

# Intermittent rate coding and cue-specific ensembles support working memory

<https://doi.org/10.1038/s41586-024-08139-9>

Received: 20 November 2023

Accepted: 1 October 2024

Published online: 6 November 2024

Open access

 Check for updates

Matthew F. Panichello<sup>1,2</sup>✉, Donatas Jonikaitis<sup>1,2</sup>, Yu Jin Oh<sup>1</sup>, Shude Zhu<sup>1</sup>, Ethan B. Trepka<sup>1</sup> & Tirin Moore<sup>1</sup>✉

Persistent, memorandum-specific neuronal spiking activity has long been hypothesized to underlie working memory<sup>1,2</sup>. However, emerging evidence suggests a potential role for ‘activity-silent’ synaptic mechanisms<sup>3–5</sup>. This issue remains controversial because evidence for either view has largely relied either on datasets that fail to capture single-trial population dynamics or on indirect measures of neuronal spiking. We addressed this controversy by examining the dynamics of mnemonic information on single trials obtained from large, local populations of lateral prefrontal neurons recorded simultaneously in monkeys performing a working memory task. Here we show that mnemonic information does not persist in the spiking activity of neuronal populations during memory delays, but instead alternates between coordinated ‘On’ and ‘Off’ states. At the level of single neurons, Off states are driven by both a loss of selectivity for memoranda and a return of firing rates to spontaneous levels. Further exploiting the large-scale recordings used here, we show that mnemonic information is available in the patterns of functional connections among neuronal ensembles during Off states. Our results suggest that intermittent periods of memorandum-specific spiking coexist with synaptic mechanisms to support working memory.

Working memory allows us to retain and manipulate information on short timescales, and it is central to complex cognitive processing and adaptive behaviour<sup>6,7</sup>. Foundational work in the 1970s showed that working memory is associated with sustained neuronal spiking activity in primate prefrontal cortex<sup>8</sup>. Subsequent studies demonstrated that the persistent spiking of many neurons is specific to a remembered cue<sup>9–11</sup>. Persistent activity has been observed during both spatial and feature-based working memory tasks<sup>12–14</sup>, as well as within many cortical and subcortical brain structures<sup>15–20</sup>. In addition to non-human primates, it has also been observed in multiple animal models<sup>21</sup> as well as in humans<sup>22,23</sup>. Combined, this evidence has established persistent spiking as the dominant model of working memory<sup>2</sup>.

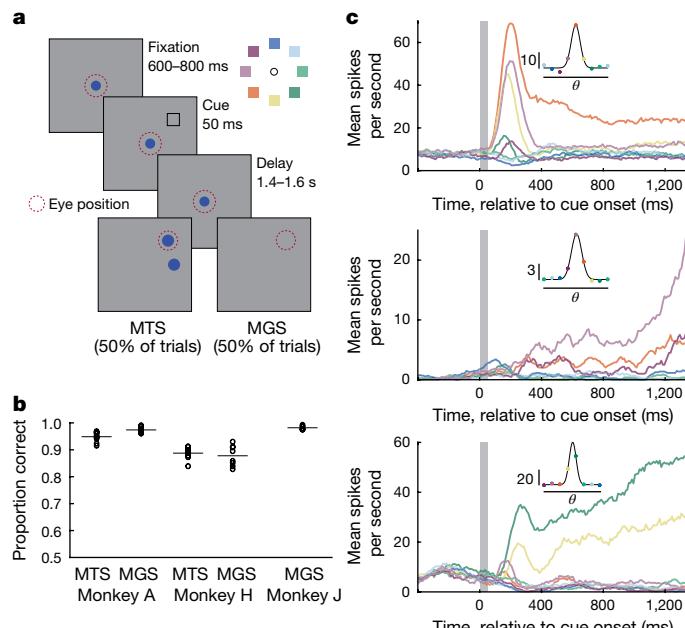
In spite of the predominance of the persistent-spiking model of working memory, an alternative class of models has received increased attention in recent years. This class of models proposes that working memory is supported by ‘activity-silent’, synaptic mechanisms rather than persistent activity<sup>4,5,24</sup>. Specifically, information held in working memory is stored by the pattern of short-term plastic changes initiated by a particular memory cue. Proof-of-principle simulations demonstrate that short-term plasticity (STP) can maintain information in the absence of persistent spiking<sup>4,24</sup>. Evidence of such latent traces has been reported using a variety of methods<sup>25–29</sup>. For example, STP, as inferred from functional connectivity, has been shown to correlate with cross-trial serial biases<sup>25</sup> and the maintenance of task sets<sup>30</sup> during working memory. However, it remains unknown whether cue-specific synaptic mechanisms operate during canonical working

memory delays, when the maintenance of memorandum information is most critical.

Nevertheless, synaptic models of working memory can potentially address key shortcomings of the persistent-spiking model. For one, persistent activity has been reported to be modest, or even absent, in some cases<sup>27–29,31</sup>, and to vary with task demands<sup>26,32</sup>. Second, and more importantly, delay-period activity can be highly variable on single trials<sup>33</sup>, prompting some to question the utility of persistent spiking as a reliable mechanism for memory maintenance<sup>3–5</sup>. In addition, the high-gamma component of prefrontal local field potentials appears bursty, rather than persistent, during memory delays, suggesting that population spiking may be similarly irregular<sup>34,35</sup>. In principle, a synaptic mechanism could eliminate, or at least minimize, disruptions in memory maintenance due to spiking irregularities. However, the relative contributions of spiking and synaptic mechanisms to working memory remain largely unresolved.

To address the above questions, we studied the activity of neurons within the lateral prefrontal cortex (areas 8 and 9/46) in three monkeys (A, HandJ; Methods). The monkeys were trained to perform one or two variants of a spatial working memory task (Fig. 1a). In both variants, the monkey was first presented with a brief (50 ms) spatial cue at one of eight possible locations while fixating a central spot. Following the cue, the monkey maintained fixation during a memory delay (1,400–1,600 ms). In one task (match-to-sample, MTS)<sup>36</sup>, two targets appeared after the delay and the monkey was rewarded for making an eye movement to the target appearing at the previously cued location. In the

<sup>1</sup>Department of Neurobiology and Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. <sup>2</sup>These authors contributed equally: Matthew F. Panichello, Donatas Jonikaitis. <sup>✉</sup>E-mail: mfp2@stanford.edu; tirin@stanford.edu



**Fig. 1 | Persistent, trial-averaged neuronal responses during spatial working memory.** **a**, Delayed MTS and MGS tasks. On each trial, the animal was presented with a cue at one of eight possible locations (inset). Following a memory delay period, the animal received fluid reward for making an eye movement to the previously cued location. **b**, Proportion correct for MTS and MGS tasks. Circles denote individual sessions, lines represent mean across  $n = 25$  sessions. **c**, Trial-averaged peristimulus time histograms for three example prefrontal neurons showing canonical persistent activity during the memory delay period. Colours denote different cue locations. Insets show cue-location ( $\theta$ ) tuning functions for each neuron during the memory delay, with scale bar (spikes per second) and gaussian fit (black trace).

second task (memory-guided saccade, MGS)<sup>9</sup>, no targets appeared after the delay and the monkey was rewarded for making an eye movement to the previously cued (blank) location. All three monkeys achieved excellent performance (Fig. 1b). Trials of both task type were randomly interleaved and were pooled for subsequent analyses.

### High-density recordings from prefrontal cortex

As expected from previous studies<sup>9,10</sup>, we observed a substantial proportion of prefrontal neurons with cue-specific memory delay activity (mean, 44% per session). Trial-averaged responses of these neurons suggest that their firing rates are sufficient to encode the remembered cue during the memory delay (Fig. 1c). However, trial averaging can obscure the high variability of spiking activity on single trials<sup>3</sup>. Thus, for individual neurons, cue information may be unreliable at times during the delay. Nonetheless, it could be that lapses in cue information for some neurons in the population are compensated by the continued activity of other neurons encoding the same memorandum. Alternatively, these lapses could be coordinated such that cue information fails to persist throughout the memory delay across the entire neuronal population. To distinguish between these possibilities, it is crucial to simultaneously measure the activity of large populations of neurons and to examine their activity on single trials, to evaluate the contribution of persistent spiking to working memory.

In the past few years, high-density, silicon probes, most notably Neuropixels probes (IMEC, Inc.), have revolutionized large-scale electrophysiological recordings in the mouse brain<sup>37,38</sup>. More recently, these probes were adapted for use in non-human primates<sup>39</sup>. We used these probes to obtain recordings from large, dense populations of prefrontal neurons in monkeys performing the spatial working memory

tasks (Fig. 2a). Our Neuropixels recordings ( $n = 25$  sessions) typically yielded hundreds of single and multi-units in each session (mean,  $329 \pm 46$ ;  $n = 8,225$  total; Methods). In addition to memory delay neurons, these recordings allowed us to capture the spatial distribution of multiple functional classes of neurons (Extended Data Fig. 1). For example, neurons selective to multiple task components (for example, visuomotor neurons) tended to be more closely spaced than those selective to only one.

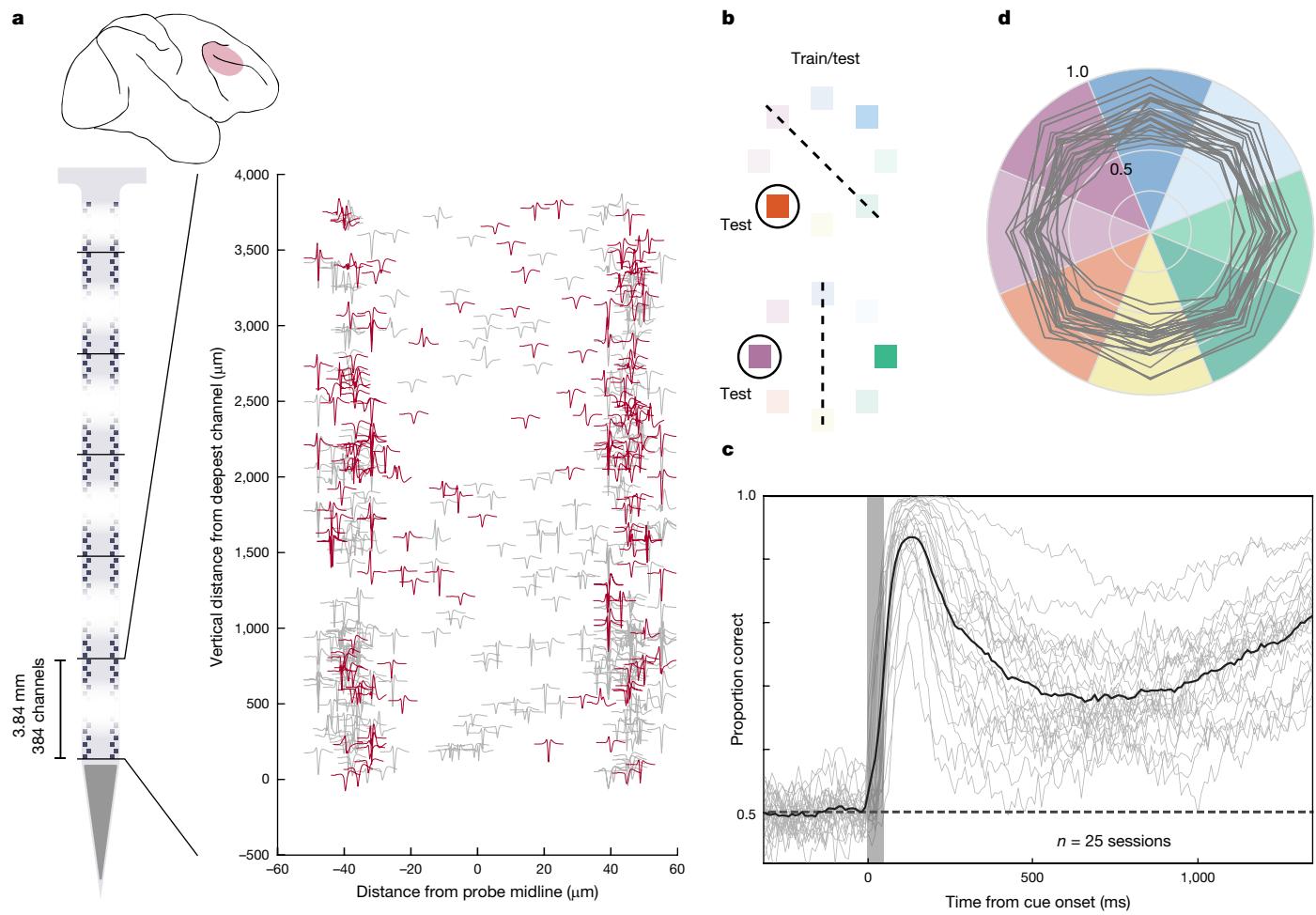
Most importantly, these recordings allowed us to quantify the information collectively conveyed by local populations of neurons about the remembered cue location. To do this, we used a leave-one-out, binary classification procedure. For each trial and time point within a session, we trained logistic regression models to discriminate the test location from its opposite location across the trial duration (Fig. 2b). For these and subsequent results, analyses of the memory delay were confined to the period from 500 to 1,400 ms following cue appearance, to avoid the influence of visually evoked responses (Methods). Across recording sessions, mean classification accuracy was significantly above chance throughout the memory delay (all  $P < 0.001$ , sign-rank) (Fig. 2c). Moreover, for each individual session, mean classification across the memory delay exceeded chance performance (all  $P < 0.001$ , sign-rank), with accuracies ranging from 59 to 89%. Last, classification accuracy was similar across cue locations (range, 67–74%, all  $P < 0.001$ , sign-rank) (Fig. 2d).

### Stability of cue information in firing rates

The persistence of cue information in the averaged classification accuracy during the memory delay, however robust, may nonetheless belie memory dynamics occurring on single trials. In particular, any lack of persistence on single trials could be obscured in the trial-averaged accuracy. To investigate this, we adapted techniques recently used to study value coding<sup>40</sup> to examine the single-trial dynamics of cue information during working memory. Specifically, we analysed the confidence of the classifier described above—the posterior probability assigned to the correct class at test, which provides a time-resolved index of the amount of cue information in population spiking during each trial (Methods). In each monkey, classifier confidence correlated with reaction time on correct trials (Extended Data Fig. 2a). Clearly, if cue information persists on single trials, confidence values should remain stably above chance (0.5) throughout the memory delay.

On the contrary, we found that confidence failed to persist through the memory delay on single trials. Instead, lapses in classifier confidence were evident throughout the memory delay and across trials within each recording session (Fig. 3a, Extended Data Fig. 3). At the start of each trial, confidence was consistently high during the visual response to the cue. However, following the disappearance of the cue, confidence often returned to chance multiple times during the memory delay. During single trials, periods of high confidence were interrupted by sharp transitions to low confidence (Fig. 3b). Lapses in confidence were not associated with microsaccades (Extended Data Fig. 4). Furthermore, these transitions between high and low confidence did not appear aligned across trials (Fig. 3a,b, Extended Data Fig. 3).

Given the apparent fluctuations between high and low confidence, we next sought to determine whether single-trial confidence was best described by one or two means. Our null hypothesis was that fluctuations in confidence reflect random perturbations around a single mean. We formalized this by fitting a single beta distribution, which is used to model the behaviour of random variables on the interval [0, 1], to the histogram of confidence values from each session (Methods). The alternative, two-mean model describes confidence using a mixture of two beta distributions, reflecting two discrete states. Indeed, we found that the cross-validated, two-state model outperformed the single-state model in 19 of 25 recording sessions; this was due to the inability



**Fig. 2 | High-density neuronal recordings from lateral prefrontal cortex.** **a**, Top, location of recordings in lateral prefrontal cortex. Bottom left, schematic of the Neuropixels NHP probe, highlighting the contiguous block of 384 active channels near the probe tip. Bottom right, spike waveform templates for 480 single and multi-units extracted from a single example recording session in monkey A, shown at their measured location on the probe surface. Units plotted in red showed selectivity for cue location during the delay period.

**b**, Leave-one-trial-out training procedure. For each trial and time point, a classifier was trained on the remainder of trials to discriminate the same cue location as the test trial from the opposite cue location. **c**, Mean classification accuracies (proportion of trials correct) for cue location from  $n = 25$  sessions, relative to cue onset. Traces indicate individual sessions. **d**, Mean proportion correct, averaged across the memory delay (+500 to +1,400 ms relative to cue onset) by cue location. Grey traces show individual sessions.

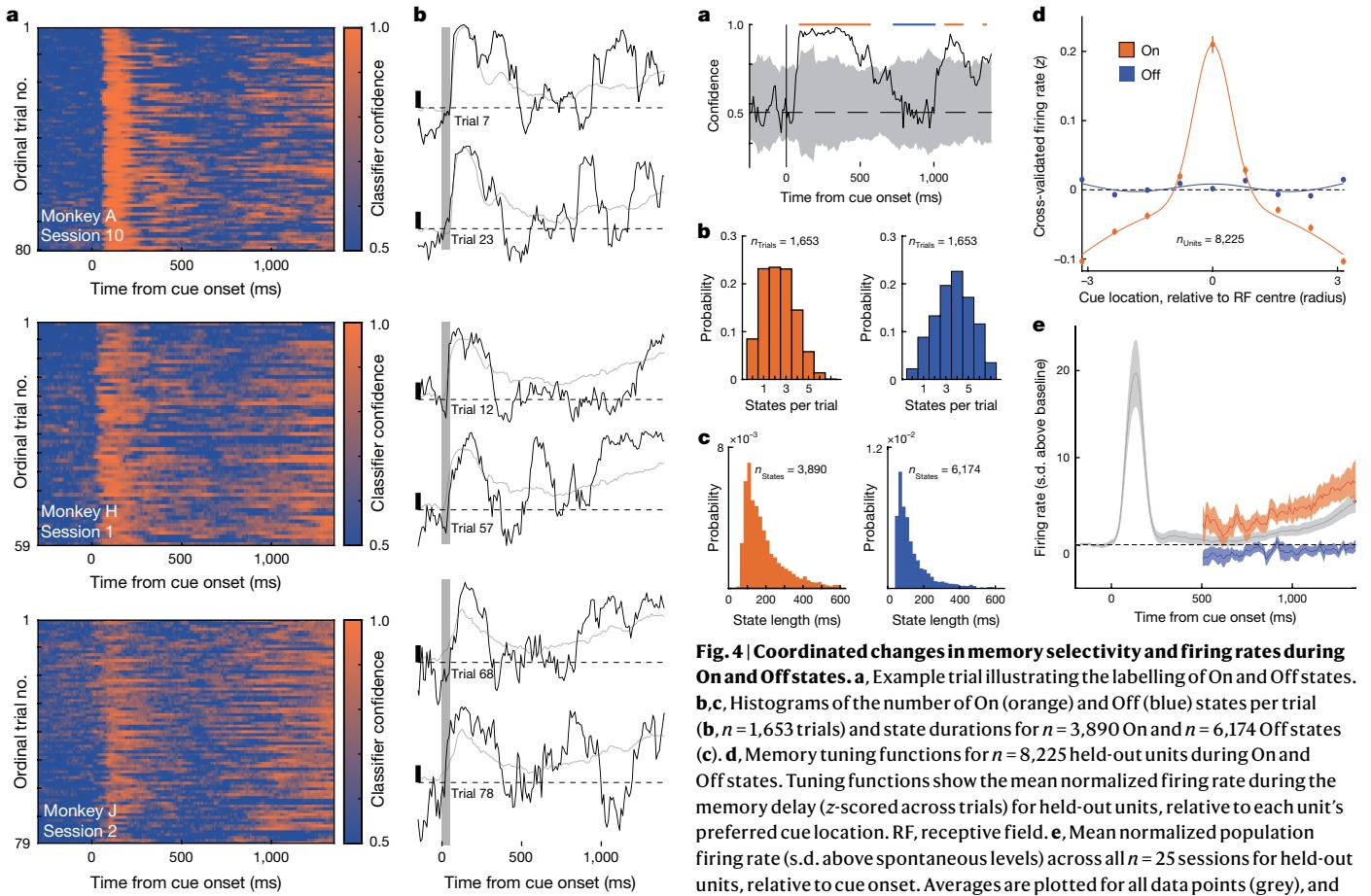
of the single-state model to capture the broad distribution of confidence values (Extended Data Fig. 5). This asymmetry in model performance was significant across recording sessions ( $\chi^2(1) = 7$ ,  $P = 0.009$ ). Similar results were obtained when using mixtures of Gaussians (Methods).

### Coordinated, intermittent rate-coding of memoranda

Having identified evidence of two discrete states, we next sought to label them on individual trials. We repeated the above classification procedure 50 times, shuffling condition labels on each iteration, to obtain a null distribution of confidence values for each trial (Fig. 4a and Methods). Across the cue and delay epochs, we labelled contiguous time points in which confidence was significantly greater than the null as ‘On’ states and labelled contiguous non-significant time points ( $P > 0.20$ ) as ‘Off’ states. During each trial, we observed a mean of  $2.35 \pm 0.04$  (median, 2.0) On states and a mean of  $3.74 \pm 0.04$  (median, 4.0) Off states from the cue period until the end of the memory delay (Fig. 4b). Periods in which confidence was significantly below the null (confidently incorrect) were rare (mean,  $0.11 \pm 0.01$  per trial). The mean duration of On states was  $192.4 \pm 2.1$  ms (median, 150), and the mean duration of Off states was  $146.1 \pm 1.7$  ms (median, 100)

(Fig. 4c). State (On or Off) at the time of the go cue predicted behavioural performance and reaction times in all three animals (Extended Data Fig. 2b,c). Importantly, classifiers trained only on Off states and tested on held-out Off states did not perform reliably above chance (Extended Data Fig. 6a), suggesting that Off states reflect time periods with no reliable information, and not a less frequent, second coding scheme that our original classifier failed to capture. In addition, we were not able to identify any predictive relationship between the phase of the local field potential along a range of frequency bands and On and Off states (Extended Data Fig. 6b), suggesting that these fluctuations in confidence are distinct from the rhythmic sampling of attention described previously<sup>41</sup>. Finally, estimates of background noise<sup>42</sup> did not differ between On and Off states, indicating that they were not associated with fluctuations in recording quality (Extended Data Fig. 6c).

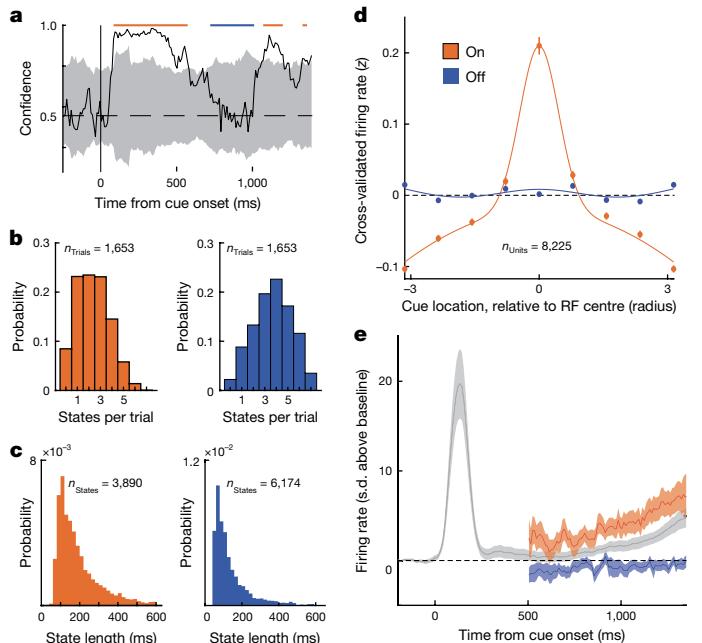
Having labelled On and Off states in this way, we next assessed their explanatory power and examined how they were reflected in the activity of individual neurons during the memory delay using cross-validation. To do this, we used half of the neurons recorded during each session to label states as On and Off and then examined the activity of neurons in the remaining (held-out) half of the population from the same sessions. We then repeated this process, switching training and test labels,



**Fig. 3 | Single-trial dynamics of memory signals in population spiking activity.** **a**, Single-trial classifier confidence, relative to cue onset, for all trials from the ‘preferred’ cue condition for three sessions. **b**, Confidence (black traces) for six example trials drawn from the three sessions in **a**. Grey traces show trial-averaged confidence values. Black scale bars denote a 0.10 increment in confidence; dashed lines denote  $y = 0.50$ .

allowing an unbiased analysis of all 8,225 units (Methods). Indeed, we found that activity differed markedly between On and Off states in two ways. First, spatial tuning during the memory delay, a hallmark of spatial working memory<sup>43,44</sup>, depended heavily on state. During On states, held-out neurons were strongly tuned to the location of the remembered cue (Fig. 4d and Extended Data Fig. 7). By contrast, during Off states, spatial tuning was virtually eliminated (Fig. 4d and Extended Data Fig. 7). Accordingly, there was a significant interaction between cue location and state (On or Off) on firing rates (two-way repeated measures analysis of variance (ANOVA),  $P < 0.001$ ). Whereas cue location explained an average of 8.2% of variance in firing rate during On states, it explained only 0.7% during Off states, a 12-fold decrease. Thus, Off states were associated with a pronounced loss of spatial tuning at the level of individual neurons.

Second, average firing rates during the memory delay also depended on state. During On states, held-out neurons exhibited firing rates above spontaneous levels (mean  $5.8 \pm 0.4$  Hz precue baseline) during the memory delay ( $P < 0.001$ , sign-rank; Fig. 4e). By contrast, during Off states, firing rates were statistically indistinguishable from spontaneous levels ( $P = 0.192$ ) and significantly lower than On states ( $P = 0.001$ ). Thus, Off states were associated not only with a loss of spatial selectivity but also with a collapse of firing rates to spontaneous levels. Together, these results show how transitions between On and Off states, derived from confidence values, reflect changes in



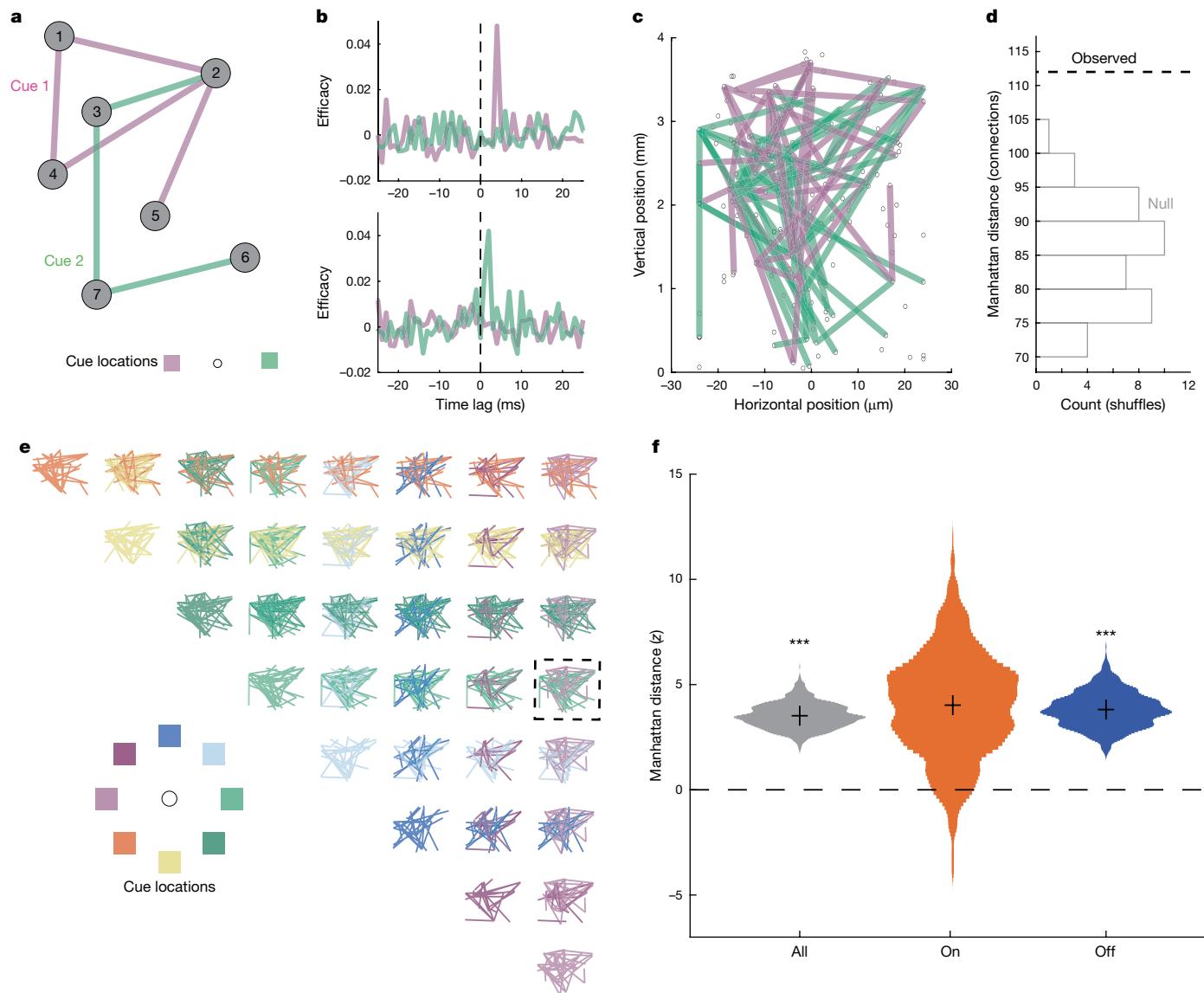
**Fig. 4 | Coordinated changes in memory selectivity and firing rates during On and Off states.** **a**, Example trial illustrating the labelling of On and Off states. **b,c**, Histograms of the number of On (orange) and Off (blue) states per trial (**b**,  $n = 1,653$  trials) and state durations for  $n = 3,890$  On and  $n = 6,174$  Off states (**c**). **d**, Memory tuning functions for  $n = 8,225$  held-out units during On and Off states. Tuning functions show the mean normalized firing rate during the memory delay (z-scored across trials) for held-out units, relative to each unit’s preferred cue location. RF, receptive field. **e**, Mean normalized population firing rate (s.d. above spontaneous levels) across all  $n = 25$  sessions for held-out units, relative to cue onset. Averages are plotted for all data points (grey), and also separately for firing rates extracted from On and Off states during the memory delay. **d,e**, Error bars denote mean  $\pm$  s.e.m.

basic firing-rate properties of individual neurons. Furthermore, these observations show that transitions between On and Off states during the memory delay are coordinated across neurons within the local population.

### Cue-specific neuronal ensembles during memory delay

The fact that cue information carried by neuronal firing rates is periodically lost during memory delay suggests that persistent activity may not be sufficient to support working memory. Therefore, we next considered evidence in favour of synaptic models. These models propose that, rather than persistent activity, working memory is instead represented by cue-specific networks of neurons<sup>4,5,24</sup> (Fig. 5a)—that is, cue-specific neuronal ensembles should be a signature of working memory. Thus, we next looked for cue-specific cell assemblies during the memory delay. As in our analyses of firing-rate dynamics, we leveraged the Neuropixels recordings to measure functional connections among the very large numbers of simultaneously recorded neuronal pairs (mean,  $52,314 \pm 11,252$  pairs per session). Specifically, we examined neuronal cross-correlations to assess their dependence on working memory.

To do this, we computed the cross-correlogram (CCG) between all neuronal pairs in each experimental session and for each cue condition based on activity during the memory delay. CCGs were computed, thresholded and firing rate normalized using established methods<sup>45,46</sup> (Fig. 5b, Extended Data Fig. 8 and Methods). We focused on CCGs with peaks at low latency (below 10 ms (refs. 45,47)), non-zero time lags, which are those most consistent with synaptic connections<sup>48</sup>.



**Fig. 5 | Cue-specific neuronal ensembles during memory delay.** **a**, Cartoon depicting the synaptic model of working memory. In the absence of cue-selective firing rates, information persists in the cue-specific patterns of potentiated connections (pink and green lines) among neurons (grey nodes). **b**, Examples of pairwise CCGs computed during memory delay following two different cues. **c**, CCG-derived connectivity maps during memory delay for two different cue locations measured in one session. Lines are drawn between neuronal pairs exhibiting significant CCGs. **d**, Differences (dashed line) between the two connectivity maps, quantified as Manhattan distance:

the sum of cue-specific connections. Comparison of this metric to a null distribution derived from condition-shuffled data (grey-outlined bars) yielded a  $z$ -score (in this case,  $z = 2.85$ ). **e**, Pairwise comparisons of connectivity maps across all cue locations. Dashed box shows comparison depicted in **c**. **f**, Mean normalized Manhattan distance (black crosses), using data from all  $n = 25$  sessions during the entire memory delay (grey,  $P = 1.2 \times 10^{-5}$ ), during On states only (orange,  $P = 0.129$ ) and during Off states only (blue,  $P = 2.5 \times 10^{-5}$ ). Violin plots show bootstrap across sessions. \*\*\* $P < 0.001$ , two-sided sign-rank.

In our recordings, we observed CCGs with low-latency, suprathreshold peaks ('significant CCGs') of this type in  $1.36 \pm 0.22\%$  of neuronal pairs, which totalled 12,740 significant CCGs across all sessions. Consistent with previous work<sup>46</sup>, the probability of observing a significant CCG decreased with the distance between neurons ( $r(39) = -0.42$ ,  $P = 0.007$ ). Significant CCGs were found at all penetration sites. Next, we compared the pattern of CCGs across cue conditions. Figure 5c shows an example of a comparison of significant CCGs computed for two cue conditions (0 and 180°) in one recording session. In this example session, the two conditions exhibited highly dissimilar ensembles of functionally connected neurons. To quantify the dissimilarity, we counted the number of condition-unique CCGs, yielding the Manhattan distance, which could then be compared with a null distribution derived from condition-shuffled data (Fig. 5d). We then

repeated this procedure for all possible pairs of conditions in each session (Fig. 5e). Across sessions, mean Manhattan distance was significantly greater than that predicted by chance ( $P < 0.001$ , sign-rank; Fig. 5f and Extended Data Fig. 7). This effect remained significant when confined to a comparison of firing-rate-matched conditions ( $P < 0.001$ , sign-rank; Extended Data Fig. 9 and Methods). Thus, the ensemble of functionally connected neurons significantly depended on the remembered cue.

If the observed cue-specific ensembles support working memory, these should be evident when cue information in firing rates is absent. To test this, we repeated the above analysis separately for On and Off states. Notably, Manhattan distance was not significantly greater than chance during On states ( $P = 0.129$ , sign-rank; Fig. 5f and Extended Data Fig. 7). However, during Off states, when spatial tuning was

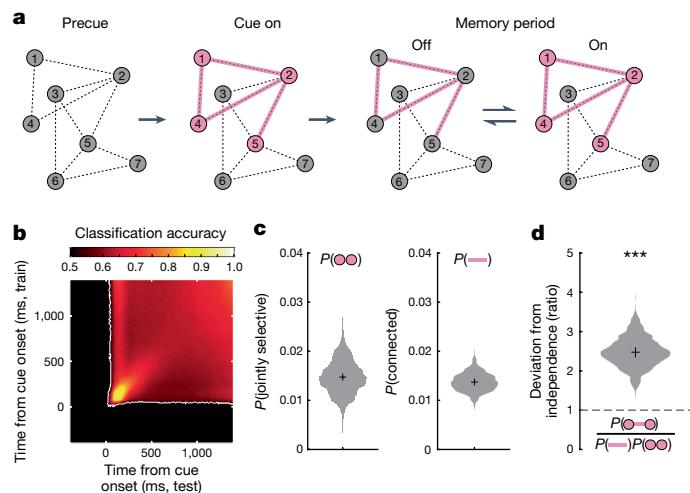
virtually absent and firing rates transitioned to spontaneous levels (Fig. 4d,e), Manhattan distance was significantly greater than chance ( $P < 0.001$ , sign-rank). This pattern of effects remained when confined to comparisons of firing-rate-matched conditions ( $P = 0.062$  and 0.001 for On and Off states, respectively, sign-rank; Extended Data Fig. 9 and Methods). Thus, even in the absence of persistent memory delay activity, memoranda information was reflected in the cue-specific ensembles of functionally connected neurons.

How might spiking and activity-silent coding work together to support memory? Synaptic models of working memory propose that evoked responses to a memory cue potentiate synapses between cue-selective neurons through STP (Fig. 6a). During the subsequent memory delay, this evoked response relaxes to spontaneous levels. However, the cue-specific pattern of potentiated synapses remains. Consequently, cue-specific elevations in firing rate may nonetheless re-emerge owing to non-specific fluctuations in extrinsic or intrinsic activity. Thus, across a memory delay, multiple transitions between spiking and activity-silent modes (On and Off) may occur<sup>4,24</sup>.

This account makes two testable predictions. First, it predicts a stable memory code—that is, despite the interruption of Off states, the same pattern of cue-selective spiking activity should support memory during On states throughout the memory delay. Indeed, similar to previous studies<sup>49</sup>, we found that the trial-averaged performance of classifiers trained on cue-evoked responses successfully generalized across the entire memory delay (Fig. 6b). Demixed principle component analysis<sup>50</sup> and analysis of single neurons also supported a stable memory code (Extended Data Fig. 10). Second, there should be a correspondence between spiking responses evoked by the cue and functional connectivity observed during memory delay. Thus, two neurons exhibiting evoked responses to a given cue should tend to be functionally connected during memory. In our data, when considering any single cue location,  $1.47 \pm 0.4\%$  of neuronal pairs responded preferentially to that cue location shortly after cue onset (average percentage across the eight possible cue locations). In addition, an average of  $1.4 \pm 0.2\%$  of neuronal pairs exhibited significant CCGs to that cue location during memory delay (Fig. 6c). If these proportions are independent, their conjunction should occur at a rate equal to their product. On the contrary, we found that pairs showing spiking responses and significant CCGs to the same cue were observed 2.5 times more than expected ( $P < 0.001$ , sign-rank; Fig. 6d). Thus, neurons that responded jointly during the evoked response were more likely to be functionally connected during memory delay, consistent with synaptic models. Overall, 55% of neurons showed a cue-specific evoked response, 81% were involved in a cue-specific functional connection and 10.4% were involved in such jointly selective and connected pairs.

## Discussion

These results elucidate the role of persistent and synaptic mechanisms in supporting working memory. By measuring the spiking of large, local populations of prefrontal neurons and resolving the dynamics of mnemonic information on single trials, we found that this information does not persist through the memory delay. Instead, cue-specific spiking activity was intermittent and was characterized by stochastically occurring, discrete transitions between robust coding of memoranda (On states) and complete lapses in such coding (Off states). Notably, these transitions were coordinated across large, local populations of neurons. Complementary to spike-rate-based coding, patterns of functional connectivity also carried information about the remembered cue during Off states. These patterns in functional connectivity are consistent with synaptic models of working memory in which cue-specific patterns of potentiated synapses facilitate the re-emergence of memory information in spike rates following silent epochs<sup>4,5,24</sup>.



**Fig. 6 | Stability of spike coding and cue specificity of neuronal ensembles in synaptic model predictions.** **a**, Cartoon depicting the interplay between spiking and silent mechanisms during working memory from synaptic models of working memory. Following the before-cue period, the memory cue evokes a distinct pattern of activity among a network of neurons (circles) and a distinct pattern of STP, temporarily facilitating connectivity (lines) among cue-associated neurons. During memory delay, even in the absence of persistent-spiking activity (Off states), cue information persists in the distinct pattern of connections. During the On state, non-specific drive reignites spiking activity among cue-associated neurons. Bidirectional arrows denote stochastic transitions between Off and On states. **b**, Mean classifier accuracy ( $n = 25$  sessions) plotted over time points on which the classifier was trained and tested. The block structure indicates good generalization across time. White line denotes the boundary of above-chance classification ( $P < 0.001$ , cluster-mass test, corrected for multiple comparisons). **c**, Left, probability ( $P$ ) that both of a pair of neurons responded preferentially to the same cue during the evoked response (0–400 ms after cue). Right, probability that both of a pair of neurons exhibited significant CCG to a particular cue during memory delay. **d**, Proportion of pairs that were both jointly selective and connected, divided by the proportion expected. Violin plots indicate bootstrap across  $n = 25$  sessions. \*\*\* $P < 0.001$ , two-sided sign-rank versus one-sided ( $P = 1.2 \times 10^{-5}$ ).

Our observations describe the dynamics of local populations of neurons in the primate lateral prefrontal cortex. Future work should explore these dynamics in other structures known to exhibit memory delay activity—for example, parietal cortex<sup>19</sup>—and on broader spatial scales, including the possible propagation of On states across the cortical surface<sup>51</sup> or distributed representations supporting memory<sup>52,53</sup>. Nevertheless, two facts suggest that the regions targeted by the present study play a key role in working memory. First, focal inactivation of lateral prefrontal cortex has long been known to severely impair performance in spatial working memory tasks<sup>54–56</sup>, including selective inactivation of memory delay activity<sup>57</sup>. Second, in the context of our study, single-trial confidence at the onset of the behavioural response phase predicted behaviour.

Similar, large-scale electrophysiological approaches should be used to assess the relative roles of spiking and putative synaptic mechanisms across a range of working memory tasks. For example, such roles may be rather different for spatial and object-based working memory, given the apparent differences between the two in the robustness of memory delay spiking activity across brain areas<sup>58</sup>. Furthermore, theoretical work suggests that the relative contributions of these two mechanisms may be related to the level of manipulation of remembered information required by a task<sup>59</sup>. The spatial tasks used here demand a relatively straightforward sensory-to-motor transformation, and so could rely more on synaptic mechanisms.

Empirical verification of how task demands sculpt memory representations will lead to a richer understanding of the mechanistic basis of working memory.

Finally, it is interesting to note that, even within a single experimental session and task condition, the relative proportions of On and Off states showed a fair degree of heterogeneity across trials. Recent electrophysiological studies provide evidence of coordinated fluctuations in local neuronal activity in alert animals<sup>60</sup>. Furthermore, these coordinated fluctuations appear related to moment-to-moment changes in global arousal states, and also predict psychophysical performance in non-human primates<sup>61,62</sup>. Neuromodulatory inputs are among the many possible mechanisms that may contribute to coordinated fluctuations in neuronal activity<sup>63</sup>. Dynamics in the local tone of neuromodulators may be sufficient to induce transitions in local cortical states<sup>60</sup>. Within lateral prefrontal cortex, dopaminergic tone is known to play a key role in the maintenance of memory delay activity<sup>64,65</sup>. Thus, examining the contribution of dopaminergic tone to the variability in the dynamics of memoranda coding could prove illuminating.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-024-08139-9>.

1. Constantinidis, C. et al. Persistent spiking activity underlies working memory. *J. Neurosci.* **38**, 7020–7028 (2018).
2. Wang, X.-J. 50 Years of persistent activity: quo vadis? *Trends Neurosci.* **44**, 888–902 (2021).
3. Lundqvist, M., Herman, P. & Miller, E. K. Working memory: delay activity, yes! Persistent activity? Maybe not. *J. Neurosci.* **38**, 7013–7019 (2018).
4. Mongillo, G., Barak, O. & Tsodyks, M. Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
5. Stokes, M. G. ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* **19**, 394–405 (2015).
6. Baddeley, A. D. in *The Oxford Handbook of Memory* (eds. Tulving, E. & Craik, F. I. M.) 77–92 (Oxford Univ. Press, 2000).
7. Ehrlich, D. B. & Murray, J. D. Geometry of neural computation unifies working memory and planning. *Proc. Natl Acad. Sci. USA* **119**, e2115610119 (2022).
8. Fuster, J. M. & Alexander, G. E. Neuron activity related to short-term memory. *Science* **173**, 652–654 (1971).
9. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic coding of visual space in the monkey’s dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989).
10. Fuster, J. M. Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J. Neurophysiol.* **36**, 61–78 (1973).
11. Wimmer, K., Nykamp, D. Q., Constantinidis, C. & Compte, A. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* **17**, 431–439 (2014).
12. Armstrong, K. M., Chang, M. H. & Moore, T. Selection and maintenance of spatial information by frontal eye field neurons. *J. Neurosci.* **29**, 15621–15629 (2009).
13. Rao, S. C., Rainer, G. & Miller, E. K. Integration of what and where in the primate prefrontal cortex. *Science* **276**, 821–824 (1997).
14. Wilson, F. A., Scalaidhe, S. P. & Goldman-Rakic, P. S. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* **260**, 1955–1958 (1993).
15. Chelazzi, L., Miller, E. K., Duncan, J. & Desimone, R. A neural basis for visual search in inferior temporal cortex. *Nature* **363**, 345–347 (1993).
16. Glimcher, P. W. & Sparks, D. L. Movement selection in advance of action in the superior colliculus. *Nature* **355**, 542–545 (1992).
17. Hikosaka, O., Sakamoto, M. & Miyashita, N. Effects of caudate nucleus stimulation on substantia nigra cell activity in monkey. *Exp. Brain Res.* **95**, 457–472 (1993).
18. Miyashita, Y. & Chang, H. S. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* **331**, 68–70 (1988).
19. Snyder, L. H., Batista, A. P. & Andersen, R. A. Coding of intention in the posterior parietal cortex. *Nature* **386**, 167–170 (1997).
20. van Kerkoerle, T., Self, M. W. & Roelfsema, P. R. Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nat. Commun.* **8**, 13804 (2017).
21. Inagaki, H., Fontolan, L., Romani, S. & Svoboda, K. Discrete attractor dynamics underlying selective persistent activity in frontal cortex. *Nature* **566**, 212–217 (2017).
22. Harrison, S. A. & Tong, F. Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458**, 632–635 (2009).
23. Vogel, E. K. & Machizawa, M. G. Neural activity predicts individual differences in visual working memory capacity. *Nature* **428**, 748–751 (2004).
24. Lundqvist, M., Herman, P. & Lansner, A. Theta and gamma power increases and alpha/beta power decreases with memory load in an attractor network model. *J. Cogn. Neurosci.* **23**, 3008–3020 (2011).
25. Barbosa, J. et al. Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat. Neurosci.* **23**, 1016–1024 (2020).
26. Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K. & Postle, B. R. Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* **24**, 61–79 (2012).
27. Rose, N. S. et al. Reactivation of latent working memories with transcranial magnetic stimulation. *Science* **354**, 1136–1139 (2016).
28. Sprague, T. C., Ester, E. F. & Serences, J. T. Restoring latent visual working memory representations in human cortex. *Neuron* **91**, 694–707 (2016).
29. Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working memory guided behaviour. *Nat. Neurosci.* **20**, 864–871 (2017).
30. Fujisawa, S., Amarasingham, A., Harrison, M. T. & Buzsáki, G. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat. Neurosci.* **11**, 823–833 (2008).
31. Shafi, M. et al. Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* **146**, 1082–1108 (2007).
32. Watanabe, K. & Funahashi, S. Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nat. Neurosci.* **17**, 601–611 (2014).
33. Compte, A. et al. Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J. Neurophysiol.* **90**, 3441–3454 (2003).
34. Lundqvist, M. et al. Gamma and beta bursts underlie working memory. *Neuron* **90**, 152–164 (2016).
35. Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L. & Miller, E. K. Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nat. Commun.* **9**, 394 (2018).
36. Hasegawa, R. P., Peterson, B. W. & Goldberg, M. E. Prefrontal neurons coding suppression of specific saccades. *Neuron* **43**, 415–425 (2004).
37. Jun, J. J. et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).
38. Steinmetz, N. A. et al. Neuropixels 2.0: a miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**, eabf4588 (2021).
39. Trautmann, E. M. et al. Large-scale high-density brain-wide neural recording in nonhuman primates. Preprint at bioRxiv <https://doi.org/10.1101/2023.02.01.526664> (2023).
40. Rich, E. L. & Wallis, J. D. Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.* **19**, 973–980 (2016).
41. Fiebelkorn, I. C., Pinsk, M. A. & Kastner, S. A dynamic interplay within the frontoparietal network underlies rhythmic spatial attention. *Neuron* **99**, 842–853 (2018).
42. Quiroga, R. Q., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* **16**, 1661–1687 (2004).
43. Rao, S. G., Williams, G. V. & Goldman-Rakic, P. S. Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. *J. Neurophysiol.* **81**, 1903–1916 (1999).
44. Rao, S. G., Williams, G. V. & Goldman-Rakic, P. S. Destruction and creation of spatial tuning by disinhibition: GABA(A) blockade of prefrontal cortical neurons engaged by working memory. *J. Neurosci.* **20**, 485–494 (2000).
45. Siegle, J. H. et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86–92 (2021).
46. Trepka, E. B., Zhu, S., Xia, R., Chen, X. & Moore, T. Functional interactions among neurons within single columns of macaque V1. *eLife* **11**, e79322 (2022).
47. González-Burgos, G., Barrionuevo, G. & Lewis, D. A. Horizontal synaptic connections in monkey prefrontal cortex: an *in vitro* electrophysiological study. *Cereb. Cortex* **10**, 82–92 (2000).
48. Ostojic, S., Brunel, N. & Hakim, V. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J. Neurosci.* **29**, 10234–10253 (2009).
49. Spaak, E., Watanabe, K., Funahashi, S. & Stokes, M. G. Stable and dynamic coding for working memory in primate prefrontal cortex. *J. Neurosci.* **37**, 6503–6516 (2017).
50. Kobak, D. et al. Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).
51. Shi, Y.-L., Steinmetz, N. A., Moore, T., Boahen, K. & Engel, T. A. Cortical state dynamics and selective attention define the spatial pattern of correlated variability in neocortex. *Nat. Commun.* **13**, 44 (2022).
52. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
53. Schmitt, L. I. et al. Thalamic amplification of cortical connectivity sustains attentional control. *Nature* **545**, 219–223 (2017).
54. Chafee, M. V. & Goldman-Rakic, P. S. Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *J. Neurophysiol.* **83**, 1550–1566 (2000).
55. Dias, E. C. & Segraves, M. A. Muscimol-induced inactivation of monkey frontal eye field: effects on visually and memory-guided saccades. *J. Neurophysiol.* **81**, 2191–2214 (1999).
56. Sommer, M. A. & Tehovnik, E. J. Reversible inactivation of macaque frontal eye field. *Exp. Brain Res.* **116**, 229–249 (1997).
57. Acker, L., Pino, E. N., Boyden, E. S. & Desimone, R. FEF inactivation with improved optogenetic methods. *Proc. Natl Acad. Sci. USA* **113**, E7297–E7306 (2016).
58. Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R. & Haynes, J.-D. The distributed nature of working memory. *Trends Cogn. Sci.* **21**, 111–124 (2017).
59. Masse, N. Y., Yang, G. R., Song, H. F., Wang, X.-J. & Freedman, D. J. Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nat. Neurosci.* **22**, 1159–1167 (2019).

60. Harris, K. D. & Thiele, A. Cortical state and attention. *Nat. Rev. Neurosci.* **12**, 509–523 (2011).
61. Davis, Z. W., Muller, L., Martinez-Trujillo, J., Sejnowski, T. & Reynolds, J. H. Spontaneous travelling cortical waves gate perception in behaving primates. *Nature* **587**, 432–436 (2020).
62. Engel, T. A. et al. Selective modulation of cortical state during spatial attention. *Science* **354**, 1140–1144 (2016).
63. Noudoost, B. & Moore, T. The role of neuromodulators in selective attention. *Trends Cogn. Sci.* **15**, 585–591 (2011).
64. Sawaguchi, T. & Goldman-Rakic, P. S. D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science* **251**, 947–950 (1991).
65. Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V. & Arnsten, A. F. T. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* **10**, 376–384 (2007).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

# Article

## Methods

### Subjects

Three adult male rhesus monkeys (*Macaca mulatta*), aged 11, 12 and 8 years, participated in the experiment. Monkeys A, H and J weighed 11, 14 and 12 kg, respectively. All surgical and experimental procedures were approved by the Stanford University Institutional Animal Care and Use Committee and were in accordance with the policies and procedures of the National Institutes of Health.

### Behavioural task

Stimuli were presented on a VIEWPixx3D monitor positioned at a viewing distance of 60 cm using Psychtoolbox and MATLAB (v.R2022a, MathWorks). Eye position was monitored at 1 kHz using an Eyelink 1000 eye-tracking system (SR Research). On each trial, the animals were presented with a cue at one of eight possible locations and reported this location after a brief memory delay to receive a fluid reward. Cues were square frames (green for monkey A, black for monkey H, white for monkey J) measuring 1° of visual angle on a side, and presented at 5–7° of eccentricity (depending on the session).

Monkeys initiated behavioural trials by fixating a central fixation spot presented on a uniform grey background. After the monkeys had maintained fixation for 600–800 ms (randomly selected on each trial), a cue appeared for 50 ms at one of eight possible locations separated by 45° around fixation. Cue presentation was followed by a delay period that varied randomly from 1,400 to 1,600 ms. Following the delay period, the fixation spot disappeared and the animal was presented with one of two possible response screens. On MTS trials, two targets appeared (filled blue circles, radius 1° of visual angle (DVA)): one at the previously cued location and the other at one of the seven remaining non-cued locations. On MGS trials, no targets appeared. In either case, the animals received a reward of juice for making an eye movement to within 5 DVA of the previously cued location and then maintaining fixation for 200 ms. MTS and MGS trials were randomly interleaved such that the animals could not predict the trial type. Monkey J was trained on, and performed only, the MGS task. The animals had to maintain their gaze within either 3 DVA (monkey A) or 2 DVA (monkeys H and J) from fixation throughout the trial until the response stage. The intertrial interval was 300–1,000 ms following each correct response. Failure to acquire fixation, fixation breaks and incorrect responses were not rewarded and were followed by a 2,000 ms intertrial interval.

### Surgical procedures and recordings

Monkeys were implanted with a titanium headpost to immobilize the head, and with a titanium chamber to provide access to the brain (see ref. 12 for full details). In a previous study<sup>66</sup>, we identified the frontal eye field based on its neurophysiological characteristics and ability to evoke saccades with electrical stimulation at low currents<sup>67</sup>. Here we recorded from both area 8, within and anterior to the frontal eye field, and the principal sulcus (9/46) (monkey A, area 8; monkey H, areas 8 and 9/46; monkey J, area 9/46), using primate Neuropixels probes<sup>39</sup>. During each session, we pierced the dura using a screw-driven, 21-gauge pointed cannula and lowered a single probe through this cannula using a combination of custom three-dimensional printed grids and motorized drives (NAN instruments). Probe trajectories spanned several cortical columns, as inferred from the broad distribution of preferred cue locations across neurons. Recordings were allowed to settle for around 30 min before the start of the experiment, to mitigate drift. We configured probes to record from 384 active channels in a contiguous block, allowing dense sampling of neuronal activity along a 3.84 mm span.

Neuropixels filter and digitize activity at the headstage separately for the action potential bands (300 Hz high-pass filter, 30 kHz sampling frequency) and local field potential (1 kHz low-pass filter, 2.5 kHz sampling frequency). Activity was monitored during experimental sessions and saved to disk using SpikeGLX (<https://billkarsh.github.io/SpikeGLX/>).

### Data preprocessing

Spiking in the action potential band was identified and sorted offline using Kilosort3 (ref. 68). Because we were interested in population-level coding of memory, we analysed both putative single- and multi-unit clusters identified by Kilosort. Spike times were aligned to a digital trigger on each trial, indicating cue onset, and corrected for a lag in stimulus presentation estimated offline using photodiode measurements from the stimulus display and the timing of the cue-evoked response. Neurons that fired fewer than 1,000 spikes in the experimental sessions (each roughly 3 h) were excluded from further analyses. Spike times were converted to smoothed firing rates (sampling interval, 10 ms) by representing each spiking event as a delta function and convolving this time series with a 100 ms boxcar. For CCG analyses, unsmoothed spike times were binned with a width and timestep of 1 ms. Incorrect trials were rare (Fig. 1b) and were excluded from subsequent analysis. Waveform templates were localized in space using Neuropixel Utilities (<https://djoshea.github.io/neuropixel-utils/>). Local field potentials (LFPs) were sampled at 2,500 Hz. Offline, LFPs were filtered using a 2–200-Hz-bandpass, zero-phase Butterworth, notch filtered at 60 Hz and downsampled to 1,000 Hz. LFPs were transformed into the time-frequency domain using Morlet wavelets and downsampled to 100 Hz.

### Statistics and reproducibility

All statistical tests are two-sided unless otherwise specified. Key findings and brain-behaviour relationships were evident in each of the three animals (Extended Data Figs. 2 and 7), and in each of the 25 individual experimental sessions (Extended Data Fig. 3). Sample size (three animals, 8,255 neurons) was chosen based on standards in the field. Each animal was exposed to every task manipulation. Within a session, task manipulations were randomized across trials. Neurons were recorded without bias, and electrodes placed to maximize signal-to-noise of the electrophysiological signal.

### Functional subtyping

To determine the functional subtype of units<sup>69</sup> (Extended Data Fig. 1), we analysed firing rates during three time epochs: visual (0–300 ms after cue onset), memory (500–1,400 ms after cue onset) and motor (100–300 ms post-fixation offset). A unit was labelled as being selective during a given epoch if firing rates during that epoch were significantly modulated by cue location (one-way ANOVA,  $P < 0.05$  criterion). Units were then sorted into functional subclasses based on the set epochs during which each unit was selective.

### Classification of cue location

Firing-rate estimates for each unit and time point relative to cue onset were z-scored across trials before classification. We used linear classifiers to quantify the amount of information on the location of the cue in populations of simultaneously recorded units. We held out each trial for test one by one, training a logistic regression classifier (as implemented by fitclinear.m in MATLAB) to predict cue location using the population vector of firing rates. Specifically, classifiers were trained to discriminate the same cue location as the test trial from the opposite cue location, using the applicable subset of trials from the training set. Data were subsampled during training to equalize trial counts for the two conditions. A unique classifier was trained and tested for each time point relative to cue onset. ‘Classification accuracy’ reflects the proportion of correctly classified test trials (Fig. 2c); ‘classifier confidence’ is the non-thresholded value of the logistic function corresponding to the probability assigned by the classifier to the correct label at test (Fig. 3).

Cross-state classification (Extended Data Fig. 6a) was similar, except that only values from trial time points labelled On or Off (as appropriate) entered the training set or were held out as a test trial. Test confidence

values were averaged across the memory delay (500–1,400 ms after cue) to yield the final results.

Cross-temporal classification (Fig. 6b) was also similar, except that we used a split-half approach in which the classifiers for each time point were trained on half of the available population of trials and tested (cross-temporally) using the other half.

### Mixture modelling of confidence

We used a mixture-modelling approach to test whether confidence during the memory delay (500–1,400 ms after cue) was best described as drawn from a one- or a two-state distribution (Extended Data Fig. 5). To do this, for each session and cue location we modelled the probability density function of confidence values during the memory delay as either a single beta distribution,

$$p(c) = \text{Beta}(c; \alpha, \beta),$$

or a mixture of two beta distributions,

$$p(c) = w \times \text{Beta}(c; \alpha, \beta) + (1 - w) \times \text{Beta}(c; \alpha', \beta'),$$

where  $c$  is confidence,  $\alpha, \beta, \alpha'$  and  $\beta'$  parameterize beta distribution(s) and  $w$  is the mixing coefficient. The best-fitting parameters of each model were identified by maximum-likelihood estimation using gradient descent in MATLAB. We used fourfold cross-validation on the population of trials to assess the likelihood of each model on held-out test data, and then normalized by the number of trials and changed the log likelihood to base 2 to yield the cross-validated score of each model in terms of bits per trial. Finally, we subtracted these two model scores and averaged across conditions to yield the difference in model performance for each session.

Note that our choice of beta distribution here is principled: it is extremely flexible, able to demonstrate a broad range of skewness and kurtosis and naturally accommodates bounded continuous variables such as confidence<sup>70</sup>. This flexibility makes this analysis conservative, ensuring that the one-state model is capable of describing a broad range of empirical distributions. Nevertheless, results were not dependent on the exact modelling approach: similar results for our one-versus-two-state model comparison were obtained when using either Gaussian mixture or hidden Markov models.

### Analysis of microsaccades

The horizontal and vertical eye position records were convolved with a Gaussian kernel ( $\sigma = 4.75$  ms) to suppress noise before taking first derivatives, yielding the eye velocity along each dimension. We then took the root sum of squares of the horizontal and vertical velocities to obtain eye speed. We flagged peaks in this time series with a minimum peak height of  $10^\circ \text{ s}^{-1}$  and a minimum interpeak distance of 50 ms as microsaccades (ref. 71 and Extended Data Fig. 4), which were confirmed by visual inspection of the data.

### Labelling of On and Off states

To identify On and Off states (Fig. 4a), we repeated the cue classification analysis described above 50 times, randomly shuffling the labels of the training set for each test trial. This yielded, for each trial, a null distribution of 50 confidence time series (Fig. 4a). We then z-scored each time point of the true confidence time series by the mean and standard deviation of this null distribution. Individually significant (above 1.96) z-values were cluster corrected for multiple comparisons over time<sup>72</sup>. In brief, we compared the sum of contiguous individually significant z-values with that expected by chance (randomization test). Clusters with a mass greater than the 95% percentile of the null were labelled On states; contiguous z-values falling below a conservative ( $P > 0.20$ ) threshold for at least five consecutive time points were labelled Off states.

### Tuning curves

To test whether On and Off states reflected coordinated changes in tuning across the neural population, we used a split-half approach. First, firing-rate estimates for each unit and time point relative to cue onset were z-scored across trials. Then, for each session, we randomly divided the population of units in half. We used one half of the units to identify On and Off states, as described above. Next, for each unit in the held-out population, we computed mean firing rate during the memory delay for each cue location separately for On and Off states, averaging across relevant time points and across trials. This yielded, for each unit, two eight-element vectors—the On and Off tuning functions. To align tuning functions across units, the preferred cue location for each was identified as the condition in which the sum of the On and Off functions was greatest, and assigned an arbitrary value of zero degrees. Alignment of tuning curves to the maximum-valued preferred cue in this way will necessarily produce a peak at zero degrees in the average tuning function, even in the absence of true tuning. To correct for this, for each unit we also computed null On and Off tuning functions by first shuffling cue labels across trials, aligned these to the preferred cue and subtracted them from the true On and Off tuning functions (Fig. 3d).

For demonstration purposes, we fit the average On and Off tuning functions with a difference of Gaussians using gradient descent in MATLAB. Difference of Gaussians is useful for describing tuning curves that show surround suppression<sup>73</sup>.

### Population firing rates

To describe how population firing rates evolved over the course of the trial, we averaged these across all units recorded in the same session and across all trials for the preferred cue location (greatest mean classification confidence during the memory delay), yielding a single time series for each session. We then normalized this time series by the mean and standard deviation of a 400 ms baseline period (-400 to 0 ms relative to cue onset), yielding a metric of population spiking in units of standard deviations above baseline (Fig. 4e, grey traces). We repeated this analysis for the memory delay, this time including only data points labelled On or Off (Fig. 4e, orange and blue traces).

### Phase–state relationships

To determine whether the phases of different frequency components of the LFP were predictive of On and Off states, we first extracted the phase from the time-frequency representation of the LFP. Next, we identified the onset time of On and Off states during the memory delay across all trials within a session. Then, for each session, probe channel and frequency (4–60 Hz), we computed the (circular) mean phase at On state and Off state onset, and the magnitude of the angular difference between these means. If phase is predictive of state, this difference should be larger than that expected by chance. Accordingly, we obtained a null distribution of phase difference magnitudes by repeating this procedure 1,000 times, randomly permuting On and Off labels for each phase measurement on each iteration and used this null distribution to generate a z-score metric. Z-scores were averaged across channels, yielding a phase–state metric for each session and frequency of interest. Finally, we tested whether these scores were greater than zero for each frequency of interest (cluster-corrected randomization test).

### Standard deviation of background noise

To ensure that changes in recording quality could not account for the presence of On and Off states in our recordings, we measured the standard deviation of background noise<sup>42</sup> in the action potential band (0.3–10 kHz). Specifically, for each On and Off state identified using the

# Article

non-parametric procedure described above, noise standard deviation was estimated as

$$\sigma_n = \text{median}\left(\frac{|x|}{0.6745}\right),$$

where  $x$  is the time series of raw action potential band values recorded during the state. Noise estimates were averaged across all On and Off states within each session.

## Demixed principal components analysis

We used demixed principal components analysis to decompose population activity into different components reflecting cue location, time and their interaction. As with our classification-based analyses, we applied demixed principal components to the smoothed firing rates from each session, focusing on activity during the delay period (500–1,400 ms after cue). The proportion of variance explained and components were extracted as described in ref. 50.

## Single-neuron ANOVA

We downsampled the smoothed firing rates for each neuron to 100 ms steps and modelled firing rate during the delay (500–1,400 ms after cue) as a linear combination of cue location, firing rate and their interaction, to estimate the proportion of variance explained by each of these terms.

## CCG analysis

To characterize functional connectivity among units, we computed cross-correlations between spike trains of all pairs of simultaneously recorded neurons with mean firing rates greater than 1 Hz. CCGs were computed separately for each cue location. Following previous studies<sup>46</sup>, to mitigate firing-rate effects, we normalized cross-correlation for each pair of neurons by the geometric mean of their firing rates for the cue location condition under consideration. The CCG for a pair of neurons ( $j, k$ ) in condition  $c$  was therefore

$$\text{CCG}(\tau)_{j,k,c} = \frac{\sum_{i=1}^M \sum_{t=\tau+1}^N x_j^i(t-\tau) \times x_k^i(t)}{\sqrt{\sum_{i=1}^M \sum_{t=\tau+1}^N x_j^i(t-\tau)^2 \times \sum_{i=1}^M \sum_{t=\tau+1}^N x_k^i(t)^2}},$$

where  $M$  is the number of trials collected for cue location  $c$ ,  $N$  is the number of time bins within a trial,  $\tau$  is the time lag between the two spike trains and  $x_k^i(t)$  is 1 if neuron  $j$  is fired in time bin  $t$  of trial  $i$ , but zero otherwise.

To correct for correlation due to stimulus locking or slow fluctuations in population response, we subtracted a jittered CCG from the original. This jittered CCG reflects the expected value of the CCG computed from all possible jitters of each spike train within a given jitter window<sup>74,75</sup>. The jittered spike train preserves both the poststimulus time histogram (PSTH) of the original spike train across trials and the spike count in the jitter window within each trial. As a result, jitter correction removes the correlation between PSTHs (stimulus locking) and those on time scales longer than the jitter window (slow population correlations). We chose a 25 ms jitter window, following previous work<sup>45,46,76,77</sup>.

We classified a CCG as significant if the peak of the jitter-corrected CCG occurred within 10 ms of zero and was more than seven standard deviations above the mean of a high-lag baseline period ( $100 > |\tau| > 50$ )<sup>45</sup>. Zero-lag CCGs were excluded from the analyses reported here, although their inclusion yielded statistically indistinguishable results.

All CCGs were estimated using spike trains during the memory delay (500–1,400 ms after cue) to avoid the influence of visually evoked responses. CCG analyses specific to On and Off states (Fig. 5f) were computed by first setting  $x(t)$  to zero for all time points not identified as On or Off (respectively), and then repeating the analysis described above.

## Manhattan distance

To determine whether patterns of functional connectivity differed according to the contents of memory, we compared the graphs of significant CCGs across cue locations in a pairwise manner (Fig. 5c–f). For each session and cue location, we represented the results of our CCG analyses as a graph in which nodes were units. The edge (connection) between each pair of units was assigned a weight of 1 if the pair had a significant CCG, and zero otherwise. Then, for each possible pair of cue locations, we computed the Manhattan distance, the number of edges with a weight that differed across the two graphs. Finally, we averaged this metric across all 28 possible pairs of conditions, yielding one summary statistic per session.

To normalize this mean Manhattan distance for comparison across sessions, we shuffled the cue location labels within each pair of neurons for each pair of conditions under consideration across trials, and repeated the entire analysis pipeline 50 times (25 for analyses specific to On and Off states), from CCG estimation through Manhattan distance calculation. We then  $z$ -scored the mean Manhattan distance for each session by this null distribution and compared these  $z$ -scores to zero (Fig. 5f).

Note that CCGs among both single and multi-units have been widely used as a measure of functional connectivity<sup>78–84</sup>. Indeed, CCGs based on multi-unit activity may be more sensitive in detection of correlations in spiking than similar analyses of single-neuron pairs<sup>78,85,86</sup>. Nonetheless, the presence of multi-units in our dataset does limit the conclusions that might be drawn about the specific neuronal subtypes involved in the cue-dependent ensembles that we observe—for example, putative pyramidal versus non-pyramidal neurons.

## Firing-rate-matched control

The geometric mean firing rate of pairs of units varied significantly across the eight cue locations (one-way ANOVA,  $P = 0.002$ ; Fig. 6a). Geometric mean firing rates were statistically indistinguishable, however, across cue locations 1–4 ( $P = 0.332$ ) and 5–8 ( $P = 0.884$ ). Therefore, we repeated the analysis of Manhattan distance described above, this time computing it among only cue locations 1–4 and 5–8 (Extended Data Fig. 6b), to yield a firing-rate-matched variant of the analysis presented in Fig. 5f.

## Joint selectivity

To determine the selectivity of units during the evoked response, we averaged each unit's cue-locked firing rate over time (0–400 ms after cue onset), yielding an  $n_{\text{Trials}} \times 1$  vector of firing rates. We then performed one-way ANOVA to evaluate the relationship between cue location and firing rate. If the effect of cue location was significant ( $P < 0.05$ ), the unit was deemed selective to cue location, and the location at which it had the greatest mean firing rate was labelled the preferred location. Pairs of units were deemed jointly selective if they were selective for the same cue location.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The data underlying this study are available at Dryad (<https://doi.org/10.5061/dryad.kkwh70sct>)<sup>87</sup>.

## Code availability

The code underlying this study is available at GitHub (<https://github.com/panichem/SingleTrialDynamics>).

66. Jonikaitis, D., Noudoost, B. & Moore, T. Dissociating the contributions of frontal eye field activity to spatial working memory and motor preparation. *J. Neurosci.* **43**, 8681–8689 (2023).
67. Bruce, C. J., Goldberg, M. E., Bushnell, M. C. & Stanton, G. B. Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. *J. Neurophysiol.* **54**, 714–734 (1985).
68. Pachitariu, M., Sridhar, S., Pennington, J. & Stringer, C. Spike sorting with Kilosort4. *Nat. Methods* **21**, 914–921 (2024).
69. Bruce, C. J. & Goldberg, M. E. Primate frontal eye fields. I. Single neurons discharging before saccades. *J. Neurophysiol.* **53**, 603–635 (1985).
70. VerKulen, J. & Smithson, M. Mixed and mixture regression models for continuous bounded responses using the beta distribution. *J. Educ. Behav. Stat.* <https://doi.org/10.3102/1076998610396895> (2016).
71. Bair, W. & O'Keefe, L. P. The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Vis. Neurosci.* **15**, 779–786 (1998).
72. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
73. Sceniak, M. P., Hawken, M. J. & Shapley, R. Visual spatial characterization of macaque V1 neurons. *J. Neurophysiol.* **85**, 1873–1887 (2001).
74. Harrison, M. T. & Geman, S. A rate and history-preserving resampling algorithm for neural spike trains. *Neural Comput.* **21**, 1244–1258 (2009).
75. Smith, M. A. & Kohn, A. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* **28**, 12591–12603 (2008).
76. Jia, X., Tanabe, S. & Kohn, A. Gamma and the coordination of spiking activity in early visual cortex. *Neuron* **77**, 762–774 (2013).
77. Zandvakili, A. & Kohn, A. Coordinated neuronal activity enhances corticocortical communication. *Neuron* **87**, 827–839 (2015).
78. deCharms, R. C. & Merzenich, M. M. Primary cortical representation of sounds by the coordination of action-potential timing. *Nature* **381**, 610–613 (1996).
79. Eckhorn, R. et al. Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biol. Cybern.* **60**, 121–130 (1988).
80. Engel, A. K., König, P., Gray, C. M. & Singer, W. Stimulus-dependent neuronal oscillations in cat visual cortex: inter-columnar interaction as determined by cross-correlation analysis. *Eur. J. Neurosci.* **2**, 588–606 (1990).
81. Gray, C. M., Engel, A. K., König, P. & Singer, W. Synchronization of oscillatory neuronal responses in cat striate cortex: temporal properties. *Vis. Neurosci.* **8**, 337–347 (1992).
82. Gray, C. M. & Singer, W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl Acad. Sci. USA* **86**, 1698–1702 (1989).
83. Luczak, A., McNaughton, B. L. & Harris, K. D. Packet-based communication in the cortex. *Nat. Rev. Neurosci.* **16**, 745–755 (2015).
84. Tanaka, H., Tamura, H. & Ohzawa, I. Spatial range and laminar structures of neuronal correlations in the cat primary visual cortex. *J. Neurophysiol.* **112**, 705–718 (2014).
85. Bedenbaugh, P. & Gerstein, G. L. Multiunit normalized cross correlation differs from the average single-unit normalized correlation. *Neural Comput.* **9**, 1265–1275 (1997).
86. Roy, S. & Alloway, K. D. Synchronization of local neural networks in the somatosensory cortex: a comparison of stationary and moving stimuli. *J. Neurophysiol.* **81**, 999–1013 (1999).
87. Panichello, M. F. Intermittent rate coding and cue-specific neuronal ensembles support working memory. *Dryad* <https://doi.org/10.5061/dryad.kkwh70sct> (2024).

**Acknowledgements** We thank D. Lopes and S. Cital for veterinary assistance, M. Solyali and B. Schneeweis for machining assistance and T. Engel, S. Druckmann, A. Soltani and X. Chen for helpful comments. This work was supported by NIH grant nos. EY014924 and NS116623, and by a Ben Barres Professorship to T.M.

**Author contributions** M.F.P. and T.M. conceived the project. D.J. and T.M. designed the tasks. D.J. and M.P. trained the animals. T.M., D.J. and M.F.P. carried out surgical procedures. M.F.P. and D.J. collected data. M.F.P. and J.Y.O. analysed the data. S.Z. and E.B.T. provided code for computing CCGs. T.M. supervised the work. T.M. acquired funding. M.F.P. and T.M. wrote the first draft. All authors reviewed and edited the manuscript.

**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-024-08139-9>.

**Correspondence and requests for materials** should be addressed to Matthew F. Panichello or Tirin Moore.

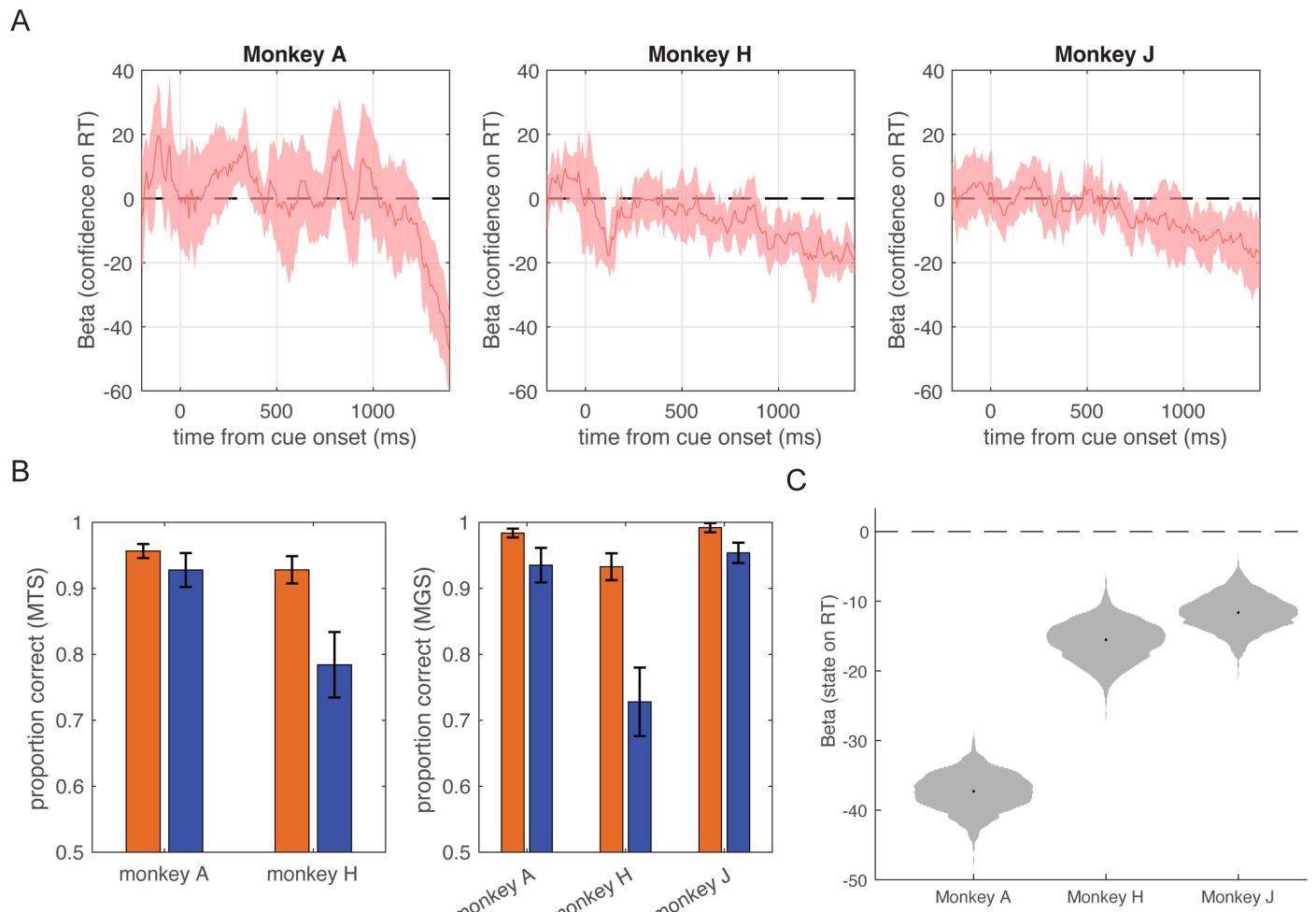
**Peer review information** *Nature* thanks Erin Rich, Eelke Spaak and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



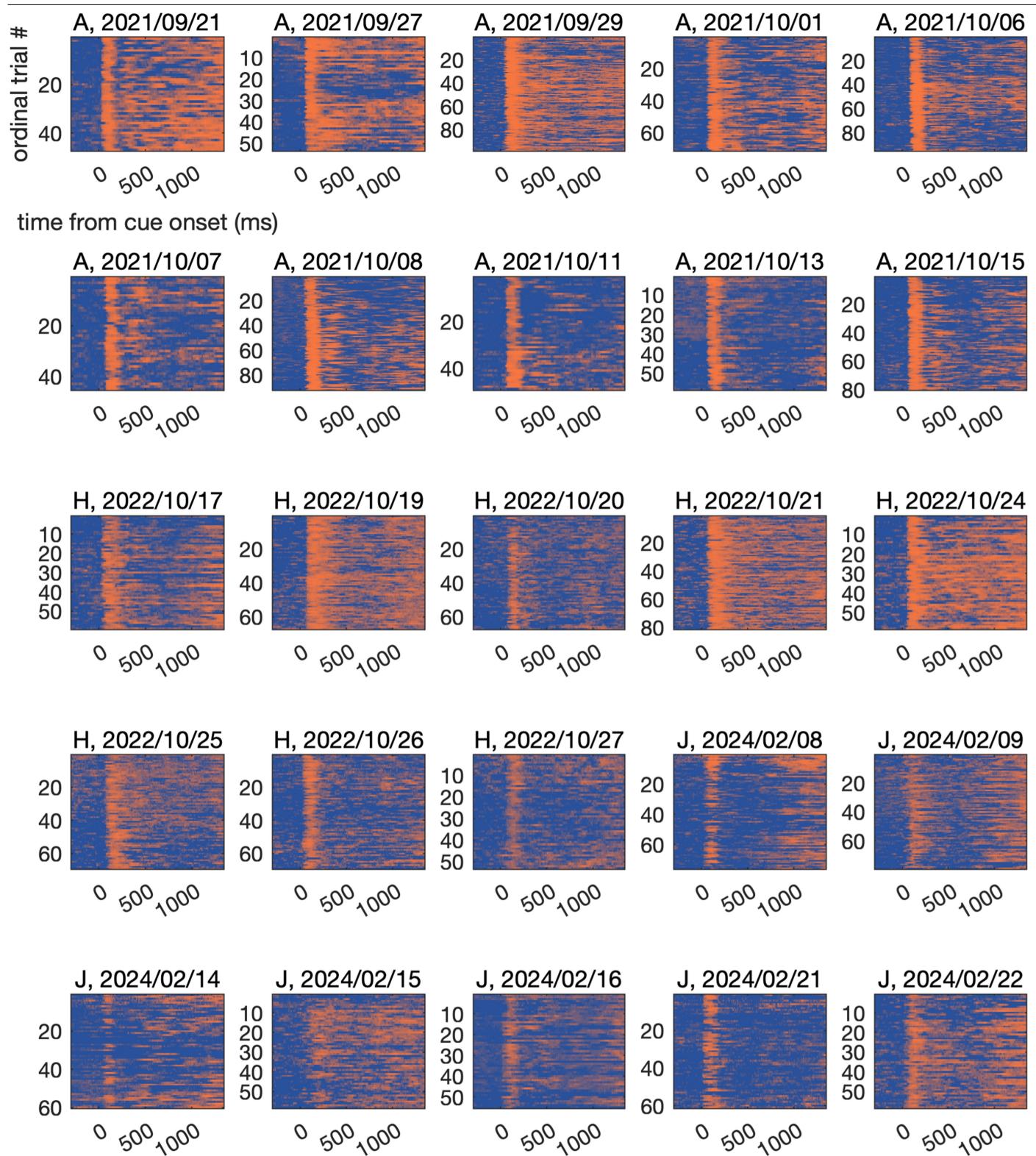
**Extended Data Fig. 1 | Normalized spatial distance between different functional classes of neurons within lateral prefrontal cortex.** Seven functional classes of neurons were defined according to selective activity to three task components: visual, delay, and motor. Neurons were defined as having a given functional property based on the presence of significant selectivity across cue conditions within the visual, delay and motor epochs of the task. The plot shows the mean distance between different classes of neurons; means were calculated and normalized by each session's total mean distance. Top-left to bottom-right diagonal elements show the mean distance within each functional class. Interestingly, we did not observe a relationship between depth and different functional subclasses (not shown), although this result should be interpreted with caution due to our non-perpendicular electrode penetrations.



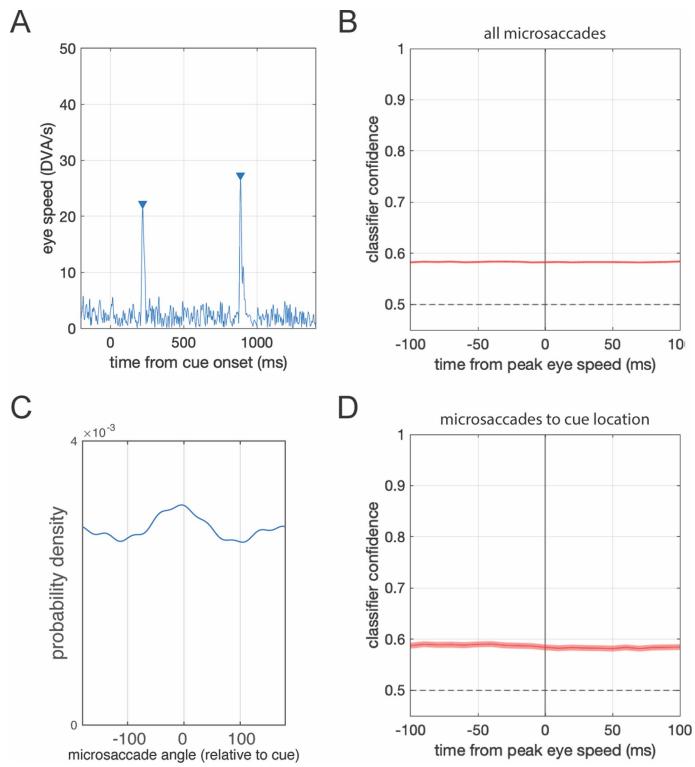
**Extended Data Fig. 2 | Confidence and state predict behavior.** (A) Regression coefficient relating classifier confidence to reaction time in milliseconds (average across sessions). Cue location and task (MTS/MGS) were included as co-regressors. Shaded area: 95% confidence intervals. (B) Mean proportion correct in the delayed match-to-sample (MTS) and memory-guided saccade (MGS) tasks, binned by state (orange = On, blue = Off) immediately after

fixation offset. Error bars are 95% confidence intervals (binomial). (C) Regression coefficient relating state (On/Off) immediately after fixation offset to reaction time in milliseconds. Cue location and task (MTS/MGS) were included as co-regressors. Negative values indicate that On states are associated with reduced reaction times. Violin plots show bootstrap across sessions.

## Article

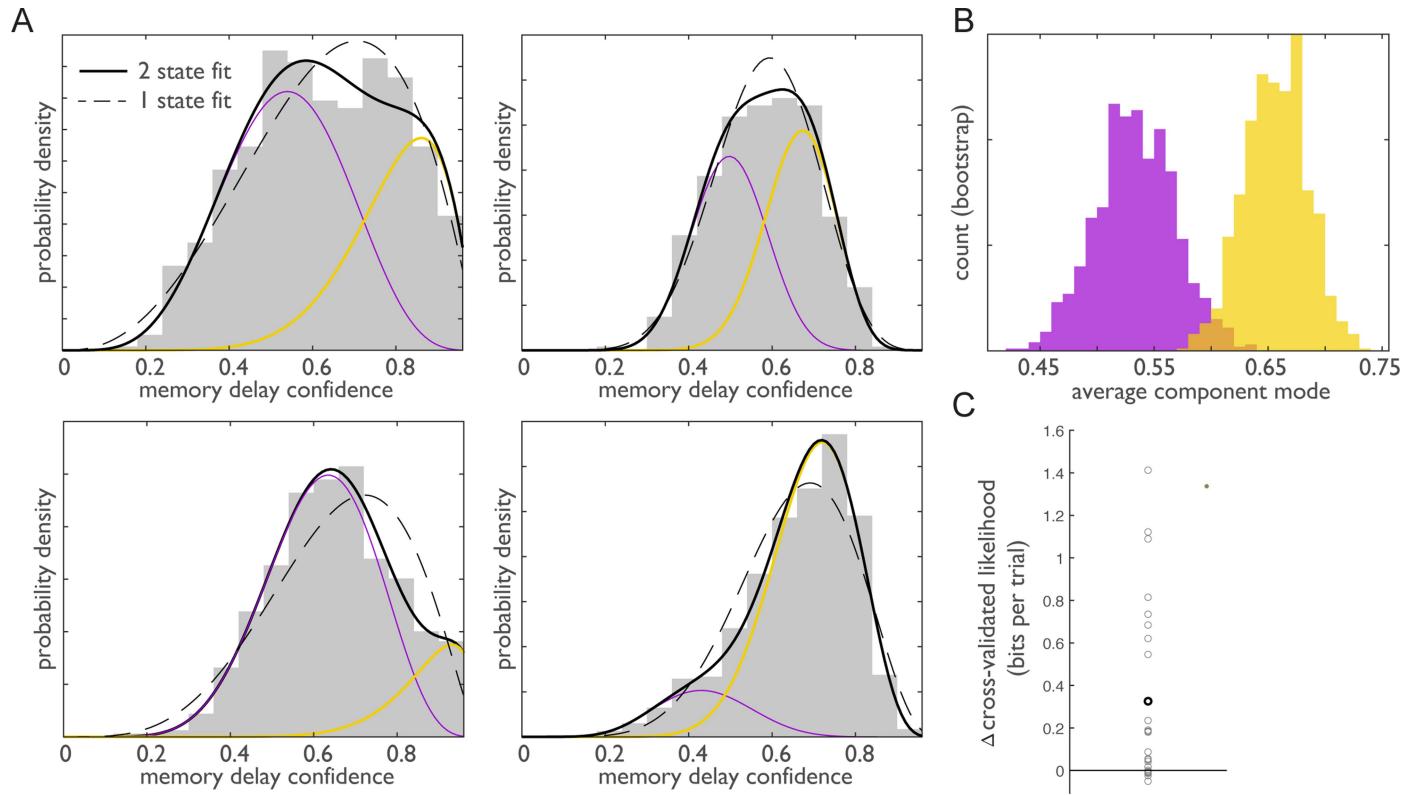


**Extended Data Fig. 3 | Single-trial classifier confidence, relative to cue onset, for all trials from the most preferred cue condition for each session.** Color scale as in Fig. 2a.



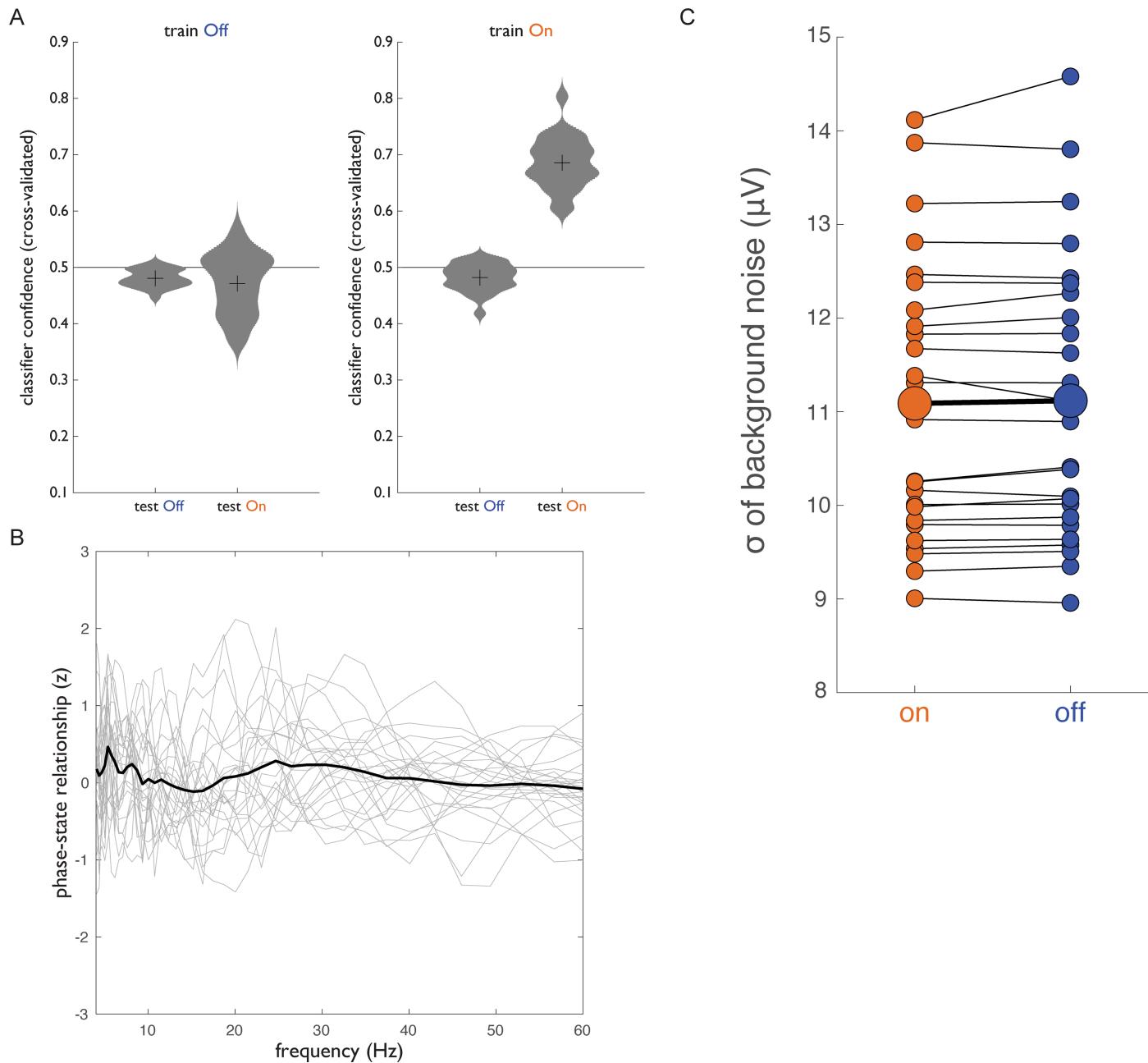
**Extended Data Fig. 4 | Classifier confidence was not affected by microsaccades.** (A) Example trial showing eye speed data and microsaccade identification. Microsaccades were identified as peaks in eye speed >10 DVA/s. (B) Mean classifier confidence during the memory delay, locked to microsaccades (average across  $N = 8,910$  microsaccades). Shaded area (small) indicates 95% confidence intervals. (C) Probability density of microsaccade angles during the memory delay, relative to the cued location. (D) Mean classifier confidence during the memory delay, locked to microsaccades to the cued location ( $\pm 20$  degrees).

# Article



**Extended Data Fig. 5 | Mixture modeling of behavior.** (A) Histogram of confidence values during memory delay time points (one cue location) from three example sessions. Dashed line shows the best fitting beta distribution. Solid line shows the best fitting mixture of two beta distributions. Purple and yellow lines show the individual component distributions for the two-state model. (B) Histogram showing average proportion-weighted component

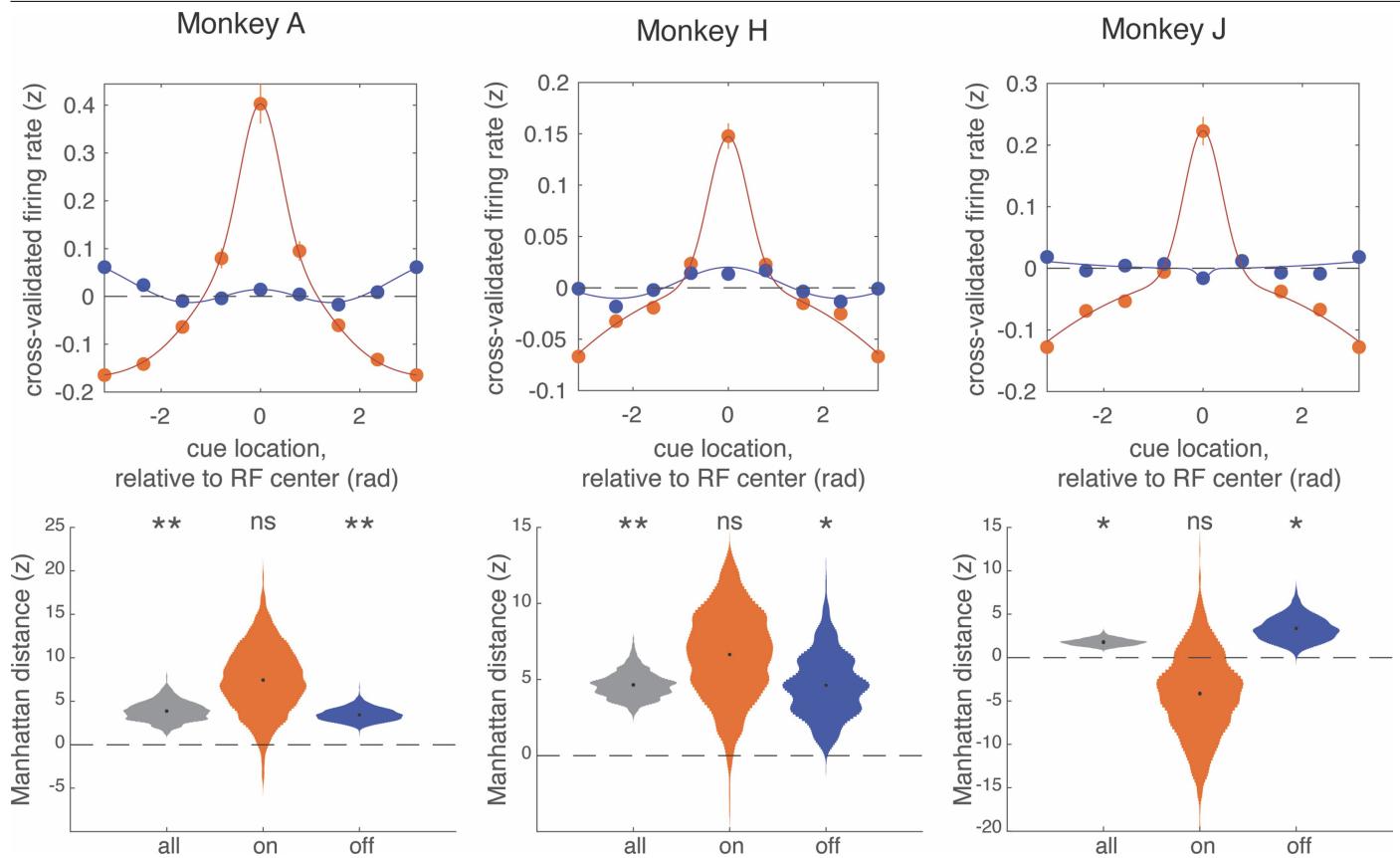
modes (bootstrap). (C) Cross-validated model comparison results for 2-state vs 1-state fits for all sessions ( $N = 25$ ). The 2-state model outperforms the 1-state ( $p = 0.009$ , chi-squared). Light circles show individual session scores; dark circle shows mean across sessions; black line shows equivalent model performance.



**Extended Data Fig. 6 | Off states were not explained by an alternate population code, rhythmic activity, or background noise.** (A) Cross-validated classifier confidence, binned by training and test state. Left: classifier confidence when trained on Off states and tested on Off and On states. Right: classifier confidence when trained on On states and tested on Off states. Violin plots show bootstrap across sessions. (B) Mean phase-state relationship

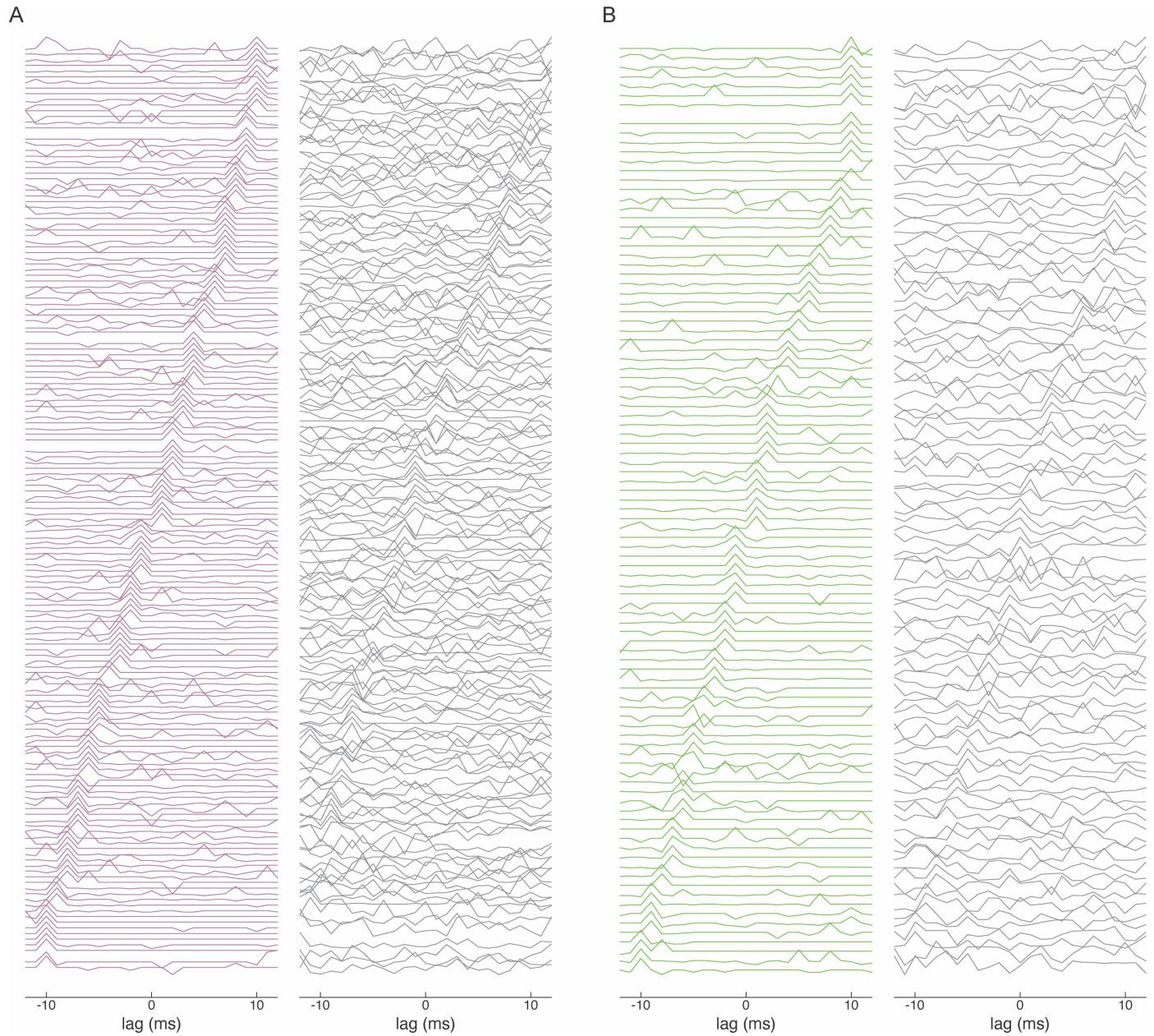
(z-scored) as a function of frequency. Gray traces show individual sessions ( $N = 25$ ). Values reflect absolute angular difference between the mean phase of On and Off states, normalized by the null (randomization test). (C) Standard deviation of the background noise, in microvolts, averaged across all On and Off states. Small circles: individual session means; large circles: grand means. Background noise did not differ across On and Off states ( $p = 0.339$ , sign-rank).

# Article



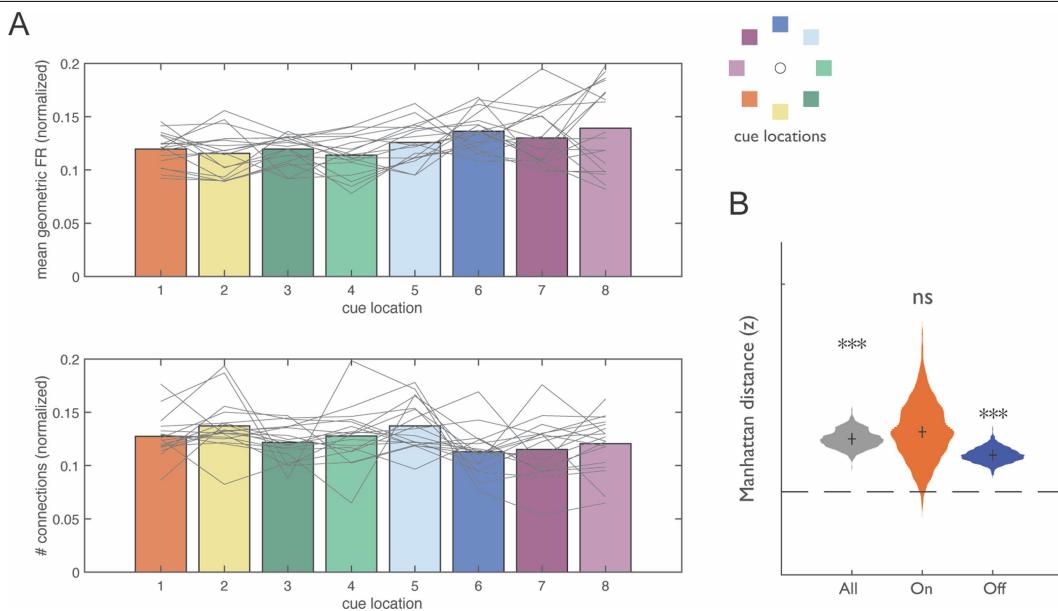
**Extended Data Fig. 7 | Results for each animal.** Top: Single-animal memory tuning functions for held-out units during On and Off states. Tuning functions show the mean normalized firing rate during the memory delay ( $z$ -scored across trials) for held-out units, relative to each unit's preferred cue location.

Error bars (small) denote SEM. Bottom: Single-animal mean normalized Manhattan distance, using data from the entire memory delay (gray), only during On states (orange), and only during Off states (blue). Violin plots show bootstrap across sessions. Asterisks and 'ns' denote significance.



**Extended Data Fig. 8 | Example cross-correlograms (CCGs) with significant and non-significant peaks.** (A) All CCG functions from the “left” cue condition with a significant non-zero lag peak (pink), along with an equal number of

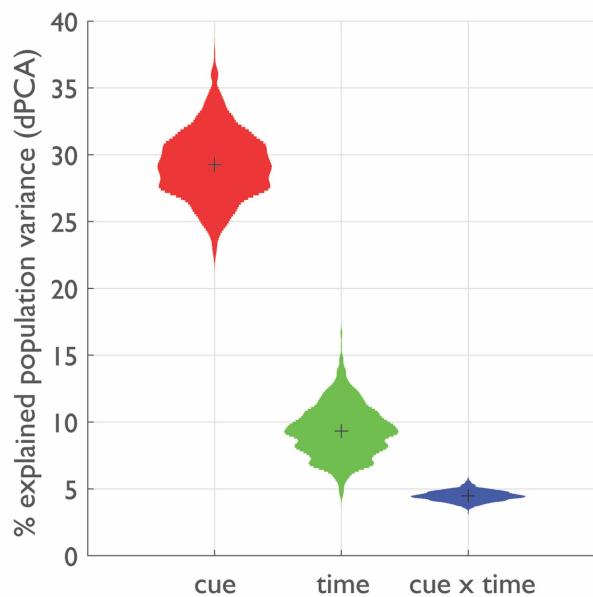
randomly selected “non-significant” CCGs (grey) from that same cue condition. (B) Same as in A, for the “right” cue condition. All CCGs in each subplot are arranged by the lag of their maximum value.



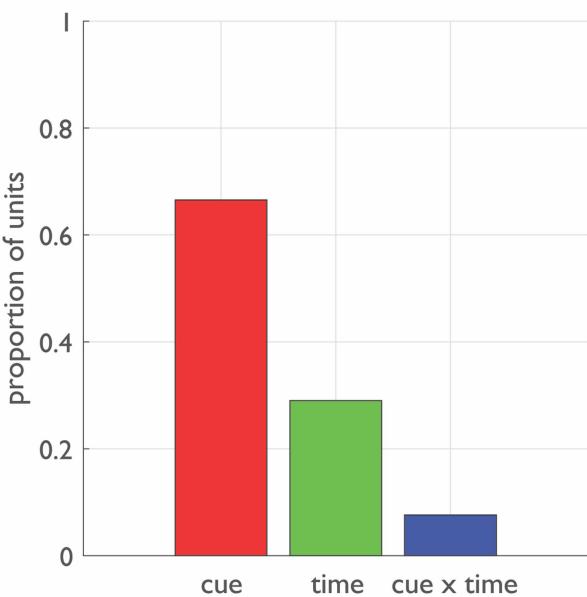
**Extended Data Fig. 9 | Firing rate matching for CCG analysis.** (A) Top: Normalized geometric mean firing rate (gFR) of all pairs of neurons during the memory delay. Bottom: Normalized number of neuronal pairs with significant CCGs. Gray lines show individual sessions; bars show mean across sessions; colors indicate cue location (lower inset). (B) Mean normalized Manhattan

distances restricted to a comparison among cue conditions (1-4 and 5-8) in which gFR was equal (see Methods). Shown are means of data from all sessions during the entire memory delay (gray), only during On states (orange), and only during Off states. Violin plots show bootstrap across sessions.

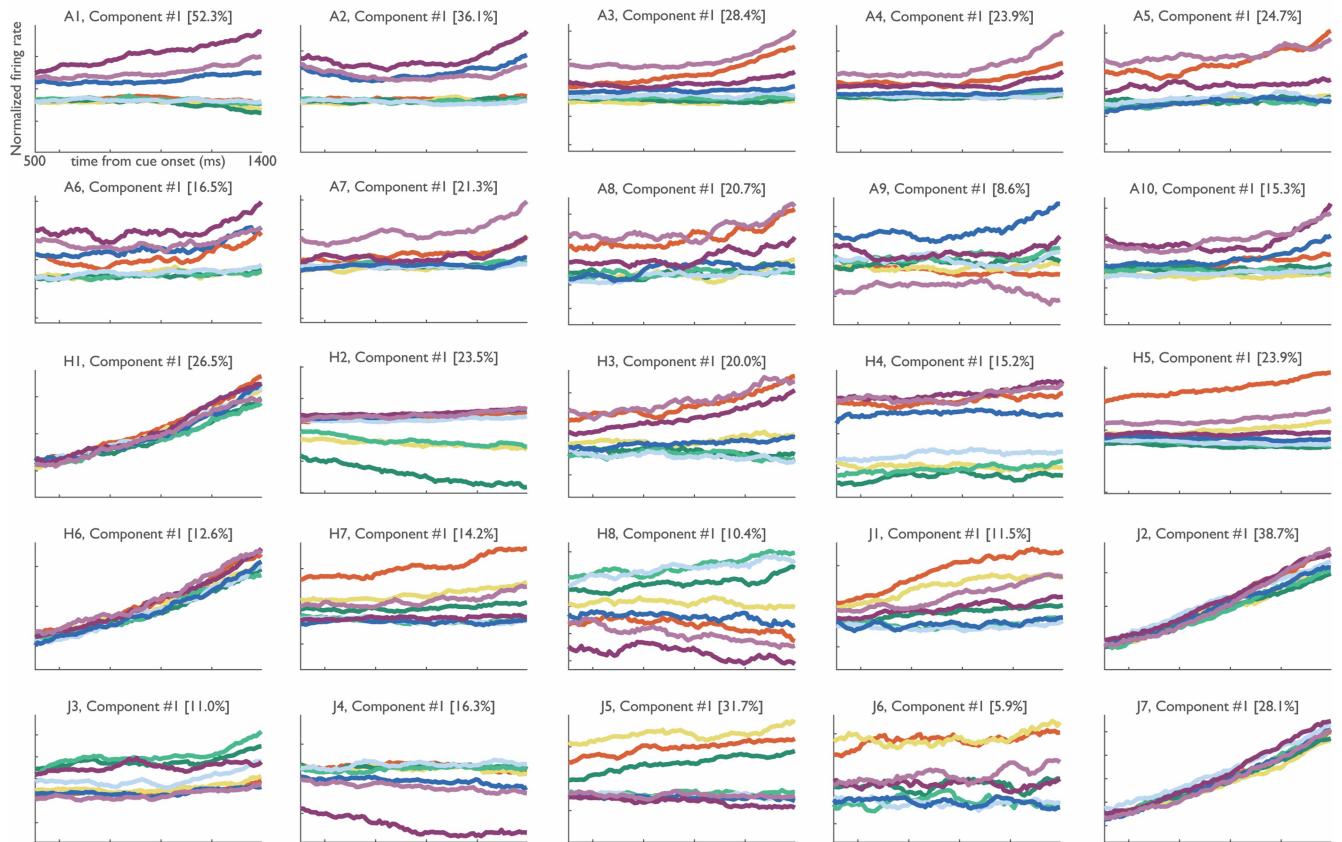
A



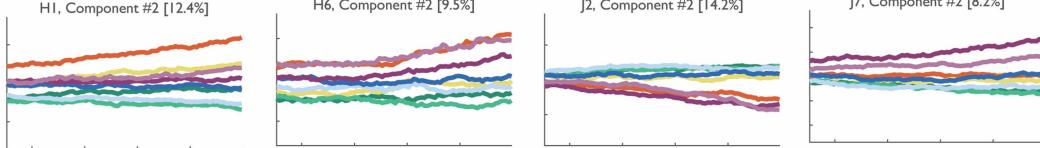
D



B



C

**Extended Data Fig. 10** | See next page for caption.

## Article

**Extended Data Fig. 10 | Trial-averaged rate coding is predominantly stable over the memory delay.** (A) Proportion of variance explained by each of the task components (dPCA). Violin plots reflect bootstrap across sessions. (B) Stable coding was also evident in the demixed components identified by dPCA. As with conventional PCA, each of these components can be understood as a linear combination (or, critically, a linear readout) of the population of trial-averaged single-neuron PSTHs. For 21 of the 25 total experimental sessions, the dPCA component explaining the most variance in the data encoded the cue in a

time-invariant manner. The largest component from the remaining 4 sessions was a cue-invariant ramp unrelated to mnemonic coding; for these sessions (C), the second component captured time-invariant information about the cue. Consistent with the analysis of explained variance, none of these components displayed sequential coding or other time-varying signals. (D) Proportion of neurons displaying a significant effect of cue, time, and their interaction during the memory delay.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	Stimuli were presented and behavioral responses registered using Psychtoolbox (version 3) and Matlab R2022a. Code for stimulus presentation is available at <a href="https://github.com/panichem/SingleTrialDynamics">https://github.com/panichem/SingleTrialDynamics</a> . Neuropixels data were acquired using SpikeGLX. Waveforms were sorted using Kilosort3.
Data analysis	Data were analyzed using built-in functions and custom code written in Matlab. Built-in functions are noted in the associated methods section (e.g., fitclinear for logistic regression). Equations for non-standard statistics are provided in the methods. Code for custom functions are provided in the manuscript as a reference to the original source or are available on GitHub ( <a href="https://github.com/panichem/SingleTrialDynamics">https://github.com/panichem/SingleTrialDynamics</a> ).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data underlying this study are available on Dryad (<https://doi.org/10.5061/dryad.kkwh70sct>).

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

n/a

Reporting on race, ethnicity, or other socially relevant groupings

n/a

Population characteristics

n/a

Recruitment

n/a

Ethics oversight

n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

As detailed in the manuscript, a total of 8,225 single and multi-unit neurons were recorded from 3 animal subjects across 25 experimental sessions. The number of subjects (3) is above the standard in the field (Fries & Maris, 2022). The number of neurons is in line with the current state-of-the-art (Trautmann et al., 2023).

Data exclusions

Neurons with very low firing rates were excluded from further analysis. Exclusion criteria are detailed in the methods.

Replication

Independent experiments were performed in 3 animals. All data were analyzed, except for the exclusion criteria described above. There were no failed replication attempts (i.e., no animals failed to learn the task and no neural recordings were excluded).

Randomization

Each animal was exposed to every task manipulation. Within a session, task manipulations were randomized across trials. Neurons were recorded without bias, with electrodes placed to maximize signal-to-noise of the electrophysiological signal.

Blinding

All animals were assigned to a single experimental group, and so blinding was not necessary or possible. However, experimenters were blinded to experimental conditions during recording of neurons and during sorting of waveforms into single neurons.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

**Materials & experimental systems**

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern
<input checked="" type="checkbox"/>	Plants

**Methods**

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

**Animals and other research organisms**

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	Subjects were three adult male rhesus macaques ( <i>Macaca mulatta</i> ), ages 8, 11, and 12.
Wild animals	No wild animals were used in the study
Reporting on sex	As monkey electrophysiology studies typically report findings from only a few subjects (Fries & Maris, 2022), isolating effects due to gender is typically not possible. All research subjects in the present study are male.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All experimental procedures were approved by the Stanford University Animal Care and Use Committee and were in accordance with the policies and procedures of the National Institutes of Health.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

**Plants**

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a