Materials and Methods

We used the same two adult male rhesus macaques (Macaca mulatta) in the main experiment as well as the ketamine and saline experiments (age: 10, 9; weight: 12, 10 kg). The oculomotor delayed response task labelled ODR1 was recorded from the same animals, while ODR2 was recorded from two different male macaques using one multielectrode Utah array implanted in each animal (Leavitt, 2017b, Leavitt, 2018).

Ethics statement. Animal care and handling (i.e., basic care, animal training, surgical procedures, and experimental injections) were pre-approved by the University of Western Ontario Animal Care Committee. This approval ensures that federal (Canadian Council on Animal Care), provincial (Ontario Animals in Research Act), regulatory bodies (e.g., CIHR/NSERC), and other national standards (CALAM) for the ethical use of animals are followed. The oculomotor delayed response task experiment complied with Canadian policies and regulations and was preapproved by the McGill University Animal Care Committee (Leavitt, 2017b, Leavitt, 2018). Regular assessments for physical and psychological well-being of the animals were conducted by researchers, registered veterinary technicians, and veterinarians.

Task Description. Each trial begins when the virtual environment appears on screen with one target location cued (Fig. 1d, red vertical bar). The cue remains on screen for 3 seconds. The disappearance of the cue is followed by a 2 second memory delay period. During both the cue and memory periods (or, "epochs"), the subject is free to visually explore the space, but navigation is disabled. After 2 seconds, navigation is enabled, and the subject navigates to the remembered location using a joystick. The subject correctly completes a trial by reaching the target location. (See movie S1.) They are not required to stop at the target, but only a small radius of space will trigger a correct response and a juice reward ($\frac{1}{2}$ grape-apple juice, $\frac{1}{2}$ water). The responses take varying amounts of time depending on the trial and the target condition; however, only 2 seconds of navigation were included in the analysis. This length of time was chosen to ensure the subject had not yet reached the target location and received a reward. During the trial, if after 10 seconds of navigation the subject has yet to reach the target location, the trial times out and is considered incorrect. At the completion of a trial, the subject is virtually "teleported" into a black box (blank, black screen) to await the next trial. The inter-trial-intervals are variable, but around 500ms on average. No additional visual or auditory cues accompany the reward or boundaries between task epochs.

The virtual task environment was developed using Unreal Engine 3 development kit, utilizing Kismet sequencing and UnrealScript (UDK, May 2012 release; Epic Games). Details about this platform and the recording setup can be found in Doucet et al., 2016. Movement speed through the environment was fixed. Target locations within the virtual arena were arranged in a 3 × 3 grid and spaced 290 unreal units apart (time between adjacent targets is approximately 0.5 seconds). The perception-navigation control variation of the task (PNC) was identical to the working memory version except that the targets remained onscreen through the trial. In this control, navigation was guided by perception, and WM was not required.

Experimental setup. Animals performed the task in an isolated room with no illumination other than the monitor. The room contained no AC power lines and was radiofrequency (RF) shielded. The task was presented on a computer LDC monitor positioned 80 cm from the subjects' eyes (27" ASUS, VG278H monitor, 1024 × 768 pixel resolution, 75 Hz refresh rate, screen height equals 33.5 cm, screen width equals 45 cm). Eye positions were monitored using a video-oculography system with sampling at 500 Hz (EyeLink 1000, SR Research). Stimulus presentation was controlled through a custom computer program (through Unreal Engine 3). Subjects were seated in a standard enclosed primate chair (Neuronitek) during the experiment and were delivered juice through an electronic reward integration system (Crist Instruments). Prior to the experiments, subjects were implanted with custom fit, PEEK cranial implants which housed the head posts and recording equipment (Neuronitek). See Blonde et al., 2018 for more information. The head posts were attached to the primate chair for head fixation. The experimental setup for the oculomotor delayed response task is outlined in both Leavitt et al. 2017b and Leavitt, 2018.

Ketamine injection. The ketamine doses were titrated so they did not induce visible behavioral changes in the animals (i.e., nystagmus or somnolence). An intramuscular injection of ketamine (Narketan, 0.25, 0.4, or 0.8 mg/kg) was administered in the hamstring muscles by a registered veterinary technician. Ketamine injections were spaced at least two days apart to allow for washout of the drug. Saline administration was conducted identically with a fixed 0.25 mg/kg dose (See Roussy et al. 2021 for more details).

Surgical procedure. Custom PEEK implants which housed recording hardware and a headpost were developed and implanted in each animal (Blonde et al., 2018). Brain navigation for surgical planning was conducted using Brainsight (Rogue Research Inc.) (Extended Data Fig. 10a). Two 10×10, microelectrode Utah arrays (96 channels, 1.5 mm in length and separated by at least 0.4 mm) (Blackrock Neurotech) were chronically implanted in each animal. Electrodes were implanted in the left LPFC (anterior to the arcuate sulcus and on either side of the posterior end of the principal sulcus) (Extended Data Fig. 10b, c). Arrays were impacted approximately 1.5 mm into the cortex. Reference wires were placed beneath the dura and a grounding wire was attached between screws in contact with the pedestal and the border of the craniotomy. Surgical procedures were conducted under general anesthesia induced by ketamine and maintained using isoflurane and propofol.

For the oculomotor delayed response task data, a 96-channel Utah array was implanted in each monkey's left LPFC in the same region that electrodes were implanted for recording during performance of the virtual working memory task (Extended Data Fig. 8b). Detailed surgical methods can be found in Leavitt et al. (2017b, 2018).

ODR Task Details.

ODR1. The oculomotor delayed response task performed by NHP B and NHP T was separated into four epochs: fixation, stimulus presentation, delay, and response. The fixation period duration was 300ms (NHP B) or 500ms (NHP T). The cue was presented for 1000 ms. The cue then disappeared and the subject maintained fixation for an additional 3000ms (NHP B) or 1500ms (NHP T), after which the fixation point disappeared and the subjects responded by making a saccade to the remembered location.

ODR2. The oculomotor delayed response task performed by NHP JL and NHP F was also separated into four epochs: fixation, stimulus presentation, delay, and response. The animal began a trial by fixating on a fixation dot and by pressing a lever. The duration of the fixation period was either 482, 636, or 789 milliseconds. A sine-wave grating target then appeared at 1 of 16 randomly selected locations positioned in a 4x4 grid for 505 ms. This was followed by a delay period ranging from 494–1500 ms. The fixation point was removed, cueing the animal to make a saccade to the location of the previously presented target and then to release the lever (see Leavitt et al. 2017b, 2018 for more details).

Analysis

Task performance. Correct trials are trials in which the animal reaches the correct target location within 10 seconds. An incorrect trial occurs if the animal does not reach the target location within 10 seconds. Percent of correct trials is calculated as the number of correct trials divided by the total number of trials. Response time (Fig. S1a) was calculated for correct trials as the time from the start of navigation to the time in which the animal reaches the correct target location.

The optimal trajectory analysis (Fig. S1b) was calculated for correct trials. It is calculated as the real length of the animal's trajectory to correct target location divided by the length of the optimal trajectory (i.e., the Euclidean distance from the start position to the target location).

For incorrect trials, we calculated the distance from the animal's final position to the correct target location (Fig. S2d). Distance values were modified from arbitrary 'Unreal' units (the unit system in Unreal Engine Development Kit, Unreal Engine 3, Epic Games) to 'Unreal' units divided by the distance between two targets to increase interpretability. A new value of 1 would represent 290 unreal units (the distance between two adjacent targets).

Eye behavior. Percent of eyes on screen (Fig. S2a) measures the number of eye data points falling on the screen divided by the total number of eye data points. Off screen data points occur when the animal looks off screen or closes their eyes (as occurs during blinking).

Eye data was classified into fixations and saccades based on a method outlined in Corrigan et al., 2017 that was developed for use in a similar virtual environment. The percent of fixations on target (Fig. S2c) was calculated by the number of fixation events falling within a trial's target location divided by total number of fixation events. We used a linear classifier (SVM) (Libsvm 3.14, Fan et al., 2008) with 5-fold cross validation to predict target location from eye fixation position data (Fig. S2d,e).

The main sequence (Fig. S2f) was calculated by separating saccades into bins of 3° of amplitude, starting at 2° and computing the medians for each bin. The proportion of single units tuned for eye position in both retinocentric and spatiocentric reference frames was calculated using a quadrant binning pattern for a 40°×30° field (Fig. S@g). A bin had to have at least ten saccades to be acceptable and sessions had at least three acceptable bins.

Spike processing. Neuronal data was recorded using a Cerebus neuronal Signal Processor (Blackrock Microsystems) via a Cereport adapter. The neuronal signal was digitized (16 bit) at a sample rate of 30 kHz. Spike waveforms were detected online by thresholding at 3.4 standard deviations of the signal. The extracted spikes were semi-automatically resorted with techniques utilizing Plexon Offline Sorter (Plexon Inc.). Sorting results were then manually supervised. Multiunits consisted of threshold-crossing events from multiple neurons with action potential-like morphology that were not isolated well enough to be classified as a well-defined single unit (for spike sorting example see Extended Data Fig. 10d, e). We collected behavioral data across 20 working memory sessions (eight in animal T, twelve in animal B) and neural data across 17 sessions. This yielded a total of 3950 units recorded: 2578 single neurons (346 in animal T, 2232 in animal B) and 1372 multiunits (512 in animal T, 860 in animal B). We collected behavioral data across 18 ketamine-working memory sessions (nine in animal T, nine in animal B) and neuronal data from 17 ketamine-working memory sessions with one session from animal T removed due to incomplete synchronization of neuronal data during the recording. This yielded a total of 2906 units recorded during ketamine-working memory sessions: 1814 single neurons (259 in animal T, 1555 in animal B) and 1092 multiunits (533 in animal T, 559 in animal B).

Spike density function. Spike density functions (SDFs) were generated by convolving the spike train with a Gaussian kernel (standard deviation=100 ms).

Time consistent neurons. To quantify time consistency of neurons, we created SDFs combined between electrode arrays over the entire trial time using neurons with firing rates above 0.5 Hz. SDFs were created for each condition that contained at least five trials. We calculated the peak firing time for each neuron in the population, then calculated the standard deviation of the peak firing time for each neuron over all trials in a condition. Finally, we created a probability distribution from the standard

deviation values (Fig 2e). Correct and incorrect trials were considered separately. We then shuffled the peak firing times for each neuron from trial to trial and created a shuffled probability distribution. We calculated the difference in mean values between the real and shuffled distributions to get the mean difference value (Fig 2f,g). To calculate the standard deviation values plotted in Extended Data 3b, we calculated trial-trial standard deviation of peak spike time for the target condition in which each neuron fired the most consistently during correct trials (i.e., lowest deviation). The same conditions were used for shuffled data and for incorrect trials.

We repeated the same analysis for both ODR tasks (Fig. S8d,e) with the following slight variations:

ODR1: This data was collected from the same animals and electrodes as our naturalistic VR task. This task contained 16 targets (Extended Data Fig. 8a) and was a variation of a traditional ODR task in which the fixation point changes location across trials, resulting in many different task conditions with varying combinations of fixation location and target location. For this reason, we grouped trials with target locations within the same quadrant (same direction saccade). To match the task structure of the VR task, we did not use data from the fixation period.

ODR2: This data was collected from NHPs JL and F (Extended Data Fig. 7b) using one Utah array implanted in the left LPFC (same region as NHP B and T). This task contained 16 targets with a consistent central fixation point (Extended Data Fig. 8a). To match the task structure of the VR task, we did not use data from the fixation period. Since the ODR2 task had jittered delay epoch timing, we used trials with delay periods > 1000 ms and included the first 1000 ms of the epoch.

Sequence Representation. Each trial was represented as a vector of spike times, $\vec{S}$, with each component given by the time of maximum spike density of the corresponding neuron. For a recording session with N neurons and T trials, this process results in a set of T vectors of length N. Let t_n represent the within-trial time t at which neuron n reached its maximum spike density. Then $\vec{S}_n = t_n$. When the neurons were sorted in increasing order of t_n values, clear bands of spike bursts are visible which span both arrays and recording length (Fig 1g,h).

Time-Boundary Cells. Cells were considered time-selective if they contributed to the sequence at the same time, t_n, across trials of all 9 target conditions (i.e., their position in the sequence is consistent across trials and not related to target condition - see Fig. S5a,b for examples). To quantify this for each neuron, we computed the full-width half-max of the distribution of spike times across all trials in a recording session (see Fig. S5c). If the width fell below a threshold, that cell was labelled time-selective. We used thresholds of 1s for NHP B and 2s for NHP T; however, we note that a 2D scan across thresholds (height and width) suggests the specific thresholds chosen does not significantly impact results.

These time-selective cells were then used to define neural boundaries. We pooled the spike times t_n of all time-selective cells across all recording sessions for each subject. The histograms of these spike times reveal two clear peaks, anticipating the delay and navigation epochs (Fig. 3e). These peaks represent the times within the trial at which time-selective cells contribute most often to single-trial sequences. The peak times were used to define the neural time boundaries used in the following analyses. Note that no neural boundaries were found for the perception-navigation control task (Extended Data 6a), so the task boundaries were used for comparisons to that control task (Extended Data 4d).

Epoch-specific sequences. To consider a sequence specific to a given task epoch, the contribution of any cells with spike time t_n outside the desired epoch were set to NaN. In the main analysis, the neural boundaries were used to define these epochs (Fig. 3,4,5). However, results are not significantly impacted if the task boundaries were used to define the epochs instead (Fig S4a-c, S7a-c, S9g).

Dimensionality Reduction. For each recording session with T trials, a correlation matrix was created by computing the Pearson correlation coefficient between each pair of vectors $\vec{S}$ representing single trial sequences. Let ρ be the resulting $T \times T$ correlation matrix, with $\rho_{(i,j)} := \text{corr}(\vec{S}_i, \vec{S}_j)$ for trials $i, j \in [1, T]$. This matrix captures the similarity of neural sequences across trials. High-valued blocks of this matrix represent groups of of trials in which similar sequences occurred.

The correlation matrix ρ was then projected onto the eigenvectors corresponding to the eigenvalues with largest modulus to generate a low-dimensional summary of the data in 3D-space. This approach is similar to spectral clustering on the correlation matrix; however, rather than performing kmeans clustering on the eigenvectors, we treat them as axes of a "similarity space" into which we project the matrix of correlation values. The details of this projection are as follows:

The eigenvalues $\lambda_1 ... \lambda_T$ of ρ with corresponding eigenvectors $v_1 ... v_T$ were computed, and labelled such that $|\lambda_1| \leq |\lambda_2| \leq ... \leq |\lambda_T|$. Note that since ρ is a real-valued symmetric matrix, $\lambda_i \in \mathbb{R}$ for all $i \in [1, T]$. Let Q be the matrix with columns given by the first three eigenvectors of ρ: $Q = [v_1, v_2, v_3]$. Define the $T \times 3$ projection matrix $P := \rho Q$. This projection matrix is used to define a "similarity space" in which $(P_{(j,1)}, P_{(j,2)}, P_{(j,3)})$ describes the (x, y, z) coordinates of a representation of one trial j in 3-space.

The points in this projection each correspond to one trial, and their positions are determined by the relative similarity of the corresponding sequences. The centroids of the clusters corresponding to each trial condition were then determined, and the matrix of Euclidean distances between the centroids was computed. (See Extended Data 7e for an example.) In this way, the spiking data from each recording session was reduced to a 9×9 distance matrix (Fig 3f).

We note that considering higher dimensional projections does not significantly increase cluster separation in the projection (see Extended Data 7d).

Correlation Analysis. Visuospatial trajectories were computed for each trial by transforming the movement trajectories (in a 3D "Unreal" coordinate system) into 2D screen coordinates using a perspective projection matrix. Average visuospatial trajectories were computed for each recording session by averaging the screen coordinates of correct trial paths to each target in the virtual environment (excluding outlier trajectories with z-score > 1 of mean Frechet distance to other trajectories in the group).

A 9×9 distance matrix was then created for each recording session by computing the Frechet distance between each pair of average visuospatial trajectories (Alt Godau, 1995)(See Fig 3d for example trajectories and Fig.3g for example distance matrix). Intuitively, the Frechet distance can be understood as the shortest leash required to walk your dog if you followed one trajectory and your dog followed the other. This distance measure reflects differences in the curvature and separation between trajectories as well as the end positions. Further, since trajectories are computed from the behavior during each session separately, it reflects behavioural differences between subjects and differences across days.

The Spearman correlation coefficient between the trajectory distance matrix and the centroids distance matrix was computed for each recording session. This provides a measure for how much the organization of the centroids in 3-space reflects the behavior during the task. To compare correct versus incorrect trials (Fig. 3h, Extended Data 4c), and WM versus the perception-guided navigation control (Extended Data 4d), the relevant groups of trials were used to define separate sets of centroids. The correlation analysis was performed for each set of centroids, and the resulting correlation values were compared.

This procedure was repeated for several different aspects of the task structure, to determine which was best reflected by neural sequence structure (see Extended Data 4f). Average movement or "World" trajectories are trajectories in the 3D "Unreal" coordinate system, before the perspective projection was applied to create the visuospatial, or "Screen" trajectories. "Optimal" trajectories refer to the shortest path from the start location to each target location. "Target location" refers to the position of the target in the virtual arena. In all cases, the distances between the 9 conditions were computed using the relevant distance measure (Frechet for trajectories, Euclidean for target locations), to create the 9×9 distance matrix used in the analysis.

Ablation Test. A percentage of cells (from 10-90 percent) were randomly excluded (i.e. the corresponding sequence components S_j were set to NaN), and the correlation analysis was performed. This was repeated for 100 iterations at each percent of cells removed. The average correlation across iterations was reported (Extended Data 4e).

A similar analysis was used to compare the sequence contributions of cells that are classically tuned for target location during the delay versus untuned cells. Here, tuned and untuned population sequences were considered separately by setting the opposite population to NaN, and performing the correlation analysis. Tuned and untuned populations of matched size were considered: the population with fewer cells defined the size, N, of the matched population, and N cells were randomly selected from the other population to define the matched sequences. This random selection was repeated for 100 iterations and the resulting correlations were averaged (see Additional Statistics).

Decoding Analysis. All decoding analyses from Figure 4 (trial epoch, target location, target column, target row) were performed using variants of a simple classifier based on distances within the low-dimensional projection of the data described above (see Dimensionality Reduction for projection details). Decoding proceeded as follows:

The data set to be decoded is the set of points in 3-space, $\{P_j = (x_j, y_j, z_j)\}$, corresponding to the set of trials j in a given recording session. These data points have ground-truth labels $\{g_j\}$ (which could be trial epoch, target location, target column, or target row depending on the application).

Training and test sets were determined using 5-fold cross-validation. For each of the 5 iterations, the training set was used to determine condition centroids (i.e.,the coordinates of points P_j corresponding to the same target were averaged, following the procedure outlined in Dimensionality Reduction). The distances between each point P_j in the test set and each condition centroid were then computed. Each trial j was assigned a label c_j according to the condition that minimized this distance. Decoding accuracy was computed as the fraction of points in the test set assigned to the correct condition (i.e., the fraction of trials for which $c_j = g_j$), and the average accuracy across the 5 iterations was reported.

Since the 9 targets in the task are positioned as a 3x3 grid in the virtual arena, each (row, column) pair defines a target location. In Figure 4d, "combined target" decoding was performed by first decoding the target row and column (as described above), then using the predicted row and column to determine which of the 9 target locations each trial corresponds to. (Note that in this case, only row-selective and column-selective cells were included in the sequences to decode row and column respectively. See below for definition of selectivity.) In the same figure, "All Cell Target" decoding was performed as described above, by predicting the target location directly from the full sequences.

A. Multinomial Regression. An additional decoding analysis was performed to provide a preliminary comparison between persistent coding and the NAS code in this data. Cells are considered to display persistent delay activity if they satisfy two criteria: (1) the cell must display target-selective tuning for at least half of the delay period and (2) the cell must display a significantly higher firing rate for tuned targets during delay relative to baseline (with baseline rates determined from the inter-trial intervals). More specifically, the cell must display target-selective tuning in at least two not necessarily consecutive time windows of 500 ms during the delay, and in at least one of these time windows, the cell's delay firing must be elevated compared to baseline. With this definition, we find a set of persistent cells in this task, consistent with previous literature. We then train two multinomial regression (MNR) models. The first predicts trial condition using the delay activity of cells identified as persistent on each trial. The second predicts trial condition using the coordinates of the sequence projection for each trial (see Dimensionality Reduction for details). See Fig. S11c for the mean decoding accuracy of the two models (using 10-fold cross-validation).

Determining Neural Selectivity. Neural selectivity was determined from the set of times at which each cell contributed to the population sequence across trials of different conditions. **Time-selective cells** contributed at consistent times across all 9 target conditions (as described above). The contributions of **row-selective** and **column-selective** cells demonstrated structure that depended on the row or column of the target. To determine whether this was the case for a particular cell, the effect size was computed between the distributions of its spike times for each pair of conditions. The maximum effect size across condition pairs was considered. (For example, a cell with large effect size between the front and back rows could be considered row-selective, even if spike times for the front and middle rows were similar. The spike-times of such a cell would distinguish between the front and back row of a target, but make less distinction between the front 2 rows.) These effect sizes define a continuous measure of selectivity across cells. Cells with selectivity above a threshold percentile were labelled "selectiv". Thresholds were determined separately for each subject, so that the population of selective cells maximizes decoding performance of the condition for which they are selective (see Fig. S5d-g for this procedure).

Supplementary Figures

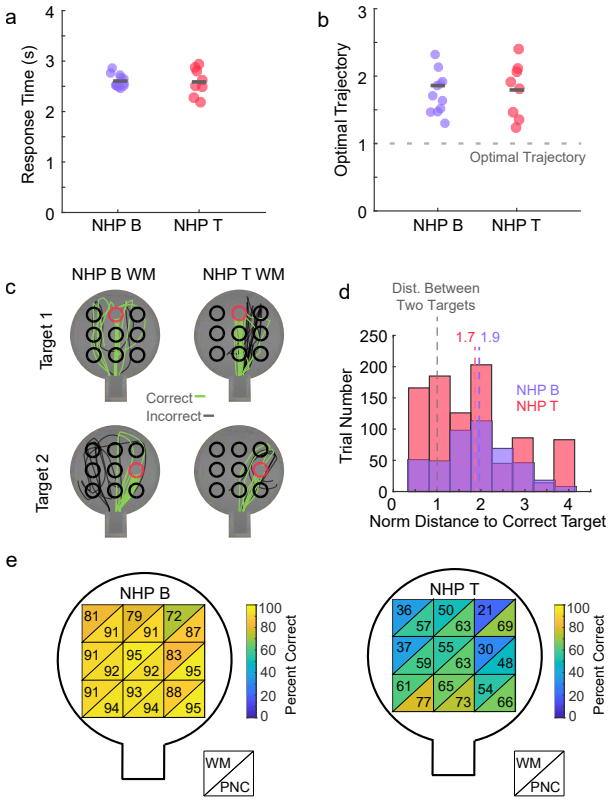

Fig. S1. Behaviour — Task Performance (**A**) Response time for correct trials. Black lines indicate the median values for each animal. Dots represent data for each session. (**B**) Optimal trajectory for correct trials. Calculated by real trajectory length/ Euclidean distance between the start location and target location. The dashed gray line indicates the optimal value of 1. Black lines represent median values and dots represent data for each session. (**C**) Example animal trajectories for two target locations. Red circles indicate the target, green lines represent correct trials and gray lines represent incorrect trials. (**D**) Distance from target for incorrect trials. Calculated as the Euclidean distance from the animal's end location to the correct target location. Since 'Unreal' units are arbitrary, distance values are normalized by the distance (in 'Unreal' units) between two targets. A normalized distance of 1 indicated by the dashed gray line is the distance between two target centers. Dashed purple and red lines represent median values. (**E**) Percent of correct trials to each target location, split by subject and task (WM compared to perception-navigation control (PNC)). Targets are depicted as arranged in the virtual circular arena. Panels 'a' and 'b' have been adapted from Roussy et al., 2022.

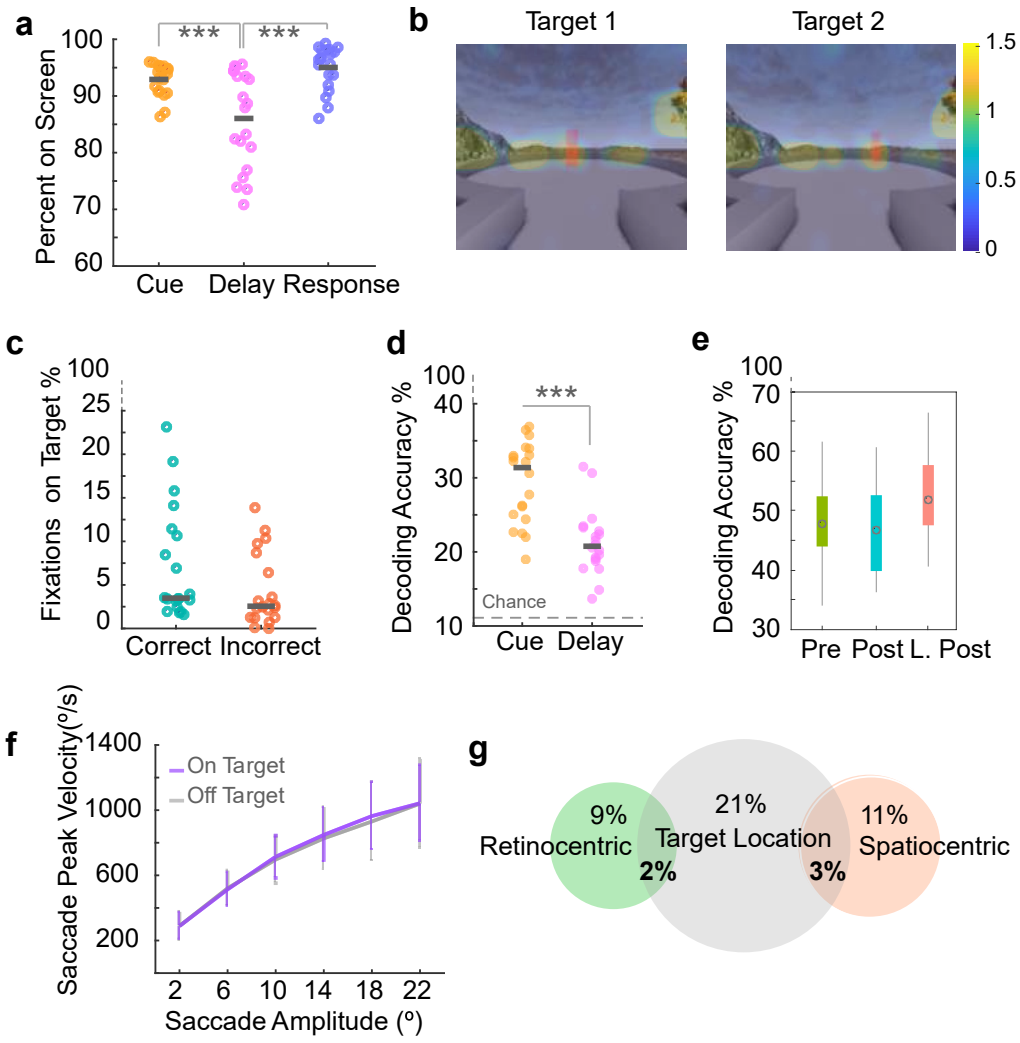

Fig. S2. Behaviour — Eve Movement (**A**) Percent of eye data points falling on screen for different task epochs. Black lines represent mean values for each group and dots represent data from each session. (**B**) Heat maps of eye fixation position on screen during delay for two target examples. Eye fixation is concentrated in task relevant areas on screen but is not primarily concentrated on the target location.(**C**) Percentage of total fixations during the delay epoch that land within the target location for correct and incorrect trials. The black lines represent median values and the dots represent data for each session. (**D**) Decoding accuracy for predicting target location from eye fixation location during the cue and delay epochs. The black lines represent median values and the dots represent data for each session. The gray line indicates chance performance for 9 classes. (**E**) Decoding accuracy for decoding target location from eye fixation position during the delay period for different ketamine injection periods. Decoding is shown for 3 classes (33.33% chance). The gray circles indicate median values. (**F**) Main sequence values during the delay period for saccades that fall on and off target location.(**G**) Proportion of neurons tuned for saccade landing position in retinocentric and spatiocentric reference frames as well as the proportion of neurons tuned for target location during the delay period. Overlapping sections indicate neurons that are selective for both remembered target location and saccade position. Panels 'f' and 'g' have been adapted from Roussy et al, 2021. Panels 'b', 'c', 'd', and 'e' have been adapted from Roussy et al., 2022. *p < 0.05, **p < 0.01, ***p < 0.001.

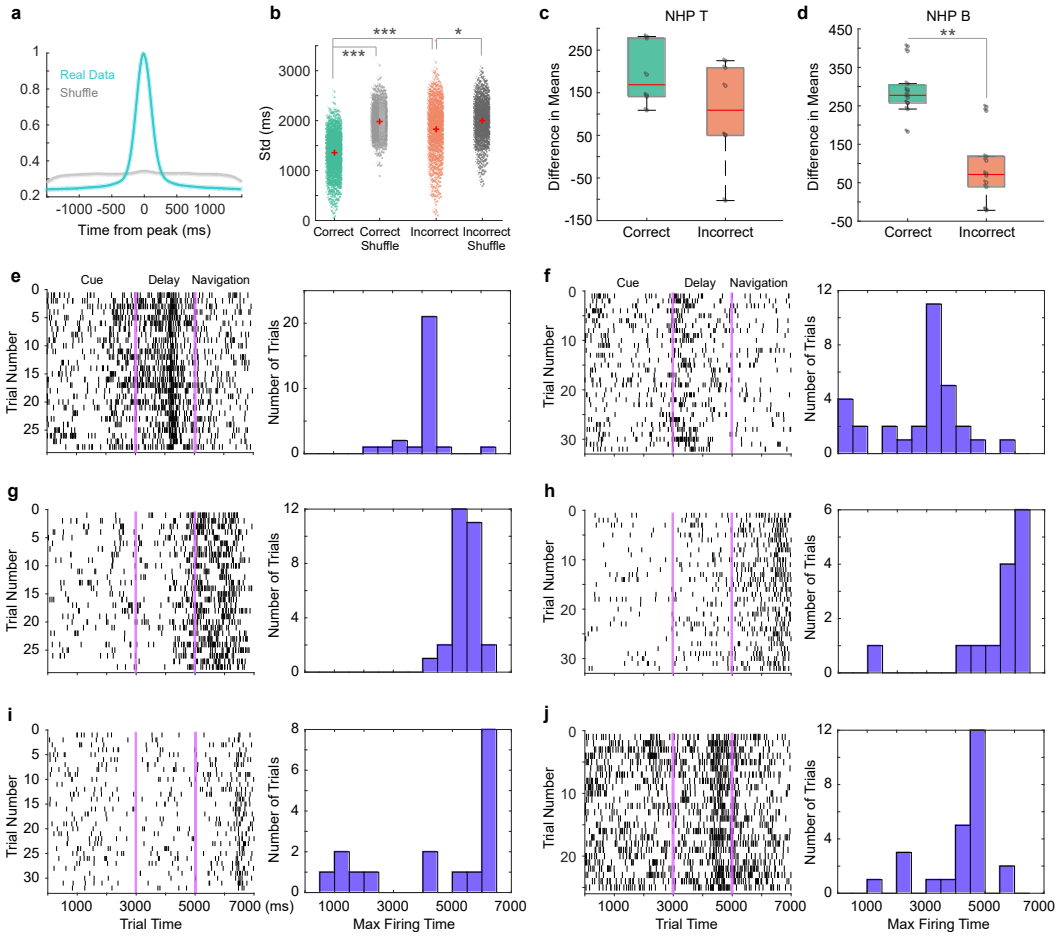

Fig. S3. Time-Consistent Neurons (**A**) Alignment of peak firing times for all neurons (blue) compared to shuffled peak firing times (grey). Here, we circularly shifted the normalized SDF of each neuron in each trial so that their maximum firing times were aligned (x axis = 0). We then averaged the aligned SDFs to produce the blue distribution, which demonstrates the timescale of these elevations in firing. The grey distribution results from first shuffling the spikes of the neurons before repeating this procedure. (**B**) Trial-trial standard deviation in max firing time for each neuron across sessions during correct and incorrect trials and for shuffled correct and incorrect data. The red crosses indicate group means. (**C**) Difference in means between real and shuffled distributions of standard deviations of max firing times across trials. Presented for correct and incorrect trials for NHP T. The red lines represent median values. Dots represent data for individual sessions. (**D**) Difference in means between real and shuffled distributions of standard deviations of max firing times across trials. Presented for correct and incorrect trials for NHP B. The red lines represent median values. Dots represent data for individual sessions. (**E-J**) Example single neurons. Left column represents the activity of one example neuron over trial time over all trials of a certain condition. Pink lines separate task epochs. Right column shows a histogram of max firing times per trial. *p < 0.05, **p < 0.01, ***p < 0.001.

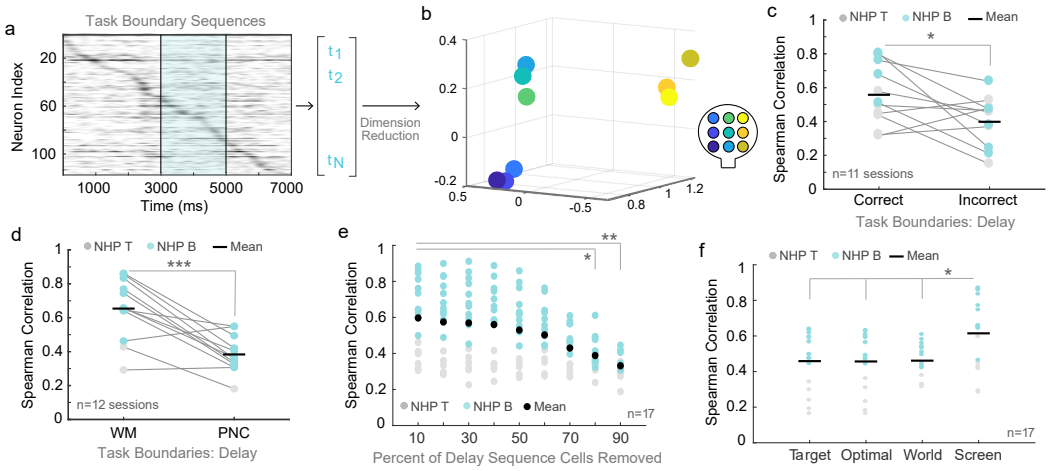

Fig. S4. Correlation Controls (**A**) Depiction of single-trial sequence representation using the delay epoch boundaries defined by the task structure. (**B**) The resulting sequence centroids for one recording session, depicted alongside the target locations in the VR arena. (**C**) Correlations between delay sequence centroids and visuospatial trajectories for sessions with both correct and incorrect trials for all 9 target locations. (Note that panels a-c correspond to Fig. 3a-c, using task epoch boundaries as opposed to neural epoch boundaries.) (**D**) Correlations between delay sequence centroids and visuospatial trajectories for correct trials during the WM task compared to the perception-navigation control task (PNC).(**E**) Ablation test, reproducing the correlation analysis for the WM memory period sequences after randomly removing a percentage of cells. Here, each dot represents the mean of 10 iterations for a single recording session. Black dots represent mean values across sessions. (**F**) Comparison of correlations between neural sequence centroids and different task variables: “Target” denotes the locations of the 9 targets; “Optimal” denotes the set of shortest paths from the start location to each target location; “World” denotes the set of trajectories in 3D virtual space the subjects took while navigating to the targets; “Screen” denotes the world trajectories as they appear through a perspective projection onto the computer screen. *p < 0.05, **p < 0.01, ***p < 0.001.

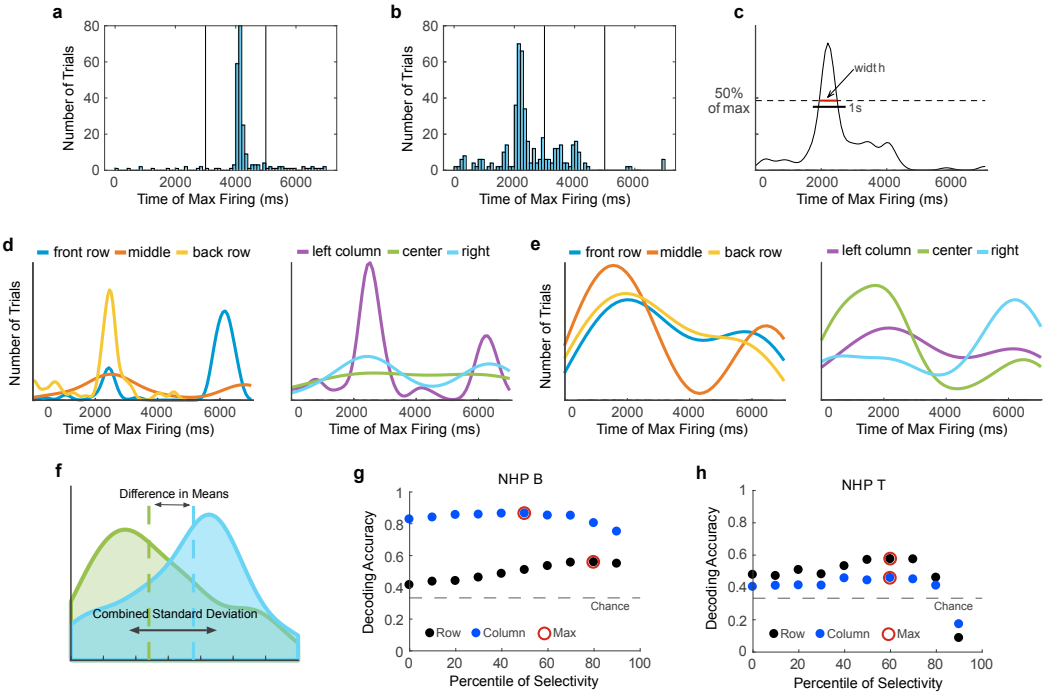

Fig. S5. Cell Selectivity (**A-B**) Two example time-selective cells. The histograms plot the times at which these cells contribute to the single-trial sequence across trials of all 9 target conditions. Vertical black lines indicate the start and end of the memory period. (**C**) Determining time selectivity. The distribution of firing times is plotted for the example cell in 'b'. A cell is considered time selective if the width (red line) of the distribution at 50% of its maximum is less than a threshold (black line). Threshold values were 1s for NHP B (as depicted), and 2s for NHP T. (**D**) Example row-selective cell. Left: distributions of max firing times for targets in each row. We see distinct peaks in the yellow and blue distributions, i.e., this neuron contributes to the sequence around 3000ms for targets in the front row, and around 6000ms for targets in the back row. Right: distributions of max firing times for targets in each column. This cell shows no clear separation in firing time by column. (**E**) Same as 'd', for a column-selective cell. Here, we see separation between the firing times for the center and right columns (green and blue distributions). (**F**) Schematic: the effect size between each pair of rows (or each pair of columns) is computed, as the difference in means of the two distributions divided by the combined standard deviation. (**G**) Thresholds for selectivity for each subject, determined by best decoding accuracy. For each percentile of selectivity (maximum effect size), sub-sequences of cells exceeding this percentile for row and column selectivity were used to decode row and column respectively. Decoding accuracy is plotted as a function of selectivity. Maximum decoding accuracy is used to define the thresholds for row and column selectivity used for each subject. These thresholds determine the cells which contribute to the row and column decoding in Fig. 4d.

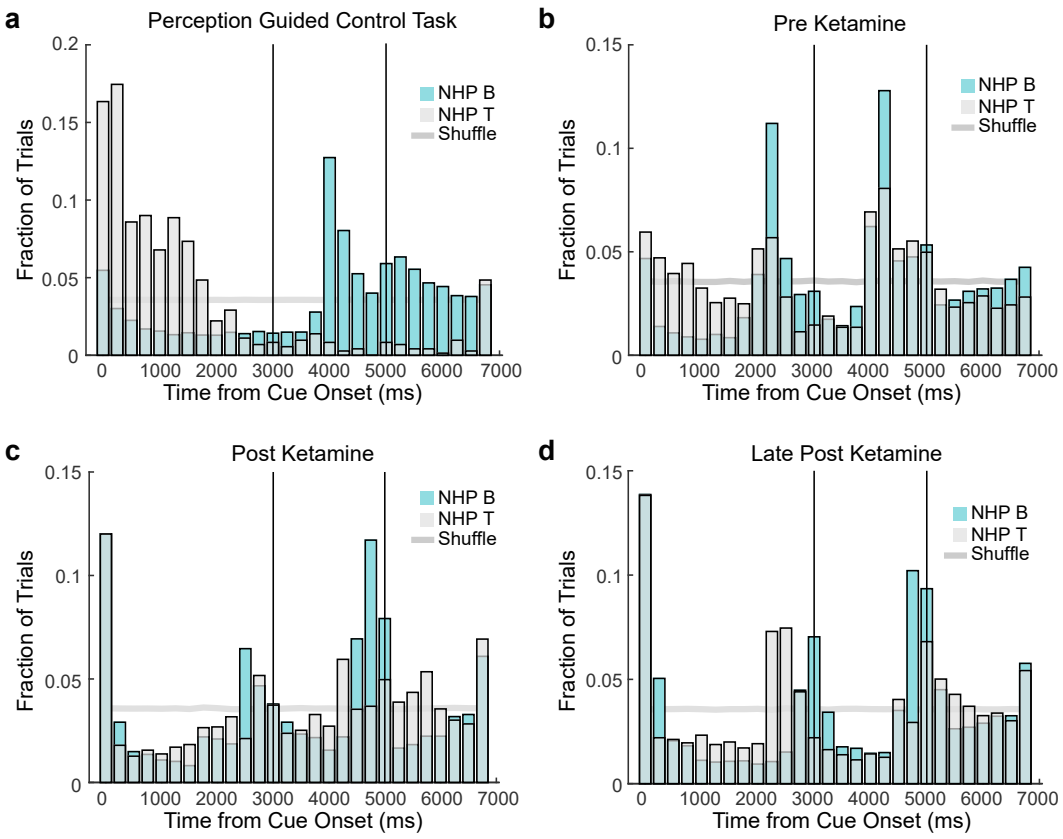

Fig. S6. Neural Boundaries (**A-D**) Times at which time-selective cells contribute to single-trial sequences across trials of all 9 target conditions. The shuffle control was created by randomly shuffling spike times of time-selective cells within trials. (**A**) Neural boundaries are not present during the perception control task, in which WM is not required. (**B**) Neural boundaries prior to ketamine administration.(**C**) Neural boundaries post ketamine administration. (**D**) Neural boundaries 1 hour after ketamine administration.

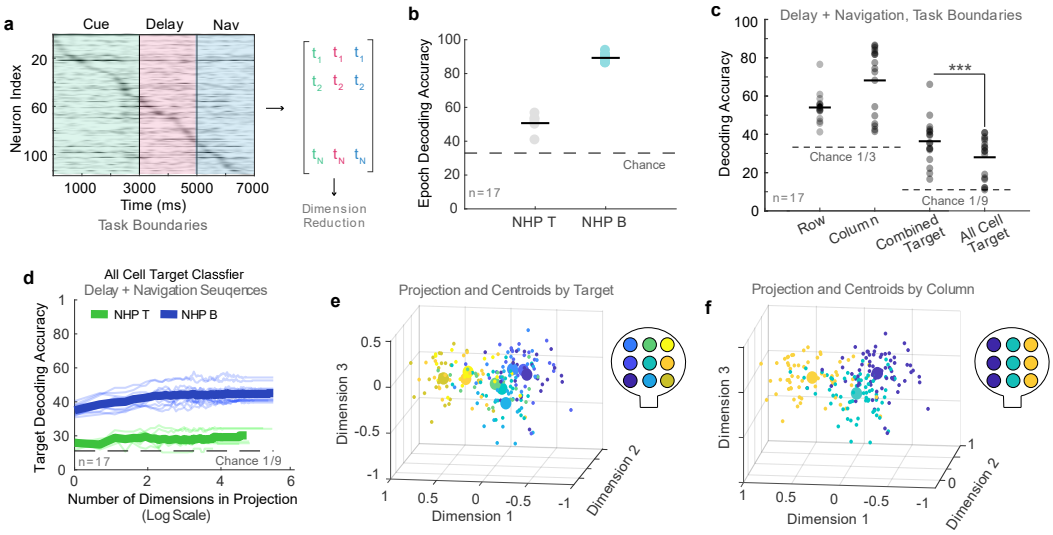

Fig. S7. Decoding Analyses (**A**) Depiction of single-trial sequence representations for cue, delay, and navigation epochs, using the task epoch boundaries, as marked by vertical black lines. (**B**) Epoch decoding accuracy using task boundaries. (See Fig 4c for accuracy using neural epoch boundaries.) (**C**) Row, column and target decoding accuracy using the task boundary delay and navigation sequences. (See Fig 4d for accuracy using neural boundaries.) (**D**) Target decoding accuracy as a function of the number of dimensions included in the projection, depicted alongside corresponding targets in the virtual arena. (**E**) Example projection coloured by target location. (**F**) Example projection coloured by target column, depicted alongside corresponding targets in the virtual arena..

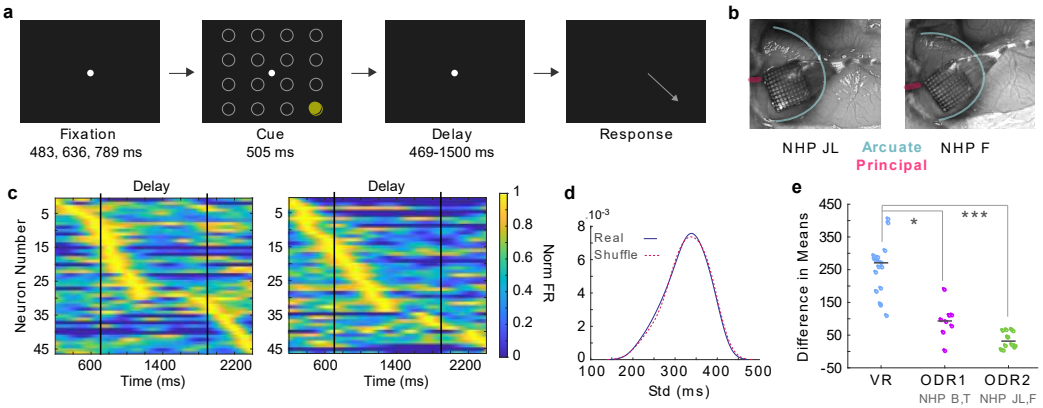

Fig. S8. Temporal organization of neural activity during ODR task (**A**) Depiction of ODR task with 16 targets. (**B**) Surgical images showing location of Utah arrays implanted in left LPFC of NHP JL and NHP F. (**C**)Two trial examples of simultaneously recorded population activity. Normalized firing rate for each neuron is arranged by max firing time. Task timing for delay epoch is marked by black lines. (**D**) Real and shuffled distributions of max firing time trial-to-trial standard deviations for an example session. (**E**) Difference in the means between real and shuffled distributions for the virtual reality task and the ODR task. The VR task and ODR1 task are from NHP B and NHP T, while ODR2 is from NHP JL and NHP F. Gray lines indicate median values and dots represent data per session. *p < 0.05, **p < 0.01, ***p < 0.001.

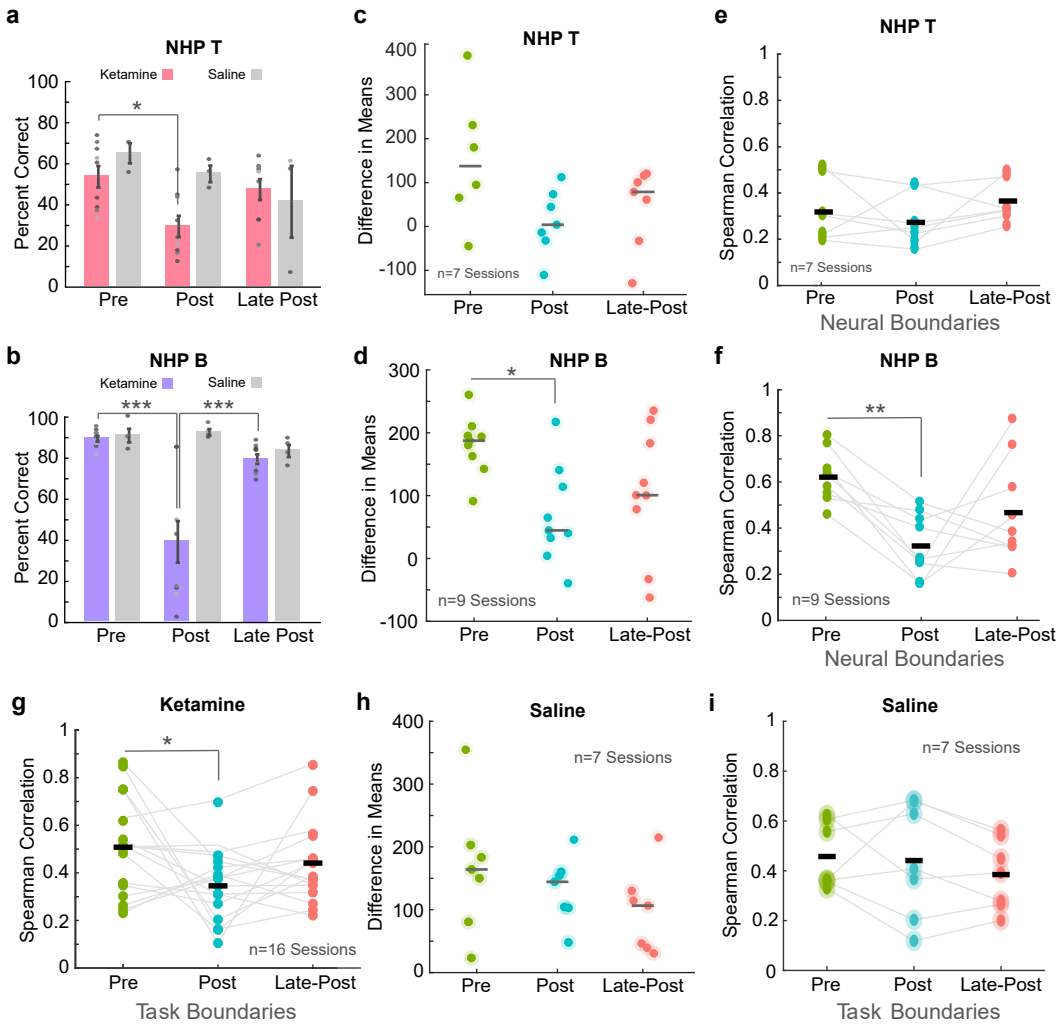

Fig. S9. Ketamine and Saline Control Analyses (**A**) Percent of correct trials for ketamine and saline sessions over injection periods for NHP T. Dots represent data for individual sessions and error bars are SEM. (**B**) Percent of correct trials for ketamine and saline sessions over injection periods for NHP B. Dots represent data for individual sessions and error bars are SEM. (**C**)Difference in the means between real and shuffled distributions of standard deviation values for neuron max firing time between trials. Presented for each ketamine injection period for NHP T. Dots represent data for each session. Black lines indicate the mean across sessions. (**D**) Difference in means between real and shuffled distributions for each ketamine injection period for NHP B. Dots represent data for each session. Black lines indicate the mean across sessions. (**E**) Correlation between the distances between condition cluster centroids and the distances between visuospatial trajectories, using neural boundary delay sequences. Correlation values are presented for ketamine injection periods for NHP T. Dots represent data per session, and lines indicate injection periods from the same session. Black lines indicate the mean across sessions. (**F**) Correlation between the distances between condition cluster centroids and the distances between visuospatial trajectories, using neural boundary delay sequences. Correlation values are presented for ketamine injection periods for NHP B. Dots represent data per session, and lines indicate injection periods from the same session. Black lines indicate the mean across sessions. (**G**) Correlation values for ketamine injection periods for both subjects using delay sequences with task epoch boundaries. Black lines indicate the mean across sessions. (See Fig. 5e for correlations using neural boundaries.)(**H**) Difference in mean values between real and shuffled distributions of max firing time standard deviations for saline injection periods. Dots represent data for each session and black lines indicate the mean across sessions (NHP T and NHP B combined). (**I**) Correlation between the distance between condition cluster centroids and distance between target trajectories. Correlation values are presented for saline injection periods for both animals combined, with balck lines indicating the mean across sessions. *p < 0.05, **p < 0.01, ***p < 0.001.

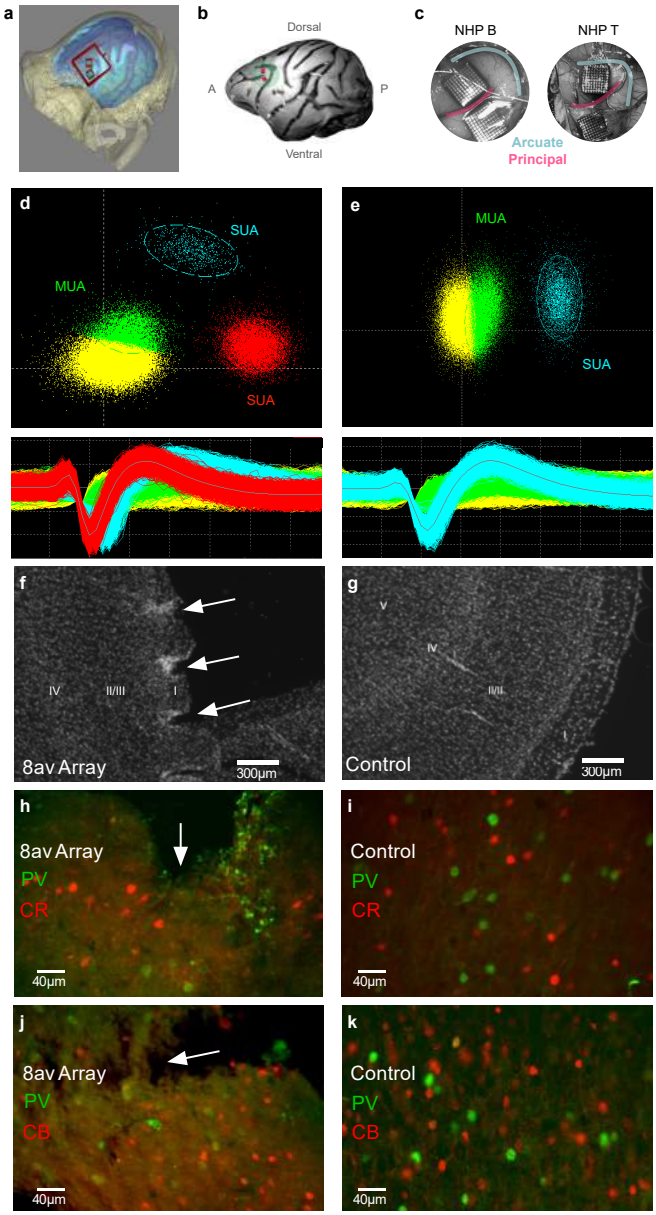

Fig. S10. Neural Recording Setup (**A**) Graphic of presurgical planning procedure showing 3D reconstructed skull and brain based on CT and MRI scans. Electrode array positioning is illustrated in blue with a red outline. The craniotomy is outlined by the larger red box. (**B**) 3D modeled brain with electrode array placement in pink. (**C**)Surgical images of array implantation in NHP B and NHP T. (**D**) Example of spike sorting for one electrode channel. Upper panel represents PCA space and the lower panel represents individual threshold crossing event waveforms. The blue and red clusters represent what we would classify as a single unit. The green cluster would be classified as a multiunit. (**E**) Example of spike sorting for one electrode channel. The blue cluster would represent a single unit. The green cluster would represent multiunit activity. This figure is modified from Roussy et al., 2021). (**F,G**) f,g, Histology showing cortical layers relative to array location (f), compared to healthy adjacent cortex as a control (g). Arrows indicate array placement. (**H,I**) Histology staining for Parvalbumin (green) and Calretenin (red) in the implanted subject (h) compared to healthy adjacent cortex as a control (i). Arrows indicate array placement. (**J,K**) Histology staining for Parvalbumin (green) and Calbindin (red) in the implanted subject (j) compared to healthy adjacent cortex as a control (k). Arrows indicate array placement.

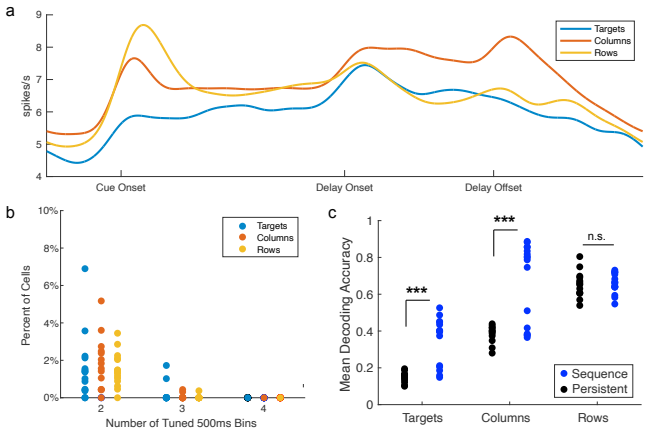

Fig. S11. Comparing persistent and sequence codes (**A**) Average response of cells identified to display persistent delay activity. **B** Percent of cells by session identified to display persistent delay activity. **C** Mean decoding accuracy by session (10-fold cross validation) for two multinomial regression models. The first predicts trial condition (target/column/row) based on the activity of persistent cells (black dots). The second predicts trial condition (target/column/row) based on the coordinates of the sequence projection obtained via dimensionality reduction (blue dots).

Supplementary Tables

240

MAIN TEXT FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
1e **Percent of Correct Trials**	NHP B, NHP T	20 WM Sessions			NHP B mean = 87% median = 85% NHP T mean = 57% median = 56%	
2g **Difference in Means**	NHP B, NHP T	17 WM Sessions	Wilcoxon Signed Rank Test	Correct and incorrect trials	*Rank* = 161 Correct mean = 253.98 median = 270.93 Incorrect mean = 93.69 median = 71.43	0.001
3h **Correlation Analysis**	NHP B, NHP T	11 WM Sessions (correct and incorrect trials for all 9 targets)	Paired T-Test	Correlation between average trajectories and centroids defined by correct vs incorrect trial sequences	t = 3.65 Correct **mean** = 0.52 median = 0.50 Incorrect **mean** = 0.27 median = 0.30	0.0040
4c **Epoch Decoding**	NHP B, NHP T	17 WM sessions	T-Test	Decoder vs chance (33%)	t = 9.53 Decoding accuracy: **mean** = 76% median = 87%	5.3779e-08

241

MAIN TEXT FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
4d **Row Decoding**	NHP B, NHP T	17 WM sessions	T-Test	Decoder vs chance (33%)	t = 13.74 Decoding accuracy: **mean** = 56% median = 55%	2.8203e-10
4d **Column Decoding**	NHP B, NHP T	17 WM sessions	T-Test	Decoder vs chance (33%)	t = 7.65 Decoding accuracy: **mean** = 71% median = 83%	9.9221e-07
4d **Target decoding**	NHP B, NHP T	17 WM sessions	Paired T-Test	Combined row + column decoder vs “all cell” sequence decoder	t = 6.30 Combined: **mean** = 40% median = 43% All Cell: **mean** = 31% median = 34%	1.0624e-05
4f **Correlation Analysis**	NHP B, NHP T	17 WM sessions 11 ODR Sessions	Rank Sum	VR compared to ODR	z=3.0775 VR: mean =0.61 **median** =0.64 ODR: mean =0.34 **median** = 0.333	0.0021

242

MAIN TEXT FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
5b **Percent correct**	NHP B, NHP T	18 Ketamine WM sessions 4 Ketamine perception sessions	Kruskal Wallis Tukey-Kramer Post Hoc	WM injection period Perception Injection Period	*H*(2,51) = 18.75 Pre, early post Pre, late-post Early-post, late-post *H*(2,9) = 0.13 <u>Pre-injection WM</u> mean = 72% **median** = 77% <u>Early Post-Injection WM</u> mean = 34% **median** = 28% <u>Late Post-Injection WM</u> mean = 63% **median** = 66% <u>Pre-Injection Perception</u> mean = 79% **median** = 84%	8.49E-05 7.7E-05 0.54 0.008 0.94

MAIN TEXT FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
					Early Post-Injection Perception mean = 84% **median** = 90% Late Post-Injection Perception mean = 85% **median** = 93%	

244

MAIN TEXT FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
5c **Difference in Means**	NHP B, NHP T	16 Ketamine WM sessions	Kruskal Wallis Tukey Kramer Post Hoc	Injection period Pre, Early post-injection Pre, Late post-injection Early post-injection, late post-injection	*F(2,44)* = 10.59 <u>Pre-injection</u> mean = 161.9 **median** = 171.6 <u>Early Post-Injection</u> mean = 40.9 **median** = 40.2 <u>Late Post-Injection</u> mean = 71.99 **median** = 100.4	0.005 0.004 0.07 0.59

245

MAIN TEXT FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
5d **Correlation**	NHP B, NHP T	17 Ketamine WM sessions	1-way ANOVA Tukey Kramer Post Hoc	Injection period Pre, Early post-injection Pre, Late post-injection Early post-injection, late post-injection	*F(2,45) = 4.93* <u>Pre-injection</u> **mean** = 0.48 median = 0.52 <u>Early Post-Injection</u> **mean** =0.26 median = 0.22 <u>Late Post-Injection</u> **mean** = 0.39 median = 0.34	0.0041 0.0029 0.3410 0.1030

246

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S1d **Distance to correct target**	NHP B, NHP T	20 WM sessions			NHP B mean=1.912 **median**=1.87 NHP T mean=1.81 **median**=1.68	
S2a **Percent Eyes on Screen**	NHP B, NHP T	20 WM sessions	1-way ANOVA Tukey Kramer Post Hoc	Epoch Cue-Delay Cue-Navigation Delay-Navigation	*F*(2,57) = 13.96 Cue **mean**= 93% median = 94% Delay **mean** = 86% median = 88% Navigation **mean** = 95% median = 97%	1.16E-05 0.0008 0.47 1.4E-05
S2c **Percent Fixations on Target**	NHP B, NHP T	20 WM sessions	Wilcoxon Rank Sum Paired T-Test	Correct-incorrect	*Rank* = 482 t(38) = 1.66 Correct mean = 7.2% **median** = 3.5% Incorrect mean = 4% **median** = 2.6%	0.05 0.1

247

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S2d **Decoding Accuracy**	NHP B, NHP T	20 WM sessions	Wilcoxon Rank Sum	Cue and delay epochs	*Rank = 562* <u>Cue</u> mean = 29.3 **median** = 31.4 <u>Incorrect</u> mean = 21.2 **median** = 20.8	4.2E-05
S2e **Main Sequence**	NHP B, NHP T	20 WM sessions	T-Test	Saccade on and off-target	*Cohen's d*	0 time bins, $p < 0.05$ 0 time bins > 0.2
S3b **Trial-Trial Std**	NHP B, NHP T	11 WM sessions	1-way ANOVA	Correct, incorrect, correct shuffle, and incorrect shuffle	*f(3,40) = 79.1*	7.3E-17
			Tukey Kramer Post-Hoc	Correct-correct shuffle		3.77E-09
				Correct-incorrect		3.78E-09
				Correct-incorrect shuffle		3.77E-09
				Correct shuffle-incorrect		0.02
				Correct shuffle-incorrect shuffle		0.99
				Incorrect-incorrect shuffle		0.01

248

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S3c **Difference in Means**	NHP T	6 WM sessions	Wilcoxon Signed Rank Test	Correct and incorrect	*Rank = 14* <u>Correct</u> mean = 191.12 **median** =168.74 <u>Incorrect</u> mean = 99.71 **median** = 108.98	0.56
S3d **Difference in Means**	NHP B	11 WM sessions	Wilcoxon Signed Rank Test	Correct and incorrect	*Rank = 55* <u>Correct</u> mean = 288.26 **median** =277.23 <u>Incorrect</u> mean = 90.09 **median** = 71.43	0.002
S4c **Correlation analysis using task boundaries**	NHP B, NHP T	11 WM sessions	Paired T-Test	Correct and incorrect trials	*t = 2.28* <u>Correct</u> **mean** = 0.56 median = 0.46 <u>Incorrect</u> **mean** = 0.39 median = 0.38	0.04

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S4d **Correlation analysis using task boundaries**	NHP B, NHP T	12 WM sessions, 12 perception sessions	Paired T-Test	WM task versus perception-guided navigation task	*t = 5.10* <u>WM</u> **mean** = 0.65 median = 0.66 <u>Perception</u> **mean** = 0.38 median = 0.36	3.4180E-04

250

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S4e **Ablation test - correct trials**	NHP B, NHP T	17 WM Sessions	ANOVA Tukey-Kramer post hoc 10-20 -30 -40 -50 -60 -70 -80 -90 20-30 -40 -50 -60 -70 -80 -90 30-40 -50 -60 -70 -80 -90 40-50 -60 -70 -80 -90 50-60 -70 -80 -90 60-70 -80 -90 70-80 -90 80-90	Correlation analysis after randomly removing a percentage of cells. (Mean across 10 iterations for each session is plotted.)	F(8,144) = 5.48 10 **mean** = 0.60 median = 0.62 20 **mean** = 0.58 median = 0.62 30 **mean** = 0.57 median = 0.59 40 **mean** = 0.56 median = 0.58 50 **mean** = 0.53 median = 0.55 60 **mean** = 0.50 median = 0.50 70 **mean** = 0.43 median = 0.44 80 **mean** = 0.39 median = 0.36 90 **mean** = 0.33 median = 0.32	4.8797e-06 1.0000 0.9999 0.9993 0.9590 0.7677 0.0757 **0.0069** **0.0001** 1.0000 1.0000 0.9971 0.9400 0.2026 **0.0273** **0.0006** 1.0000 0.9989 0.9632 0.2517 **0.0377** **0.0009** 0.9998 0.9848 0.3389 0.0600 **0.0017** 0.9999 0.7032 0.2303 **0.0135** 0.9339 0.5244 **0.0616** 0.9983 0.7229 0.9857

251

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S4f **Correlation analysis**	NHP B, NHP T	17 WM sessions	ANOVA Tukey-Kramer post hoc	"Target" location vs "optimal" trajectories vs movement trajectories in VR "world" coordinates vs visuospatial "screen" trajectories. Target-optimal Target-world Target-Screen Optimal-World Optimal-Screen World-Screen	*F(3,64)= 4.29* <u>Targets</u> **mean** = 0.46 median = <u>Optimal</u> **mean** = 0.46 median = 0.36 <u>World</u> **mean** = 0.46 median = <u>Screen</u> **mean** = 0.61 median = 0.36	0.008 1 1 0.02 0.99 0.02 0.03
S7b **Epoch Decoding**	NHP B, NHP T	17 WM sessions	T-Test	Compare to chance (33%)	t = 12.4181 <u>Combined Accuracy</u> mean=79% median=88% <u>NHP T</u> **mean**=60% median=60% <u>NHP B</u> **mean**=90% median=90%	1.2507E-09

252

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S7c **Decoding using task boundaries**	NHP B, NHP T	17 WM sessions	T-Test	Row decoding vs chance (33%)	t = 8.9743 Row **mean**= 48% median= 47%	1.2109E-07
				Column decoding vs chance (33%)	t = 7.7598 Column **mean**= 66% median= 74%	8.2176e-07
				Target decoding vs chance (11%)	t = 8.1505 Target **mean**= 31% median= 30%	4.3509e-07
S8e **ODR** **Difference in means**	NHP B, NHP T	17 WM sessions 12 ODR Sessions	Kruskal Wallis Tukey Kramer Post-Hoc	VR and ODR VR-ODR1 VR-ODR2 ODR1-ODR2	*H(2,34)* = 27.2 VR mean=253.98 **median**=270.9 ODR1 mean=91.53 **median**=93.2 ODR2 mean=34.85 **median**=31.6	1.2E-06 0.01 1.02E-06 0.29

253

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S9a **Percent correct**	NHP T	8 Ketamine Sessions	2-way ANOVA	Injection period drug (ketamine, saline)	Drug: $F(1,30)=2.8$ Injection Period: $F(2,30)=2.9$ Interaction: $F(2,30)=2.2$	0.107 0.066 0.135
			Tukey Kramer Post Hoc	Ket pre-injection, Sal pre-injection		0.894
				Ket pre-injection, Ket early post-injection		0.037
				Ket pre-injection, Sal early post-injection		1
				Ket pre-injection, Ket late post-injection		0.961
				Ket pre-injection, Sal late post-injection		0.864
				Sal pre-injection, Ket early post-injection		0.027
				Sal pre-injection, Sal early post-injection		0.972
				Sal pre-injection, Ket late post-injection		0.581

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
				Sal pre-injection, Sal late post-injection		0.487
				Ket early post-injection, Sal early post-injection		0.198
				Ket early post-injection, Ket late post-injection		0.205
				Ket early post-injection, Sal late post-injection		0.873
				Sal early post-injection, Ket late post-injection	**<u>Ketamine</u>** <u>Pre</u> **mean**=54, median=60 <u>Early-Post</u> **mean**=30, median=24 <u>Late-Post</u> **mean**=47, median=56	0.980
				Sal early post-injection, Sal late post-injection	**<u>Saline</u>** <u>Pre</u> **mean**=65, median=70 <u>Early-Post</u> **mean**=55, median=56 <u>Late-Post</u> **mean**=42, median=57	0.904
				Ket late post-injection, Sal late post-injection		0.993

255

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S9b **Percent Correct**	NHP B	9 Ketamine sessions	2-way ANOVA	Injection period drug (ketamine, saline)	Drug: $F(1,33)=12.9$ Injection Period: $F(2,33)=6.7$ Interaction: $F(2,33)=9.5$	0.001 0.0034 0.0005
			Tukey Kramer Post-Hoc	Ket pre-injection, Sal pre-injection		1
				Ket pre-injection, Ket early post-injection		0
				Ket pre-injection, Sal early post-injection		0.99
				Ket pre-injection, Ket late post-injection		0.76
				Ket pre-injection, Sal late post-injection		0.988
				Sal pre-injection, Ket early post-injection		0.0001
				Sal pre-injection, Sal early post-injection		1
				Sal pre-injection,Ket late post-injection		0.827

256

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
				Sal pre-injection, Sal late post-injection		0.98
				Ket early post-injection, Sal early post-injection		0
				Ket early post-injection, Ket late post-injection		0.0001
				Ket early post-injection, Sal late post-injection		0.0006
				Sal early post-injection, Ket late post-injection	**<u>Ketamine</u>** <u>Pre</u> **mean**=89, median=90 <u>Early-Post</u> **mean**=39, median=43	0.74
				Sal early post-injection, Sal late post-injection	<u>Late-Post</u> **mean**=80, median=83	0.96
				Ket late post-injection, Sal late post-injection	**<u>Saline</u>** <u>Pre</u> **mean**=91, median=90 <u>Early-Post</u> **mean**=93, median=92 <u>Late-Post</u> **mean**=84, median=84	0.1

257

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S9c **Difference in Means**	NHP T	7 Ketamine Sessions	Kruskal Wallis	Injection period	*H(2,17) = 3.46* Pre mean=137.7, **median**=94.9 Early-Post mean=9.6, **median**=0.59 Late-Post mean=34.98, **median**=70.09	0.1775
S9d **Difference in Means**	NHP B	9 Ketamine Sessions	Kruskal Wallis Tukey Kramer Post-Hoc	Injection Period Pre, early-post Pre, late-post Early-post, late-post	*H(2,24)* = 7.24 Pre mean=180.47, **median**=187.5 Early-Post mean=68.8, **median**=44.63 Late-Post mean=104.88, **median**=100.8	0.03 0.02 0.24 0.53

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S9e **Correlation analysis** **Neural Boundaries**	NHP T	7 Ketamine Sessions	1-way ANOVA	Injection Period	F*(2,24)* = *0.82* Pre **mean**= 0.32 median= 0.30 Early-Post **mean**= 0.28 median= 0.25 Late-Post **mean**= 0.36 median= 0.33	0.45
S9f **Correlation analysis** **Neural Boundaries**	NHP B	9 Ketamine Sessions	1-way ANOVA Tukey Kramer Post-Hoc	Injection Period Pre, early-post Pre, late-post Early-post, late-post	F*(2,24)* = *7.52* Pre **mean**= 0.62 median= 0.63 Early-Post **mean**= 0.32 median= 0.27 Late-Post **mean**= 0.47 median= 0.38	0.0029 0.0020 0.1345 0.1665

259

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S9g **Correlation analysis** **Task Boundaries**	NHP B, NHP T	16 Ketamine Sessions	1-way ANOVA Tukey Kramer Post-Hoc	Injection Period Pre, early-post Pre, late-post Early-post, late-post	F*(2,24)* = *3.32* <u>Pre</u> **mean**= 0.51 median= 0.51 <u>Early-Post</u> **mean**= 0.35 median= 0.36 <u>Late-Post</u> **mean**= 0.44 median= 0.41	0.0451 0.0360 0.5484 0.2950
S9h **Difference in Means (Saline)**	NHP B, NHP T	7 Saline Sessions	Kruskal Wallis	Injection Period	*H(2,18)* = *2.07* <u>Pre</u> mean=165.43, **median**=164.1 <u>Early-Post</u> mean=136.27, **median**=144.5 <u>Late-Post</u> mean=96.76, **median**=106.6	0.36

260

SUPPLEMENTARY FIGURE STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S9i **Correlation Analysis (Saline)** **Task Boundaries**	NHP B, NHP T	7 Saline Sessions	1-way ANOVA Tukey Kramer Post-Hoc	Injection Period Pre, early-post Pre, late-post Early-post, late-post	F*(2,18)* = *0.34* Pre **mean**= 0.46 median= 0.36 Early-Post **mean**= 0.44 median= 0.40 Late-Post **mean**= 0.38 median= 0.39	0.7181

Fig.	**NHP**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
S11c **Multinomial regression model accuracy**	NHP B, NHP T		1-way ANOVA	Persistent cell based MNR vs NAS based MNR		
Targets		15 WM sessions			Persistent: mean=0.15 median=0.13 NAS: mean=0.35 median= 0.40	**4.3e-29**
Columns		16 WM sessions			Persistent: mean=0.38 median=0.39 NAS: mean=0.67 median= 0.75	**1.3e-27**
Rows		16 WM sessions			Persistent: mean=0.65 median=0.65 NAS: mean=0.66 median= 0.66	0.66

261

ADDITIONAL STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
Observed correlations compared to chance	NHP B, NHP T	17 WM sessions	Paired T-Test	Neural boundary delay sequence-trajectory correlation compared to mean of shuffled null obtained by shuffling target labels 100 times.	t = -4.23 Observed: **mean=** 0.5440 median= 0.4965 Shuffled: **mean=** 0.3227 median= 0.3190	6.31E-04
Single vs multiple contributions	NHP B, NHP T	17 WM sessions	Paired T-Test	Correlation analysis in which each cell can contribute to the sequence for each epoch, to test the assumption that each cell contributes only once to a single trial sequence.Here, the correlation of neural boundary delay sequences are considered.	*t =0.2473* Single **mean** = 0.54 median =0.50 Multiple **mean** = 0.55 median =0.62	0.08
Percent of cells that are "time selective"	NHP B, NHP T	17 WM sessions	WM: mean=4.8%, median=4.7% Perception: mean=6.2%, median=3.4% PreKet: mean=5.0%, median=3.6% PostKet: mean=2.8%, median=2.3% LatePost Ket: mean=3.5%, median=2.6%			

ADDITIONAL STATISTICS						
Fig.	**Subj.**	**Data**	**Stat Test**	**Comparison**	**Stat-Value**	**P-Value**
Delay-Navigation sequence correlation	NHP B, NHP T	17 WM sessions	Test whether observed absolute correlation value of delay sequence to same-trial navigation sequence is larger than the 99th percentile of correlation values between the delay sequence and 25 randomly drawn different trial navigation sequences.	Sequences from same trial vs from different trials. Absolute value correlation from delay sequences to 1) same-trial nav sequences, vs 2) mean across different-trial nav sequences.		Delay sequences are less correlated to same-trial navigation sequences than different-trial navigation sequences on 96.31% of trials.

Supplementary Videos

Supplementary video 1: An example trial of the WM task can be found at [https://www.youtube.com/watch?v=nZDYJw2aFLQ]. Note that the video is played back at an increased speed. During the trial, a cue location (red vertical bar) is presented for 1 second, then disappears for a 2 second memory delay, after which the subject responds by navigating to the remembered location using a joystick.