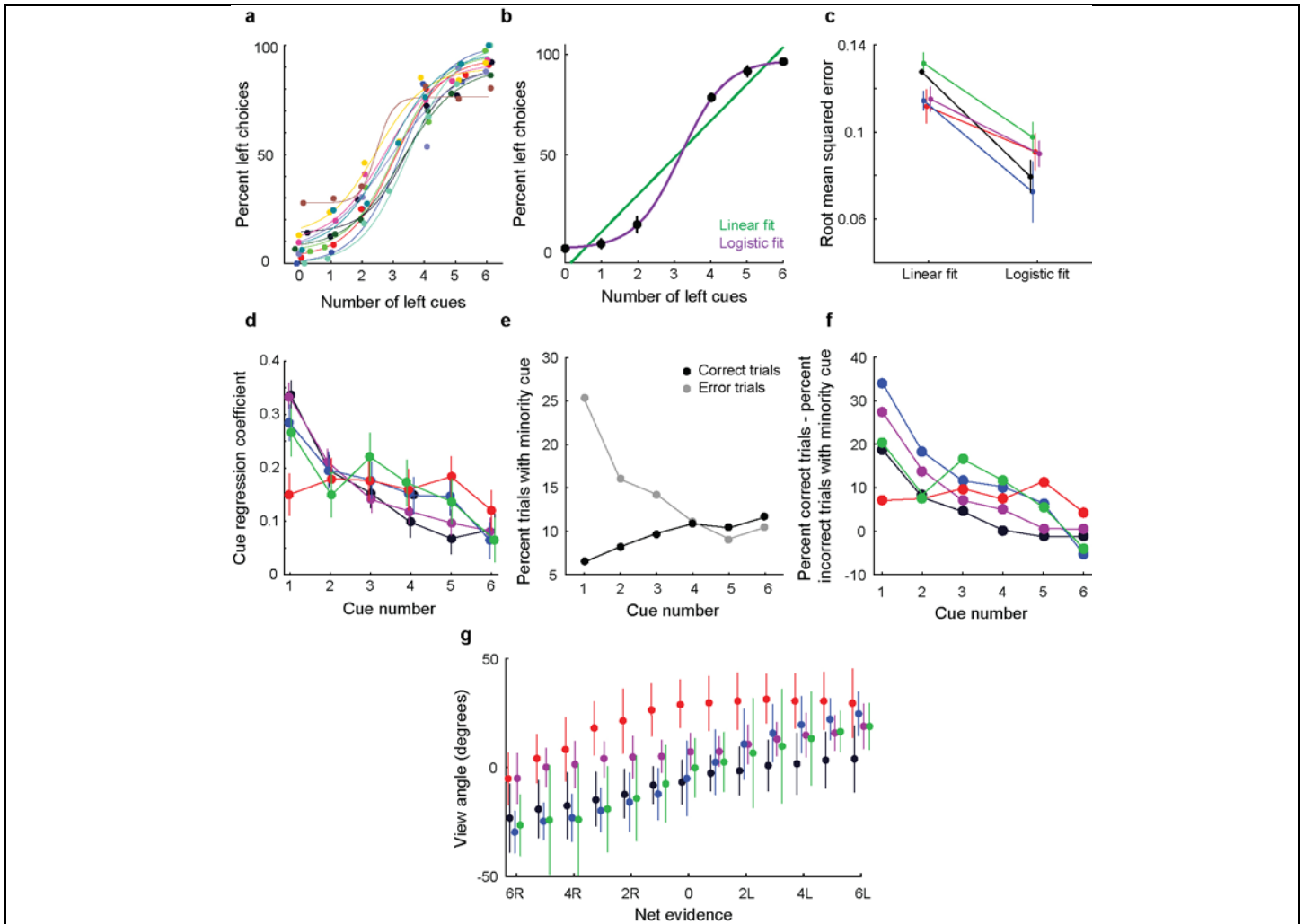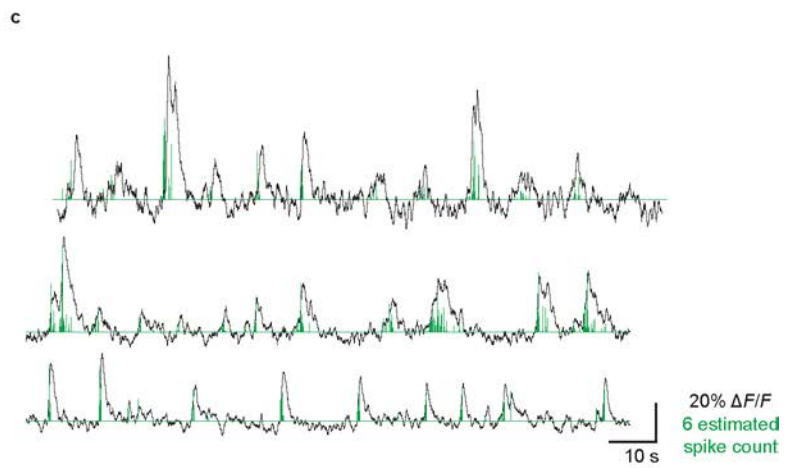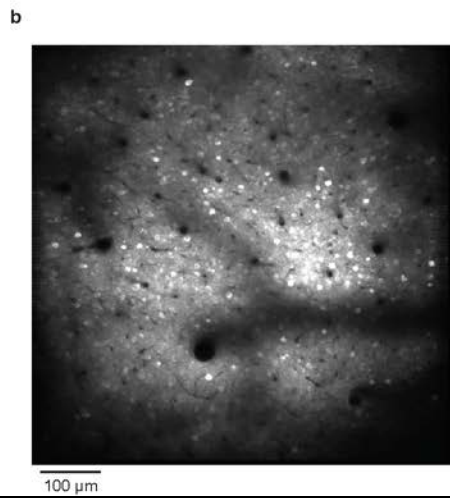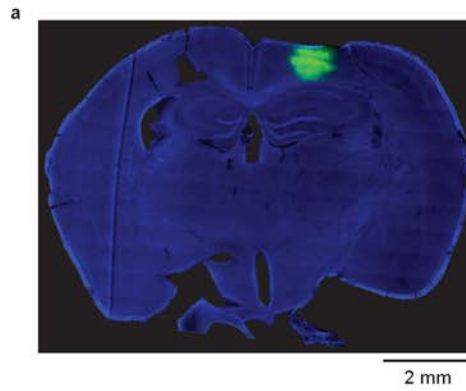Supplementary Figure 1

Behavioral Training.

a, Mazes used for behavioral training. Asterisks indicate reward location. Only some example mazes are shown (for example, right choice and not left choice maze in maze 1). **b,** Distribution of net evidence corresponding to different difficulties used in training the final task (maze 8; see **d**). **c,** Screen captures of the virtual environment at cue 1, cue 6, and the turn in maze 8. **d,** Behavioral performance across sessions for three example mice. Colors correspond to the maze colors indicated in **a**. Shapes correspond to the net evidence probabilities in **b**.

Supplementary Figure 2
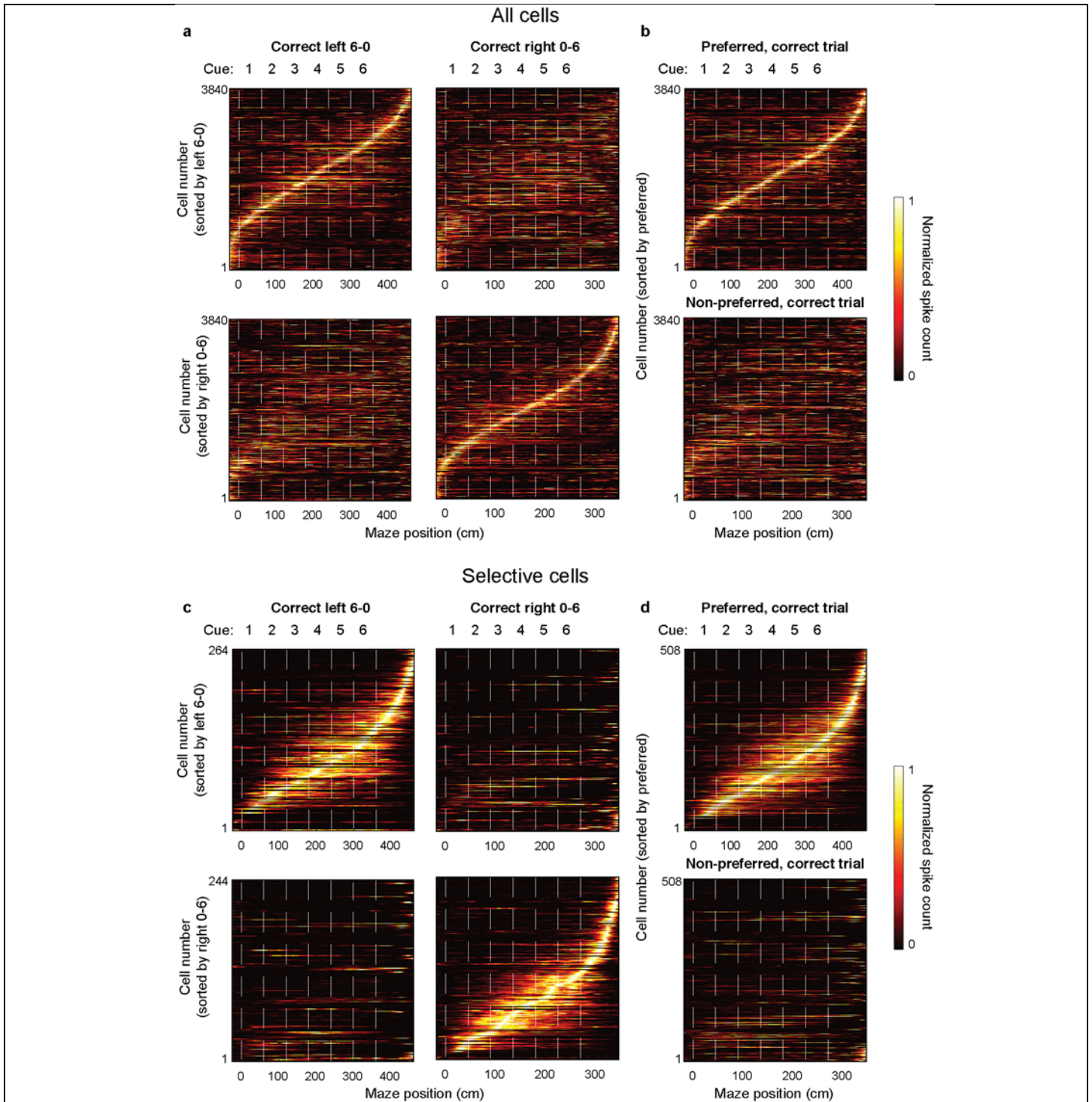
Behavioral analysis of evidence accumulation.

a, Behavioral performance on each of the 11 imaging sessions, fit with a logistic function. **b,** Example performance from a single mouse across seven behavioral sessions fit by a linear (green) and logistic (purple) model (Methods M.2.4.1). **c,** Across mice, the logistic model fit the data significantly better than the linear model (p < 0.05 for all mice, two-sample Student's t-test), suggesting that mice used more than one piece of evidence per trial to make a choice. Error bars represent mean ± s.e.m. across datasets. Mice are colored the same as in Fig. 1c. **d,** Multivariate linear regression in which the mouse's choice was the response variable and the six cue identities were separate explanatory variables. Regression coefficients for five mice (7-12 sessions each) are shown. Four of the five mice weighted early cues more than late cues. Error bars indicate confidence intervals. **e-f,** Fraction of correct (black) and error (gray) trials containing a minority cue (a cue indicating the incorrect choice) at each cue position, for a single mouse (**e**) and as the difference of the error and correct points (**f**) for five mice. **g,** Relationship between net evidence and view angle for each mouse combined across all cue positions.

a,

2 mm

b,

100 μm

c,

20% ΔF/F
6 estimated
spike count
10 s

Supplementary Figure 3

Example imaging field of view and activity traces.

a, Example histology image of GCaMP6m-expressing neurons in the PPC. **b,** Example two-photon image of GCaMP6m-expressing neurons in layer 2/3 of the PPC. **c,** Example ΔF/F traces (black) and deconvolved estimated spike counts (green) (Methods M.3.4).

All cells

a

Correct left 6-0 | Correct right 0-6

b

Preferred, correct trial

Non-preferred, correct trial

Selective cells

c

Correct left 6-0 | Correct right 0-6

d

Preferred, correct trial

Non-preferred, correct trial

Supplementary Figure 4

Mean population activity patterns in PPC for all cells and selective cells.

a, Normalized mean activity across correct left 6-0 (left) and correct right 0-6 (right) trials for all neurons pooled across all datasets (n = 3840 cells from 11 datasets, 5 mice). Traces were normalized to the peak of each cell's activity on either correct left 6-0 (top) or correct right 0-6 (bottom) trials, averaged, and sorted by the peak's maze position. **b,** Same as in **a**, except for on preferred (top) or non-preferred (bottom) correct 6-0 trials. Cells were sorted according to each cell's activity in its preferred condition. Preferred trial type was determined

for each cell individually based on which trial type had higher mean activity. **c-d,** Same as **a-b,** but only for selective cells. Selective cells were defined as all cells with choice classification accuracy above 70%.
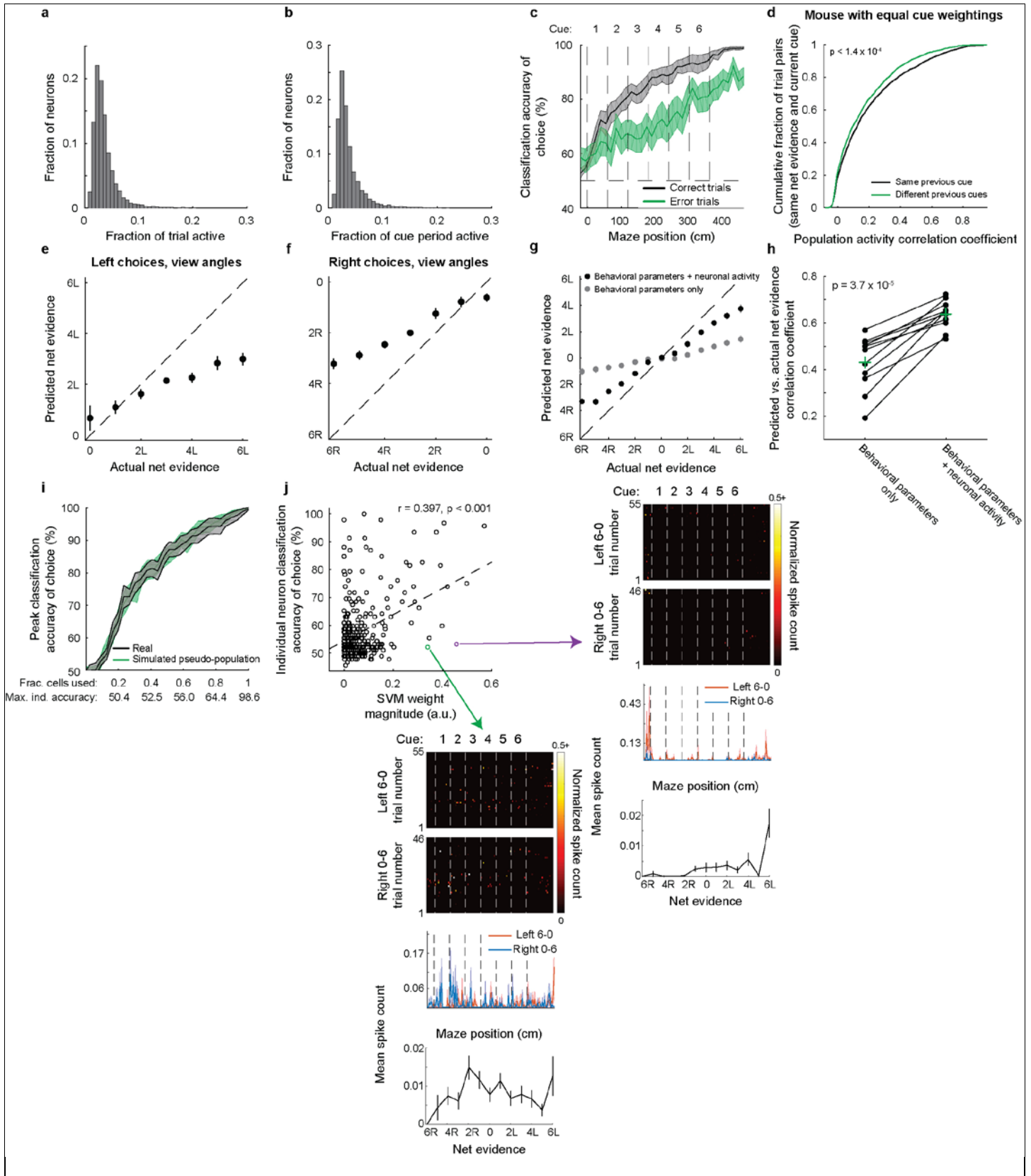
Supplementary Figure 5

Analyses of single neuron- and population-level representations of task-relevant features.

a-b, Histogram of the fraction of the entire trial (**a**) and cue period (cues 1-6) (**b**) neurons were active (n = 3840 neurons from 5 mice). **c,** SVM classification accuracy (mean ± s.e.m., n = 11 datasets) for choice based on population activity on correct and error trials. Independent classifiers were trained and tested at each maze position. **d,** Same as Fig. 6a except for a mouse with equal cue weightings. Cumulative distribution of the pairwise trial-trial population activity correlation coefficients for epochs with the same (black) or different (green) previous cues, keeping net evidence and epoch constant (e.g. *LRLXXX* vs. *RLLXXX* trials at cue 3) (p < 1.4 x 10^{-4}, two-sample KS test, n = 2 datasets; mouse colored as red in Fig. 1c, Supplementary Fig. 2b, d). This analysis tested if neuronal activity at a given epoch contained information about the previous epoch's cue, independent of maze epoch and net evidence. **e-f,** SVR classifiers for net evidence performed on trials with nearly identical (±2.5°) view angles on left choice (**e**) and right choice (**f**) trials. **g,** Actual net evidence vs. net evidence predicted by an SVR classifier trained on behavioral parameters only (gray) or both behavioral parameters and neuronal population activity (black) (Methods M.4.4.3). Error bars represent mean ± s.e.m. across datasets (n = 11). Across mice the predicted vs. actual net evidence correlation coefficient was significantly higher for the model with behavioral parameters and neuronal activity than for the model with behavioral parameters only (p < 0.001 relative to shuffled net evidence labels). Net evidence therefore appeared decodable beyond information provided by view angle. **h,** Data from (**g**) shown for individual datasets. Green crosses represent means across datasets (n = 11; p = 3.7 x 10^{-5}, two-sample Student's t-test). **i,** Peak classifier accuracy for choice for classifiers constructed with increasing numbers of neurons, added from least to most selective (based on histograms from Fig. 2c). Real data are shown in black and a simulated pseudo-population is shown in green. To create the pseudo-population, trial identities were shuffled (within a trial-type category) independently for each neuron to break neuron-neuron correlation structure but to preserve each neuron's activity within the trial (Methods M.4.7.2). Shaded error bars represent mean ± s.e.m. across datasets, and max individual neuron classification accuracies/correlations were the mean across datasets. **j,** Individual neurons' choice classification accuracy as a function of the magnitude of the weight placed on each neuron by a linear SVM choice classifier trained on all neurons. The population classifier reached a peak accuracy of 100%. While neurons with higher individual classification accuracy were weighted more strongly, the SVM still weighted some neurons with low individual accuracy. Single trial activity on left 6-0 and right 0-6 trials for two example neurons with relatively high weight are also shown. These two neurons illustrate two ways that neurons with low individual selectivity can contribute to a population code. The left neuron (green) is active on both trial types with high variability, but slightly more so on right trials. The right neuron (purple) is primarily active on left trials, but is only active on a small subset of trials (see Methods M.4.4.4). Top panels: each row is an individual trial. Bottom panels: mean ± s.e.m. For each net evidence condition (e.g. 2L), the mean spike count was calculated by combining the activity at all cue epochs matching the given net evidence.
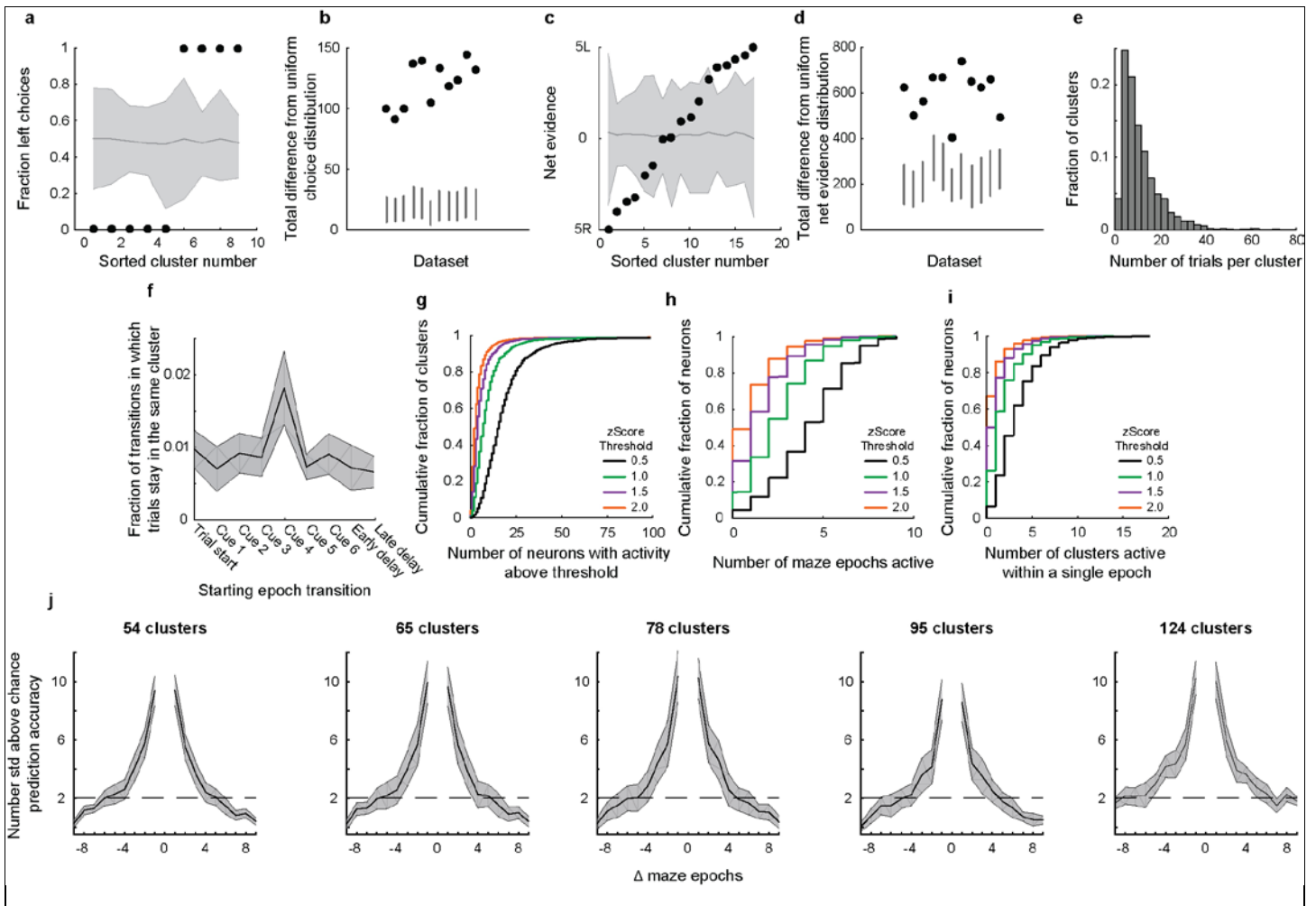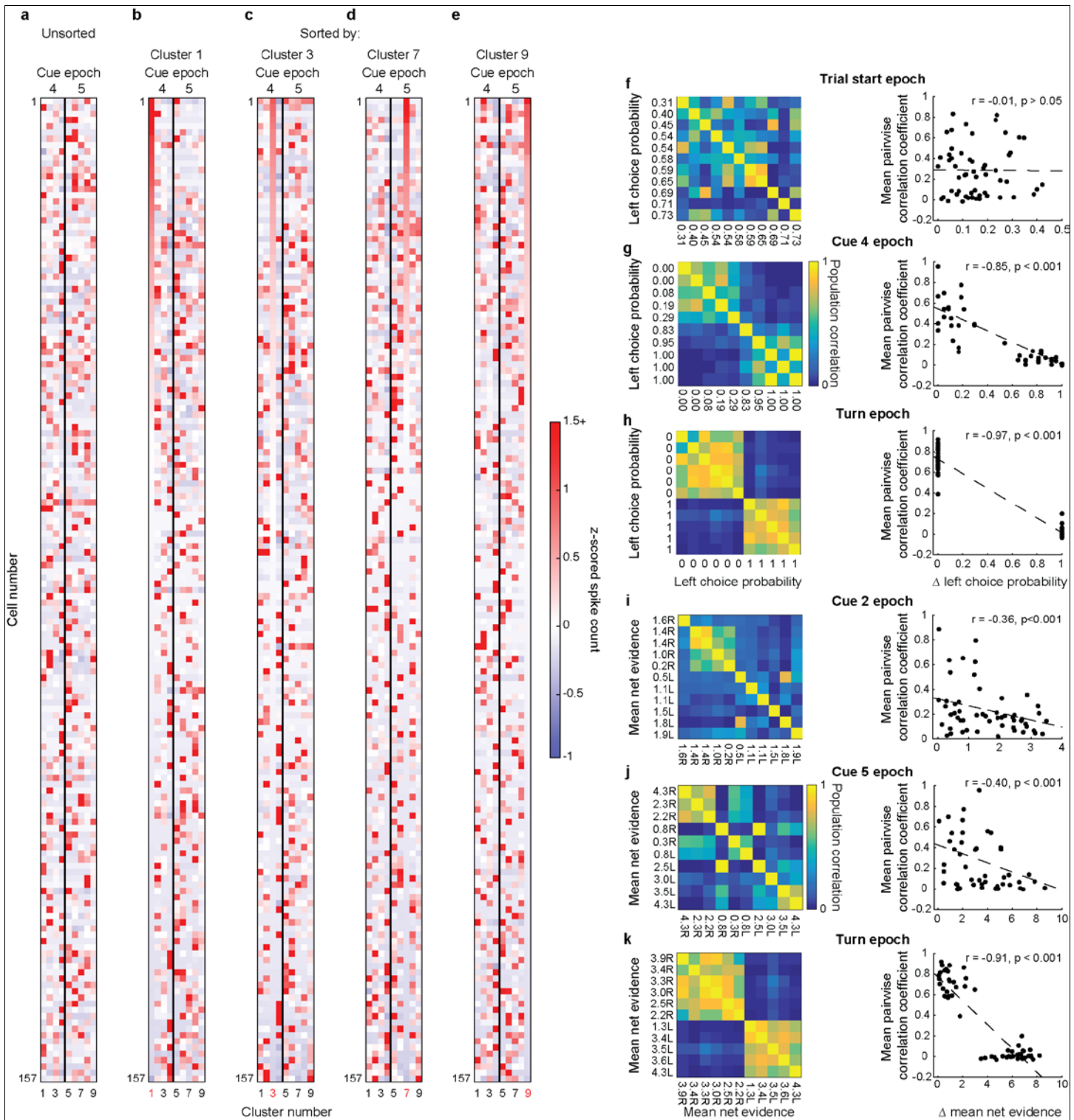
Supplementary Figure 6

Characterization of behavioral and neuronal patterns across clusters.

a, Fraction of trials in each cluster in the turn epoch that were left choice trials for an example dataset. Clustering revealed neuronal activity patterns related to behavioral choices. Gray area indicates the median and 99% confidence intervals of the shuffled distribution of trial assignments to clusters. **b,** Comparison of the total difference from a uniform distribution for the real data (circles) to the 99% confidence intervals of the corresponding shuffle for each dataset (lines). The total difference was calculated as the summed absolute difference from the shuffle median across clusters. **c-d,** Same as in **a-b,** but for net evidence during the fifth cue. **e,** Distribution of trials per cluster across all epochs and datasets (n = 2457 clusters). **f,** Cluster self-transition probabilities for clustering performed using all epochs together. Transition probabilities were considered from one epoch to the next epoch. Low self-transition probabilities suggested that activity patterns changed over the time of consecutive epochs. Error bars represent mean ± s.e.m across datasets. **g,** Cumulative distribution of the number of neurons active in each cluster for different z-score activity thresholds. **h,** Cumulative distribution of the number of maze epochs in which a neuron was active in at least one cluster for different z-score activity thresholds. **i,** Cumulative distribution of the number of clusters in which a neuron was active within a single epoch for different z-score activity thresholds. **j,** For a given trial based on the current cluster identity, the accuracy of predicting the clusters occupied by that trial in the past and future epochs did not depend greatly on the clustering preference parameters (percentile of the distance matrix used for clustering; 1st, 10th, 30th, 50th, 70th from left to right) and, hence, numbers of clusters. Cluster numbers are the mean number of clusters for each preference parameter across datasets. Error bars represent mean ± s.e.m. across datasets.
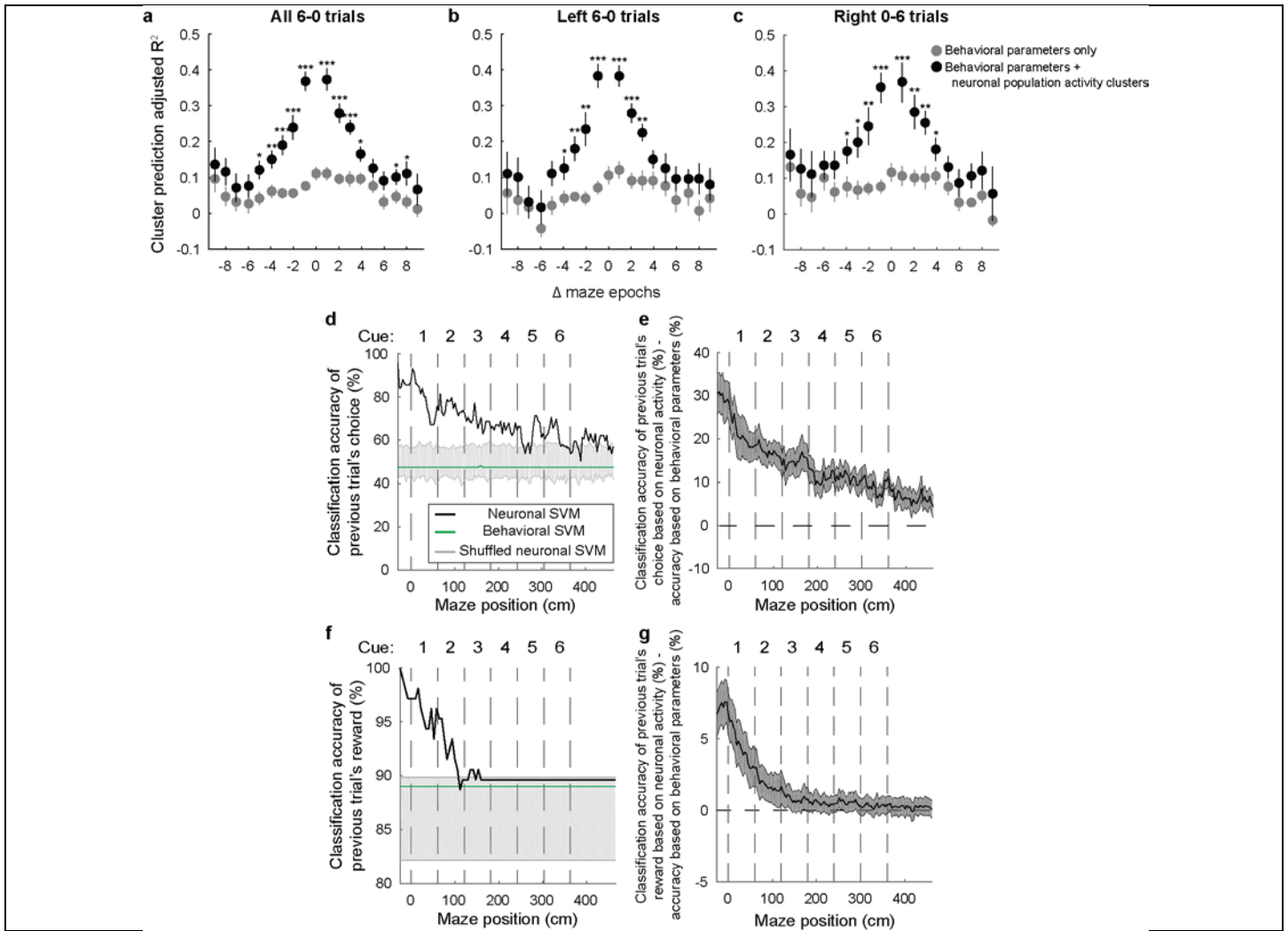
Supplementary Figure 7

**Visualizations of neuronal activity across clusters.**

a-e, Mean z-scored spike count for individual neurons across clusters comprised only of correct left 6-0 trials at two adjacent epochs (Cues 4 and 5) from a single dataset. These plots demonstrate that the activity across clusters and epochs featured largely different patterns of active neurons. Neurons were either unsorted (**a**) or sorted according to their activity in clusters 1, 3, 7, or 9 (**b-e**). Neurons whose mean z-scored activity was less than 0.001 in all of the displayed clusters were excluded for display purposes (these neurons were active during a different trial epoch). Clusters were generated from correct left 6-0 trials. **f-h,** Left panels: Matrix of population activity correlations between each pair of cluster centers sorted according to the cluster's left choice probability at three different maze epochs. For each cluster, the population activity was calculated as the mean activity vector across trials for each cluster. Right panels: Population activity correlation between each pair of clusters as a function of their difference in left choice probability. **i-k,** Same as in **f-h**, but for net evidence.

Supplementary Figure 8

Contribution of behavioral variability to trial-trial variability and classification of the previous trial's outcome.

a-c, Our ability to predict the past and future population activity pattern based on the current population activity pattern could not be explained by behavioral variability. We performed a multivariate logistic regression to predict a trial's cluster identity at a given epoch based on only the behavioral parameters at another epoch (gray) or both the behavioral parameters and the cluster identity at another epoch (black). To allow for a binary classifier, we only included those trials whose cluster identity contained either the most or second most trials during the prediction epoch (Methods M.4.7.3). Consistently, the model based on both behavioral parameters and the previous cluster identity outperformed the model based on only behavioral parameters. This analysis was performed on left 6-0 trials (b) and right 0-6 trials (c) separately, and pooled together for all 6-0 trials (a). The behavioral parameters used were x/y position, x/y treadmill velocity, and view angle. Separate models were fit for each combination of previous and future cluster identities and combined based on the number of maze epochs between them (Δmaze epochs). Adjusted R^2 values were used to compare the predictive power of models with different numbers of explanatory variables. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, two-sample Student's t-test. **d,** Comparison of a neuronal activity-based SVM (black), behavioral parameter-based SVM (green), and the 99% confidence interval of a neuronal activity-based SVM with shuffled labels (gray) for the previous trial's choice for a single dataset. The behavioral parameter-based SVM could not discriminate the previous trial's choice. Classifiers were trained to distinguish the mouse's choice on the previous trial independently at each bin in the current trial. **e,** Difference between the classification accuracy of the neuronal activity-based SVM and the behavioral parameter-based SVM for the previous trial's choice. Error bars represent mean ± s.e.m. across datasets. **f-g,** Same as in (**d-e**), but with classifiers for whether or not the previous trial was rewarded.
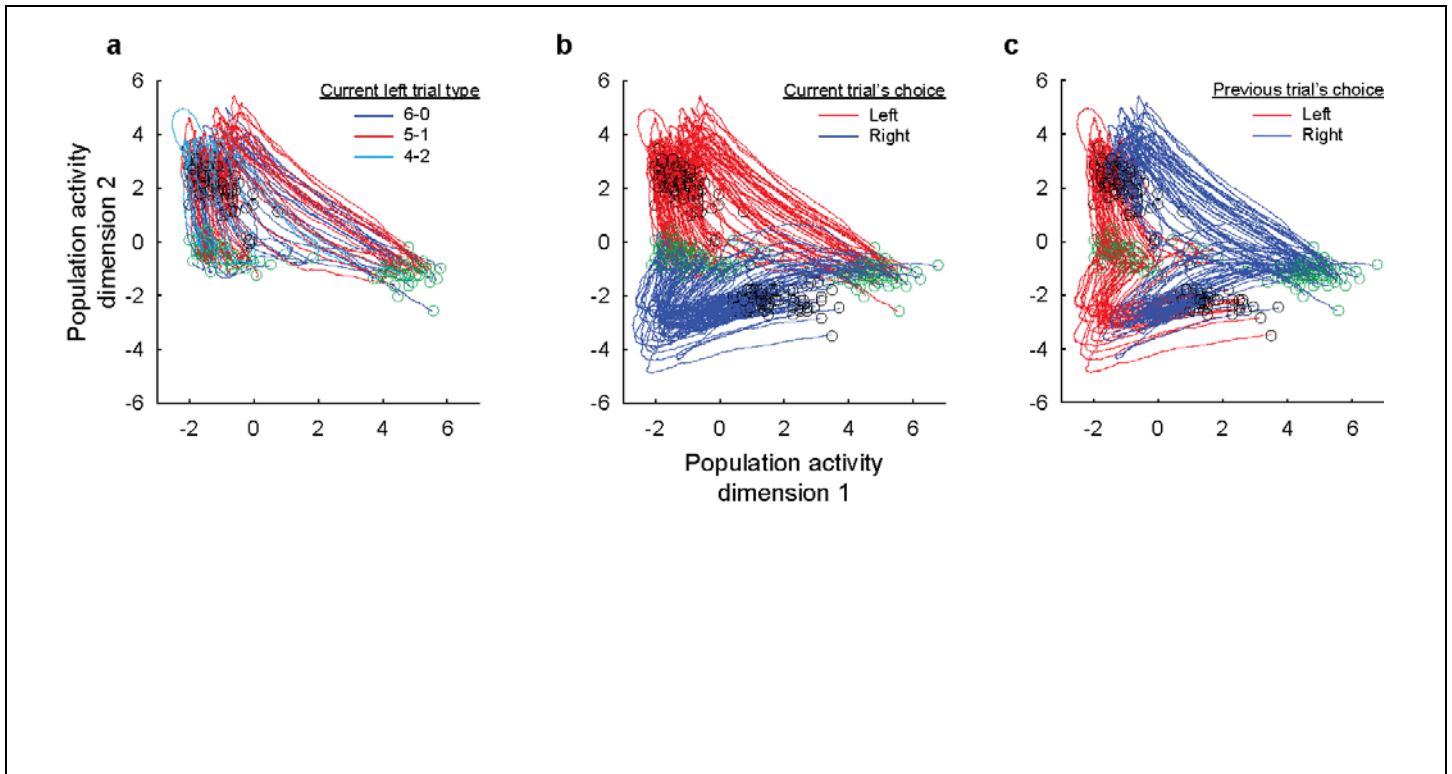
Supplementary Figure 9

Visualizations of trial trajectories.

a-c, Trajectories of correct trials colored by the current trial type (**a**), the current trial's choice (**b**), and the previous trial's choice (**c**). Trials with the same choice but different trial types were highly overlapping (**a**), while trials with different choices were highly different (**b**). Much of the variance within a choice could be explained by the outcome of the previous trial (**c**). Green and black circles mark the trial start and trial end, respectively. For visualization purposes, the dimensionality of the data was reduced using factor analysis.
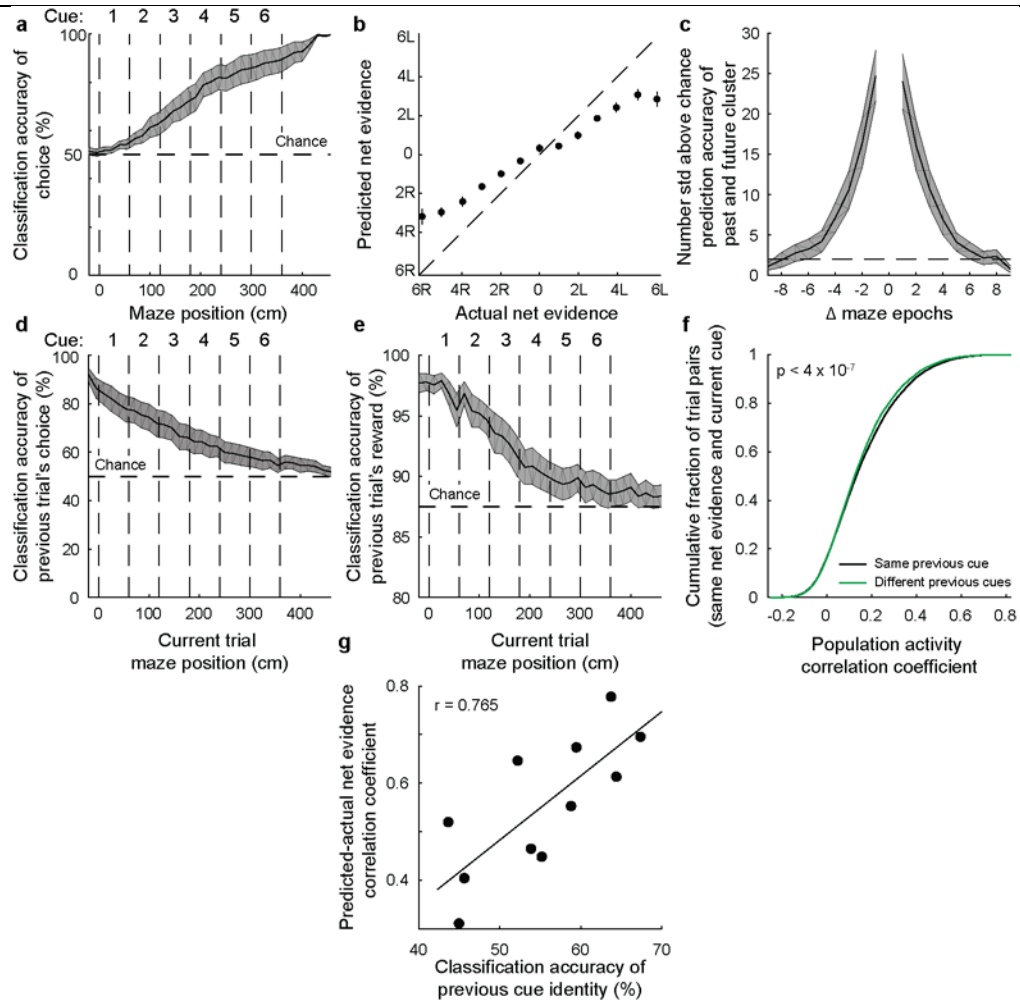
Supplementary Figure 10

Main results re-analyzed using ΔF/F values.

a, Classification accuracy for choice as a function of maze position (SVM, radial basis function kernel). Independent classifiers were trained and tested at each maze position. Error bars represent mean ± s.e.m. across datasets. Compare to Fig. 2e. **b,** Actual net evidence vs. net evidence predicted by a SVR classifier. Error bars represent mean ± s.e.m. across datasets. Compare to Fig. 2f. **c,** For a given trial based on the current epoch's cluster identity, the accuracy of predicting the clusters occupied by that trial in the past and future epochs, compared to shuffled assignments of trials to clusters. Error bars represent mean ± s.e.m. across datasets. Compare to Fig. 4e. **d-e,** Classification accuracy as in (**a**), but for previous trial's choice and for whether the previous trial was rewarded (**e**). Compare to Fig. 5b-c. **f,** Cumulative distribution of the pairwise trial-trial population activity correlation coefficients for trials with the same (black) or different (green) previous cues given the same maze epoch and same net evidence (p < 4 x 10⁻⁷, two-sample KS test). Compare to Fig. 6a. **g,** Relationship between classification accuracy of the previous cue and the classification accuracy of net evidence across datasets (r = 0.76, p < 0.001). Compare to Fig. 6b.