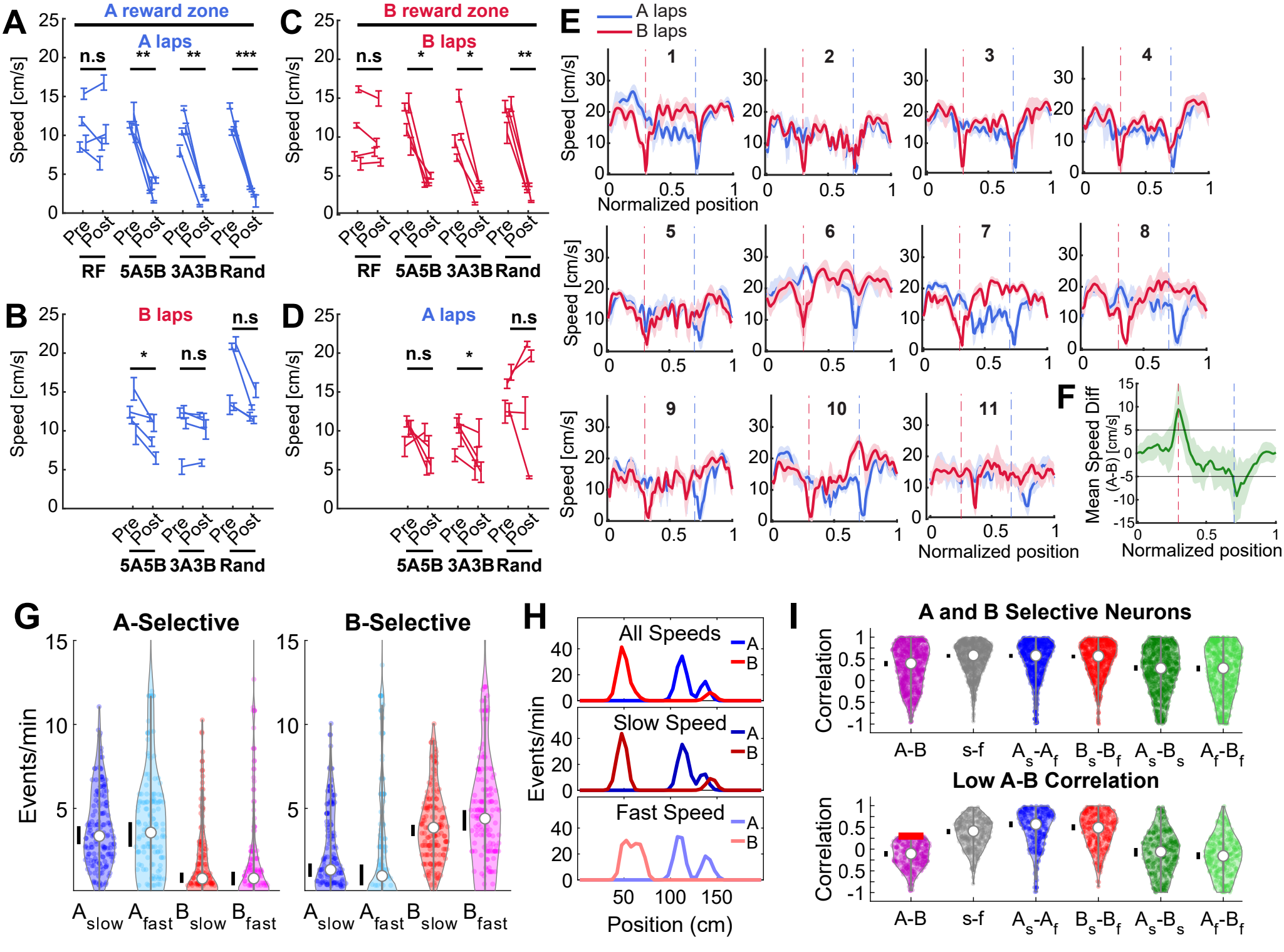
Cell Reports, Volume *41*

Supplemental information

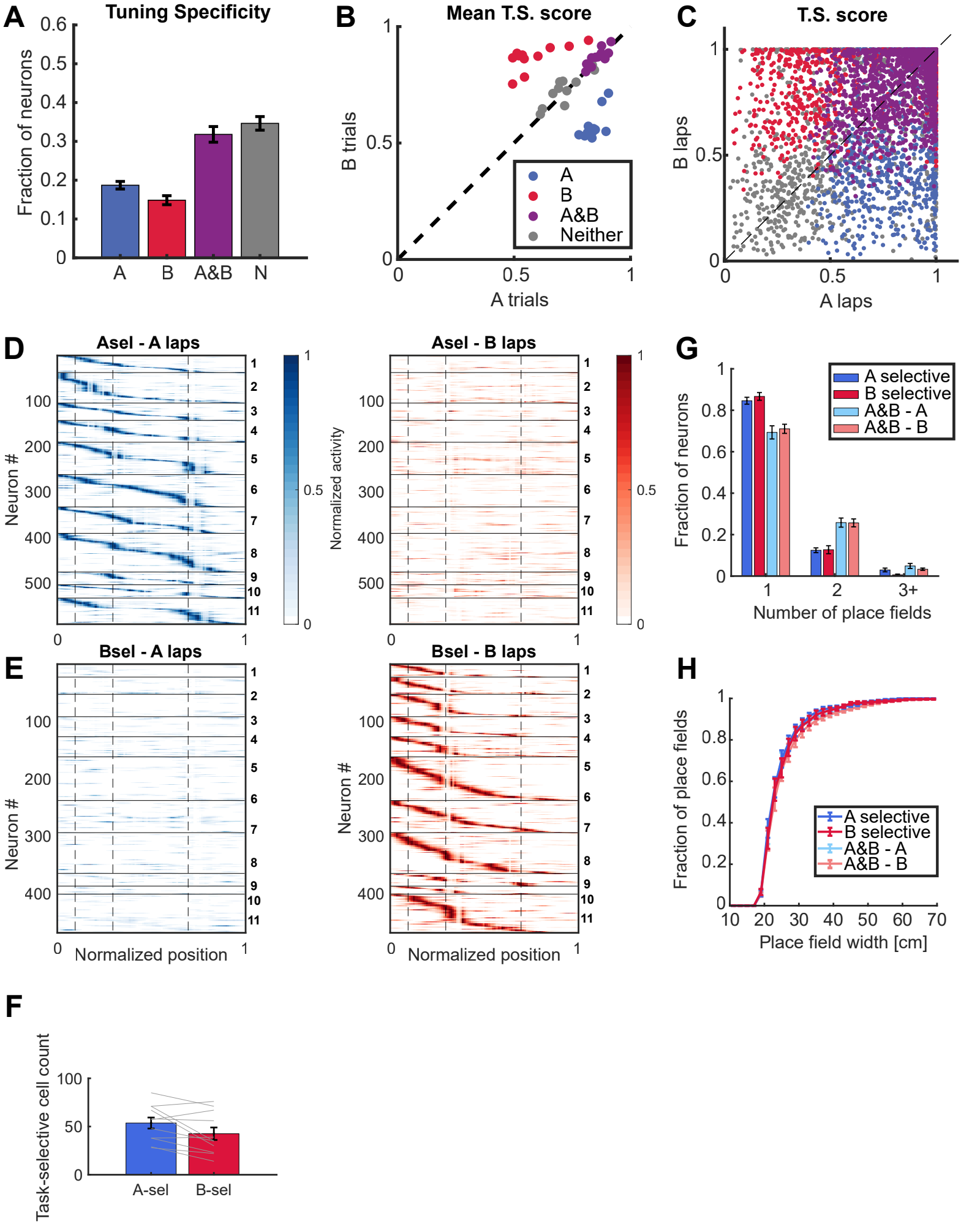
Task-selective place cells show behaviorally driven dynamics during learning and stability during memory recall

Roland Zemla, Jason J. Moore, Maya D. Hopkins, and Jayeeta Basu

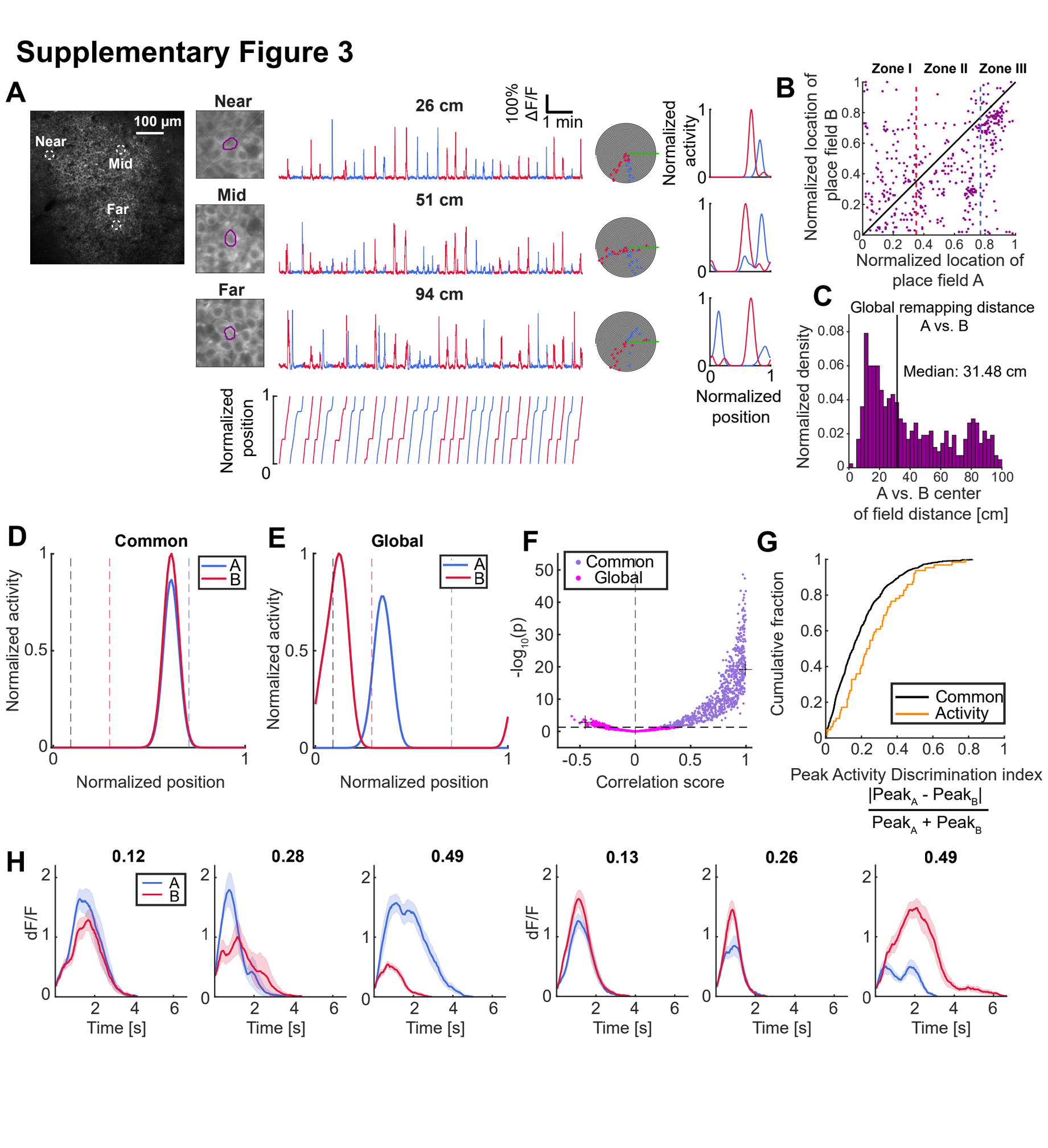
Supplementary Figure 1



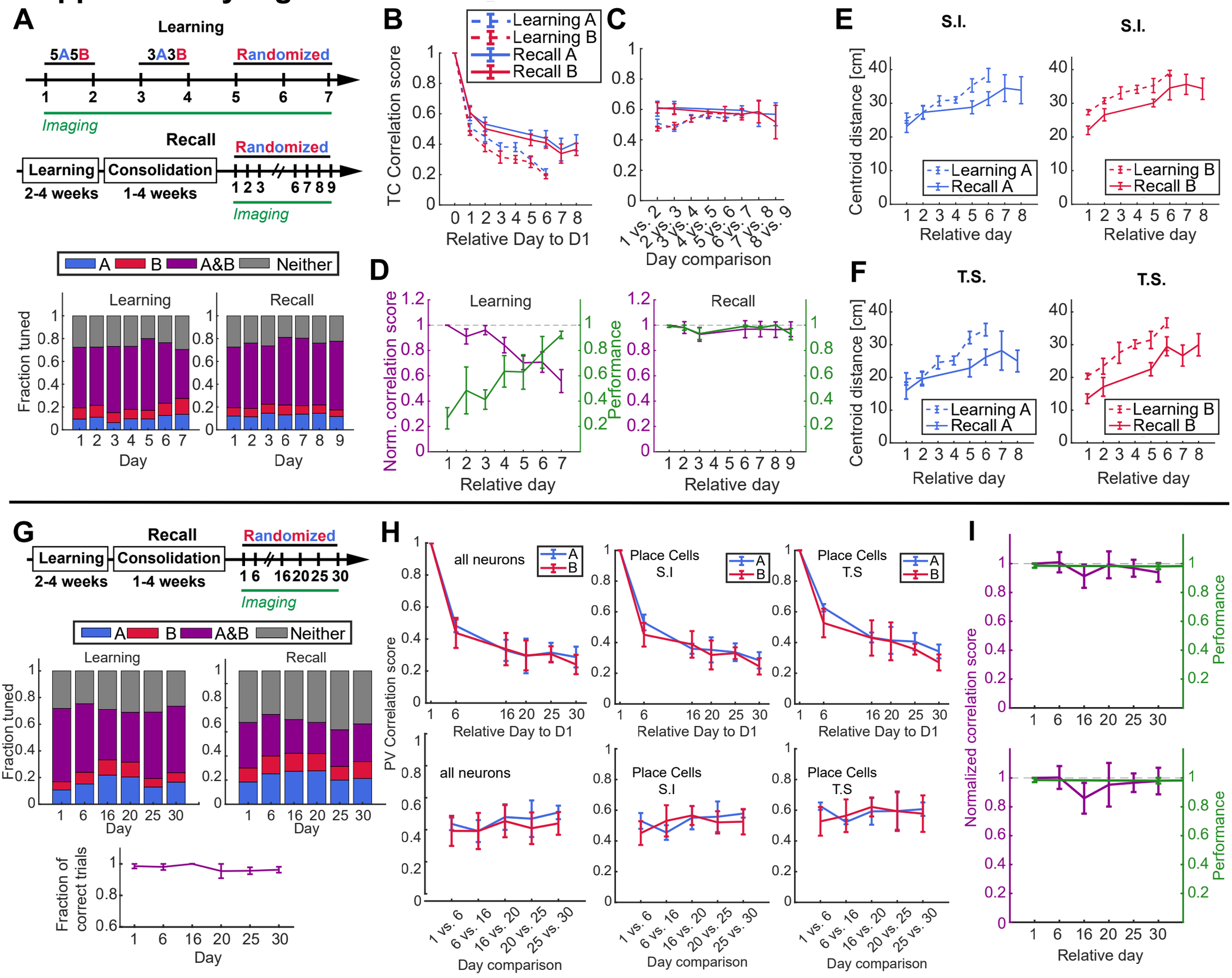
Supplementary Figure 2



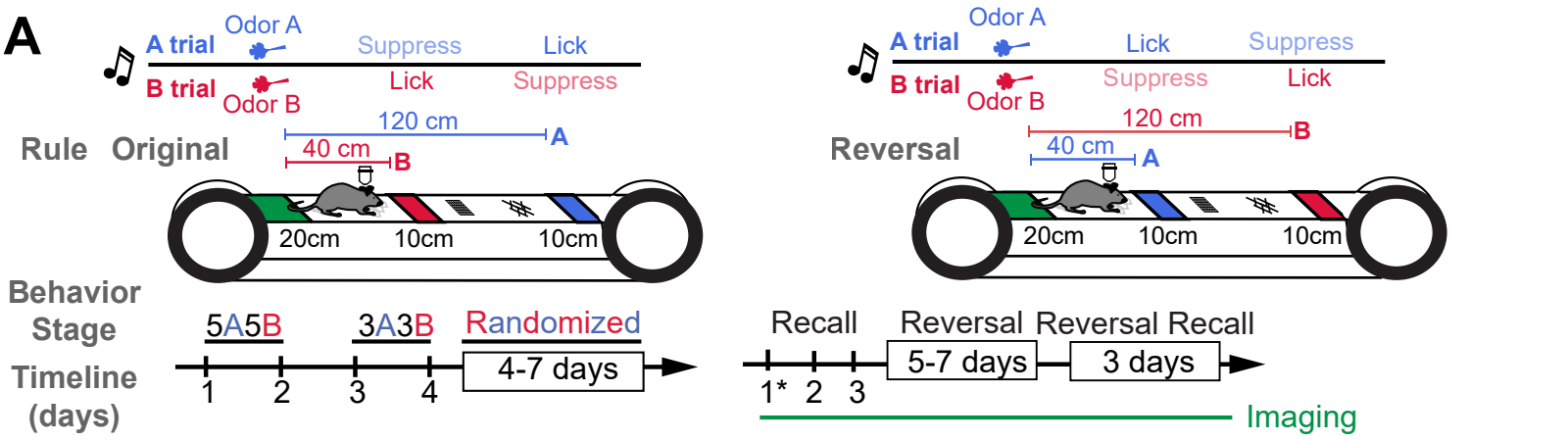
Supplementary Figure 3

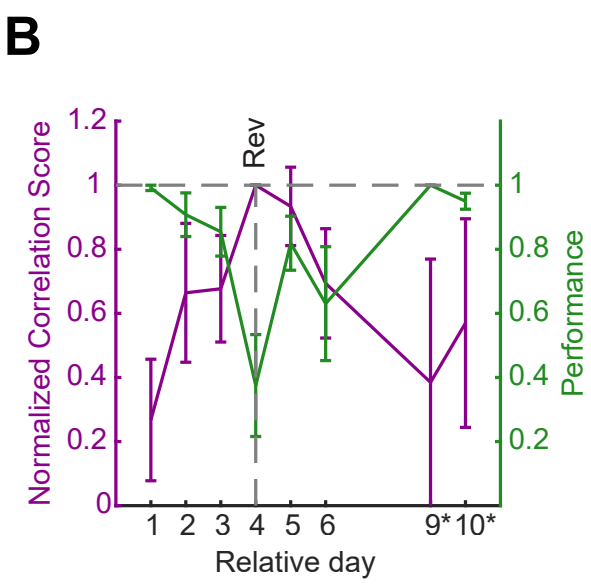


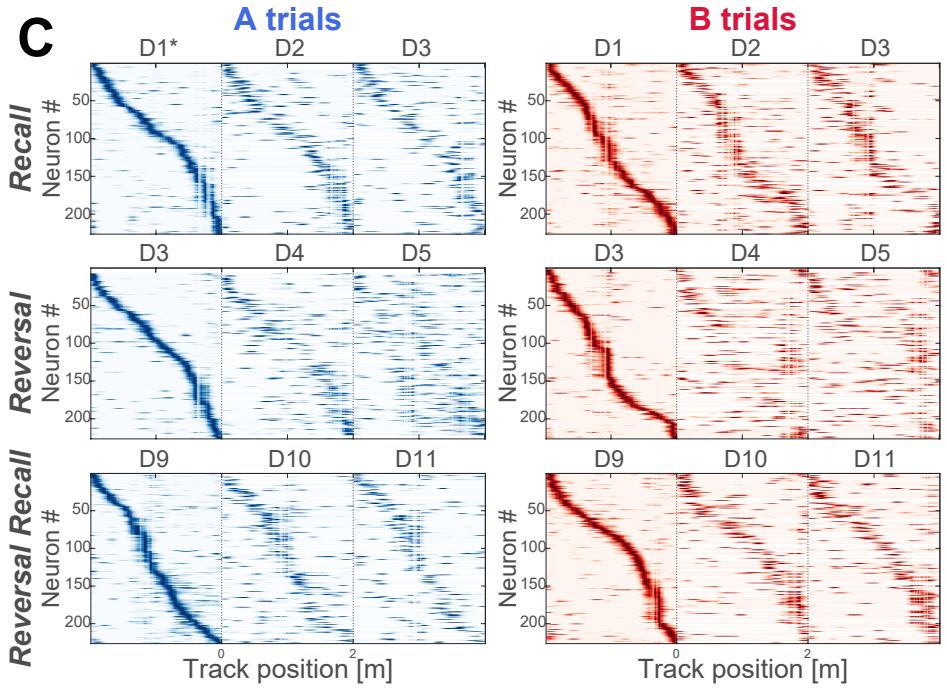
Supplementary Figure 4

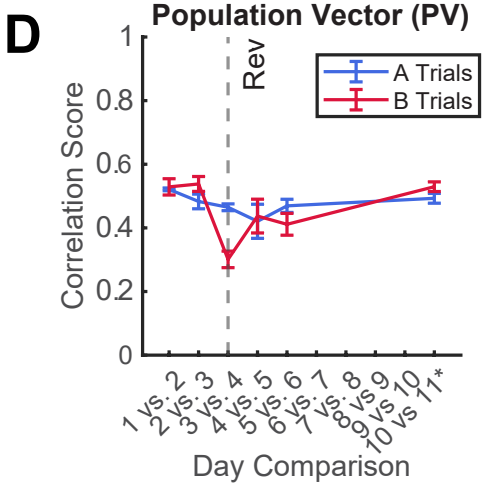


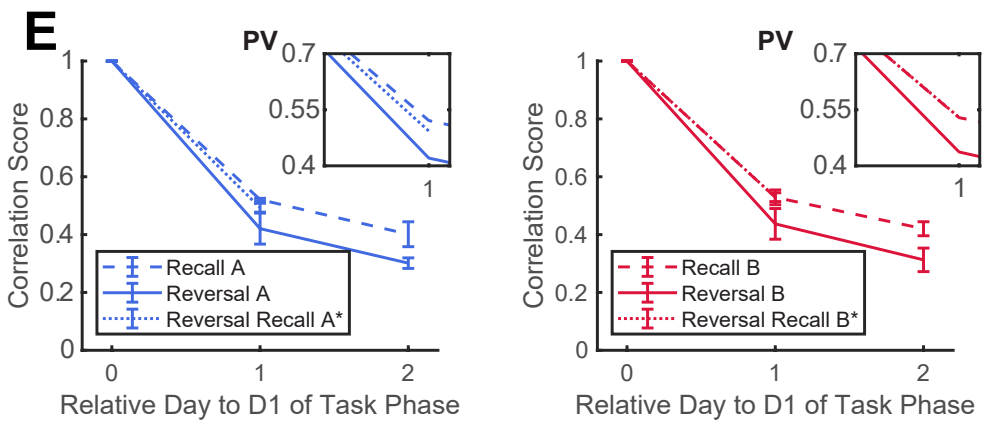
Supplementary Figure 5

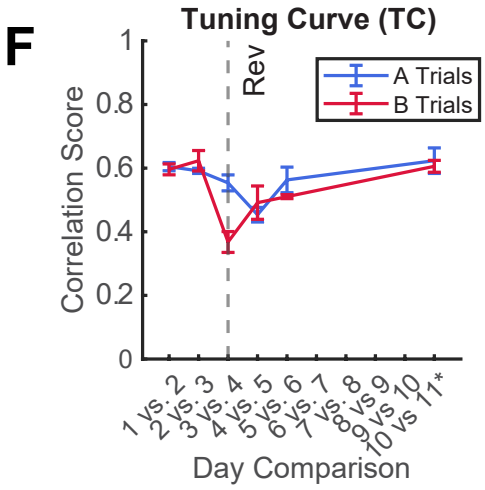


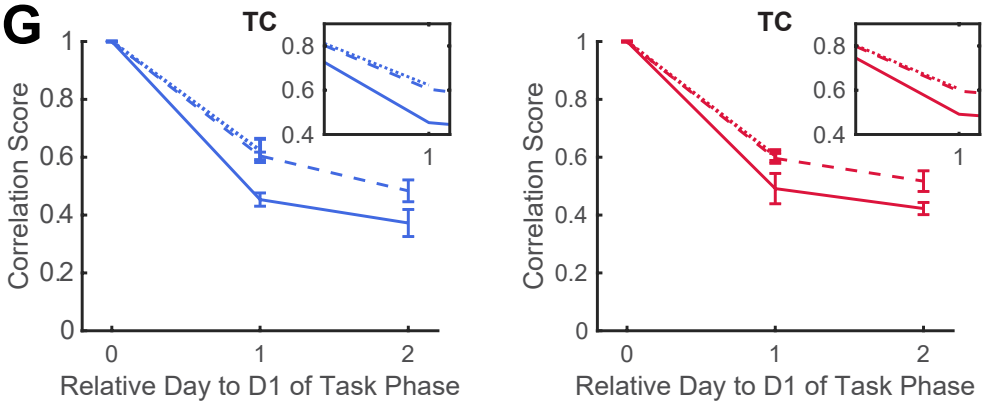




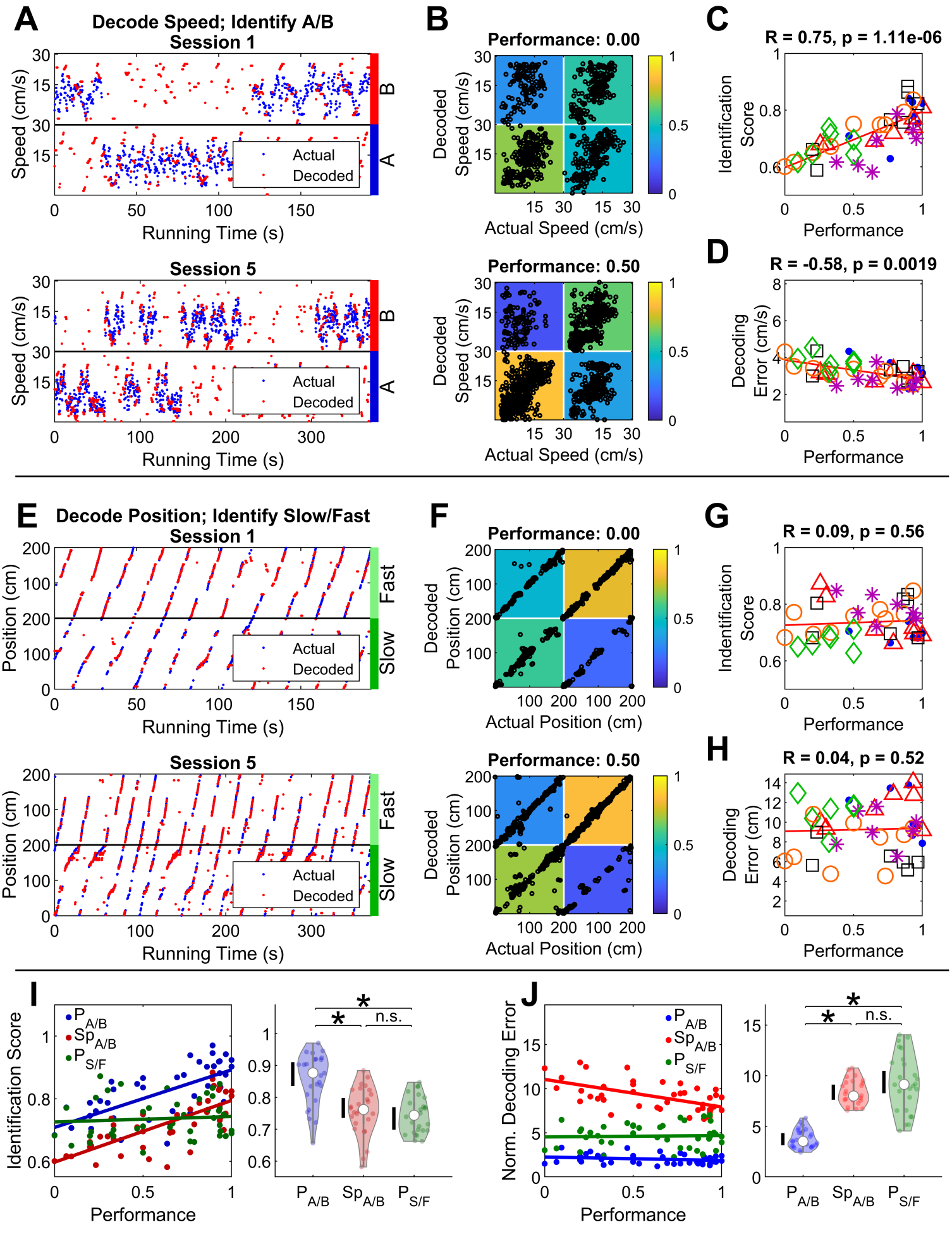




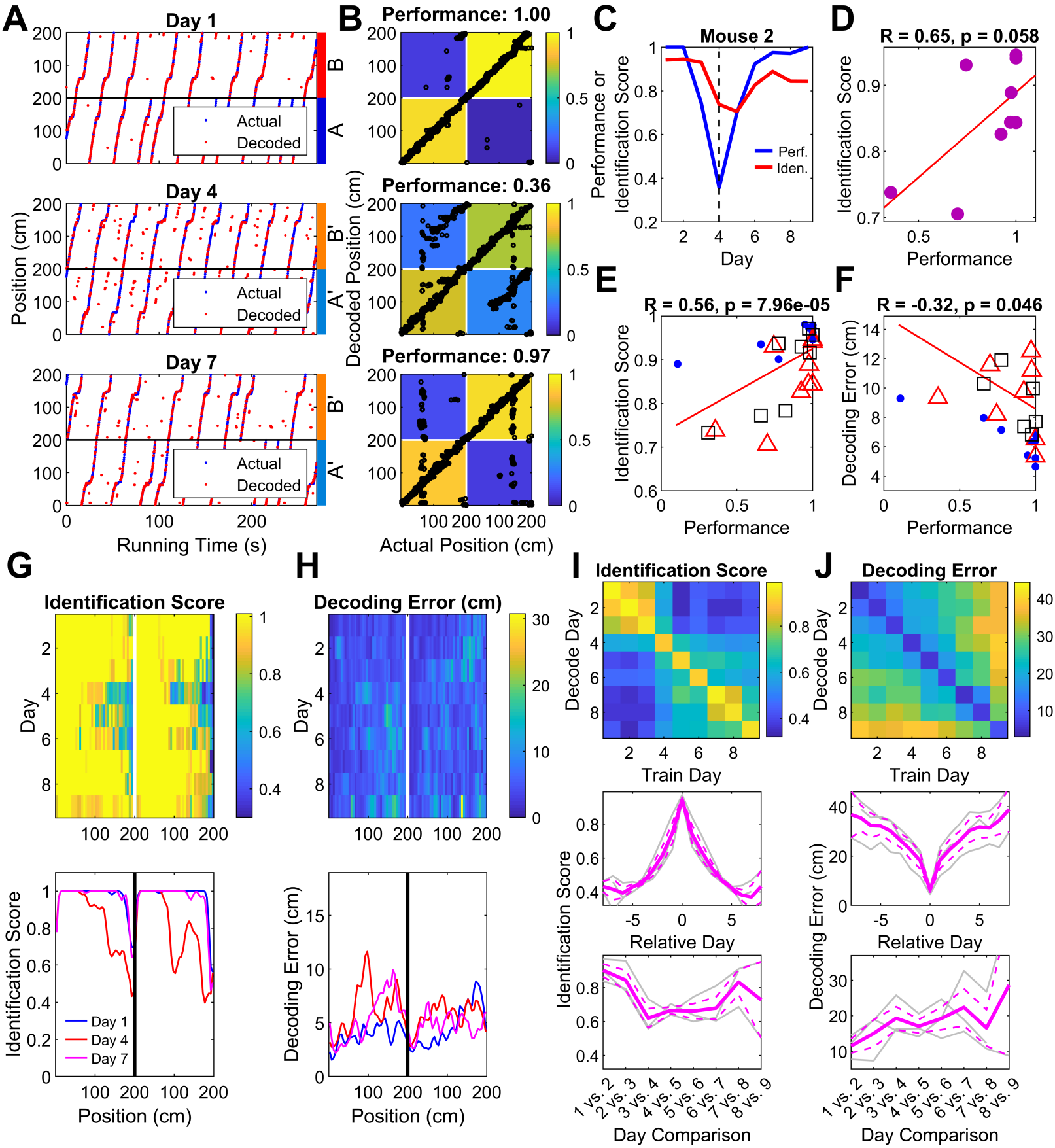




Supplementary Figure 6



Supplementary Figure 7



Supplemental Data Figure Legends:

Supplementary Fig. 1: Peri-reward zone speed across training stages and speed of each animal along track during learned behavior. Trial-selective neurons are not determined by speed differences. Related to Figure 1.

Analysis of speed in the area immediately prior to reward zone entry and within the zone across progressive training stages revealed a consistent decline in mean speed within trial-appropriate reward zones. This was consistent across all animals. Further analyses demonstrate that A-trial and B-trial selective cells are not driven by speed differences, but rather contextual cues.

(**A**) Animal speed in the pre- and post- A reward zone area across training stages on A laps. A statistically significant drop in speed was observed on all training stages with no difference in speed observed during random foraging (Pre- vs. post- zone speed, A reward zone, A laps, RF: 11.18 ± 1.57 vs. 10.56 ± 2.22, paired *t*-test, t_3 = 0.557, *P* = 0.617; Random: 11.47 ± 0.78 vs. 2.5 ± 0.38, paired *t*-test, t_3 = 13.699, ****P* < 0.001, n = 4 mice).

(**B**) In contrast, no significant trend was observed in the A reward zone on B trial laps with no significant difference on the last training stage (A laps, 5A5B: 12.17 ± 1.23 vs. 9.45 ± 1.19, paired *t*-test, t_3 = 3.60, **P* = 0.037; Random: 17.12 ± 2.25 vs. 12.87 ± 0.86, paired *t*-test, t_3 = 2.701, *P* = 0.074, n = 4 mice). Similarly, a significant speed decrease in reward zone B was observed on B trials during the training stages (**C**), with no such trend observed on A trial laps (**D**); Pre- vs. post- zone speed, B reward zone, B laps, RF: 10.21 ± 2.18 vs. 9.55 ± 1.8, paired *t*-test, t_3 = 1.037, *P* = 0.376; Random: 12.73 ± 0.84 vs. 2.9 ± 0.51, paired *t*-test, t_3 = 9.377, ***P* = 0.003, n = 4 mice; A laps, 5A5B: 9.97 ± 0.69 vs. 7.12 ± 1.09, paired *t*-test, t_3 = 1.632, *P* = 0.201; Random: 14.69 ± 1.25 vs. 14.28 ± 3.93, paired *t*-test, t_3 = 0.140, *P* = 0.898, n = 4 mice). Speed was calculated over a 2 second interval along the length of track immediately prior to zone entry and 2 second interval following zone entry across all laps. Error bars represent mean ± s.e.m. Paired t-tests were used to calculated significance.

(**E**) For animals used to analyze the trial-selective and remapping properties of place cells in Figs. 2 and 3, all n = 10 mice (11 FOVs) showed similar spatial bin speeds across the track with a characteristic decline in speed in the trial-appropriate peri-reward zones. The mean animal speed was ~15-20 cm/s. Dotted lines represent the spatial bin where the A (blue) or B (red) trial reward zone begins. Solid lines represent mean spatial bin speed across the respective trials laps and shaded area is the s.e.m. in that spatial bin. The track was subdivided into 100 spatial bins for analysis.

(**F**) There was no substantial difference between animal speed between A and B trials (less than 5 cm/s difference) except near the trial-specific reward zones where the animal was expected to slow down. Data shown as mean ± s.e.m. Pre- and vs. post-reward zone speed n= 10 animals, paired t-test A-trials, P>0.5, and B-trials P>0.5.

(**G**, left) Calcium transient rate for neurons selective only in A trials (not in B trials): A_{slow}: 3.35, [2.86, 3.94]; A_{fast}: 3.56, [2.71, 4.18]; B_{slow}: 0.82, [0.56, 1.17]; B_{fast}: 0.83, [0.41, 1.22]; $A_{slow} > B_{slow}$: $p=1.61x10^{-16}$; $A_{fast} > B_{fast}$: $p=3.06x10^{-11}$; $A_{fast} > A_{slow}$: p=0.53; $B_{fast} > B_{slow}$: p=0.89, N=154 neurons for all statistics in (A).

(**G**, right) Calcium transient rate for neurons selective only in B trials (not in A trials): A_{slow}: 1.34, [0.92, 1.72]; A_{fast}: 0.97, [0.41, 1.65]; B_{slow}: 3.85, [3.35, 4.02]; B_{fast}: 4.39, [3.67, 4.89]; $A_{slow} < B_{slow}$: $p=7.87x10^{-15}$; $A_{fast} < B_{fast}$: $p=1.23x10^{-7}$; Afast > Aslow: p=0.06; Bfast > Bslow: p=0.07, N=163 neurons for all statistics in (B).

(**H**) Sample activity maps for a single neuron with different place fields in A and B trials. The locations of these place fields do not appreciably change as a function of running speed.

(**I**, top) For neurons significantly tuned in A and B trials, correlation coefficients between maps in: A vs B: 0.40, [0.33 0.45], n=679 neurons; slow vs fast: 0.58, [0.53 0.61], n=670 neurons; A_{slow} vs A_{fast}: 0.57, [0.52 0.61], n=619 neurons; B_{slow} vs B_{fast}: 0.56, [0.51 0.59], n=631 neurons; A_{slow} vs B_{slow}: 0.28, [0.23 0.35], n=644 neurons; A_{fast} vs B_{fast}: 0.28, [0.21 0.35]; A-B vs Slow-Fast: $p=3.97x10^{-25}$, n=572 neurons.

(**I,** bottom) For neurons significantly tuned in A and B trials with A-B correlation < 0.3, correlation coefficients between maps in: A vs B: -0.11, [-0.17 -0.45], n=305 neurons; slow vs fast: 0.41, [0.34 0.46], n=303 neurons; A_{slow} vs A_{fast}: 0.57, [0.50 0.64], n=281 neurons; B_{slow} vs B_{fast}: 0.49, [0.43 0.57], n=290 neurons; A_{slow} vs B_{slow}: -0.06, [-0.17 0.03], n=286 neurons; A_{fast} vs B_{fast}: -0.16, [-0.23 -0.07]; A-B vs Slow-Fast: $p=6.57x10^{-45}$, n=267 neurons. Throughout the figure, data are taken from day 6 of the learning phase, and designated significant if they satisfied the spatial information tuning selection criterion.

Supplementary Fig. 2: Fraction of task-selective neurons according to tuning specificity scoring of spatial tuning and place field properties of task-selective and shared place cells. Related to Figure 2.

(**A**) Fraction of place cells tuned according to tuning specificity (T.S.) showed distinct distributions between trials similar to that observed when classified according to spatial information scoring (T.S. fraction: Friedman test, χ_3^2 = 27.33, ***P < 0.001). More A- than B-selective neurons were generally present using the T.S. score (0.19 ± 0.01 vs. 0.15 ± 0.01, paired Wilcoxon signed-rank test, W_{10} = 48, *P = 0.032). Both A- and B-selective were fewer in number compared to A&B neurons using the T.S. score (A vs. A&B: 0.19 ± 0.01 vs. 0.32 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = -66, **P = 0.003; B vs. A&B: 0.15 ± 0.01 vs. 0.32 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = -66, **P = 0.003). Error bars represent mean ± s.e.m.

(**B**) Mean tuning specificity score for each class of neurons on A and B laps during each session.

(**C**) Tuning specificity scores from all imaged neurons (5158 neurons).

(**D**) A-trial selective place cell maps separated by animal/FOV (numerals on right side separated by continuous horizontal black lines) show A-selective place cells are present across all animals. Plot on the right shows the absence of spatial tuning on B trials.

(**E**) B-selective place cells are also observed across all animals.

(**F**) Bar plot showing that there are more trial-selective A place cells compared to B selective place cells (A- vs. B-selective place counts: 53.6 ± 5.7 vs. 42.5 ± 6.5 place cells, paired *t*-test, t_7 = -2.43, **P* = 0.035). Error bars indicate mean +/- s.e.m.

(**G**) Most task-selective place cells had single place fields on both A and B trials with a higher fraction of B-selective place cells having single place fields compared to place fields on B trials for task shared place cells. A-selective place cells were less likely to have two place fields relative to task shared place cells on A trials (Place field count, single, A vs. B: 0.85 ± 0.02 vs. 0.87 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = -14, *P* = 0.577; A vs. A&B-A: 0.85 ± 0.02 vs. 0.69 ± 0.03, paired Wilcoxon signed-rank test, W_{10} = 66, ***P* = 0.003; A vs. A&B-B: 0.85 ± 0.02 vs. 0.71 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = 66, ***P* = 0.003, Place field count, double, A vs. B: 0.12 ± 0.01 vs. 0.13 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = -10, *P* = 0.7; A vs. A&B-A: 0.12 ± 0.01 vs. 0.26 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = -66, ***P* = 0.003; A vs. A&B-B: 0.12 ± 0.01 vs. 0.26 ± 0.02, paired Wilcoxon signed-rank test, W_{10} = -66, ***P* = 0.003; Place field count, triple, A vs. B: 0.03 ± 0.01 vs. 0.01 ± 0, paired Wilcoxon signed-rank test, W_{10} = 28, **P* = 0.031; A vs. A&B-A: 0.03 ± 0.01 vs. 0.05 ± 0.01, paired Wilcoxon signed-rank test, W_{10} = -37, *P* = 0.064; A vs. A&B-B: 0.03 ± 0.01 vs. 0.03 ± 0.01, paired Wilcoxon signed-rank test, W_{10} = -64, ***P* = 0.006, *n* = 11 FOV from 10 mice).

(**H**) The distribution of place field widths did not differ between selective and shared place cells (Place field width, A vs. B: 26.37 ± 0.27 vs. 27.44 ± 0.35, 2-sample Kolmogorov-Smirnov test, $D_{695, 542}$ = 0.09, *P* = 0.069, *n* = 695 vs. 542 neurons; A vs. A&B-A: 26.37 ± 0.27 vs. 26.85 ± 0.14, 2-sample Kolmogorov-Smirnov test, $D_{695, 3655}$ = 0.03, *P* = 0.765, *n* = 695 vs. 3655 neurons; A vs. A&B-B: 26.37 ± 0.27 vs. 27.45 ± 0.14, 2-sample Kolmogorov-Smirnov test, $D_{695, 3533}$ = 0.05, *P* = 0.286, *n* = 695 vs. 3533 neurons, B- vs. A&B-A: 27.44 ± 0.35 vs. 26.85 ± 0.14, 2-sample Kolmogorov-Smirnov test, $D_{542, 3655}$ = 0.1, ***P* = 0.002, *n* = 542 vs. 3655 neurons; B- vs. A&B-B: 27.44 ± 0.35 vs. 27.45 ± 0.142-sample Kolmogorov-Smirnov test, $D_{542, 3533}$ = 0.04, *P* = 0.765, *n* = 542 vs. 3533 neurons, A&B-A vs. A&B-B: 26.85 ± 0.14 vs. 27.45 ± 0.14, 2-sample Kolmogorov-Smirnov test, $D_{3655, 3533}$ = 0.06, ****P* < 0.001, *n* = 3655 vs. 3533 neurons, 11 FOV from 10 mice). All data shown as mean ± s.e.m.

Supplementary Fig. 3: Distribution of globally remapping cells and their classification by spatial tuning curve correlation. Related to Figure 3.

Global remapping place cells had place fields that remapped across the entire length of the track with the majority of place fields showing mean centroid shifts of less than 30 cm.

(**A**, left) Imaging fields of view with three examples of globally remapping place cells categorized according to the distance between their place field centroids on A vs. B trials.

(A, right) Examples of three globally remapping neurons with increasing place field centroid differences. Magnified FOVs correspond to the neurons labeled on the left with calcium traces, polar event plots, and spatial tuning curves, respectively, shown on the right.

(**B**) Scatterplot of the place field centroids for all globally remapping neurons on A vs. B trials. Place cells with place fields on A laps tend to shift to earlier locations on the lap on B trials, particularly in Zone II of the track. Dashed red line indicates location of the start of reward zone B, while dashed blue line indicates the start of reward zone B.

(**C**) Distribution of place field centroid differences between A and B laps. The centroid difference is skewed toward shorter distances with a median of ~31 cm.

(**D**) A non-remapping place cell with a common place field that does not shift between lap trials and demonstrates a positive spatial tuning curve correlation value.

(**E**) A remapping place cell with a shifting place field between trials shows a negative spatial tuning curve correlation indicating remapping

(**F**) Scatterplot of the Pearson's correlation value (r) of spatial tuning curves against the negative logarithm of the associated p values for each place cell having a significant place field in both A and B trials. Place cells that had either negative and/or a correlation p-value < 0.05 were considered global remapping place cells (magenta) while the remaining cells were classified as common (purple). Cross mark on the right side of the scatter plot indicates common place cell shown in (A), while cross mark on the left side shows global remapping place cell in (B). Dashed horizonal line is the negative logarithm of cutoff p-value = 0.05 and vertical dashed line represents a correlation score of 0. The list of all correlation scores and corresponding p-values can be found in Table 1.

(**G**) The activity discrimination index validates the presence of a distinct, activity-remapping subpopulation of common place cells that share the same place field on A and B trials, but differ by their trial-specific in-field place activity. A cumulative density plot showing a significant difference in the distribution of activity discrimination scores between in-field place field activity on A vs. B trials calculated according to the formula shown below (Common vs activity remapping index score: 0.19 ± 0.01 vs. 0.26 ± 0.02,

2-sample Kolmogorov-Smirnov test, $D_{700, 64}$ = 0.24, **P = 0.001, n = 700 vs. 64 neurons, 11 FOV from 10 mice).

(**H**) Mean in-field calcium transients aligned to the onset time of the transients for activity remapping place cells which show higher mean in-field activity on A trials (left three plots) and those which higher in-field mean activity on B trials (right three plots). Neurons are shown according to increase activity index score from left to right.

Supplementary Fig. 4: Characterization of spatial stability and learning-dependent spatial tuning decorrelation using spatial information criterion and trial-specific mean place field centroid distance; long-term stability of place maps during recall. Related to Figure 4.

Using the spatial information (S.I.) tuning selection criterion, the results of learning-dependent decorrelation and stability of recall maps were similar to those we observed when using the tuning specificity (T.S.) spatial tuning criterion.

Place maps remained stable beyond the one-week imaging interval with a slow decay in map correlation relative to first imaging day on subsequent weeks.

(**A,** top) Schematic illustrating the accelerated training regimen and imaging schedule for odor-cued spatial navigation. Mice were advanced to each stage of task training after 2 days rather than 1 week (see Methods). Below is the timeline for memory recall imaging experiments following learning.

(**A**, bottom) The fraction of place cells tuned in each trial did not show any significant trend during learning (left) (Fraction of tuned place cells, A trials, learning: one-way RM mixed effects analysis, effect of training day, $F_{6, 21.93}$ = 2.34, P = 0.067; B-trials: one-way RM mixed effects analysis, effect of training day, $F_{6, 27}$ = 1.55, P = 0.201, n = 6 mice) or recall (A trials, recall: one-way RM mixed effects analysis, effect of training day, $F_{6, 24}$ = 0.39, P = 0.88; B trials: one-way RM mixed effects analysis, effect of training day, $F_{6, 24}$ = 0.72, P = 0.637, n = 5 mice).

(**B**) The correlation of spatial tuning curves for S.I.-tuned neurons relative to the first day of imaging similarly showed a significantly steeper decline during learning compared to recall on A and B trials (TC correlation, A trials: two-way RM mixed effects analysis, effect of time, $F_{3, 22.23}$ = 57.28, ***P < 0.001, effect of behavior, $F_{1, 8.93}$ = 8.08, *P = 0.019, interaction between time and behavior, $F_{3, 22.23}$ = 3.46, *P = 0.034; B trials: two-way RM mixed effects analysis, effect of time, $F_{3, 22.64}$ = 57.54, ***P < 0.001, effect of behavior, $F_{1, 9.14}$ = 17.28, **P = 0.002, interaction between time and behavior, $F_{3, 22.64}$ = 2.02, P = 0.14, n = 6 learn cohort, 5 recall cohort mice; Day 2 vs. 7, learning, A trials: 0.47 ± 0.01 vs. 0.2 ± 0.02, paired t-test, t_2 = 8.1, *P = 0.015; B trials: 0.46 ± 0.01 vs. 0.19 ± 0.02, paired t-test, t_2 = 8.55, *P = 0.027, n = 3 mice; Day 2 vs. 7, recall, A trials: 0.6 ± 0.05 vs. 0.44 ± 0.05, paired t-test, t_4 = 5.89, **P = 0.008; B trials: 0.61 ± 0.03 vs. 0.41 ± 0.04, paired t-test, t_4 = 11.71, ***P < 0.001, n = 5 mice; Day 7, learning vs. recall, A trials: 0.22 ± 0.02 vs. 0.44 ± 0.05, unpaired t-test, t_7 = -3.6, **P = 0.009; B trials: 0.19

± 0.02 vs. 0.41 ± 0.04,unpaired *t*-test, t_7 = -5.21, ***P* = 0.002, *n* = 4 learn cohort, 5 recall cohort mice).

(**C**) The same pattern of stability emerged when S.I. tuned place cells were used for the neighboring day correlation analysis during learning and recall. Learning and recall plots split for ease of visualization (Neighboring session TC correlation, A trials: two-way RM mixed effects analysis, effect of time, $F_{2, 14.33}$ = 2.93, *P* = 0.086, effect of behavior, $F_{1, 9.2}$ = 3.7, *P* = 0.086, interaction between time and behavior, $F_{2, 14.33}$ = 5.47, **P* = 0.017; B trials: two-way RM mixed effects analysis, effect of time, $F_{2, 13.92}$ = 0.81, *P* = 0.466, effect of behavior, $F_{1, 8.61}$ = 4.55, *P* = 0.063, interaction between time and behavior, $F_{2, 13.92}$ = 6.72, ***P* = 0.009, *n* = 6 learn cohort, 5 recall cohort mice; Days 1 vs. 2 Vs. Day 6 vs. 7, learning, A trials: 0.47 ± 0.01 vs. 0.56 ± 0.02, paired *t*-test, t_2 = -15.6, ***P* = 0.004; B trials: 0.46 ± 0.01 vs. 0.55 ± 0.01, paired *t*-test, t_2 = -6.36, **P* = 0.024, *n* = 3 mice; Days 1 vs. 2 Vs. Day 6 vs. 7, recall, A trials: 0.6 ± 0.05 vs. 0.59 ± 0.03, paired *t*-test, t_4 = 0.32, *P* = 0.764; B trials: 0.61 ± 0.03 vs. 0.57 ± 0.04, paired *t*-test, t_4 = 1.4, *P* = 0.234, *n* = 5 mice).

(**D**) The decorrelation between A&B tuned place cells (based on S.I. criterion) was also inversely correlated to the animals' performance during training (left) in contrast to recall session (right; Normalized A vs. B lap correlation scores, learning: Kruskal-Wallis test, H_5 = 105.83, ****P* < 0.001, *n* = 1967 neurons from 6 mice; Day 2: 0.91 ± 0.06, 1-sample Wilcoxon signed-rank test against 1, W_{316} = -10531, ***P* = 0.001, *n* = 317 neurons; Day 7: 0.56 ± 0.09, 1-sample Wilcoxon signed-rank test against 1, W_{236} = -16113, ****P* < 0.001, *n* = 237 neurons; Normalized A vs. B lap correlation scores, recall: Kruskal-Wallis test, H_5 = 4.34, *P* = 0.502, *n* = 1577 neurons from 5 mice; Day 2: 0.98 ± 0.05, 1-sample Wilcoxon signed-rank test against 1, W_{338} = 228, *P* = 0.95, *n* = 339 neurons; Day 7: 0.97 ± 0.06, 1-sample Wilcoxon signed-rank test against 1, W_{228} = -2699, *P* = 0.179, *n* = 229 neurons). All data represented as mean ± s.e.m. except for normalized A vs. B correlation scores which are represented median ± 95% C.I

(**E**) The mean centroid difference between place fields relative to day 1 of imaging increased more during learning compared to recall on A and B laps (Mean centroid difference relative to Day 1 A trials: two-way RM mixed effects analysis, effect of time, $F_{3, 22.44}$ = 28.61, ****P* < 0.001, effect of behavior, $F_{1, 8.9}$ = 2.6, *P* = 0.141, interaction between time and behavior, $F_{3, 22.44}$ = 2.67, *P* = 0.072; B trials: two-way RM mixed effects analysis, effect of time, $F_{3, 24.47}$ = 17.03, ****P* < 0.001, effect of behavior, $F_{1, 9.49}$ = 11.83, ***P* = 0.007, interaction between time and behavior, $F_{3, 24.47}$ = 0.06, *P* = 0.981, *n* = 6 learn cohort, 5 recall cohort mice; Learning vs. recall, Day 5, A trials 35.29 ± 1.97 vs. 28.8 ± 1.91, unpaired *t*-test, t_8 = 2.36, *P* = 0.09; B trials: 35.27 ± 2.1 vs. 30.07 ± 1.14, unpaired *t*-test, t_8 = 2.17, *P* = 0.119, *n* = 5 learn cohort, 5 recall cohort mice).

(**F**) The same effects was observed using tuning specificity metric (Mean centroid difference relative to Day 1, A trials: two-way RM mixed effects analysis, effect of time, $F_{3, 22.19}$ = 30.7, ****P* < 0.001, effect of behavior, $F_{1, 8.62}$ = 3.52, *P* = 0.095, interaction between time and behavior, $F_{3, 22.19}$ = 2.78, *P* = 0.065, *n* = 6 learn, 5 recall mice; B trials:

two-way RM mixed effects analysis, effect of time, $F_{3, 23.34} = 25.56$, ***$P < 0.001$, effect of behavior, $F_{1, 8.98} = 11.81$, **$P = 0.007$, interaction between time and behavior, $F_{3, 23.34} = 0.14$, $P = 0.936$, $n = 6$ learn, 5 recall mice; Learning vs. recall, Day 5, A trials: 32.16 ± 1.89 cm vs. 22.83 ± 2.64 cm, unpaired *t*-test, $t_8 = 2.88$, *$P = 0.041$; B trials: 31.45 ± 2.6 cm vs. 22.51 ± 2.01 cm, unpaired *t*-test, $t_8 = 2.72$, $P = 0.052$, $n = 5$ learn, 5 recall mice).

(**G**, top) Schematic illustrating the timeline for the 30-day cohort.

(**G**, middle) The relative fraction of place cells tuned to either trial type remained stable across weeks with no trend in relative distribution using either spatial information (S.I.; left) and tuning specificity (T.S.; right) to select spatial tuned place cells (Fraction tuned place cells [S.I.], A trials: one-way RM mixed effects analysis, effect of training day, $F_{5, 7.07} = 2.01$, $P = 0.193$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{5, 7.05} = 0.47$, $P = 0.789$, $n = 3$ mice; Fraction tuned place cells [T.S.], A trials: one-way RM mixed effects analysis, effect of training day, $F_{5, 7.13} = 1.83$, $P = 0.224$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{5, 7.54} = 0.34$, $P = 0.877$, $n = 3$ mice).

(**G**, bottom) The performance of all imaged animals (n=3) remained stable at or near 100% on all long-term recall sessions.

(**H**, top) The stability of spatial maps dropped significantly on sixth day imaging session, but remained at stable correlation score ~0.4-0.5 on sessions thereafter for all neurons (left) and those significantly tuned to space on using S.I. or T.S criteria (right; PV correlation, A trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 5.08} = 11.39$, **$P = 0.01$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 5.01} = 78.75$, ***$P < 0.001$, $n = 3$ mice; Day 6 vs. Day 30, A trials: 0.48 ± 0.04 vs. 0.29 ± 0.06, paired *t*-test, $t_1 = 3.36$, $P = 0.3$; B trials: 0.44 ± 0.09 vs. 0.24 ± 0.06, paired *t*-test, $t_1 = 20.54$, $P = 0.061$, $n = 2$ mice; TC correlation [S.I.], A trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 5.09} = 20.19$, **$P = 0.003$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 5.04} = 43.78$, ***$P < 0.001$, $n = 3$ mice; Day 6 vs. Day 30, A trials: 0.53 ± 0.05 vs. 0.28 ± 0.05, paired *t*-test, $t_1 = 7.74$, $P = 0.157$; B trials: 0.45 ± 0.08 vs. 0.24 ± 0.05, paired *t*-test, $t_1 = 14.17$, $P = 0.088$, $n = 2$ mice; TC correlation [T.S.],A trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 5.21} = 15.82$, **$P = 0.004$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 5.09} = 14.9$, **$P = 0.005$, $n = 3$ mice; Day 6 vs. Day 30, A trials: 0.63 ± 0.02 vs. 0.34 ± 0.05, paired *t*-test, $t_1 = 6.45$, $P = 0.186$; B trials: 0.53 ± 0.09 vs. 0.27 ± 0.05, paired *t*-test, $t_1 = 23.39$, $P = 0.054$, $n = 2$ mice).

(**H**, bottom) Neighboring imaging sessions remained stable as well using same analysis (Neighboring session PV correlation, A trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 4.35} = 2.69$, $P = 0.17$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 4.01} = 12.91$, *$P = 0.015$, $n = 3$ mice; Day 1 vs. 6 Vs. Day 20 vs. Day 25, A trials: 0.48 ± 0.04 vs. 0.51 ± 0.11, paired *t*-test, $t_1 = -0.43$, $P = 0.741$; B trial: 0.44 ± 0.09 vs. 0.45 ± 0.1, paired *t*-test, $t_1 = -3.3$, $P = 0.339$, $n = 2$ mice; Neighboring

session TC correlation [S.I.], A trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 4.36} = 4.56$, $P = 0.076$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 4.25} = 1.95$, $P = 0.259$, $n = 3$ mice; Day 1 vs. 6 Vs. Day 20 vs. Day 25, A trials: 0.53 ± 0.05 vs. 0.56 ± 0.1, paired *t*-test, $t_1 = -0.57$, $P = 0.672$; B trials: 0.45 ± 0.08 vs. 0.52 ± 0.07, paired *t*-test, $t_1 = -14.44$, $P = 0.086$, $n = 2$ mice; Neighboring session TC correlation [T.S.], A trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 4.7} = 0.85$, $P = 0.551$; B trials: one-way RM mixed effects analysis, effect of training day, $F_{4, 4.47} = 0.44$, $P = 0.775$, $n = 3$ mice; Day 1 vs. 6 Vs. Day 20 vs. Day 25, A trials: 0.63 ± 0.02 vs. 0.59 ± 0.12, paired *t*-test, $t_1 = 0.33$, $P = 0.96$, $n = 2$ mice; B trials: 0.53 ± 0.09 vs. 0.59 ± 0.13, paired *t*-test, $t_1 = -1.88$, $P = 0.526$, $n = 2$ mice).

(**I**) No decorrelation in place maps was observed between A and B trial laps for matching neurons across time with either place tuning criterion (Spatial Information, top; Tuning specificity, bottom) (Normalized A vs. B lap correlation scores [S.I.]: Kruskal-Wallis test, $H_4 = 6.9$, $P = 0.141$, $n = 786$ neurons from 3 mice; Day 6: 1.01 ± 0.07,1-sample Wilcoxon signed-rank test against 1, $W_{160} = 1365$, $P = 0.249$, $n = 161$ neurons; Day 25: 0.97 ± 0.06, 1-sample Wilcoxon signed-rank test against 1, $W_{172} = -1361$, $P = 0.302$, $n = 173$ neurons; Normalized A vs. B lap correlation scores [T.S.]: Kruskal-Wallis test, $H_4 = 3.68$, $P = 0.451$, $n = 251$ neurons from 3 mice; Day 6: 1 ± 0.08, 1-sample Wilcoxon signed-rank test against 1, $W_{58} = 16$, $P = 0.952$, $n = 59$ neurons; Day 25: 0.97 ± 0.06, 1-sample Wilcoxon signed-rank test against 1, $W_{61} = -319$, $P = 0.263$, $n = 62$ neurons). All data represented as mean ± s.e.m. except for normalized A vs. B correlation scores which are represented median ± 95% C.I.

Supplementary Fig. 5: Reversal learning of odor-associated reward locations induces place map reorganization. Related to Figure 4.

(**A**, top) Schematic illustrating the original odor-reward zone pairings and the reversal pairings.

(**A,** bottom) Behavioral timeline for the reversal cohort. Accelerated training for intial learning of the task. Recall phase was defined as the first day after two consecutive days of >80% performance. Mice performed OC-GOL at the recall stage for 3 days. On the fourth day the odor-reward zone locations were reversed such that odor A now predicted rewards at the near reward zone while odor B now predicted rewards at the far reward zone. All subsequent day mice were trained on this reversed association.

(**B**) Line plots showing the normalized tuning curve (TC) correlation score (magenta) between same-day A and B trials between matching neurons between any two imaging sessions relative to day 4 when reversal learning began and performance (green) as the fraction of total correct trials (Normalized A vs. B lap correlation scores relative to reversal learning: Kruskal-Wallis test, $H_6 = 96.78$, ***$P < 0.001$, $n = 514$ neurons from 3 mice; Day 1 relative to Day 4 (reversal start): 1-sample Wilcoxon signed-rank test against 1, $W_{91} = -3076$, ***$P < 0.001$, $n = 92$ neurons; Day 6 relative to Day 4: 1-sample Wilcoxon signed-rank test against 1, $W_{82} = -1202$, **$P = 0.006$, $n = 83$ neurons). Dashed vertical line denotes the first day of the rule reversal, while horizontal line indicates

default value to maximum correlation between trials. Sessions 1-6 occurred on the same day for all mice (n=3). Sessions 9 and 10, denoted by asterisks, were different days binned together due to different learning rates across mice. These sessions were binned according to performance, but all occurred within a 3-day window. For all mice, session 9 and 10 were consecutive days.

(**C**) Rate maps for a single representative mouse at different stages of the task. Recall (top row) plots the first three days of recall with each map sorted by the neuron ID for Day 1. (left) Rate maps created from A trials only. (right) Rate maps created from B trials only. Reversal (middle row) plots the first two days of rule reversal exposure (Day 4 and 5) and the last day of recall with each map sorted by the neuron ID for Day 3. Reversal recall (bottom row) plots the last three days of imaging with the rule reversal (Day 9, 10, 11) with each map sorted by the neuron ID for Day 9.

(**D**) Population vector (PV) correlation for both A (blue) and B (red) trials between neighboring sessions (PV correlation, A trials, across behavioral phases: one-way RM mixed effects analysis, effect of training day, $F_{5, 11} = 1.11$, $P = 0.408$; Days 1 vs. 2 Vs. Day 3 vs. 4, A trials: paired *t*-test, $t_2 = -5.07$, $^*P = 0.036$, $n = 3$ mice; B trials: one-way RM mixed effects analysis, effect of training day, $F_{5, 11} = 5.49$, $^{**}P = 0.009$; Days 1 vs. 2 Vs. Day 3 vs. 4, B trials: paired *t*-test, $t_2 = 3.84$, $P = 0.061$, $n = 3$ mice)

(**E**) PV correlation for A trials (left) and B trials (right) organized by phase of task paradigm. We observed a trend of greater phase-specific population map decorrelation during reversal learning compared recall on both A and B trials, but this trend was statistically insignificant (see supplementary statistics table for details). Recall (dashed line) represents the correlation for Day 1-3 of recall relative to the first day of recall. Reversal (solid line) represents the correlation for the first three days of reversal learning relative to the first day of reversal (Day 4). Reversal recall (dotted line) represents the correlation for the last two days of reversal recall (Day 10 and 11) relative to the first day of reversal recall (Day 10). Insets of PV and TC drops from Day 0 to Day 1 for each phase of task.

(**F**) Tuning curve (TC) correlation for both A and B trials using tuning specificity criterion between neighboring sessions (TC correlation, A trials, across behavioral phases: one-way RM mixed effects analysis, effect of training day, $F_{5, 9.11} = 98.33$, $^{***}P < 0.001$; Days 1 vs. 2 Vs. Day 3 vs. 4, A trials: paired *t*-test, $t_2 = 11.17$, $^{**}P = 0.007$, $n = 3$ mice; B trials: one-way RM mixed effects analysis, effect of training day, $F_{5, 11} = 89.65$, $^{***}P < 0.001$; Days 1 vs. 2 Vs. Day 3 vs. 4, B trials: paired *t*-test, $t_2 = 11.02$, $^{**}P = 0.008$, $n = 3$ mice).

(**G**) TC correlation for A trials (left) and B trials (right) organized by phase of task paradigm identical to E.

Supplementary Fig. 6: Position and trial type decoding outperforms speed decoding . Related to Figure 6.

(**A**) A population vector decoder roughly predicts the speed of an animal while discriminating between A and B trials, during both the initial training session (top) and late training session (bottom). Plotted are the decoded speeds (red) against the actual speed (blue) of the mouse as a function of time. On the late training session, the decoder for this animal accurately predicts the trial type the animal is on. Throughout panels A-D, the speeds 0-30 are represented twice, once each for A and B trials. Data from this mouse only is used in panel B.

(**B**) Confusion matrices quantifying decoding accuracy demonstrate improved accuracy in identifying trial type associated with improved performance. Each black point represents the actual speed and trial of the animal plotted against the decoder's prediction. Each matrix cell represents the number of decoded points falling into each quadrant divided by the total data points in each trial type.

(**C**) Cumulative analysis across all training sessions revealed a strong positive relationship between performance and identification score (Correlation between performance and identification score in learning cohort: two-way ANOVA, $R = 0.75$, effect of mouse $F_{5,31} = 0.99$, $P = 0.50$, effect of performance, $F_{1,31} = 36.4$, $P = 1.11 \times 10^{-6}$, effect of interaction, $F_{5,31} = 0.84$, $P = 0.53$, n = 43 sessions from 6 mice). Each point denotes a single training session and each type of mark a different animal.

(**D**) In contrast, only a weak relationship was observed between speed decoding error and task performance (Correlation between performance and decoding error in learning cohort: two-way ANOVA, $R = -0.58$, effect of mouse, $F_{5,31} = 1.78$, $P = 0.15$, effect of performance $F_{1,31} = 11.5$, $P = 0.0019$, effect of interaction, $F_{5,31} = 1.26$, $P = 0.31$, n = 43 sessions from 6 mice).

(**E**) A population vector decoder accurately predicts the position of an animal while discriminating between slow and fast running epochs, during both the initial training session (top) and late training session (bottom). Plotted are the decoded positions (red) against the actual position (blue) of the mouse as a function of time. Regardless of performance, the decoder for this animal accurately predicts the position and speed category (slow or fast) the animal is on. Throughout panels E-H, the positions 0-200 are represented twice, once each for slow and fast running epochs. Data from this mouse only is used in panel F.

(**F**) Confusion matrices quantifying decoding accuracy in identifying trial type in both early and late training sessions. Each black point represents the actual position and speed category plotted against the decoder's prediction. Each matrix cell represents the number of decoded points falling into each quadrant divided by the total data points in each trial type.

(**G**) Cumulative analysis across all training sessions revealed a non-significant relationship between performance and decoding accuracy when identifying speed category (Correlation between performance and decoding score in learning cohort: two-way ANOVA, $R = 0.09$, effect of mouse $F_{5,31} = 4.26$, $P = 0.005$, effect of performance,

$F_{1,31}$ = 0.35, P = 0.56, effect of interaction, $F_{5,31}$ = 4.22, P = 0.005, n = 43 sessions from 6 mice). Each point denotes a single training session and each type of mark a different animal.

(**H**) Additionally, only a week relationship was observed between position decoding error and task performance (Correlation between performance and decoding score in learning cohort: two-way ANOVA, R = 0.35, effect of mouse, $F_{5,31}$ = 1.80, P = 0.14, effect of performance $F_{1,31}$ = 0.42, P = 0.52, effect of interaction, $F_{5,31}$ = 1.13, P = 0.36, n = 43 sessions from 6 mice).

(**I**) Left, overlay of the relationship between performance and identification scores from the 3 decoders. $P_{A/B}$: Decode position while identifying A/B trials (Figure 6); $Sp_{A/B}$: Decode speed while identifying A/B trials (Figure S15A-D); $P_{S/F}$: Decode position while identifying Slow/Fast running epochs (Figure S15E-H). Right, only including sessions with performance > 0.5, the identification score for the $P_{A/B}$ decoder (0.88, [0.84, 0.91]) was significantly higher than either the $Sp_{A/B}$ decoder (0.76, [0.74, 0.80], $p=8.3x10^{-6}$) or the $P_{S/F}$ decoder (0.74, [0.70, 0.77], $p=9.7x10^{-5}$), which were not significantly different from each other (p=0.14). N = 26 sessions, Wilcoxon sign-rank test used for all statistics.

(**J**) Left, overlay of the relationship between performance and normalized decoding error (see Methods) from the 3 decoders. Right, only including sessions with performance > 0.5, the decoding error for the $P_{A/B}$ decoder (3.53, [3.15, 4.34]) was significantly smaller than either the $Sp_{A/B}$ decoder (8.00, [7.63, 9.12], $p=8.3x10^{-6}$) or the $P_{S/F}$ decoder (9.14, [8.17, 10.5], $p=8.3x10^{-6}$), which were not significantly different from each other (p=0.18). N = 26 sessions, Wilcoxon sign-rank test used for all statistics.

Supplementary Fig. 7: Population vector decoding for reversal learning. Related to Figure 6.

(**A**) Population vector decoding of position while identifying trial type, as in Figure 6A, for the reversal learning experiment. Days 1, 4, and 7 are plotted, with Day 4 being the first day of Reversal learning. Note the dip in performance that recovers with additional experience.

(**B**) Confusion matrices quantifying decoding accuracy as in Figure 6B, for the same days as in **A**. Note that behavioral performance drops as the accuracy of the decoder drops.

(**C**) Example plot showing that as the performance of a single mouse changes, the identification score of the decoder also changes, with a notable drop on Day 4, the first day of reversal learning.

(**D**) Same data as in (C) with performance plotted against identification score, revealing a weak positive correlation (R=0.65, p=0.06, two-sided t-test, n = 9 sessions from one mouse).

(**E**) Cumulative analysis across all training sessions revealed a strong positive relationship between performance and identification score (Correlation between performance and identification score: two-way ANOVA, R = 0.56, effect of mouse $F_{2,20}$ = 5.49, P = 0.01, effect of performance, $F_{1,20}$ = 24.4, p = 8.0×10^{-5}, effect of interaction, $F_{2,20}$ = 2.83, P = 0.08, n = 25 sessions from 3 mice). Each point denotes a single training session and each type of mark a different animal.

(**F**) In contrast, a weak relationship was observed between performance and position decoding error (Correlation between performance and decoding error: two-way ANOVA, R = -0.32, effect of mouse $F_{2,20}$ = 4.93, P = 0.01, effect of performance, $F_{1,20}$ = 4.54, p = 0.05, effect of interaction, $F_{2,20}$ = 3.83, P = 0.04, n = 25 sessions from 3 mice). Each point denotes a single training session and each type of mark a different animal.

(**G**) Top, colormap showing the change in identification score during reversal learning. Days 1-3 late in training with the original contingencies show near perfect identification. Day 4 is the first day of reversed reward contingencies. A decrease in identification score is visible, which recovers with additional training. Each row denotes the median identification score across all 3 mice. Bottom, plots of identification score for sessions 1, 4, and 7, showing an eventual recovery to baseline.

(**H**) An equivalent plot of the decoding position error averaged across all animals. Decoding error does not show a clear pattern with regards to reversal learning, indicating that the contingency reversal affects the ability to identify the trial type more than the decoding of absolute position.

(**I**) Top, Image depicting the identification score of classifiers trained on one day and tested on another (see Methods). Day 4 is the first day of reversal learning. The block organization suggests a new representation for identifying trial type was learned after reversal. Middle, mean value of successive diagonals of the above image, showing the drop in identification score with experience. Bottom, identification score on adjacent days shows a sharp dropoff after Day 4.

(**J**) As in I, but for decoding error. Note there is no block organization or obvious change in performance after reversal learning.

Supp. Table 1

Panel	Aggregation	Comparison	N	Test	df	Test statistic	p-value	p-value adjusted	Adjustment
		A Trials							
1D	by animal	Fraction of licks across stages	16 fraction licks X mice	RM 1-way ANOVA		$F_{training\_stage}(1.51, 4.52) = 12.4$	0.017	N/A	N/A
1D	by animal	RF vs Stage 3	4	Paired t-test	3	5.74	0.011	0.031	3-way Holm-Sidak
1D	by animal	Stage 1 vs 3	4	Paired t-test	3	2.49	0.088	0.17	3-way Holm-Sidak
1D	by animal	Stage 2 vs 3	4	Paired t-test	3	0.26	0.81	0.81	3-way Holm-Sidak
		B Trials							
1D	by animal	Fraction of licks across stages	16 fraction licks X mice	RM 1-way ANOVA		$F_{training\_stage}(1.45, 4.35) = 19.0$	0.0082	N/A	N/A
1D	by animal	RF vs Stage 3	4	Paired t-test	3	5.39	0.013	0.037	3-way Holm-Sidak
1D	by animal	Stage 1 vs 3	4	Paired t-test	3	1.64	0.20	0.36	3-way Holm-Sidak
1D	by animal	Stage 2 vs 3	4	Paired t-test	3	0.25	0.81	0.81	3-way Holm-Sidak
		A Trials							
1E	by animal	Fraction of correct trials across stages	16 fraction correct X mice	RM 1-way ANOVA		$F_{training\_stage}(1.00, 3.01) = 12.7$	0.038	N/A	N/A
1E	by animal	RF vs Stage 3	4	Paired t-test	3	∞	0	0	3-way Holm-Sidak
1E	by animal	Stage 1 vs 3	4	Paired t-test	3	∞	0	0	3-way Holm-Sidak
1E	by animal	Stage 2 vs 3	4	Paired t-test	3	2.03	0.14	0.14	3-way Holm-Sidak
		B Trials							
1E	by animal	Fraction of correct trials across stages	16 fraction correct X mice	RM 1-way ANOVA		$F_{training\_stage}(1.00, 3.01) = 93.4$	0.0023	N/A	N/A
1E	by animal	RF vs Stage 3	4	Paired t-test	3	31	<0.0001	0.0003	3-way Holm-Sidak
1E	by animal	Stage 1 vs 3	4	Paired t-test	3	183.5	<0.0001	0.0003	3-way Holm-Sidak
1E	by animal	Stage 2 vs 3	4	Paired t-test	3	8.38	0.0036	0.0036	3-way Holm-Sidak

Table S1. Statistical details of behavioral learning of a head-fixed, odor-cued spatial navigation task, Related to Figure 1

Supp. Table 2

Panel	Aggregation	Comparison	N	Test	df	Test statistic	p-value	p-value adj	Adjustment
2C	by animal	AUC/min difference btwn A vs B laps in RUN - A sel.	11	Paired Wilcoxon Sign Rank	10	66	0.00098	0.0029	3-way Holm-Sidak
2C	by animal	AUC/min difference btwn A vs B laps in RUN - B sel.	11	Paired Wilcoxon Sign Rank	10	-66	0.00098	0.0029	3-way Holm-Sidak
2C	by animal	AUC/min difference btwn A vs B laps in RUN - A&B	11	Paired Wilcoxon Sign Rank	10	-30	0.21	0.21	3-way Holm-Sidak
2C	by animal	AUC/min difference btwn A vs B laps in NO RUN - A sel.	11	Paired Wilcoxon Sign Rank	10	46	0.042	0.082	3-way Holm-Sidak
2C	by animal	AUC/min difference btwn A vs B laps in NO RUN - B sel.	11	Paired Wilcoxon Sign Rank	10	-62	0.0029	0.0088	3-way Holm-Sidak
2C	by animal	AUC/min difference btwn A vs B laps in NO RUN - A&B	11	Paired Wilcoxon Sign Rank	10	-32	0.17	0.17	3-way Holm-Sidak
*	*by animal*	*Transient/min difference btwn A vs B laps in RUN - A sel.*	*11*	*Paired Wilcoxon Sign Rank*	*10*	*66*	*0.00098*	*0.0029*	*3-way Holm-Sidak*
*	*by animal*	*Transient/min difference btwn A vs B laps in RUN - B sel.*	*11*	*Paired Wilcoxon Sign Rank*	*10*	*-66*	*0.00098*	*0.0029*	*3-way Holm-Sidak*
*	*by animal*	*Transient/min difference btwn A vs B laps in RUN - A&B*	*11*	*Paired Wilcoxon Sign Rank*	*10*	*-14*	*0.58*	*0.92*	*3-way Holm-Sidak*
*	*by animal*	*Transient/min difference btwn A vs B laps in NO RUN - A sel.*	*11*	*Paired Wilcoxon Sign Rank*	*10*	*34*	*0.15*	*0.38*	*3-way Holm-Sidak*
*	*by animal*	*Transient/min difference btwn A vs B laps in NO RUN - B sel.*	*11*	*Paired Wilcoxon Sign Rank*	*10*	*-64*	*0.0020*	*0.0058*	*3-way Holm-Sidak*
*	*by animal*	*Transient/min difference btwn A vs B laps in NO RUN - A&B*	*11*	*Paired Wilcoxon Sign Rank*	*10*	*-30*	*0.21*	*0.50*	*3-way Holm-Sidak*
	*Not plotted								
2D	by animal	Fraction of A,B, A&B tuned, neither - Spatial info.	11	Friedman test	3	30.0	1.42×10^{-06}		N/A
2D	by animal	Fraction of neurons tuned - A vs B - Spatial info.	11	Paired Wilcoxon Sign Rank	10	54	0.014	0.014	3-way Holm-Sidak
2D	by animal	Fraction of neurons tuned - A vs A&B - Spatial info.	11	Paired Wilcoxon Sign Rank	10	-66	0.00098	0.0029	3-way Holm-Sidak
2D	by animal	Fraction of neurons tuned - B vs A&B - Spatial info.	11	Paired Wilcoxon Sign Rank	10	-66	0.00098	0.0029	3-way Holm-Sidak

2F	pooled	Distribution of PF centroid locations - A selective (25 bins)	590	Rayleigh test of circular uniformity	~	4.72	0.0089	N/A	N/A
2F	pooled	Distribution of PF centroid locations - B selective (25 bins)	468	Rayleigh test of circular uniformity	~	89.6	$1.14x10^{-41}$	N/A	N/A
2G	pooled	Place field centroid distribution of A selective vs B selective neurons	590 vs 468	2-sample Kolmogorov–Smirnov	~	0.22	$1.45x10^{-11}$	N/A	N/A
2H	by animal	Tuning curve correlation of A vs B laps - A sel. Vs B sel neurons	11	Paired Wilcoxon Sign Rank	10	14	0.58	0.58	3-way Holm-Sidak
2H	by animal	Tuning curve correlation of A vs B laps - A sel. Vs AB neurons	11	Paired Wilcoxon Sign Rank	10	-66	0.00098	0.0029	3-way Holm-Sidak
2H	by animal	Tuning curve correlation of A vs B laps - B sel. Vs AB neurons	11	Paired Wilcoxon Sign Rank	10	-66	0.00098	0.0029	3-way Holm-Sidak

Table S2. Statistical details of task-selective place cells, Related to Figure 2

Supp. Table 3

Panel	Aggregation	Comparison	N	Test	Degrees of Freedom	Test statistic	p-value	p-value adjusted	Adjustment
3C	by animal	Fraction of remapping neuron subtypes (common, activity, global, other)	11	Friedman test	3	29.0	2.28×10^{-06}	N/A	N/A
3C	by animal	Common vs activity	11	Paired Wilcoxon Sign Rank	10	66	0.00098	0.0020	2-way Holm-Sidak
3C	by animal	Common vs global	11	Paired Wilcoxon Sign Rank	10	48	0.032	0.032	2-way Holm-Sidak
3D	pooled	Distribution of PF centroid locations in common neurons (25 bins)	700	Rayleigh test of circular uniformity	N/A	13.8	9.35×10^{-07}	N/A	N/A
3E	by animal	Zone I (B before A) vs 0	11	1-sample Wilcoxon Sign Rank	10	42	0.067	N/A	N/A
3E	by animal	Zone II (B before A) vs 0	11	1-sample Wilcoxon Sign Rank	10	66	0.00098	N/A	N/A
3E	by animal	Zone III (B before A) vs 0	11	1-sample Wilcoxon Sign Rank	10	-22	0.35	N/A	N/A

Table S3. Statistical details of remapping properties between trial types, Related to Figure 3

Supp. Table 4

Panel	Aggregation	Comparison	N	Test	df	Test statistic	p-value	p-value adj	Adjustment
		A tuned cells, Learning							
4D	by animal	Fraction across time	6	1-way RM linear mixed effects analysis	~	F{6, 22.2) = 7.24	0.00022	N/A	N/A
4D	by animal	Fraction in D1 vs D6	5	Paired t-test	4	-2.42	0.072	0.14	2-way Holm-Sidak
4D	by animal	Fraction in D1 vs D7	4	Paired t-test	3	-2.70	0.074	0.14	2-way Holm-Sidak
		B tuned cells, Learning							
4D	by animal	Fraction across time	6	1-way RM linear mixed effects analysis	~	F{6,27) = 1.36	0.27	N/A	N/A
4D	by animal	Fraction in D1 vs D6	5	Paired t-test	4	-1.25	0.28	0.48	2-way Holm-Sidak
4D	by animal	Fraction in D1 vs D7	4	Paired t-test	3	-1.28	0.29	0.48	2-way Holm-Sidak
		A&B tuned cells, Learning							
4D	by animal	Fraction across time	6	1-way RM linear mixed effects analysis	~	F{6,22.3659) = 1.38	0.27	N/A	N/A
4D	by animal	Fraction in D1 vs D6	5	Paired t-test	4	0.80	0.47	0.47	2-way Holm-Sidak
4D	by animal	Fraction in D1 vs D7	4	Paired t-test	3	1.55	0.22	0.39	2-way Holm-Sidak
		A tuned cells, Recall							
4D	by animal	Fraction across time	5	1-way RM linear mixed effects analysis	~	F{6, 28) = 0.32	0.92	N/A	N/A
4D	by animal	Fraction in D1 vs D6	5	Paired t-test	4	-1.20	0.30	0.30	2-way Holm-Sidak
4D	by animal	Fraction in D1 vs D7	5	Paired t-test	4	-1.80	0.15	0.27	2-way Holm-Sidak
		B tuned cells, Recall							
4D	by animal	Fraction across time	5	1-way RM linear mixed effects analysis	~	F{6, 28) = 0.26	0.95	N/A	N/A
4D	by animal	Fraction in D1 vs D6	5	Paired t-test	4	-1.15	0.31	0.53	2-way Holm-Sidak
4D	by animal	Fraction in D1 vs D7	5	Paired t-test	4	-0.87	0.43	0.53	2-way Holm-Sidak
		A&B tuned cells, Recall							
4D	by animal	Fraction across time	5	1-way RM linear mixed effects analysis	~	F{6, 24) = 0.83	0.56	N/A	N/A
4D	by animal	Fraction in D1 vs D6	5	Paired t-test	4	-0.81	0.46	0.46	2-way Holm-Sidak
4D	by animal	Fraction in D1 vs D7	5	Paired t-test	4	-1.39	0.24	0.42	2-way Holm-Sidak

Table S4. Statistical details of learning-induced distribution of task-specific cells, Related to Figure 4

Supp. Table 5

Panel	Aggregation	Comparison	N	Test	df	Test statistic	p-value	p-value adj	Adjustment
		A Laps							
4F	by animal	PV correlation relative to D1	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(3, 22.4)$: 50.9 $F_{behavior}(1, 9.04)$: 9.02 $F_{timeXbeh}(3, 22.4)$: 0.78	$3.73x10^{-10}$ 0.015 0.52	N/A	N/A
4F	by animal	Learning vs recall PV correlation D6	5, 5	Unpaired t-test	8	-3.81	0.0052	0.010	2-way Holm-Sidak
4F	by animal	Learning vs recall PV correlation D7	4, 5	Unpaired t-test	7	-2.78	0.027	0.027	2-way Holm-Sidak
4F	by animal	Learning PV correlation D2 vs D6	4	Paired t-test	3	13.0	0.00098	0.0020	2-way Holm-Sidak
4F	by animal	Learning PV correlation D2 vs D7	3	Paired t-test	2	5.89	0.028	0.028	2-way Holm-Sidak
4F	by animal	Recall PV correlation D2 vs D6	5	Paired t-test	4	4.54	0.010	0.010	2-way Holm-Sidak
4F	by animal	Recall PV correlation D2 vs D7	5	Paired t-test	4	7.89	0.0014	0.0028	2-way Holm-Sidak
		B Laps							
4F	by animal	Population vector correlation relative to D1	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(3, 22.4)$: 43.9 $F_{behavior}(1, 9.07)$: 9.13 $F_{timeXbeh}(3, 22.4)$: 1.78	$1.49x10^{-9}$ 0.014 0.18	N/A	N/A
4F	by animal	Learning vs recall PV correlation D6	5, 5	Unpaired t-test	8	-3.58	0.0072	0.014	2-way Holm-Sidak
4F	by animal	Learning vs recall PV correlation D7	4, 5	Unpaired t-test	7	-3.37	0.012	0.014	2-way Holm-Sidak
4F	by animal	Learning PV correlation D2 vs D6	4	Paired t-test	3	4.05	0.027	0.053	2-way Holm-Sidak
4F	by animal	Learning PV correlation D2 vs D7	3	Paired t-test	2	5.06	0.037	0.053	2-way Holm-Sidak
4F	by animal	Recall PV correlation D2 vs D6	5	Paired t-test	4	9.22	0.00077	0.0015	2-way Holm-Sidak
4F	by animal	Recall PV correlation D2 vs D7	5	Paired t-test	4	7.79	0.0015	0.0015	2-way Holm-Sidak
		A Laps							
4F	by animal	Tuning curve TS correlation relative to D1	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(3, 22.4)$: 27.7 $F_{behavior}(1, 8.85)$: 8.69 $F_{timeXbeh}(3, 22.4)$: 3.95	$1.02x10^{-7}$ 0.017 0.021	N/A	N/A
4F	by animal	Learning vs recall Tuning curve TS correlation D6	5, 5	Unpaired t-test	8	-3.89	0.0046	0.0092	2-way Holm-Sidak
4F	by animal	Learning vs recall Tuning curve TS correlation D7	4, 5	Unpaired t-test	7	-4.05	0.0049	0.0092	2-way Holm-Sidak
4F	by animal	Learning Tuning curve TS correlation D2 vs D6	4	Paired t-test	3	8.62	0.0033	0.0065	2-way Holm-Sidak

4F	by animal	Learning Tuning curve TS correlation D2 vs D7	3	Paired t-test	2	5.84	0.028	0.028	2-way Holm-Sidak
4F	by animal	Recall Tuning curve TS correlation D2 vs D6	5	Paired t-test	4	2.16	0.10	0.097	2-way Holm-Sidak
4F	by animal	Recall Tuning curve TS correlation D2 vs D7	5	Paired t-test	4	5.10	0.0070	0.014	2-way Holm-Sidak
		B Laps							
4F	by animal	Tuning curve TS correlation relative to D1	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(3, 22.5)$: 32.9 $F_{behavior}(1 ,8.81)$: 11.8 $F_{timeXbeh}(3, 22.5)$: 1.73	$2.17x10^{-8}$ 0.0077 0.19	N/A	N/A
4F	by animal	Learning vs recall Tuning curve TS correlation D6	5, 5	Unpaired t-test	8	-2.32	0.049	0.049	2-way Holm-Sidak
4F	by animal	Learning vs recall Tuning curve TS correlation D7	4, 5	Unpaired t-test	7	-4.03	0.0050	0.010	2-way Holm-Sidak
4F	by animal	Learning Tuning curve TS correlation D2 vs D6	4	Paired t-test	3	2.69	0.074	0.074	2-way Holm-Sidak
4F	by animal	Learning Tuning curve TS correlation D2 vs D7	3	Paired t-test	2	6.49	0.023	0.045	2-way Holm-Sidak
4F	by animal	Recall Tuning curve TS correlation D2 vs D6	5	Paired t-test	4	12.8	0.00021	0.00043	2-way Holm-Sidak
4F	by animal	Recall Tuning curve TS correlation D2 vs D7	5	Paired t-test	4	5.32	0.0060	0.0060	2-way Holm-Sidak
		A Laps							
4G	by animal	PV correlation neighboring days - learning vs recall, 1 vs 2, 2 vs 3, 6 vs 7	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(2, 14.3)$: 3.98 $F_{behavior}(1, 9.24)$: 5.45 $F_{timeXbeh}(2, 14.3)$: 7.37	0.042 0.044 0.0063	N/A	N/A
4G	by animal	PV correlation learning - time - (1 vs 2) vs (6 vs 7)	3	Paired t-test	2	-5.24	0.034	N/A	N/A
4G	by animal	PV correlation recall - time – (1 vs 2) vs (6 vs 7)	5	Paired t-test	4	0.55	0.61	N/A	N/A
4G	by animal	PV correlation learning vs recall - behavior - 1 vs 2	5, 5	Unpaired t-test	8	-2.60	0.031	0.062	2-way Holm-Sidak
4G	by animal	PV correlation learning vs recall - behavior -6 vs 7	4, 5	Unpaired t-test	7	-0.50	0.63	0.63	2-way Holm-Sidak
		B Laps							
4G	by animal	PV correlation neighboring days - learning vs recall, 1 vs 2, 2 vs 3, 6 vs 7	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(2, 14.2)$: 0.98 $F_{behavior}(1, 9.06)$: 1.88 $F_{timeXbeh}(2, 14.2)$: 6.18	0.40 0.20 0.012	N/A	N/A

4G	by animal	PV correlation learning - time - (1 vs 2) vs (6 vs 7)	3	Paired t-test	2	-2.42	0.14	N/A	N/A
4G	by animal	PV correlation recall - time – (1 vs 2) vs (6 vs 7)	5	Paired t-test	4	1.54	0.20	N/A	N/A
4G	by animal	PV correlation learning vs recall - behavior - 1 vs 2	5, 5	Unpaired t-test	8	-2.16	0.062	0.12	2-way Holm-Sidak
4G	by animal	PV correlation learning vs recall - behavior -6 vs 7	4, 5	Unpaired t-test	7	0.40	0.70	0.70	2-way Holm-Sidak
		A Laps							
4G	by animal	TC correlation neighboring days - learning vs recall, 1 vs 2, 2 vs 3, 6 vs 7	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(2, 14.4)$: 0.74 $F_{behavior}(1, 9.14)$: 2.32 $F_{timeXbeh}(2, 14.4)$: 6.01	0.49 0.16 0.013	N/A	N/A
4G	by animal	TC correlation learning - time - (1 vs 2) vs (6 vs 7)	3	Paired t-test	2	-5.89	0.028	N/A	N/A
4G	by animal	TC correlation recall - time - (1 vs 2) vs (6 vs 7)	5	Paired t-test	4	0.81	0.46	N/A	N/A
4G	by animal	TC correlation learning vs recall - behavior - 1 vs 2	5, 5	Unpaired t-test	8	-1.50	0.17	0.31	2-way Holm-Sidak
4G	by animal	TC correlation learning vs recall - behavior -6 vs 7	4, 5	Unpaired t-test	7	0.19	0.86	0.86	2-way Holm-Sidak
		B Laps							
4G	by animal	TC correlation neighboring days - learning vs recall, 1 vs 2, 2 vs 3, 6 vs 7	6, 5	2-way RM linear mixed effects analysis	~	$F_{time}(2, 14.0)$: 0.55 $F_{behavior}(1, 8.83)$: 4.49 $F_{timeXbeh}(2, 14.0)$: 9.32	0.59 0.06 0.0026	N/A	N/A
4G	by animal	TC correlation learning - time - (1 vs 2) vs (6 vs 7)	3	Paired t-test	2	-5.09	0.036	N/A	N/A
4G	by animal	TC correlation recall - time - (1 vs 2) vs (6 vs 7)	5	Paired t-test	4	1.71	0.16	N/A	N/A
4G	by animal	TC correlation learning vs recall - behavior - 1 vs 2	5, 5	Unpaired t-test	8	-3.68	0.006	0.012	2-way Holm-Sidak
4G	by animal	TC correlation learning vs recall - behavior -6 vs 7	4, 5	Unpaired t-test	7	0.20	0.85	0.85	2-way Holm-Sidak

Table S5. Statistical details of learning-induced place cell remapping, Related to Figure 4

Supp. Table 6

Panel	Aggregation	Comparison	N	Test	Degrees of Freedom	Test statistic	p-value
5C	pooled	TS - pooled neuron - group difference - Learn - D2,D3,D4,D5,D6,D7	179 225 211 153 160 122	Kruskal-Wallis	5	63.3	$2.51x10^{-12}$
5C	pooled	AB Corr vs 1, Day 2 - Learn	179	1-sample Wilcoxon Sign Rank	178	-3536	0.011
5C	pooled	AB Corr vs 1, Day 3 - Learn	225	1-sample Wilcoxon Sign Rank	224	-3387	0.083
5C	pooled	AB Corr vs 1, Day 4 - Learn	211	1-sample Wilcoxon Sign Rank	210	-6908	0.0001
5C	pooled	AB Corr vs 1, Day 5 - Learn	153	1-sample Wilcoxon Sign Rank	152	-7847	$8.89x10^{-13}$
5C	pooled	AB Corr vs 1, Day 6 - Learn	160	1-sample Wilcoxon Sign Rank	159	-7718	$4.89x10^{-11}$
5C	pooled	AB Corr vs 1, Day 7 - Learn	122	1-sample Wilcoxon Sign Rank	121	-4197	$8.25x10^{-08}$
5D	pooled	TS - pooled neuron - group difference - Recall - D2,D3,D6,D7,D8,D9	133 117 98 91 66 72	Kruskal-Wallis	5	7.15	0.21
5D	pooled	AB Corr vs 1, Day 2 - Recall	133	1-sample Wilcoxon Sign Rank	132	-5	1.00
5D	pooled	AB Corr vs 1, Day 3 - Recall	117	1-sample Wilcoxon Sign Rank	116	-2081	0.0047
5D	pooled	AB Corr vs 1, Day 6 - Recall	98	1-sample Wilcoxon Sign Rank	97	-831	0.14
5D	pooled	AB Corr vs 1, Day 7 - Recall	91	1-sample Wilcoxon Sign Rank	90	-1044	0.039
5D	pooled	AB Corr vs 1, Day 8 - Recall	66	1-sample Wilcoxon Sign Rank	65	49	0.88
5D	pooled	AB Corr vs 1, Day 9 - Recall	72	1-sample Wilcoxon Sign Rank	71	-638	0.073
5E	by session	Correlation between A-B map correlation and performance in Learning Cohort	44 sessions from 6 mice	t-test	42	-6.18	$2.17x10^{-7}$
5F	by session	Correlation between A-B map correlation and performance in Recall Cohort	35 sessions from 6 mice	t-test	42	-1.97	0.058

Table S6. Statistical details of map dissimilarities across learning and recall, Related to Figure 5

Supp. Table 7

Panel	Aggregation	Comparison	N	Test	Degrees of Freedom	Test statistic	p-value
6D	by session	Correlation between performance and identification score	9 sessions from 1 mouse	t-test	7	4.9481	0.0017
6E	by session	Correlation between performance and identification score	43 sessions from 6 mice	2-way ANOVA	~	$F_{mouse}(5, 31) = 1.187$ $F_{identification}(1, 31) = 29.014$ $F_{mouseXidentification}(5, 31) = 0.81$	0.34 $7.1x10^{-6}$ 0.55
6F	by session	Correlation between performance and decoding error	43 sessions from 6 mice	2-way ANOVA	~	$F_{mouse}(5, 31) = 9.46$ $F_{error}(1, 31) = 20.8$ $F_{mouseXerror}(5, 31) = 3.00$	$1.5x10^{-5}$ $7.6x10^{-5}$ 0.03
6I	by adjacent sessions	Correlation between day and identification score	28 session comparisons from 6 mice	2-way ANOVA	~	$F_{mouse}(5, 27) = 1.94$ $F_{identification}(1, 27) = 25.1$ $F_{mouseXidentification}(5, 27) = 1.85$	0.14 $1.3x10^{-4}$ 0.16
6J	by adjacent sessions	Correlation between day and decoding error	28 session comparisons from 6 mice	2-way ANOVA	~	$F_{mouse}(5, 27) = 3.75$ $F_{error}(1, 27) = 0.71$ $F_{mouseXerror}(5, 27) = 2,51$	0.02 0.41 0.07

Table S7. Statistical details of population vector decoding performance correlated with behavior, Related to Figure 6