representation of the past is used online to guide behaviour.

Under this model, the role of memory encoding in the inter-trial interval is to guide learning online: reward tags past features whose conjunction led to successful outcomes (for example, the conjunction of turning left when the light is on in the left arm). While population activity in the inter-trial interval reliably encodes features of the past throughout training, we previously showed that synchrony of the population only consistently occurs immediately before learning (Maggi et al., 2018). This suggests that the synchronisation of mPfC representations of features predicting success is correlated with successful rule-learning. Consistent with such past-encoding contributing to online learning, we show here that neither the encoding nor synchrony pattern in the inter-trial interval are carried

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

forward long-term into sleep.

By contrast, we report here representations of the present in the trial are carried forward and reactivated in sleep. Reactivation of waking activity during slow-wave sleep has been repeatedly linked to the consolidation of memories (Stickgold, 2005; Tononi and Cirelli, 2014; Sawangjit et al., 2018). Indeed, interrupting the re-activation of putative waking activity in hippocampus impairs task learning (Girardeau et al., 2009). Thus, under the dual-function model, we propose the reactivation in mPfC of mixed encodings of the present may be consolidating the conjunction of present features and choice that is going to be successful when re-used in future.

Further insight into these and other ideas here would come from stable recordings of the same population across multiple sessions, to track how encoding of the past and present evolves and is or is not reused. In particular, it would be insightful to establish if re-activated trial representations in sleep reappear in subsequent sessions.

The medial prefrontal cortex plays a key role in both short-term memory (Fujisawa et al., 2008; Jun et al., 2010; Maggi et al., 2018) and choice behaviour (Averbeck et al., 2006; Rigotti et al., 2013; Erlich et al., 2015; Hanks et al., 2015). Our finding here of independent coding of the past and the present suggest these roles can be carried out sequentially within the same mPfC neural population.

Acknowledgments

We thank Adrien Peyrache for the data, discussions, and comments on early drafts of this manuscript, Hazem Toutounji and Martin O'Neill for comments on drafts, and the Humphries' lab past and present (Abhinav Singh, Javier Caballero, Mat Evans, Francois Cinotti, Tomas Fiers) for discussion. This work was supported by the Medical Research Council [grant numbers MR/J008648/1 and MR/P005659/1]. The original data collection was supported by the EU Framework (FP6) "ICEA" grant.

Author Contributions

M.D.H and S.M. designed the analyses. S.M. analysed the data. M.D.H and S.M. wrote the manuscript.

Declaration of Interest

The authors declare no conflicts of interest.

References

A. Compte, N. Brunel, P. G.-R. . X. W. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex*, 10:910–923.

Averbeck, B. B. and Lee, D. (2007). Prefrontal neural correlates of memory for sequences. *J Neurosci*, 27:2204–2211.

Averbeck, B. B., Sohn, J.-W., and Lee, D. (2006). Activity in prefrontal cortex during dynamic selection of action sequences. *Nat Neurosci*, 9:276–282.

Baeg, E. H., Kim, Y. B., Huh, K., Mook-Jung, I., Kim, H. T., and Jung, M. W. (2003). Dynamics of population code for working memory in the prefrontal cortex. *Neuron*, 40(1):177–188.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

Barraclough, D. J., Conroy, M. L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature neuroscience*, 7:404–410.

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., and Wiener, S. I. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron*, 66(6):921–936.

Benchenane, K., Tiesinga, P. H., and Battaglia, F. P. (2011). Oscillations in the prefrontal cortex: a gateway to memory and attention. *Curr Opin Neurobiol*, 21(3):475–485.

Cavanagh, S. E., Towers, J. P., Wallis, J. D., Hunt, L. T., and Kennerley, S. W. (2018). Reconciling persistent and dynamic hypotheses of working memory coding in prefrontal cortex. *Nature communications*, 9:3498.

Constantinidis, C., Funahashi, S., Lee, D., Murray, J. D., Qi, X.-L., Wang, M., and Arnsten, A. F. T. (2018). Persistent spiking activity underlies working memory. *Journal of neuroscience*, 38:7020–7028.

Durstewitz, D., Seamans, J. K., and Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neurosience*, 3:1184–1191.

Durstewitz, D., Vittoz, N. M., Floresco, S. B., and Seamans, J. K. (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron*, 66(3):438–448.

Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D., and Brody, C. D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 4.

Euston, D. R., Gruber, A. J., and McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, 76(6):1057–1070.

Euston, D. R., Tatsuno, M., and McNaughton, B. L. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science*, 318(5853):1147–1150.

Fujisawa, S., Amarasingham, A., Harrison, M. T., and Buzsaki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat Neurosci*, 11(7):823–833.

Fusi, S., Miller, E. K., and Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. *Curr Opin Neurobiol*, 37:66–74.

Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., and Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience*, 12:1222–1223.

Guise, K. G. and Shapiro, M. L. (2017). Medial prefrontal cortex reduces memory interference by modifying hippocampal encoding. *Neuron*, 94:183–192.e8.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520:220–223.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, Berlin.

Horst, N. K. and Laubach, M. (2012). Working with memory: evidence for a role for the medial prefrontal cortex in performance monitoring during spatial delayed alternation. *Journal of neurophysiology*, 108(12):3276–3288.

Ito, H. T., Zhang, S.-J., Witter, M. P., Moser, E. I., and Moser, M.-B. (2015). A prefrontal–thalamo–hippocampal circuit for goal-directed spatial navigation. *Nature*, 522(7554):50.

Jadhav, S. P., Kemere, C., German, P. W., and Frank, L. M. (2012). Awake hippocampal sharp-wave ripples support spatial memory. *Science*, 336(6087):1454–1458.

Jones, M. W. and Wilson, M. A. (2005). Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. *PLoS Biol*, 3(12):e402.

Jun, J. K., Miller, P., Hernández, A., Zainos, A., Lemus, L., Brody, C. D., and Romo, R. (2010). Heterogenous population coding of a short-term memory and decision task. *J Neurosci*, 30(3):916–929.

Jung, M. W., Qin, Y., McNaughton, B. L., and Barnes, C. A. (1998). Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. *Cereb Cortex*, 8(5):437–450.

Karlsson, M. P., Tervo, D. G., and Karpova, A. Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science*, 338(6103):135–139.

Laskowski, C. S., Williams, R. J., Martens, K. M., Gruber, A. J., Fisher, K. G., and Euston, D. R. (2016). The role of the medial prefrontal cortex in updating reward value and avoiding perseveration. *Behavioural Brain Research*, 306:52–63.

Laubach, M., Caetano, M. S., and Narayanan, N. S. (2015). Mistakes were made: neural mechanisms for the adaptive control of action initiation by the medial prefrontal cortex. *Journal of Physiology-Paris*, 109(1):104–117.

Libby, A. and Buschman, T. J. (2019). Rotational dynamics reduce interference between sensory and memory representations. *bioRxiv*, page 641159.

Lundqvist, M., Herman, P., and Miller, E. K. (2018). Working memory: Delay activity, yes! persistent activity? maybe not. *Journal of neuroscience*, 38:7013–7019.

Machens, C. K., Romo, R., and Brody, C. D. (2010). Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. *J Neurosci*, 30:350–360.

Maggi, S., Peyrache, A., and Humphries, M. D. (2018). An ensemble code in medial prefrontal cortex links prior events to outcomes during learning. *Nature Communications*, 9.

Maingret, N., Girardeau, G., Todorova, R., Goutierre, M., and Zugaro, M. (2016). Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nature Neuroscience*, 19:959–964.

Miller, E. K. and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24(1):167–202.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

Miller, P., Brody, C. D., Romo, R., and Wang, X.-J. (2005). A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb Cortex*, 15:679–679.

Murray, J. D., Bernacchia, A., Roy, N. A., Constantinidis, C., Romo, R., and Wang, X.-J. (2017). Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 114:394–399.

Narayanan, N. S. and Laubach, M. (2008). Neuronal correlates of post-error slowing in the rat dorsomedial prefrontal cortex. *Journal of neurophysiology*, 100(1):520–525.

Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I., and Battaglia, F. P. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nature Neuroscience*, 12:919–926.

Poucet, B., Lucchessi, H., and Thinus-Blanc, C. (1991). What information is used by rats to update choices in the radial-arm maze? *Behav Processes*, 25:15–26.

Powell, N. J. and Redish, A. D. (2016). Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behaviour. *Nature Communications*, 7.

Rich, E. L. and Shapiro, M. (2009). Rat prefrontal cortical neurons selectively code strategy switches. *Journal of Neuroscience*, 29(22):7208–7219.

Rich, E. L. and Shapiro, M. L. (2007). Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *J Neurosci*, 27(17):4747–4755.

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–590.

Russo, E., Ma, T., Spanagel, R., Durstewitz, D., Toutounji, H., and Köhr, G. (2020). Coordinated prefrontal state transition leads extinction of reward-seeking behaviors. *bioRxiv*.

Sawangjit, A., Oyanedel, C. N., Niethard, N., Salazar, C., Born, J., and Inostroza, M. (2018). The hippocampus is crucial for forming non-hippocampal long-term memory during sleep. *Nature*, 564(7734):109.

Schuck, N., Gaschler, R., Wenke, D., Heinzle, J., Frensch, P., Haynes, J.-D., and Reverberi, C. (2015). Medial prefrontal cortex predicts internally driven strategy shifts. *Neuron*, 86:331–340.

Siegel, M., Buschman, T. J., and Miller, E. K. (2015). Cortical information flow during flexible sensorimotor decisions. *Science*, 348(6241):1352–1355.

Singh, A., Peyrache, A., and Humphries, M. D. (2019). Medial prefrontal cortex population activity is plastic irrespective of learning. *Journal of Neuroscience*, pages 1370–17.

Spaak, E., Watanabe, K., Funahashi, S., and Stokes, M. G. (2017). Stable and dynamic coding for working memory in primate prefrontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 37:6503–6516.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

Spellman, T., Rigotti, M., Ahmari, S. E., Fusi, S., Gogos, J. A., and Gordon, J. A. (2015). Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature*, 522(7556):309–314.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, 437(7063):1272.

Sul, J. H., Kim, H., Huh, N., Lee, D., and Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, 66(3):449–460.

Tononi, G. and Cirelli, C. (2014). Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron*, 81(1):12–34.

Wasmuht, D. F., Spaak, E., Buschman, T. J., Miller, E. K., and Stokes, M. G. (2018). Intrinsic neuronal dynamics predict distinct functional roles during working memory. *Nature communications*, 9:3499.

Wimmer, K., Nykamp, D. Q., Constantinidis, C., and Compte, A. (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience*, 17:431–439.

Young, J. J. and Shapiro, M. L. (2009). Double dissociation and hierarchical organization of strategy switches and reversals in the rat pfc. *Behavioral Neuroscience*, 123:1028–1035.

Methods

Task description and electrophysiological data

All the data in this study comes from previously published data (Peyrache et al., 2009). The full details of training, spike-sorting and histology can be found in (Peyrache et al., 2009). The experiments were carried out in accordance with institutional (CNRS Comité Opérationnel pour l'Ethique dans les Sciences de la Vie) and international (US National Institute of Health guidelines) standards and legal regulations (Certificate no. 7186, French Ministère de l'Agriculture et de la Pêche) regarding the use and care of animals.

Four Long-Evans male rats were implanted with tetrodes in prelimbic cortex and trained on a Y-maze task (Figure 1a). During each session, prelimbic activity was recorded for 20-30 minutes of sleep or rest epoch before the training phase, in which rats worked at the task for 20-40 minutes. After that, another 20-30 minutes of sleep or rest epoch recording followed. During the sleep epochs, intervals of slow-wave sleep were identified offline from the local field potential (details in (Peyrache et al., 2009; Benchenane et al., 2010)).

The Y-maze had symmetrical arms, 85 cm long, 8 cm wide, and separated by 120 degrees, connected to a central circular platform (denoted as the choice point throughout). Each rat worked at the task phase by self-initiating the trial, leaving the beginning of the start arm. Trial finished when the rat reached the end of the chosen goal arm. If the chosen arm was correct according to the current rule, the rat was rewarded with drops of flavoured milk. As soon as the animal reached the end of the chosen arm an inter-trial interval started and lasted until the rat completed its self-paced return to the beginning of the start arm.

Each rat was exposed to the task completely naïve and had to learn the rule by trial-and-error. The rules were presented in sequence: go to the right arm; go to the cued arm;

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

go to the left arm; go to the uncued arm. The light cues at the end of the two arms were lit in a pseudo-random sequence across trials, regardless the rule in place.

The recording sessions taken from the study of Peyrache and colleagues (Peyrache et al., 2009) were 53 in total. Each of the four rats learnt at least two rules, and they contributed with 14, 14, 11, and 14 sessions. The learning, rule change, and other sessions for each rat were intermingled. We used 49 sessions for most of the analysis. One session was omitted for missing position data, one for consistent choice of the right arm (in a dark arm rule) preventing decoder analyses (see below), and one for missing spike data in a few trials. An additional session was excluded for having only two neurons firing in all trials. Tetrode recordings were spike-sorted within each recording session. In the sessions we analysed here, the populations ranged in size from 4-25 units. Spikes were recorded with a resolution of 0.1 ms. Simultaneous tracking of the rat's position was recorded at 30 Hz.

Behavioural analysis

Each session was classified according to its behavioural features. The learning sessions were identified according to the original study (Peyrache et al., 2009) as the ones with three consecutive correct trials followed by a performance of at least 80% correct. The first of the three correct trials was the learning trial. Only ten sessions satisfied this criterion. We quantified this learning as a step-like change in performance by fitting a robust regression line to the cumulative reward curve before and after the learning trial. The slopes of the two lines gave us the rate of reward accumulation before (r_{before}) and after (r_{after}) the learning trial.

Eight sessions were characterised by 10 consecutive correct trials or eleven correct out of twelve trials followed by a change in the rule. The first trial with the new rule was identified as the rule change trial. The change in performance in these sessions, labelled "Rule change" sessions, was quantified with the same method above. A robust regression line was fitted to the cumulative reward curve before and after the rule change trial.

To identify other possible learning session, we fitted this piece-wise linear regression model to each trial in turn (allowing a minimum of 5 trials before and after each tested trial). We then found the trial at which the increase in slope ($r_{after} - r_{before}$) was maximised, indicating the point of steepest inflection in the cumulative reward curve. In the learning sessions, the learning trial largely agreed with this method. Amongst the remaining sessions, labelled "Other", we searched for signs of incremental learning using this method. We found 22 sessions falling in this category in addition to the 10 learning sessions. We called these sessions "minor-learning".

Decoder analysis and independence of ensembles encoding

To predict which task feature was encoded into the ensemble we trained and tested a range of linear decoders (Hastie et al., 2009; Maggi et al., 2018). In the main text we report the results obtained using a logistic regression classifier, but for robustness we also tested three other decoders: linear discriminant analysis; (linear) support vector machines; and a nearest neighbours classifier. The full details of the decoding analysis can be found in Maggi et al. (2018).

Briefly, we linearised the maze in five equally-sized sections and we computed the firing rate vector of the core population of length N for each position p, $\mathbf{r^p}$. For each trial $t = 1, \ldots, T$ and each section of the maze $p = 1, \ldots, 5$, the set of population firing rate vectors $\{\mathbf{r^p}(1), \ldots, \mathbf{r^p}(T)\}$ was used to train the decoder. Each relevant trials' task

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

information was binary labelled: outcome (labels: 0, 1), the chosen arm (labels: left, right) and the position of the light cue (labels: left, right). The same classifier was also trained to decode previous trials' task information. We used leave-one-out cross-validation and we quantified the accuracy of the decoder as the proportion of correctly predicted labels over the T held out trial intervals.

For each rat and each session, the distribution of outcomes and arm choices depended on the rats' performance, which could differ from 50%. Therefore, we trained and cross-validated the same classifier on the same data-sets, but shuffling the labels of the task features. In this way we obtained the accuracy of detecting the right labels by chance. We repeated the shuffling and fitting 50 times and we averaged the accuracy across the 50 repetitions. The results displayed in the figures showed the difference between the decoder accuracy on the original data with the accuracy of the shuffled label.

To test if the classifier results were a property of the population or were driven by a few tuned neurons, we selected the sessions in which no neurons significantly changed their firing rate according to a feature. For each neuron, we computed its firing rate on each trial $(1 \dots T)$ of a session $\{r(1), \dots, r(T)\}$. We divided the firing rate vector into two classes (right vs left direction choice, right vs left light position and rewarded vs unrewarded) for each feature. We then tested whether the firing rate was significantly changing within the feature (Kolmogorov-Smirnov test) using $p < 0.05$ as our criterion for a "tuned" neuron. Finally, for every feature, we excluded all the sessions that had at least one neuron that significantly changed its firing rate, and we ran the classifier with the remaining sessions.

To compare the decoding accuracy between trials and inter-trial intervals, we trained again the classifier using the population firing rate vectors computed on the entire maze $\{\mathbf{r}(1), \dots, \mathbf{r}(T)\}$. We then trained the classifier on all the trials. We saved the population vector of weights and we tested the model, optimised to decode trial activity, on every inter-trial interval to evaluate the accuracy in decoding retrospective inter-trial interval labels. The same procedure was used to train the linear classifier on all the inter-trial intervals to test its accuracy in decoding trials activity. The population vector of weight was also saved for this model.

The angle, θ, between the population vector of trials', w_t, and inter-trial intervals', w_I, weights was computed as $\theta = \cos^{-1}\left(\frac{w_t \cdot w_I}{\|w_t\|\|w_I\|}\right)$.

We further evaluated the independence of trial and inter-trial interval population vectors by quantifying their separability in a low dimensional space. We used principal components analysis (PCA) to project the population vectors of a session onto a common set of dimensions. To do so, we constructed the data matrix $\mathbf{X}$ from the firing rate vectors of the core population, by concatenating trials and inter-trial intervals in their temporal order $\{\mathbf{r}_t(1), \mathbf{r}_I(1), \dots, \mathbf{r}_t(T), \mathbf{r}_I(T)\}^{\mathrm{T}}$; the resulting matrix thus had dimensions of $2T$ rows and N (neurons) columns. Applying PCA to X, we projected the firing rate vectors on to the top d principal axes (eigenvectors of $\mathbf{X^T X}$) to create the top d principal components. For each set of d components, we quantified the separation between the projected trial and inter-trial interval population vectors using a linear classifier (Support Vector Machine, SVM), and report the proportion of misclassified vectors. We repeated this for between $d = 1$ and $d = 4$ axes for each session.

Reactivation of task-feature encoding in sleep

In order to quantify the reactivation of waking activity in pre- and post-session sleep, we used the population firing rate vectors computed for the decoder. We considered here the average population vector computed across all the trials in the all maze for each feature.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

In details, we quantified separately the average population firing rate vector for all the right choice trials and all the left choice trials. Similarly we did for right and left cue location and for correct and error trials. We then compare the ranked average population firing rate vector for each feature with the firing rate vector of each 1 second time bin of slow-wave sleep pre- and post-session. We used Spearman's correlation coefficient to compare them and to quantify the difference between the distributions of each feature and the slow-wave sleep pre- and post-session. Spearman's coefficient was chosen specifically to remove any effects of global rate variations across the vectors within or between epochs. In order to have a reactivation of activity in post-session sleep, we expected the distribution of Spearman CC between a feature and pre-session slow-wave sleep to be leftward shifted compare to the distribution of Spearman CC between the same feature and post-session slow-wave sleep. We quantified this shift by measuring the difference in the medians distribution between the two Spearman distributions. If the delta medians was positive then we had a higher correlation of the population firing vector with the post-session slow-wave sleep compared to the pre-session slow-wave sleep. If the delta median (called Reactivation ($M_{post} - M_{pre}$) in the text) was negative, then the population firing rate vector was more similar to the pre-session slow-wave sleep population vector. To then control for different time scales of reactivation in sleep we repeated the same procedure changing the time bin in the slow-wave sleep pre- and post-session. We used time bins from 100 ms to 10 sec every 150 ms for trials and from 10 sec to 200 sec every 2 sec for inter-trial intervals.

Testing for reinforcement-driven ensembles

The reinforcement-dependent recall of ensemble activity was identified as per Maggi et al. (2018). We firstly selected the spike-trains of the N neurons active in every trials. We convolved these spike-trains with a Gaussian ($\sigma = 100$ ms) to obtain a spike-density function f_k for the kth spike-train. All the recall analyses were repeated for different Gaussian widths ranging from 20 ms to 240 ms (Figure A.3). Each spike-train was then z-scored to obtain a normalised spike-density function f^* of unit variance: $f_k^* = (f_k - \langle f_k \rangle)/\sigma_k$, where $\langle f_k \rangle$ is the mean of f_k, and σ_k its standard deviation, taken over all the trials of a session.

We then wanted to track the changes in the co-activity pattern of the core population along the sessions. We first computed the pairwise similarity matrix between the spike density functions of each neuron in trial t, $\mathbf{S_t}$. This matrix was rectified in order to keep track only of those pair with positive co-activity pattern (Figure A.1a). To compare the co-activity patterns across trials along the sessions, we then computed the recall matrix (Figure A.1b, $\mathbf{R}$, where each entry $R(t, u)$) is the rectified correlation coefficient between the similarity matrices $\mathbf{S_t}$ and $\mathbf{S_u}$.

We grouped the entries of $\mathbf{R}$ into two groups according to the outcome of the trials. In such way we obtained two block diagonals $\mathbf{R_1}$ and $\mathbf{R_2}$ (such as $\mathbf{R_{error}}$ and $\mathbf{R_{correct}}$, as illustrated in Figure A.1b-c). We summarised the recall between groups by computing the mean of each block.

As a control and to keep comparison with our prior study of inter-trial intervals (Maggi et al., 2018), we defined a null model to dissect the contribution of differences between the error and correct trials. For each session we created a predicted recall matrix $\hat{\mathbf{R}}$, by averaging 1000 random recall matrices, each computed from shuffled spike trains. Each spike-train was shuffled by randomly re-ordering its inter-spike intervals. This shuffling was meant to destroy any specific temporal pattern of the spike train, allowing to quantify the pairwise similarity contribution due exclusively to the duration of trial. Our final residual

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

recall matrix $\tilde{\mathbf{R}} = \mathbf{R} - \hat{\mathbf{R}}$ was obtained as the difference between the Recall matrix and the average shuffled recall matrix (Figure A.1c). We report all results for this residual matrix $\tilde{\mathbf{R}}$, though differences between using $\mathbf{R}$ and $\tilde{\mathbf{R}}$ were minimal (Figure A.1d).

Data Availability

The spike-train and behavioural data that support the findings of this study are available in CRCNS.org (DOI: 10.6080/K0KH0KH5) (ref. (Peyrache et al., 2009)). Code to reproduce the main results of the paper is available at:

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

A Appendix - independence of population synchrony

Population synchrony is recalled in correct trials around learning

To further support the evidence of independent coding of the past and the present in the same mPfC population we investigated whether this independence also extended to the independence of synchrony across the population. We previously showed that these mPfC populations repeated similar patterns of population synchrony across inter-trial intervals, most reliably around the point of learning (Maggi et al., 2018). As this synchrony was triggered by reward, we suggested that reward acts as a trigger to synchronise the features of the past trial encoded within the population; in principle, this would be a powerful tool for learning, as the joint presence of those features is predictive of reward. We thus asked if these populations have such a synchrony code in the trials too. We show here that trials also have evidence of outcome-dependent synchrony in the population; in the next section we ask if these synchrony patterns are also reactivated in sleep.

We characterised population synchrony as the matrix of pairwise similarities between neurons on each trial (Figure A.1a). We then compared these similarity matrices across trials to ask if the same pattern of population synchrony is recalled on different trials. We first asked if the same pattern of synchrony was recalled on correct trials or on error trials (Figure A.1b). Other factors differing between error and correct trials could contribute to differences in correlations between neurons (though time spent along the maze was not one of them – Figure A.1e). To control for spurious correlations, we constructed a predicted Recall matrix from label-shuffled data, and subtracted this to leave the residual Recall matrix (Figure A.1c). This correction did not markedly change the differences between correct and error trials in the recall of a pattern of population synchrony (Figure A.1d).

In learning sessions, the population's synchrony pattern was consistently more similar between trials with upcoming rewards than errors (Figure A.2a), suggesting a recalled pattern of synchrony across the population precedes correct choice on a trial.

This consistent recall of a synchrony pattern across trials was specific to the combination of learning sessions and future outcomes. We found no consistent outcome-dependent recall of a synchrony pattern in Other sessions (Figure A.2a), or in trials that were immediately after a correct trial (Figure A.2b), ruling out a history-dependent effect of reward. We further observed no consistently recall of a synchrony pattern across trials conditioned on other task-related features, whether they were upcoming (Figure A.2c,d) or in previous trials (Figure A.2e,f). Finally, to ensure that the recall of a synchrony pattern was not affected by the temporal precision of the spike-train correlation used, we repeated all these comparisons for different resolutions of the Gaussian width used to convolve spike-trains before computing their pair-wise similarity. Across an order of magnitude for the temporal resolution, we still observed recall effects only for learning sessions and only when preceding the outcome (Figure A.3).

Population synchrony is also independent between trials and inter-trial intervals

Both trials (present paper) and inter-trial intervals (Maggi et al., 2018) showed outcome-dependent recall of population synchrony in mPfC. As we had established that the population codes in trials and inter-trial intervals were independent in both waking and sleep, we asked if the synchrony patterns recalled during learning were also independent between trials and inter-trial intervals.

We found that, within a learning session, the patterns of population synchrony in a

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

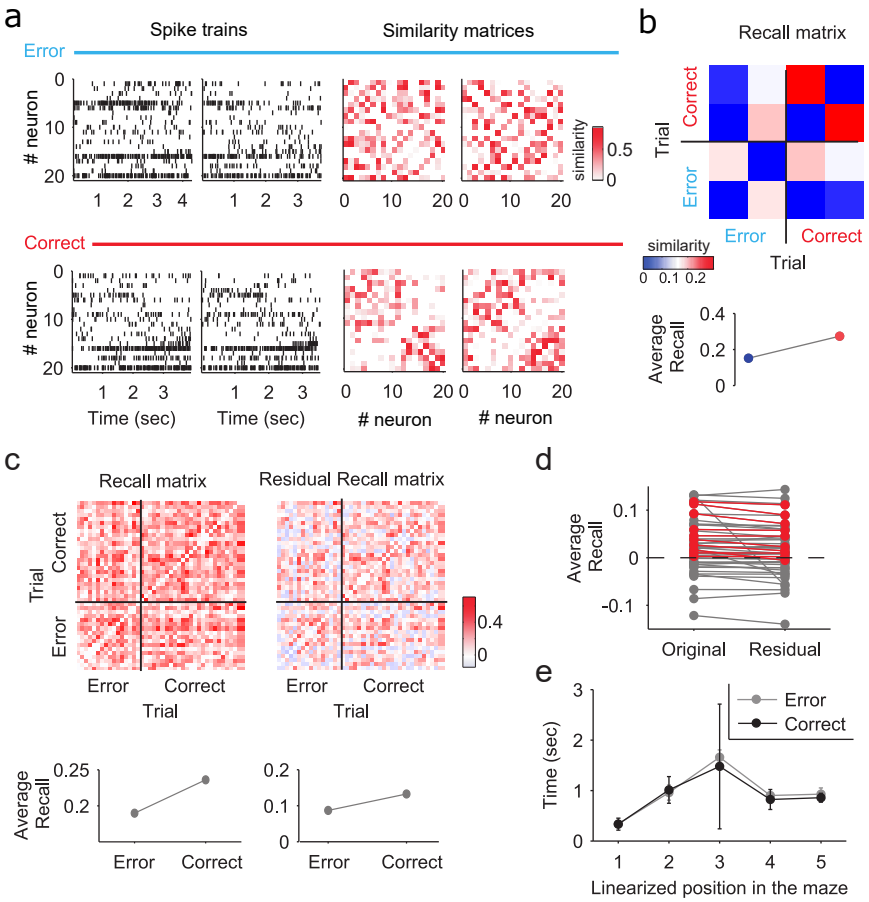

Figure A.1: Comparing population synchrony between trials.
(**a**) Example raster plots of spike trains during two error and two correct trials from the same session (left panels). Right panels: corresponding pairwise similarity matrices.
(**b**) Top, matrix of similarities between the example similarity matrices in panel (a). This "Recall" matrix shows higher values within correct trials compared to error trials, indicating that population synchrony was more similar between correct trials. Bottom: we summarise this by comparing the average recall values between error (blue) and correct (red) trials.
(**c**) To control for spurious correlations, we computed an expected Recall matrix from shuffled data, and subtracted it from the Recall matrix, to give a residual Recall matrix. Example Recall (left panel) and residual Recall (right panel) matrices for one learning session, reordered according to the trial outcome. Bottom panels show the average recall value among error and correct trials.
(**d**) The difference between average correct and error recall is compared between Recall (original) and residual Recall for all sessions. Each line is a session. Red lines are the learning sessions.
(**e**) Time spent along the maze for error and correct trials did not differ (median ± SEM).

trial and its following inter-trial interval were as different from each other as shuffled data (Figure A.4a, redrawn from Maggi et al. (2018)). The synchrony between neurons in the population thus seems as independent between the task periods as the population encoding of past and present.

To address the reactivation of synchrony patterns in sleep, we first computed two session-wide synchrony patterns for each session, one from concatenating all correct trials; the other from concatenating all error trials (Figure A.4b). We found the synchrony pattern in correct trials was a closer match to the synchrony pattern in post- than pre-training slow-wave sleep (Figure A.4c). Error trials showed no preferential match of synchrony pat-

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

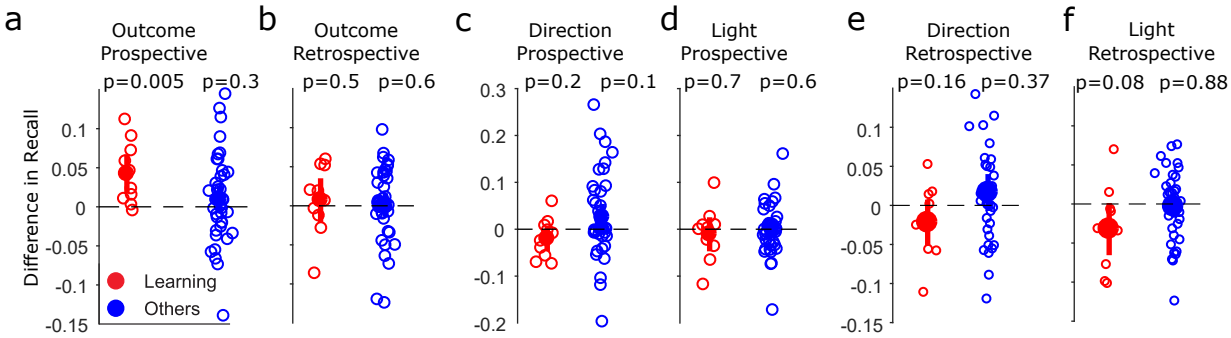

Figure A.2: The ensemble reactivation pattern is specific for prospective and not retrospective encoding.
(**a**) The difference in average recall between correct and error trials, across session types. Empty circles are single session values, while bold symbols are means ± 2SEM. P-values are from Wilcoxon sign rank test, $N = 10$ learning and $N = 39$ Other sessions.
(**b**) As for (a), with the difference in recall computed with respect to the outcome of the preceding trial.
(**c**) As for (a), with the difference in recall computed with respect to the upcoming direction choice on each trial.
(**d**) As for (a), with the difference in recall computed with respect to the cue location on each trial.
(**e**) As for (a), computed with respect to the direction chosen in the preceding trial.
(**f**) As for (a), computed with respect to the cue position in the preceding trial.

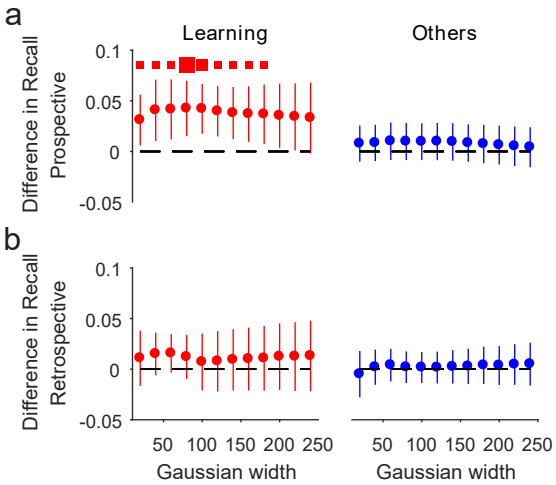

Figure A.3: The outcome-dependent ensemble activity pattern is specific for learning across multiple time scales.
Outcome-dependent delta Recall for different convolution width (from 20 ms to 240 ms every 20 ms step) for learning and other sessions shows a recall effect only for learning session in a prospective (**a**) approach and not retrospective (**b**). For each Gaussian width we show the mean ± 2SEM. Red squares on top of the panels indicate bins with difference in recall significantly better than chance (Wilcoxson sign rank test, $p < 0.05$ small squares; $p < 0.01$ medium size squares; $p < 0.005$ big squares.

tern to either sleep epoch (Figure A.4d). Repeating the same analysis for the inter-trial intervals, we found no preferential match of synchrony pattern for either sleep epoch, neither for intervals following correct outcomes (Figure A.4e), or those following error (Figure A.4f). Outcome dependent population synchrony in trials was thus recapitulated in sleep independently of the population synchrony in inter-trial intervals.

bioRxiv preprint doi: https://doi.org/10.1101/668962; this version posted August 12, 2020. 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.

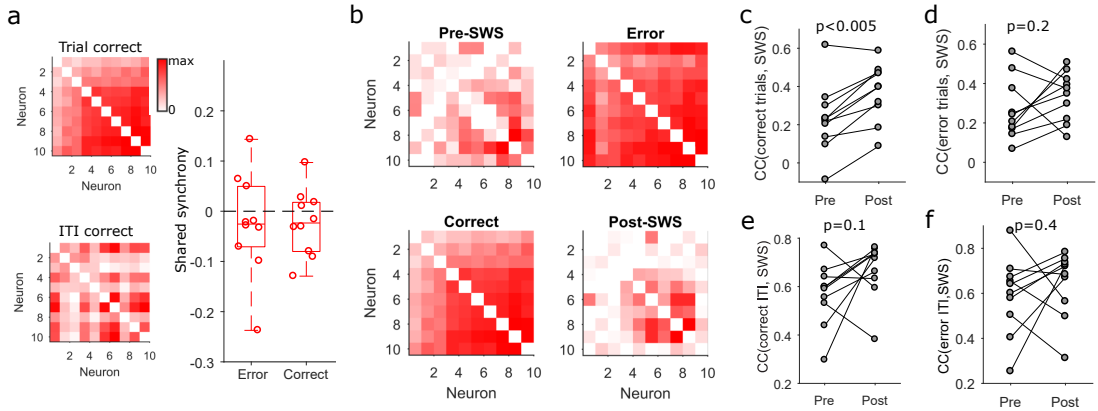

Figure A.4: Population synchrony is also orthogonal between trials and inter-trial intervals.
(**a**) Example similarity matrices for the concatenated correct trials (upper panel) and concatenated correct inter-trial intervals (lower panel) on the left. On the right, distributions of relative similarity between inter-trial intervals and preceding trials compared to a shuffle control model. Distributions around zero indicate that trials and inter-trial intervals are independent (redrawn from Maggi et al. (2018)).
(**b**) Similarity matrices of concatenated pre-session slow-wave sleep (SWS) episodes, error trials, correct trials and post-session slow-wave sleep episodes for an example learning session.
(**c**) Correlation between concatenated correct trials and pre-session slow-wave sleep episodes compared to the correlation between concatenated correct trials and post-session slow-wave sleep episodes. Each dot is a learning session. p-value from a paired t-test is shown on top.
(**d**) Correlation between concatenated error trials and pre-session slow-wave sleep episodes compared to the correlation between concatenated error trials and post-session slow-wave sleep episodes.
(**e**) As in panel (c) Correlation between concatenated correct inter-trial intervals and pre-session slow-wave sleep episodes compared to the correlation between concatenated correct inter-trial intervals and post-session slow-wave sleep episodes.
(**f**) As in panel (d) Correlation between concatenated error inter-trial intervals and pre-session slow-wave sleep episodes compared to the correlation between concatenated error inter-trial intervals and post-session slow-wave sleep episodes.