

Figure 3. Learning stages. (A) Habituation. No visual stimuli are present. Exploration of any new arm yields reward. (B) Follow the pattern. Guiding visual stimuli are present in both the initiation and goal arms. (C) Initiation arm only. Visual stimulus is only present in the initiation arm.

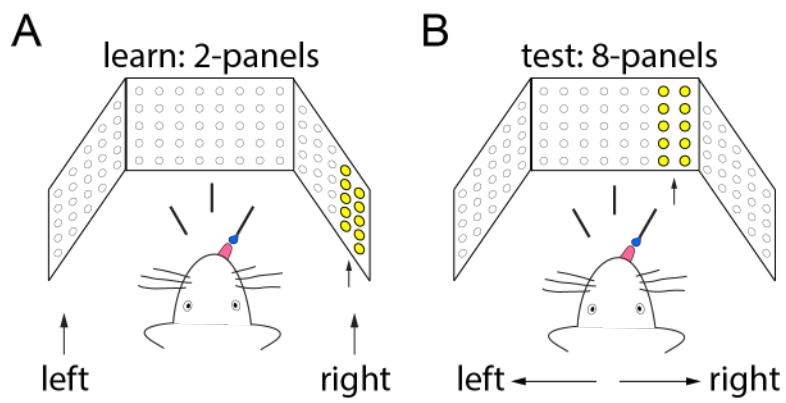


Figure 1. Visuospatial categorical 2AFC paradigm. (A) Schematic showing the categorical decision-making paradigm. After initiation, trials are followed by left or right stimulus during training. (B) During the testing phase, an intermediate stimulus location is presented for trials after initiation.



Figure 2. Visual stimuli for freely-moving behavioral paradigm. Visual gratings are displayed on tablets underneath the clear acrylic of the Y-maze and act as guiding visual stimuli. Horizontal bars encode for a left turn based correct trial, while vertical bars encode for a right turn based correct trial.

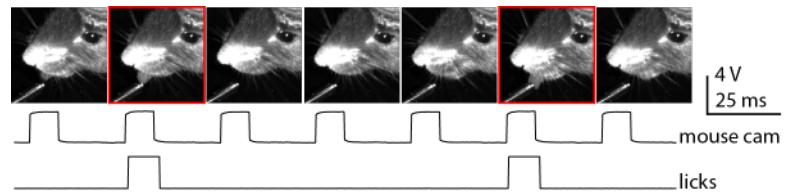


Figure 4. Video confirmation of lick detection system. Behavioral images (top) where a lick was visually observed (red box outline), compared with registered camera frames (middle), and detected licks by the Arduino (bottom).

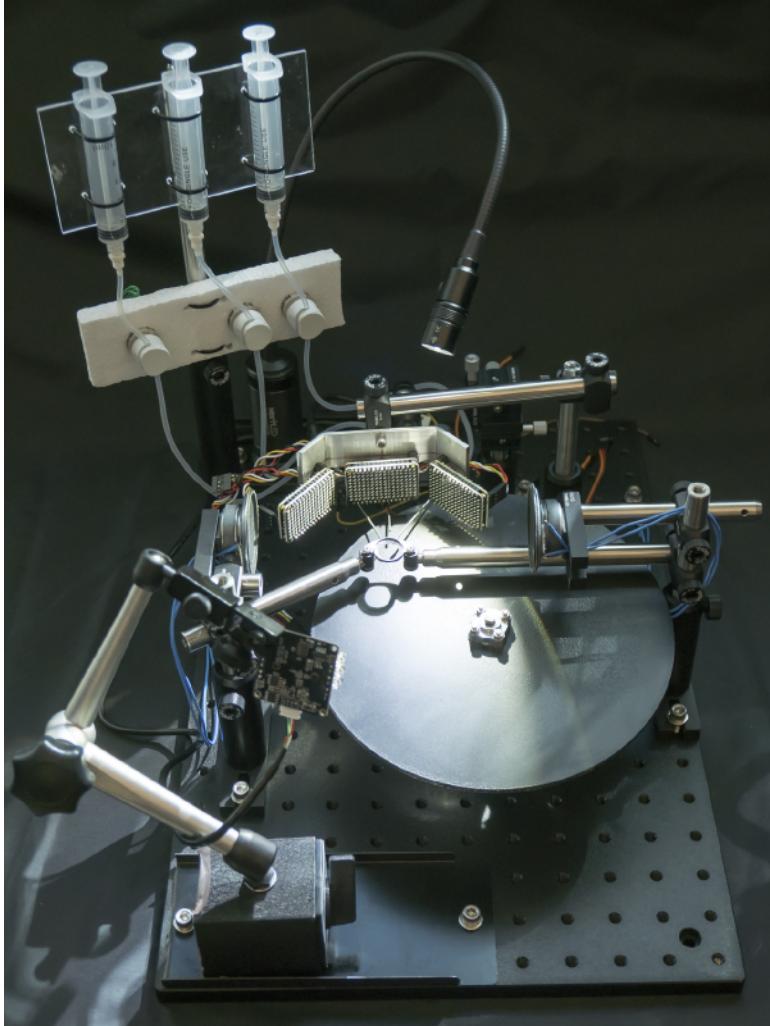


Figure 5. HERBs apparatus. Complete hardware setup consisting of water reward delivery system, LED panels for visual stimuli presentation, speakers for auditory stimuli presentation, running disc, headfixation poles, and a USB camera.

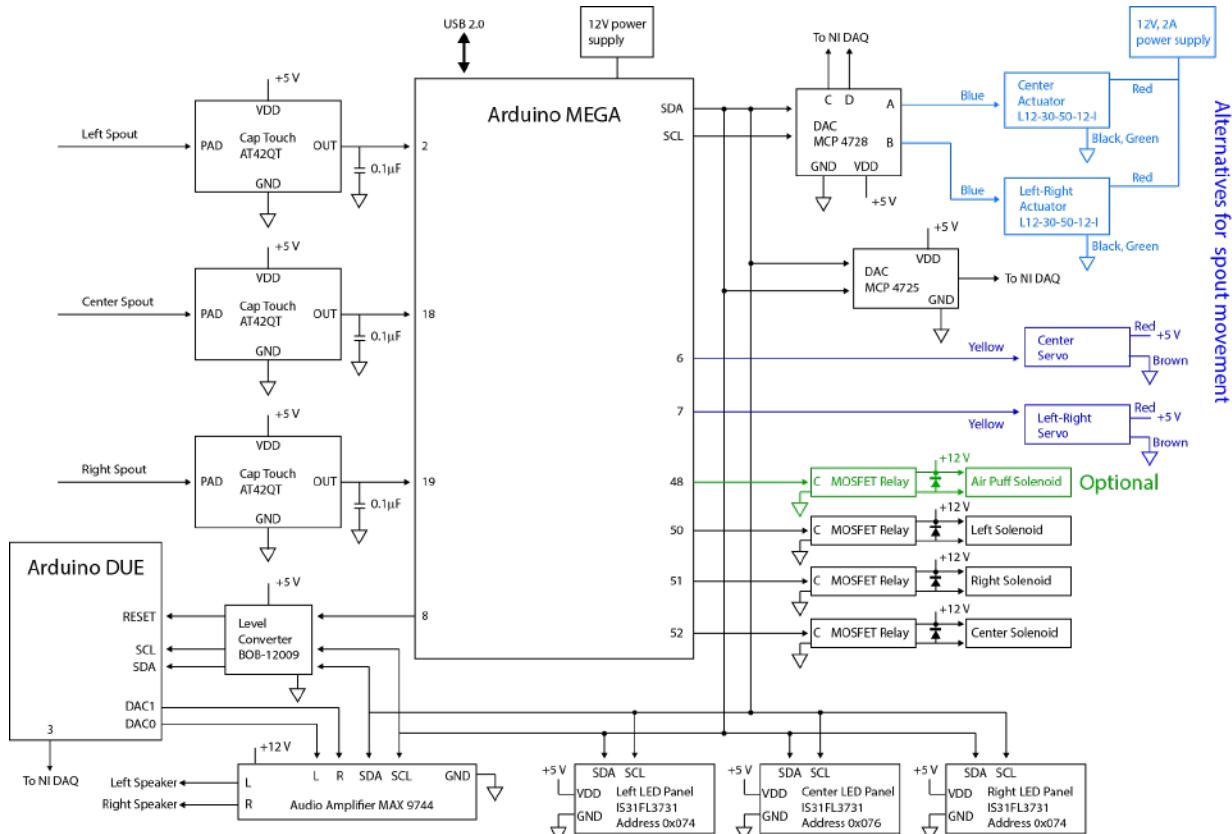


Figure 7. HERBs wiring diagram. Schematic of wire connections with labeled pinouts. Alternative wiring options for servo control are included (blue) for linear actuator usage (light blue) versus rotary servos (dark blue). Wiring is shown for air puff delivery if desired (green). Necessary wiring for visual and auditory stimuli presentation is shown (black).

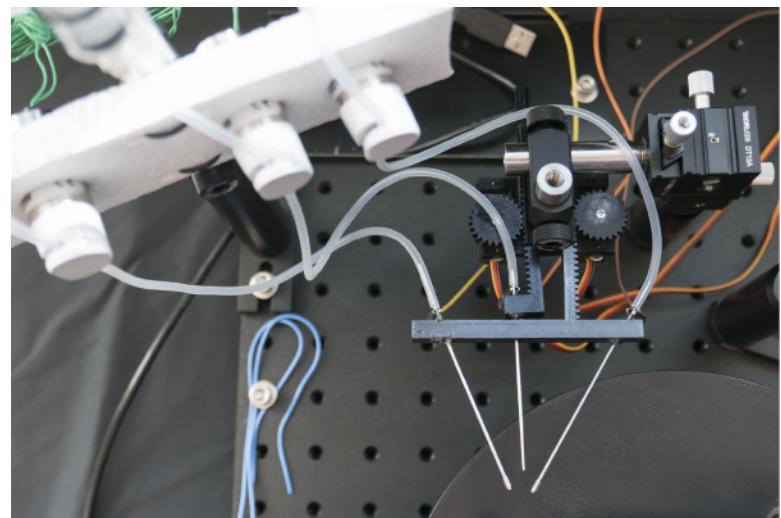


Figure 6. HERBs moving spouts and water reward setup. Solenoid based water reward delivery system and 3D printed servo to linear actuator assembly to control three movable spouts.

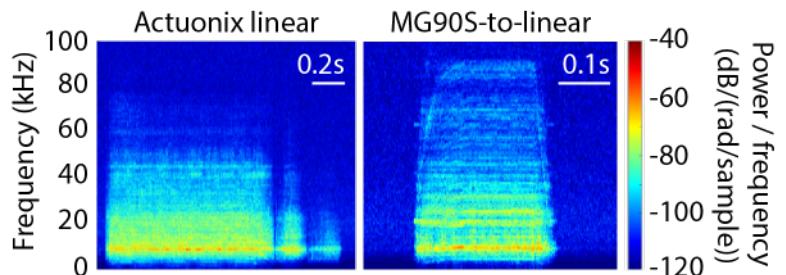


Figure 8. Linear actuator sound profile. Spectrogram of the linear actuator (Actuonix; left) and rotary servo-to-linear converter (right).

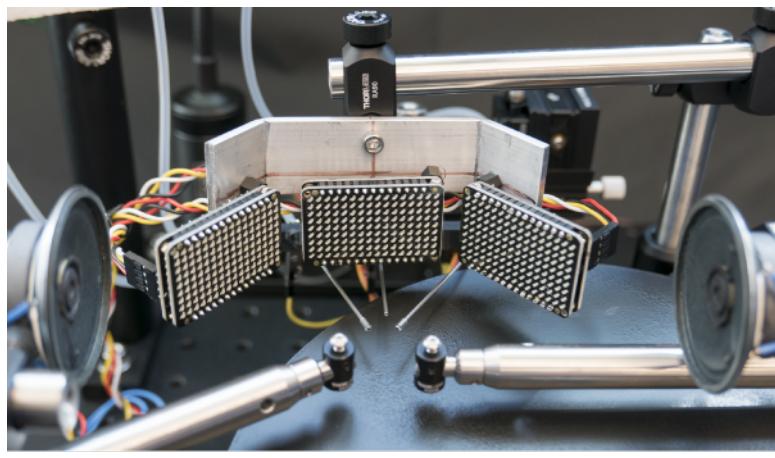


Figure 9. HERBs stimuli presentation setup. Three 16 x 9 LED panels to display visual stimuli and a speaker on each side of the head-mount holding bars.

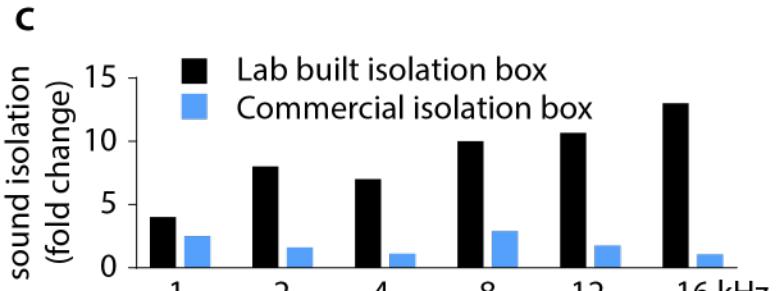
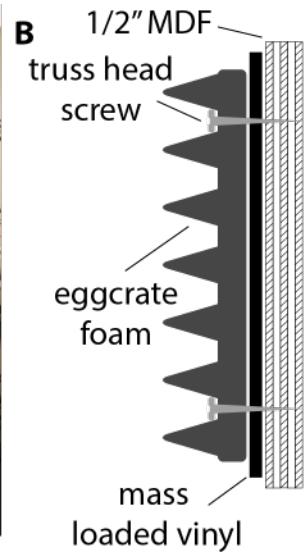


Figure 10. HERBs rig sound proofing. (A) Enclosure housing two behavioral rigs with the electronic drivers mounted on the outside wall. (B) Schematic of the sound isolation setup inside the enclosure. (C) Performance of sound isolation comparison between lab built versus commercially available affordable isolation boxes at varying tones.

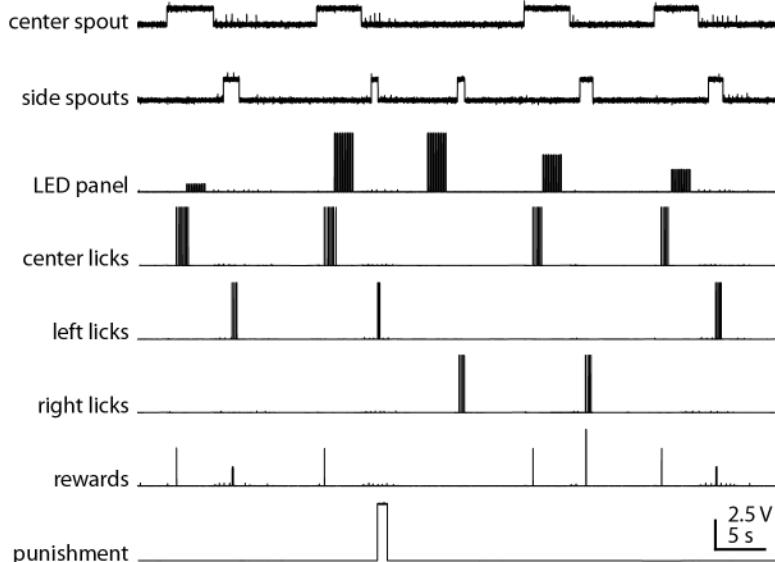


Figure 12. Signal output from HERBs. Recorded on a data acquisition device during 2AFC behavior. The “center spout” signal shows the movement of the central lick spout, with high voltage indicating the spout being in position for the animal to reach it. The “side spouts” signal shows the simultaneous movement of the left and right lick spouts, with high voltage indicating the spouts being in position for the animal to reach it. The “LED panel” shows the signal indicating when the LEDs are on, with the voltage encoding the location of the stimulus along the LED crescent. “Center licks” registers licks on the center spout. “Left licks” registers licks on the left spout. “Right licks” registers licks on the right spout. “Rewards” registers when the reward is delivered, with the voltage encoding the spout (high, right reward; medium, center reward; low, left reward). “Punishment” indicates the on and offset of the punishment tone following an incorrect choice.

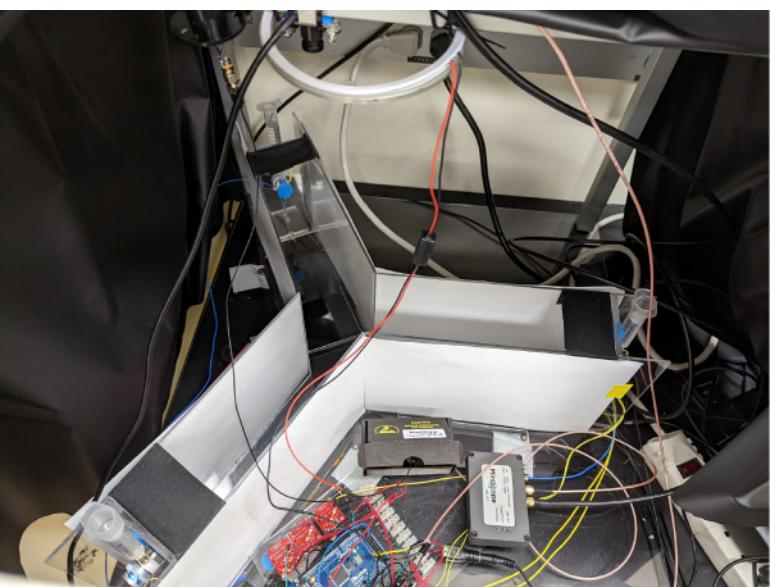


Figure 11. Hardware setup for freely-moving behavior. Image of one of three rigs used for calcium imaging during a freely-moving, two alternative forced choice task. The rig consists of a y-maze, 3 lickspouts, 3 tablets, Arduino, miniscope, DAQ board, commutator, light ring, and camera.

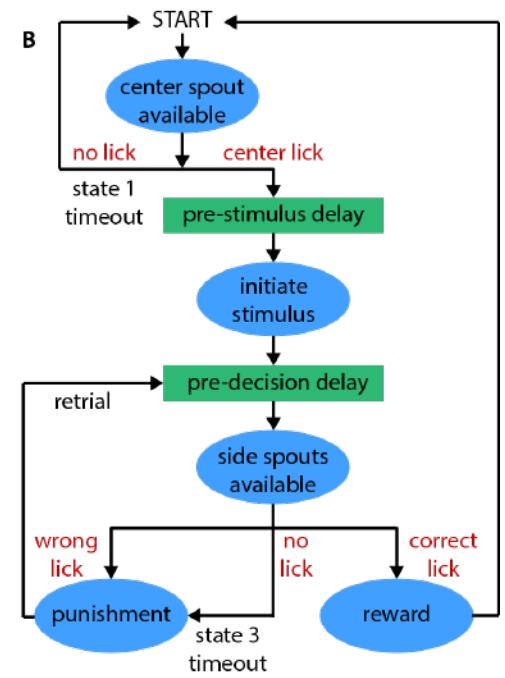
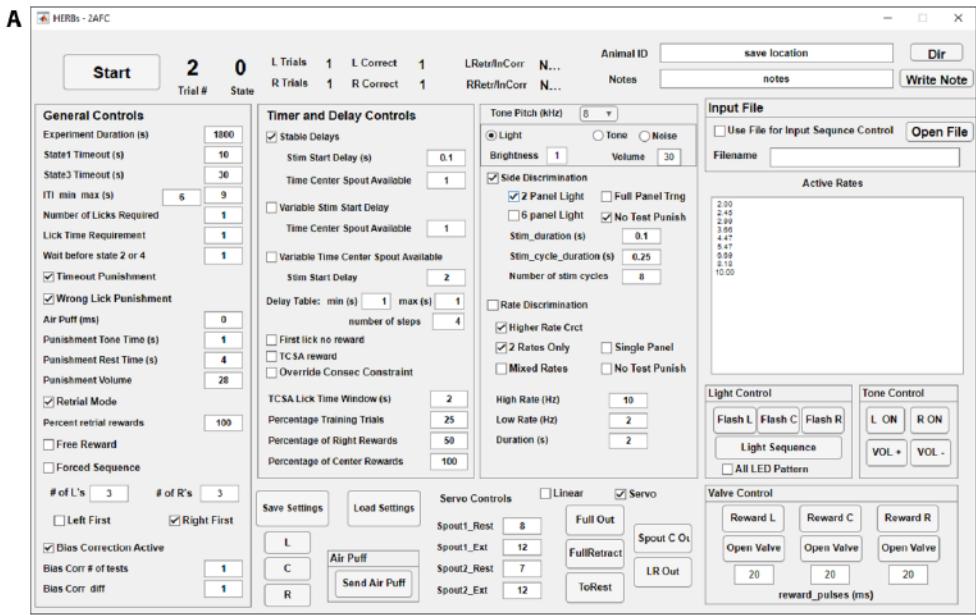


Figure 13. HERBs 2AFC GUI and state machine. (A) GUI for the 2AFC paradigm. **(B)** Simplified state machine for the 2AFC task.

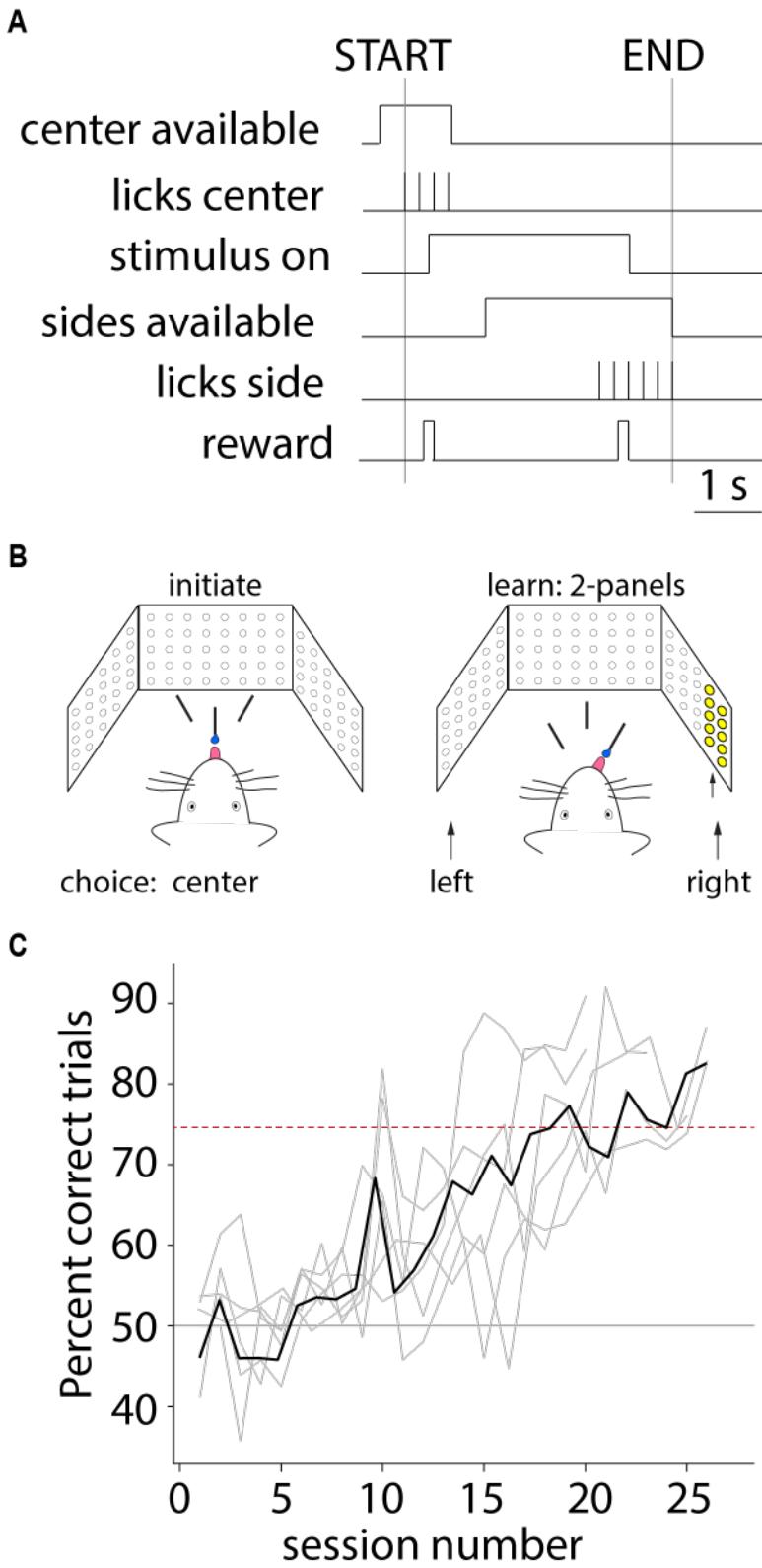


Figure 14. Visuospatial decision-making with 2AFC behavioral paradigm. (A) Schematic of trial structure. (B) Schematic of left-right visuospatial decision-making task. (C) Individual learning rates ($n = 6$) and mean (black).

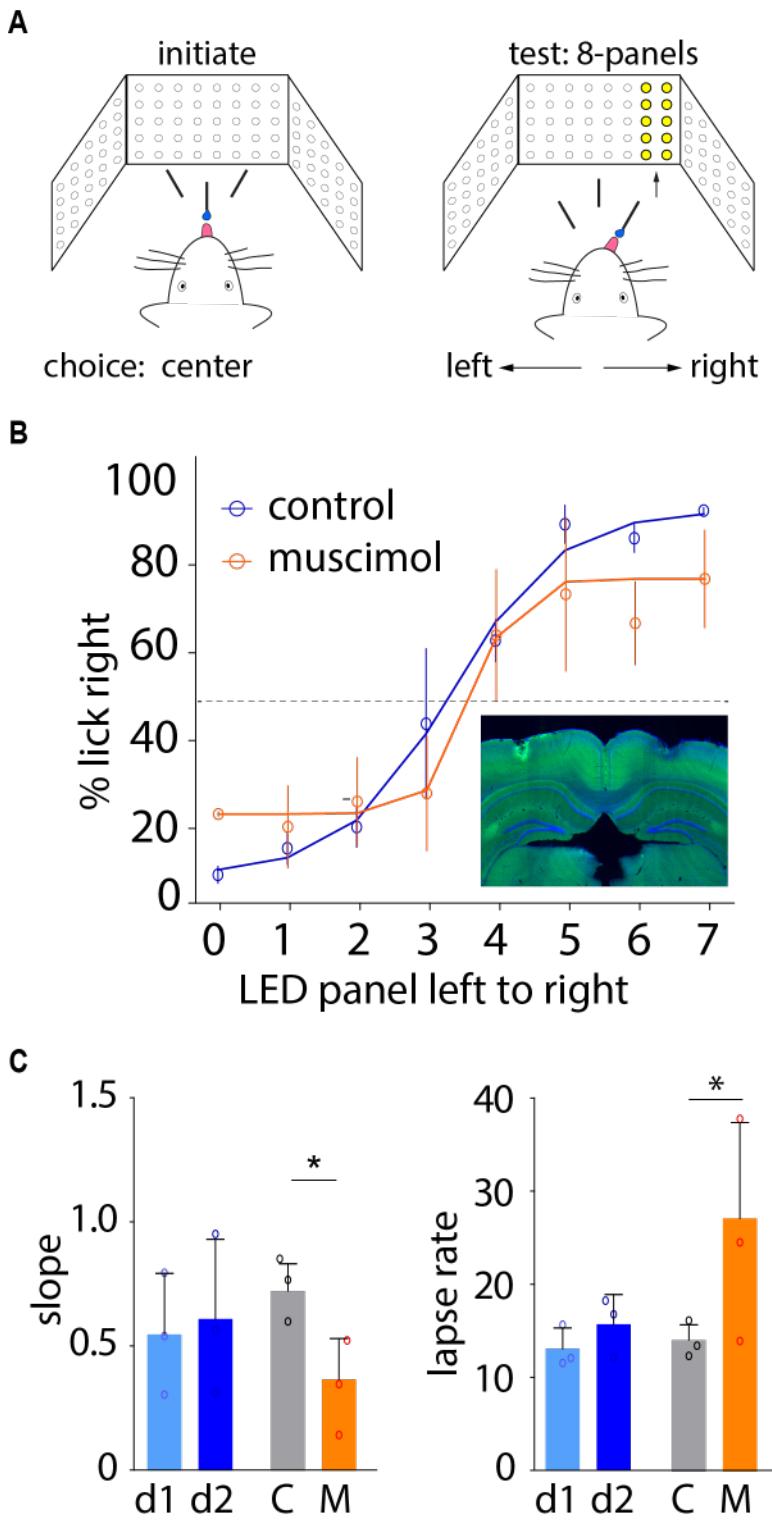


Figure 15. Categorical visuospatial decision-making with 2AFC behavioral paradigm and muscimol effect. (A) Schematic of categorical visuospatial decision-making task. (B) Psychometric curve of categorical behavioral performance after bilateral vehicle saline injection (blue) or muscimol injection (orange) injection into the posterior parietal cortex (PPC). Image of brain slice shows site of cannula, DAPI (blue), and fluorescein (green). (C) Slope and lapse rate comparisons of behavioral performance on consecutive days (d1-d2, blue, $n = 3$) or between saline injected controls (gray, $n = 3$) and muscimol injected mice (orange, $n = 3$). *

Trial #	0	correct trials limit	500	Animal ID	Y:\Ali Olyarm_ryarm\Mouse Data\stage 3\mouseID\		Dir
Start		State	0	Reward_count	0	Notes	test
General Controls		Valve Control					
Experiment Duration (s)	2400	Reward 1	Reward 2	Reward 3	Save Settings	<input type="checkbox"/> FAN / MARBLES	AllBackground
Number of Licks Required	1	Open Valve	Open Valve	Open Valve	Load Settings	<input type="checkbox"/> icecream/pizza	closescreen
Lick Time Requirement	1	35	40	45		<input type="checkbox"/> bypass	
Percent StimL	50	reward_pulses (ms)				<input type="checkbox"/> EasyMode (stage 1)	
Left correct						<input type="checkbox"/> FollowLightMode (stage 2)	
Right correct						<input checked="" type="checkbox"/> Mixstage (stage 3)	
NO STIM							
Percent Stage2							
Timeout Punishment							
Timeout Punishment (s)							
Rule Change Count							

Figure 16. Y-maze 2AFC behavioral GUI. User graphical interface for inputting the desired settings for each stage in the y-maze two alternative forced choice behavioral paradigm.

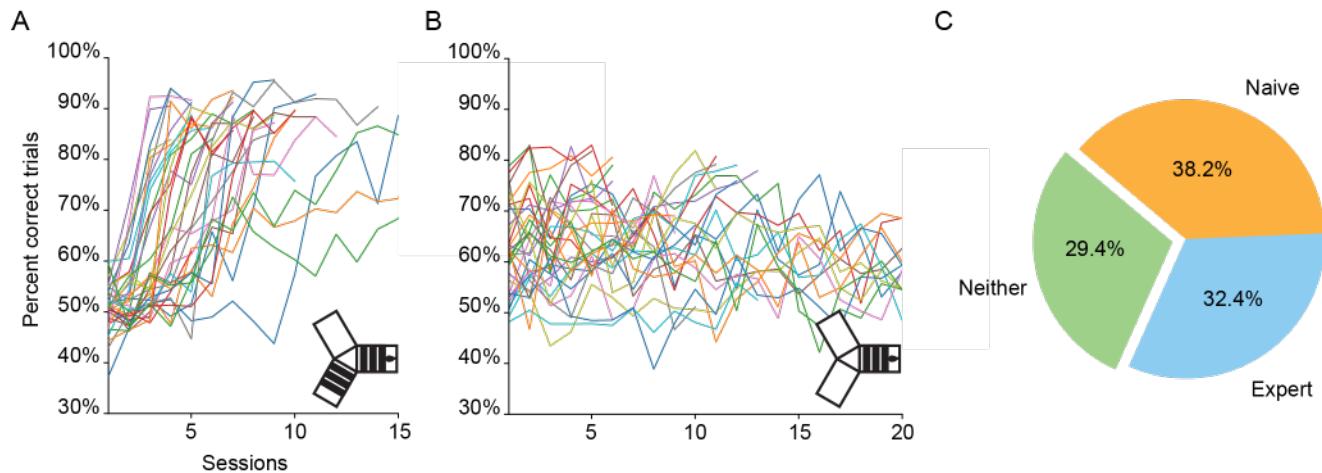


Figure 17. Behavioral learning rates for FTP and INIT stages. **(A)** Behavioral performance of each mouse in the follow the pattern stage (FTP). **(B)** Behavioral performance of each mouse in the initiation arm only stage (INIT). **(C)** Percentage of mice where the first session on the INIT stage was expert (percent correct trials > 65%), naive (40% < percent correct trials < 60%), or neither (60% ≤ percent correct trials ≤ 65%).

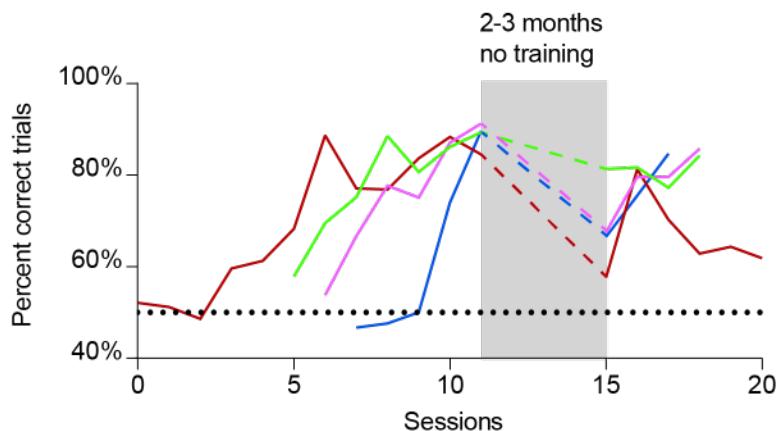
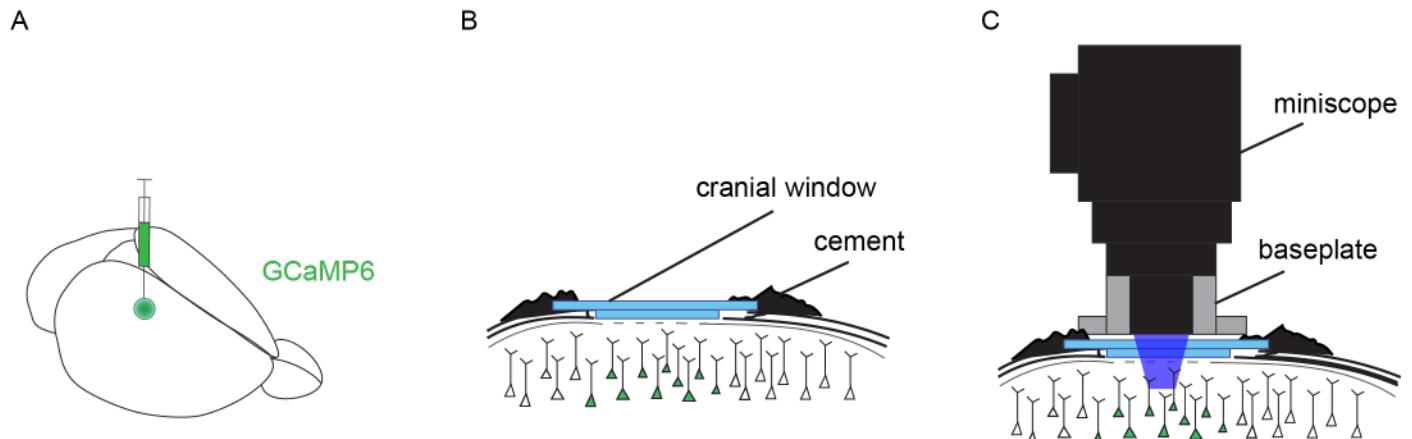


Figure 18. Long-term memory behavioral performance. Mice were trained until expert behavioral performance was achieved at the initiation arm only stage (INIT). They were then returned to their home cages for 2-3 months with no behavioral sessions. Finally, they were tested again on the INIT stage ($n = 4$, black dotted line = chance level behavioral performance).



GCaMP injection ... 2 weeks Craniotomy ... 2 weeks Baseplating
P30~P40

Figure 19. Surgery timeline. (A) Mouse is injected with GCaMP in the PPC between ages P30-P40. (B) After 2 weeks, a craniotomy is performed and a cranial window is cemented over the exposed brain. (C) After 2 more weeks, a baseplate is cemented over the cranial window to allow for miniscope mounting and imaging.

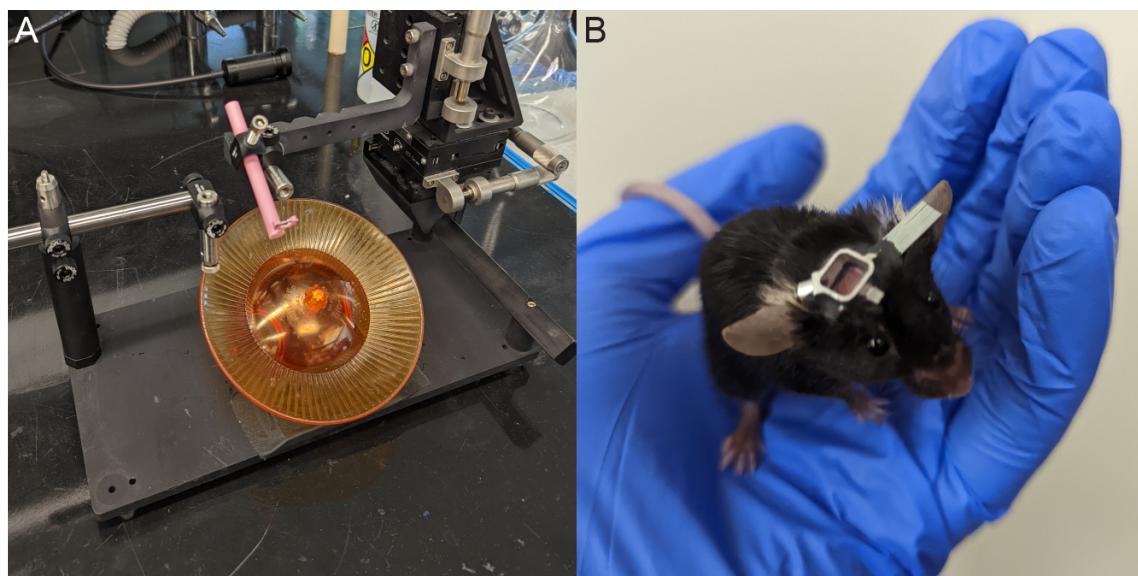


Figure 20. Preparations for calcium imaging. (A) A 3D printed miniscope holder attached to micromanipulator is held over a running wheel, adjacent to a head clamping set-up. (B) Mouse after receiving GCaMP injection, cranial window surgery, and baseplate attachment.

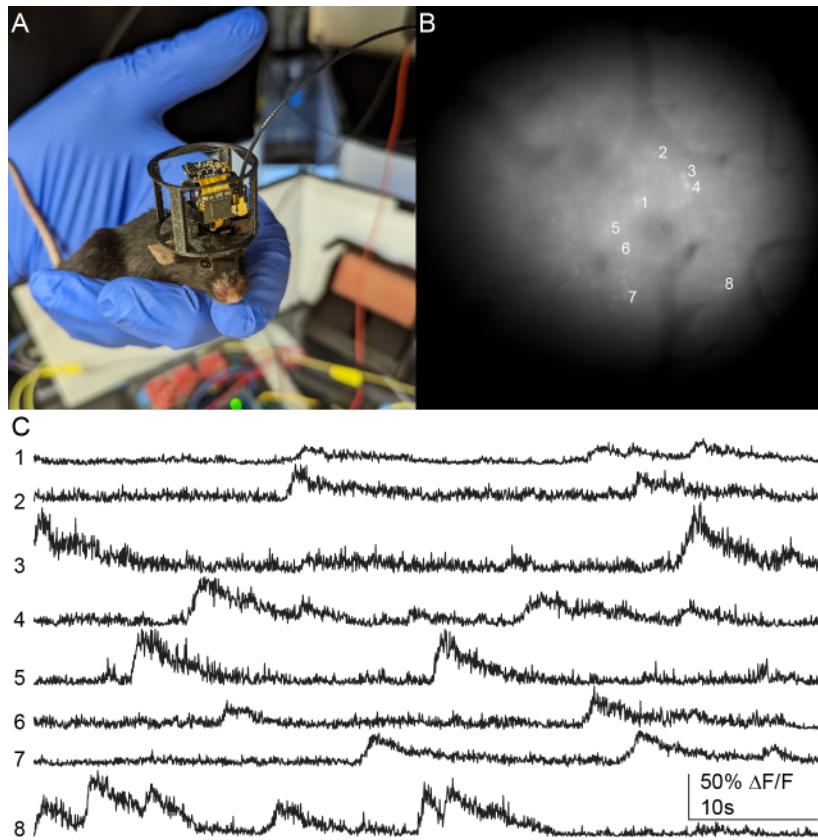


Figure 21. Miniscope recording example traces. (A) Mouse with a mounted miniscope and 3D printed scope cage. (B) Raw image from miniscope recording of the posterior parietal cortex with GCaMP injection. (C) 1 minute long calcium traces from 8 example cells.

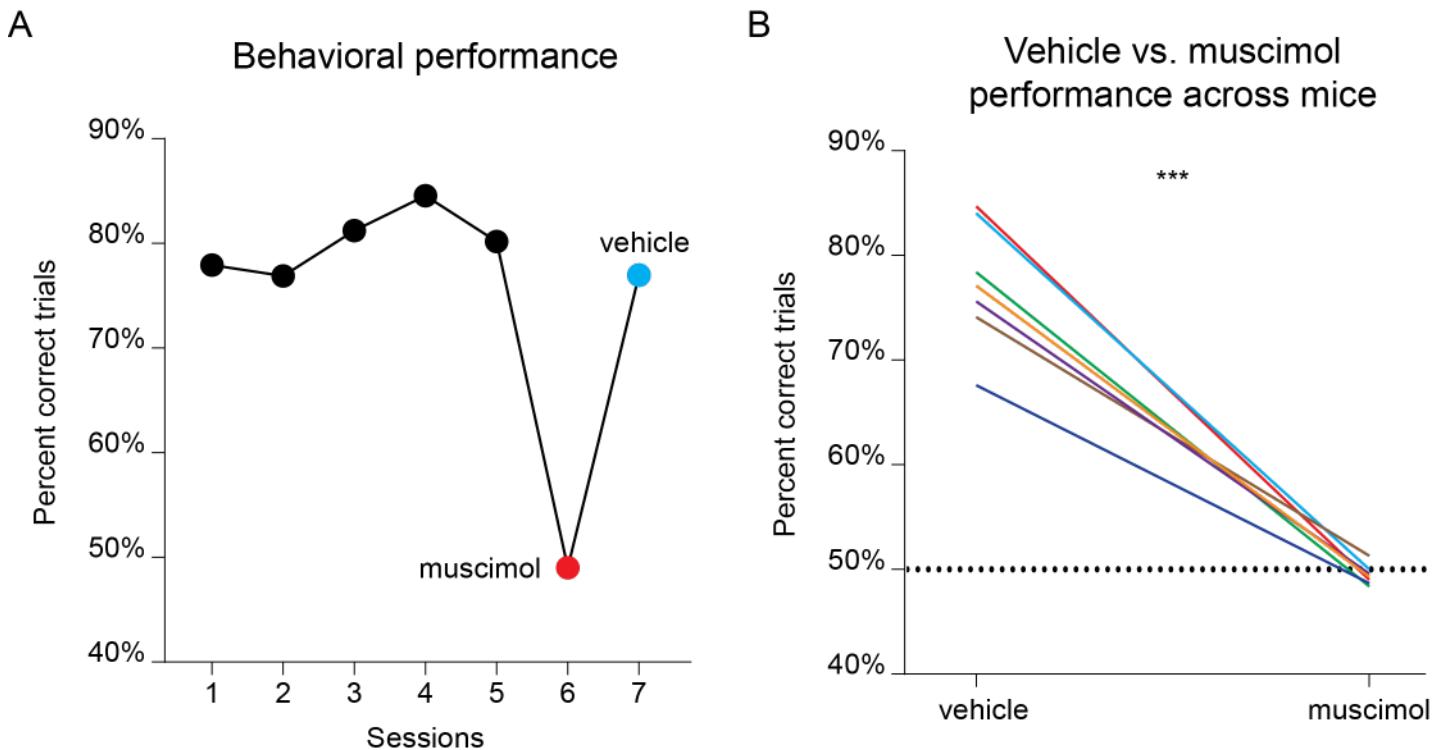


Figure 22. Effect of muscimol on behavioral performance. (A) Mouse was trained to expert proficiency on the initiation arm only stage (INIT). After at least 3 consecutive sessions on this stage with greater than 75% correct trials, muscimol was injected into the posterior parietal cortex (PPC) before the next session. Finally, saline was injected into the PPC before the following session. (B) 7 mice went through the same behavioral stages described, but the order of muscimol vs. vehicle (ACSF) days were randomized. ***: $p < 0.001$, paired t-test.

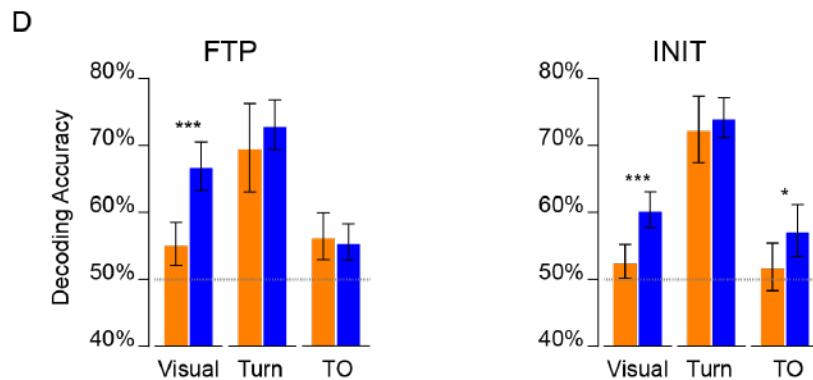
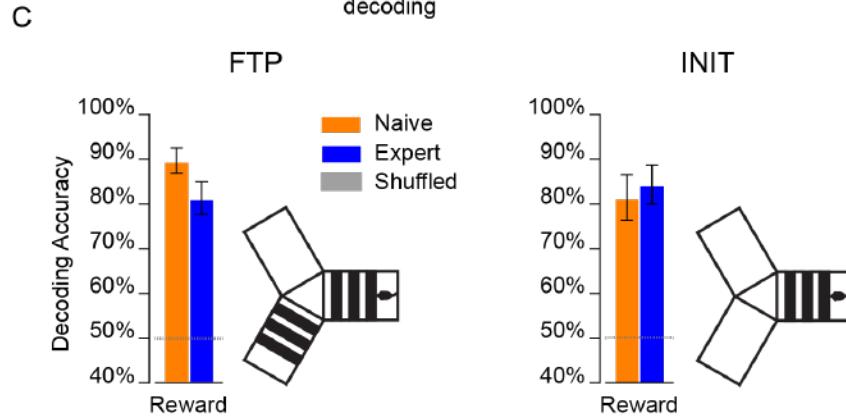
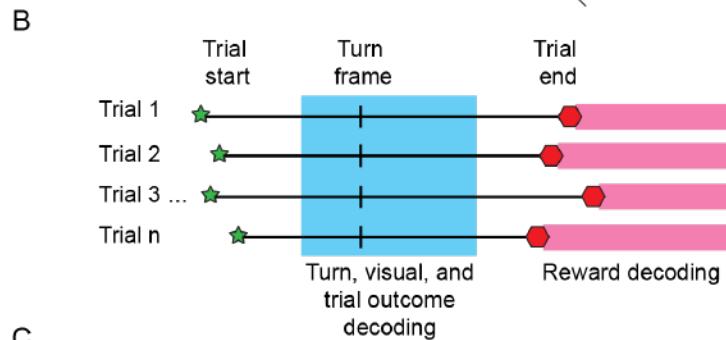
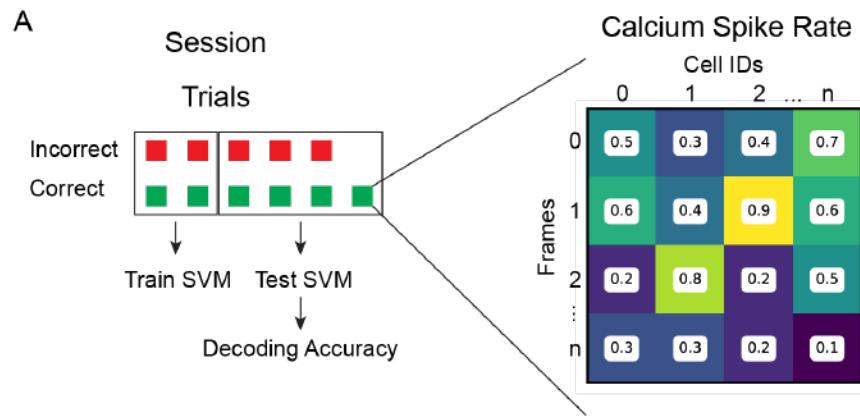


Figure 23. Decoding visual stimuli, turn direction, trial outcome, and reward using a support vector machine. (A) Schematic of information used in the support vector machine (SVM). A subset of trials belonging to two different categories, such as correct and incorrect trials, are used to train the SVM. Then, an independent group of trials is used to test the decoding accuracy. Each trial contains calcium spike rate data pertaining to each cell in the recording. The number of mice used in the datasets are as followed: FTP stage ($n_{\text{naive}} = 23$, $n_{\text{expert}} = 25$) and INIT stage ($n_{\text{naive}} = 15$, $n_{\text{expert}} = 22$) (B) Schematic of trial durations and the range of frames used per category of decoding. The trials vary in length but are aligned with either the turn frame or trial end frame depending on the category of decoding. (C) Decoding accuracy for reward during the follow the pattern (FTP) stage and initiation arm only stage (INIT). (D) FTP and INIT decoding accuracy naive vs. expert for visual stimuli, turn direction, and trial outcome. Shuffled decoding accuracy has a mean of 50.03% with an upper 95% confidence interval of the mean at 50.07% and a lower 95% confidence interval of the mean at 49.99%. *: $p < 0.05$, *** $p < 0.001$, unpaired t-test. Error bars represent 95% confidence intervals.

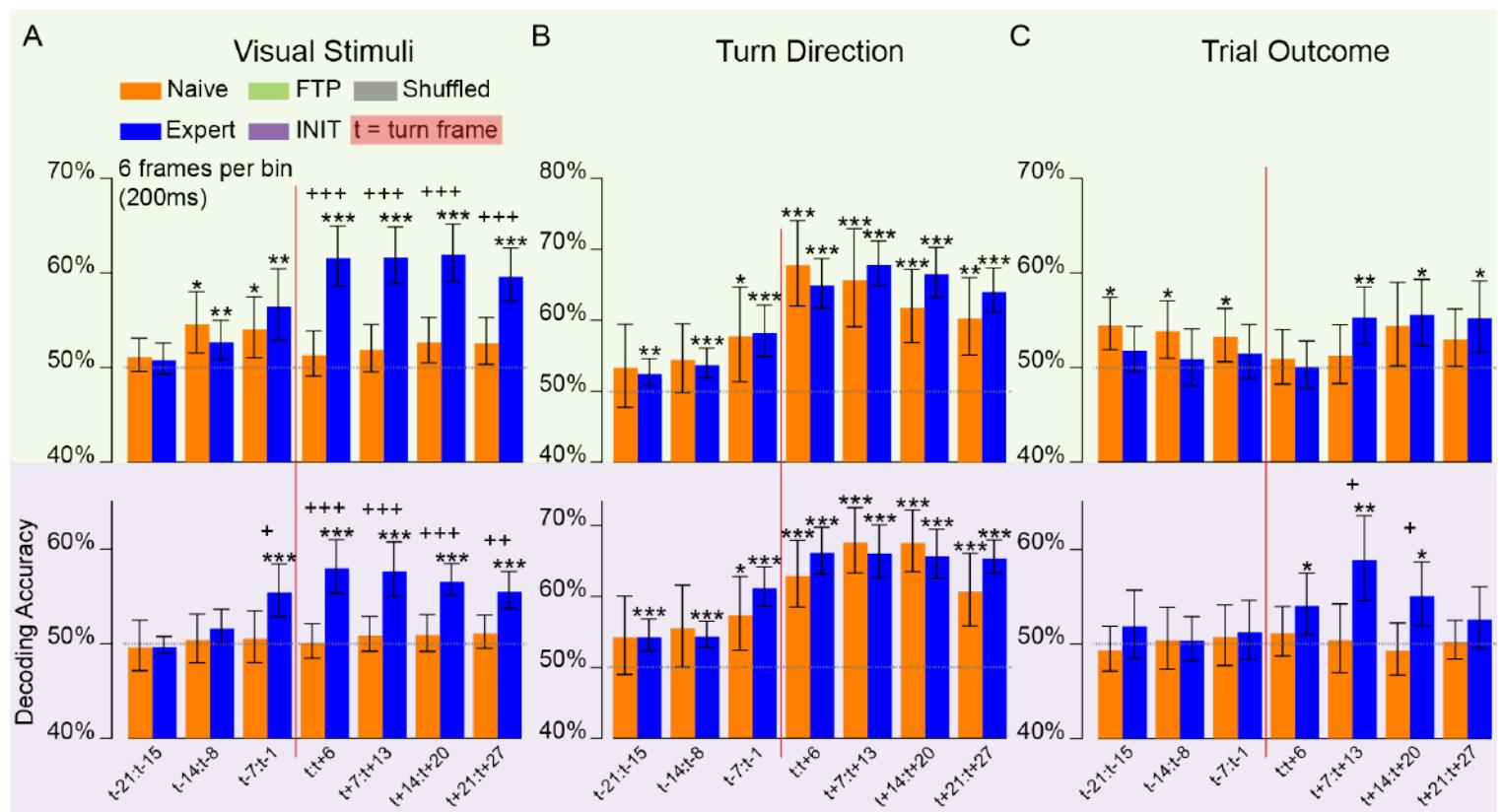


Figure 24. Time binned decoding. Decoding accuracy within 200ms timebins across a trial, for turn direction, visual stimuli, and trial outcome based on the FTP stage ($n_{\text{naive}} = 23$, $n_{\text{expert}} = 25$, green background) and INIT stage ($n_{\text{naive}} = 15$, $n_{\text{expert}} = 22$, purple background). The turn frame is denoted by the red line. A one-sample t-test was first used to test whether the decoding accuracy is greater than the shuffled decoding accuracy at 50%. **(A)** Decoding accuracy for visual stimulus identity. **(B)** Decoding accuracy for turn direction. **(C)** Decoding accuracy for trial outcome. Due to the large number of tests conducted, we accounted for multiple comparisons by computing adjusted p-values using false discovery rate (FDR). Within group, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. Between groups, +: $p < 0.05$, ++: $p < 0.01$, +++: $p < 0.001$. Error bars represent 95% confidence intervals.

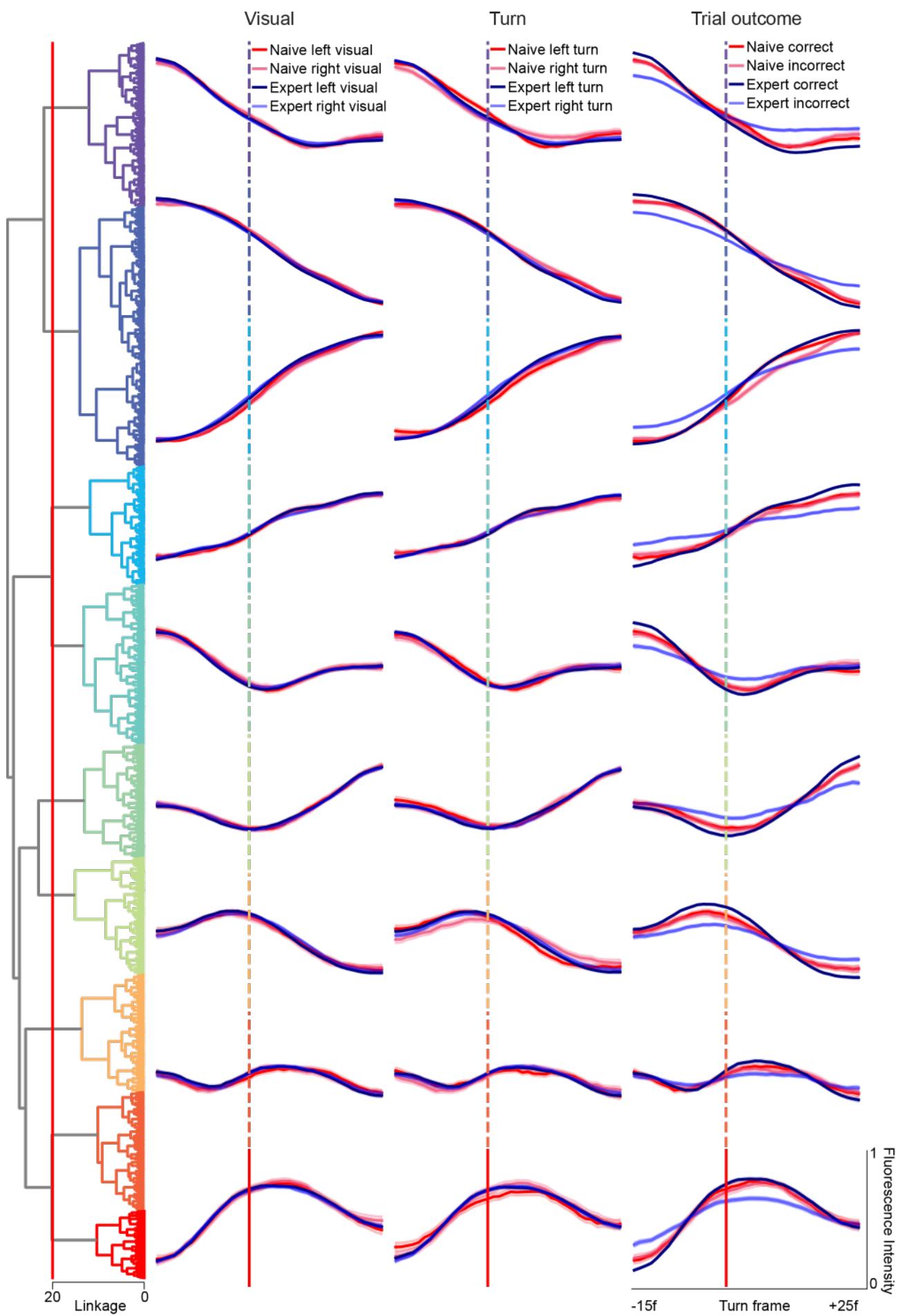


Figure 25. FTP clustering across mice. Neurons across 23 naive and 25 expert mice are used for unsupervised hierarchical clustering, based on the average trial activity for each neuron from follow the pattern trials. After the clusters are determined, the average trial activity is displayed for naive and expert trials, per trial category (visual stimuli, turn direction, and trial outcome) along with the 95% confidence intervals.

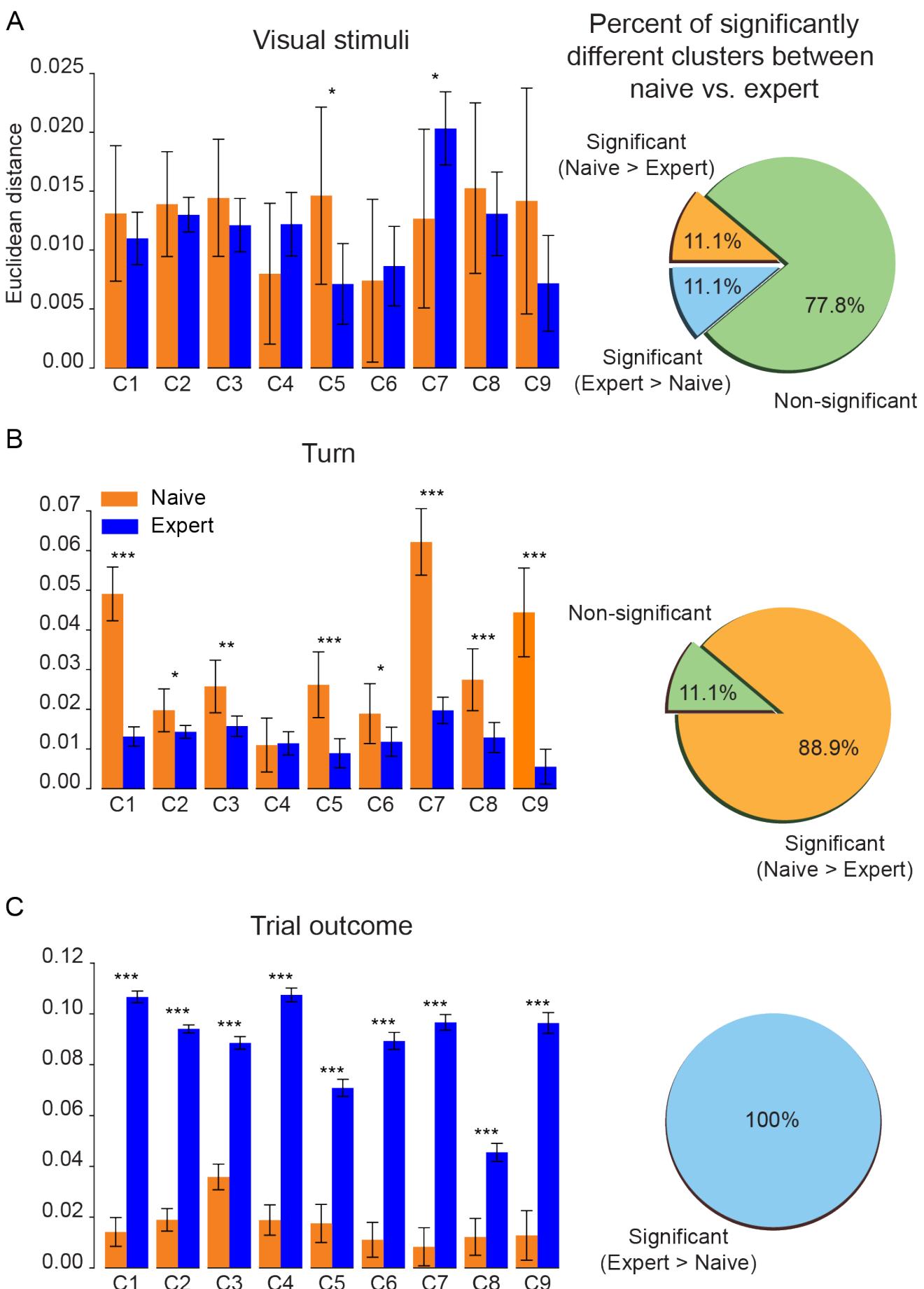


Figure 26. FTP individual cluster comparisons. (A) Visual stimuli, naive vs. expert Euclidean distance. (B) Turn direction, naive vs. expert Euclidean distance. (C) Trial outcome, naive vs. expert Euclidean distance. *: $p < 0.05$, **: $p < 0.01$, *** $p < 0.001$, unpaired t-test. Error bars represent 95% confidence intervals.

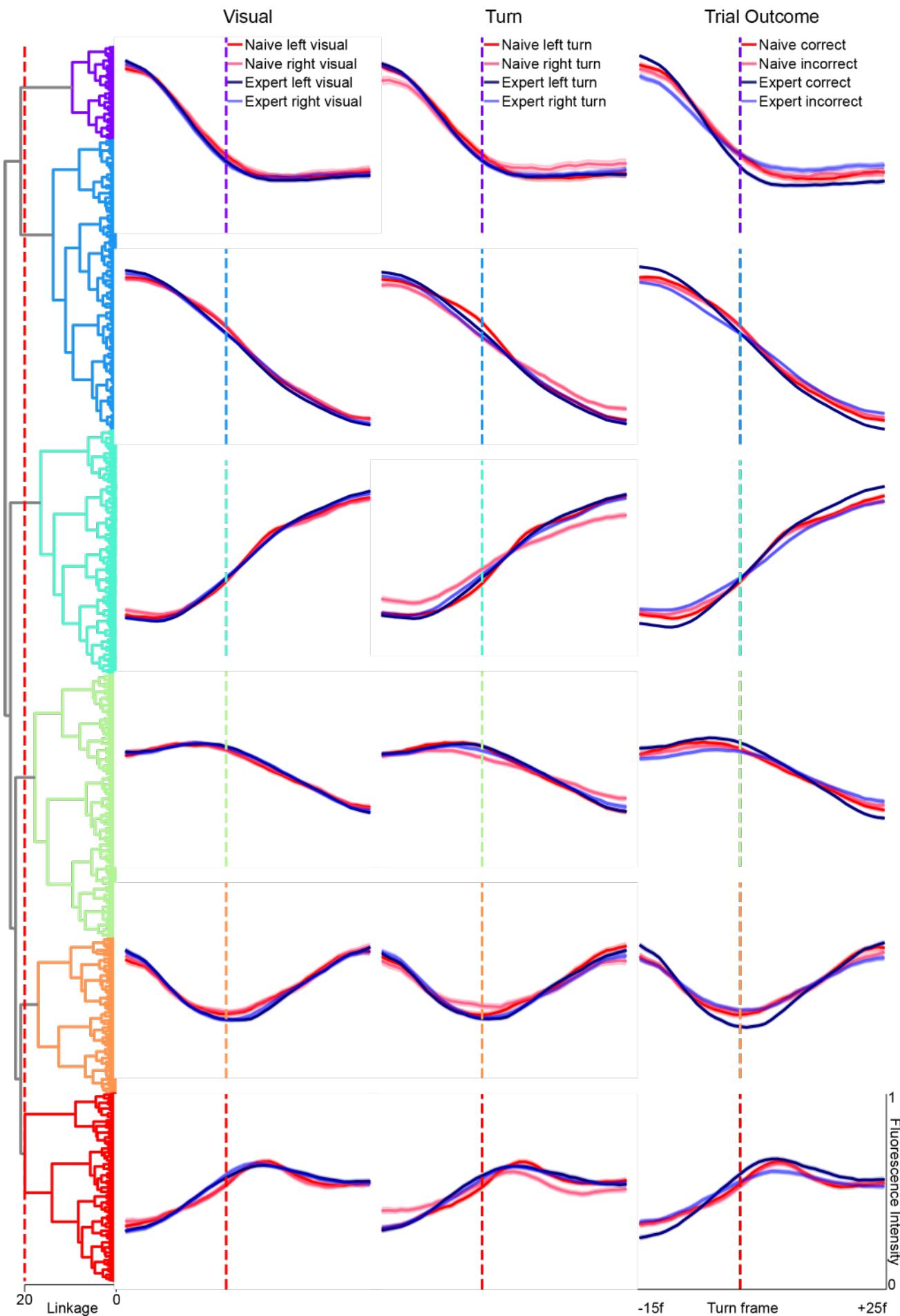


Figure 27. INIT clustering across mice. Neurons across 15 naive and 22 expert mice are used for unsupervised hierarchical clustering, based on the average trial activity for each neuron from initiation arm only trials. After the clusters are determined, the average trial activity is displayed for naive and expert trials, per trial category (visual stimuli, turn direction, and trial outcome) along with the 95% confidence intervals.

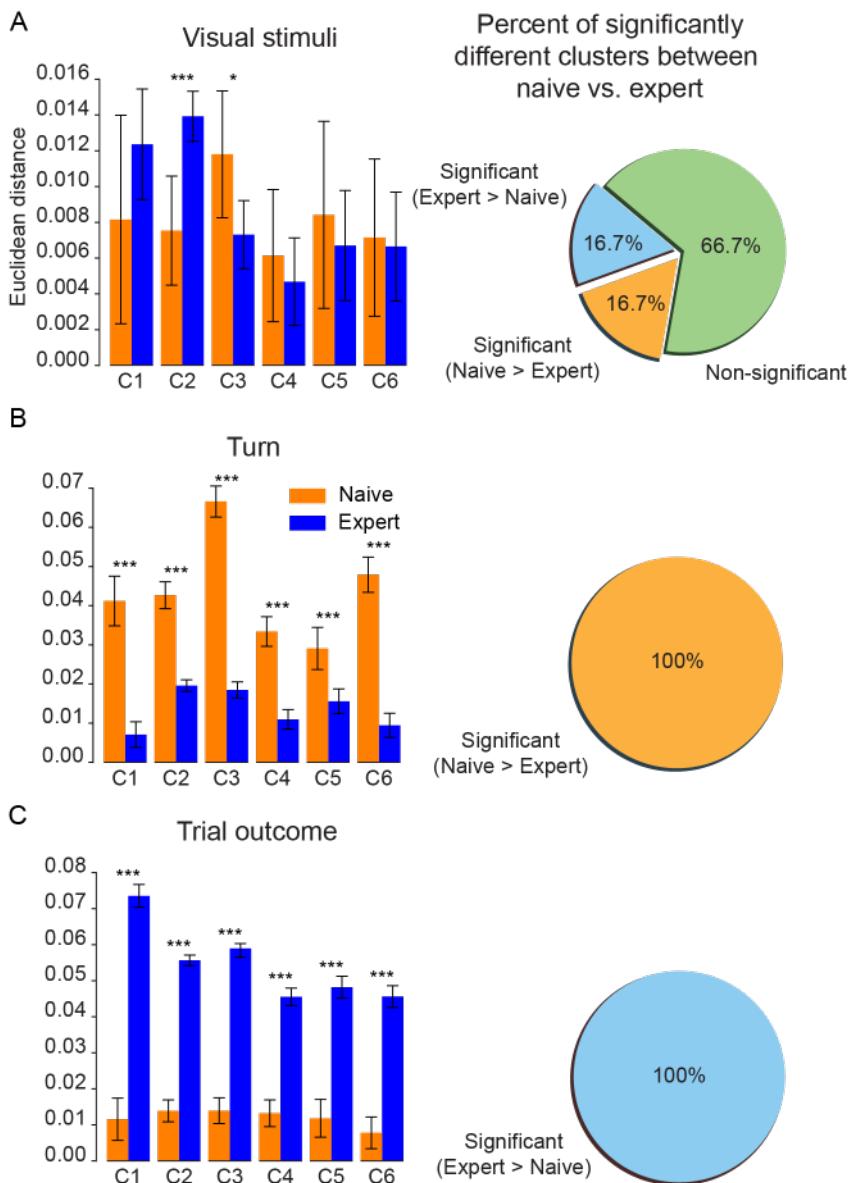


Figure 28. INIT individual cluster comparisons. (A) Visual stimuli, naive vs. expert Euclidean distance. (B) Turn direction, naive vs. expert Euclidean distance. (C) Trial outcome, naive vs. expert Euclidean distance. *: $p < 0.05$, **: $p < 0.001$, unpaired t-test. Error bars represent 95% confidence intervals.

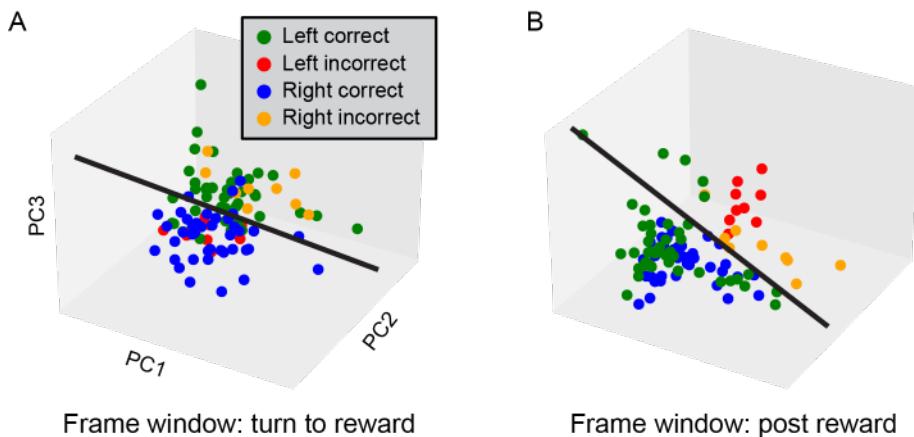


Figure 29. PCA analysis of trial types. (A) In an expert mouse, turn direction can be separated using principal component analysis when analyzing the neural activity between the turn onset and reward frames. Each trial is presented as a single point in PCA state space. (B) When the analysis is shifted to begin immediately after the reward frames in the same trials, the neuronal firing preferentially shifts to reward encoding.

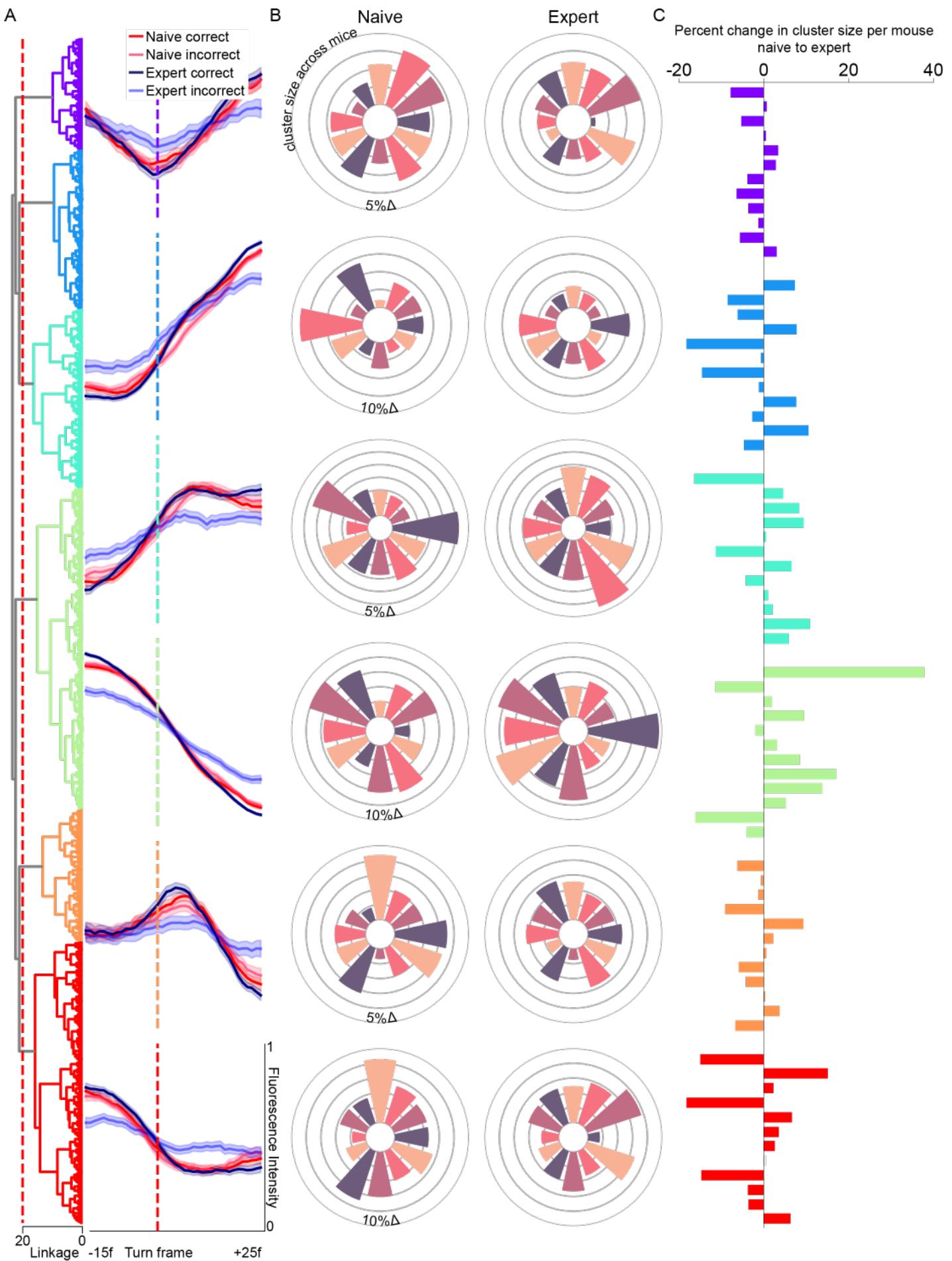


Figure 30. FTP cluster size dynamics with learning across mice. (A) Neurons across 12 mice are used for unsupervised hierarchical clustering, based on the average trial activity for each neuron during the follow the pattern stage. After the clusters are determined, the average trial activity is displayed for naive and expert trials, per trial outcome along with the 95% confidence intervals. (B) Each bar represents the percentage of cells for a single mouse within the given cluster, for both its naive session and expert session. The changes per circle are in steps of either 5% or 10%, as denoted. (C) The percent change in cluster size per mouse is determined by subtracting the percentage of cells in the given cluster for the expert session from the naive session.

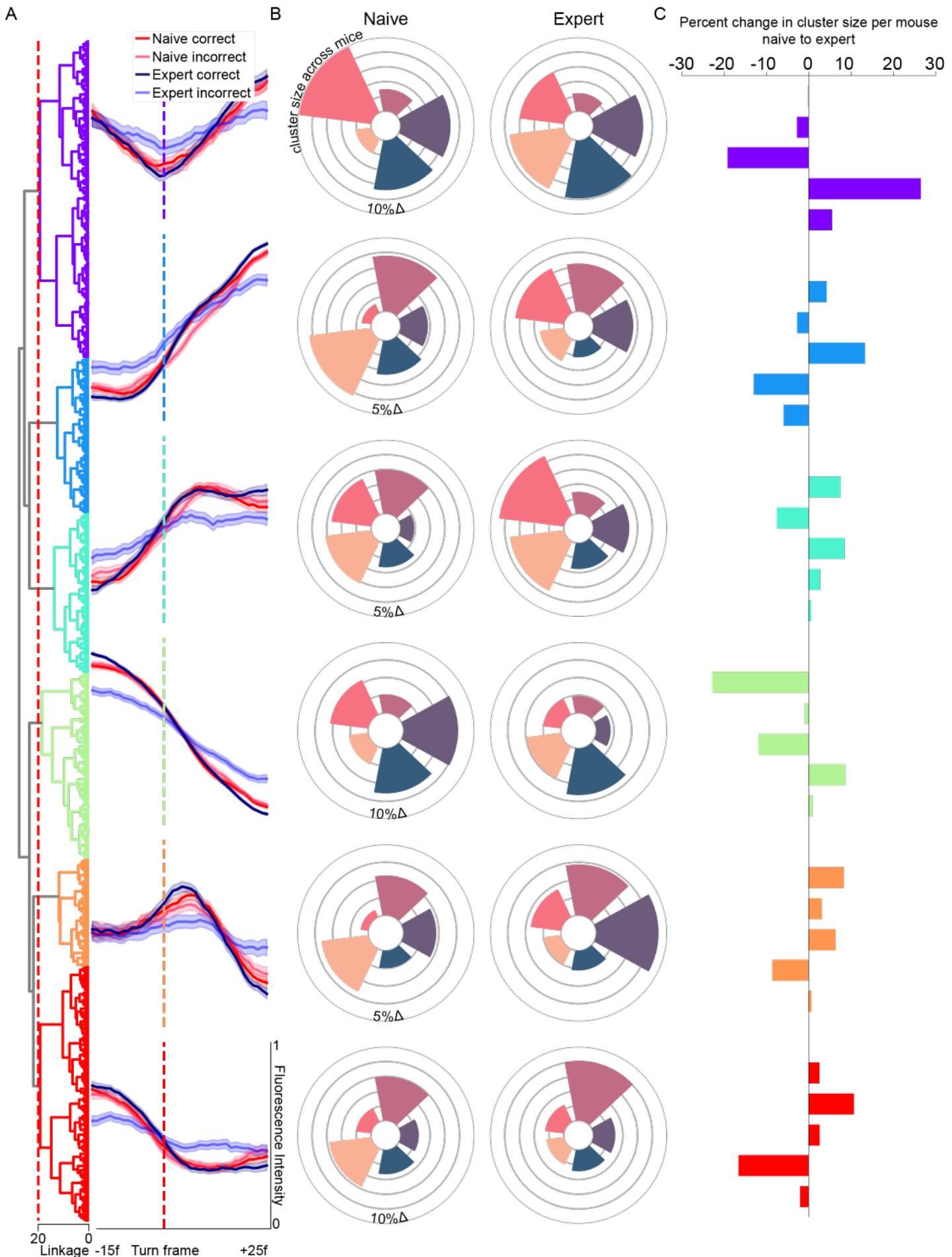


Figure 31. INIT cluster size dynamics with learning across mice. (A) Neurons across 5 mice are used for unsupervised hierarchical clustering, based on the average trial activity for each neuron during the initiation arm only stage. After the clusters are determined, the average trial activity is displayed for naive and expert trials, per trial outcome along with the 95% confidence intervals. (B) Each bar represents the percentage of cells for a single mouse within the given cluster, for both its naive session and expert session. The changes per circle are in steps of either 5% or 10%, as denoted. (C) The percent change in cluster size per mouse is determined by subtracting the percentage of cells in the given cluster for the expert session from the naive session.

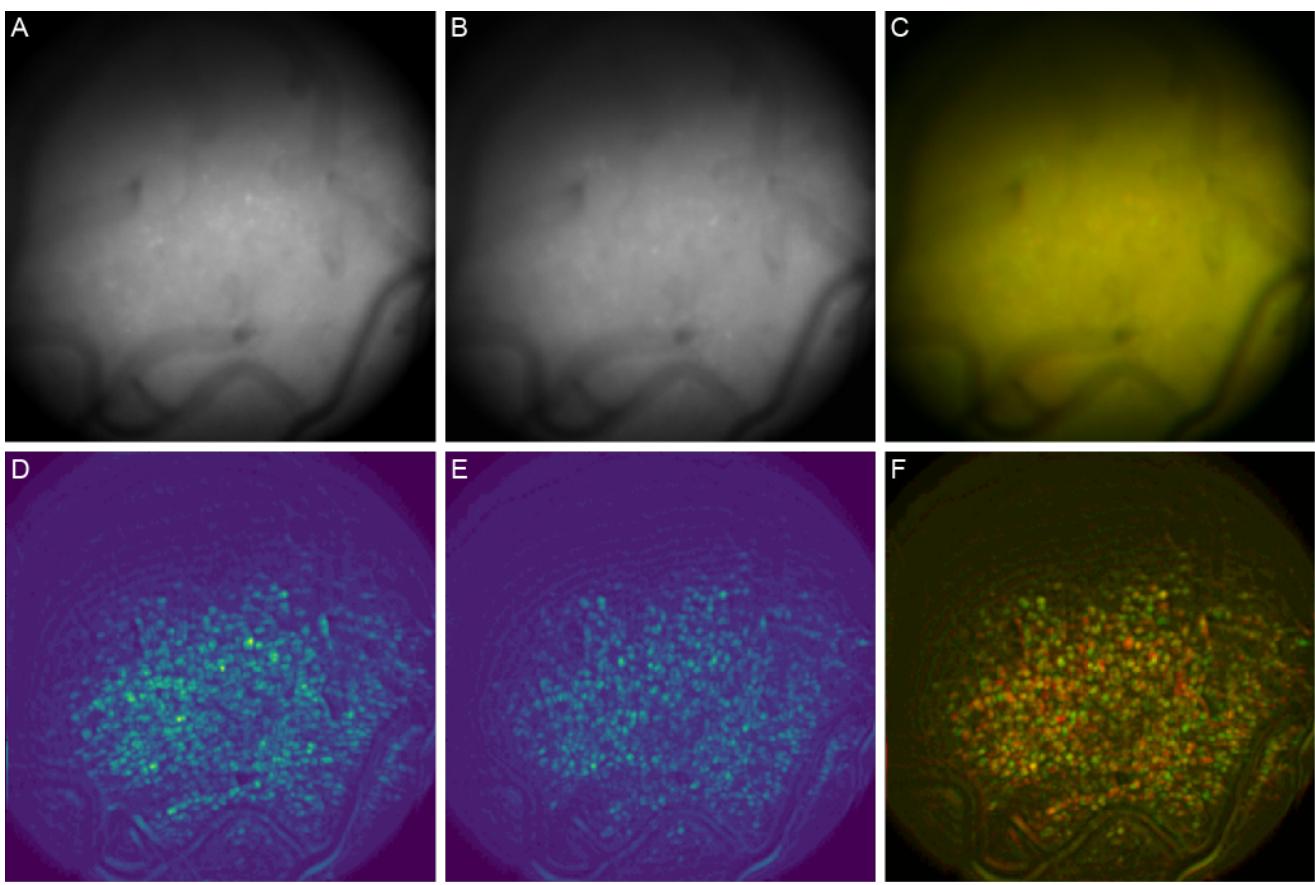


Figure 32. Cross-session image preprocessing and cell extraction. (A) Image from a raw recording with max fluorescence intensity obtained across 1000 frames (~33.3 sec). (B) Another image from a recording 6 days later in the same mouse. (C) Images from A (green channel) and B (red channel) overlaid. (D) Image from A shifted to obtain optimal overlap with the same cells from image B using Minian analysis script. (E) Image B shifted to obtain optimal overlap with the same cells from image A using Minian analysis script. (F) Images from D (green channel) and E (red channel) overlaid.

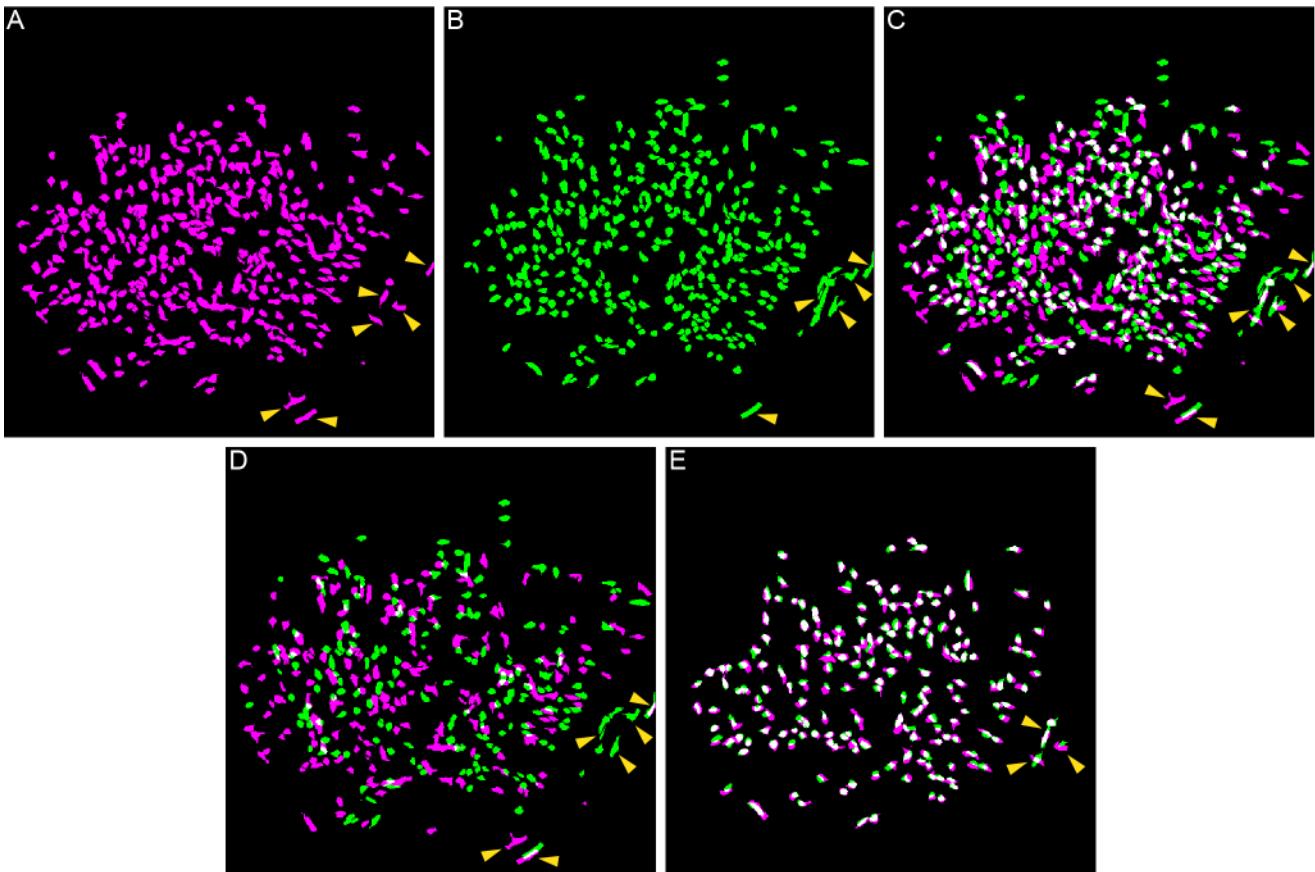


Figure 33. Cross-session cell matching. Cell identities were determined using the Minian analysis package, and cells were rejected from analysis based on size and shape parameters (indicated via yellow arrowheads). (A) All cells from recording session A. (B) All cells from recording session B, 6 days later in the same mouse. (C) All cells from recording A (purple) and all cells from recording B (green) overlaid. (D) Unmatched cells between recording A (purple) and recording B (green). (E) Matched, overlapping cells between recordings A and B (white).

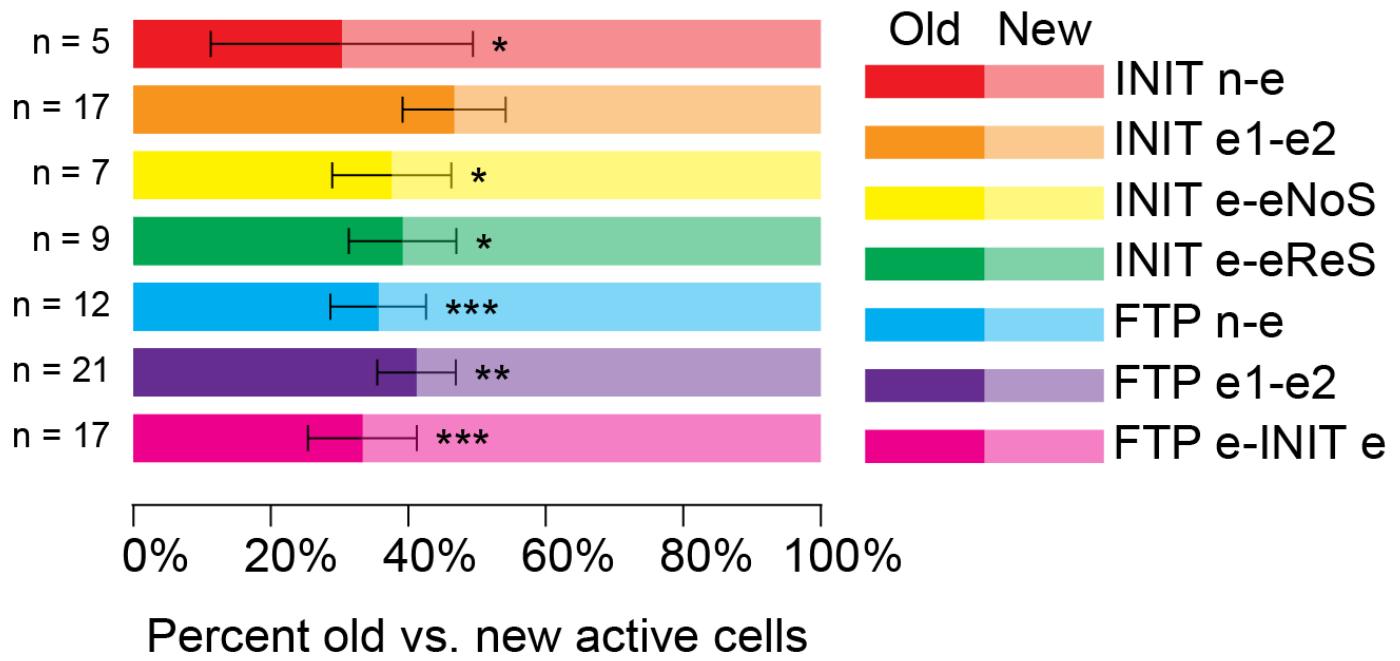


Figure 34. Active cell turnover across sessions. Analysis is done across 7 different types of paired sessions. Percent old active cells refers to the percentage of cells in the second session pairing that was also actively present in the first session. Percent new active cells refers to the percentage of cells in the second session pairing that was not actively present in the first session. INIT n-e (naive to expert INIT sessions). INIT e1-e2 (expert to expert INIT sessions). INIT e-eNoS (expert INIT session to session with visual stimuli hidden). INIT e-eReS (expert INIT session to expert INIT session with visual stimuli returned). FTP n-e (naive to expert FTP sessions). FTP e1-e2 (expert to expert FTP sessions). FTP e-INIT e (expert FTP session to expert INIT session). *: $p < 0.05$, **: $p < 0.01$, *** $p < 0.001$, paired t-test. Error bars represent 95% confidence intervals.

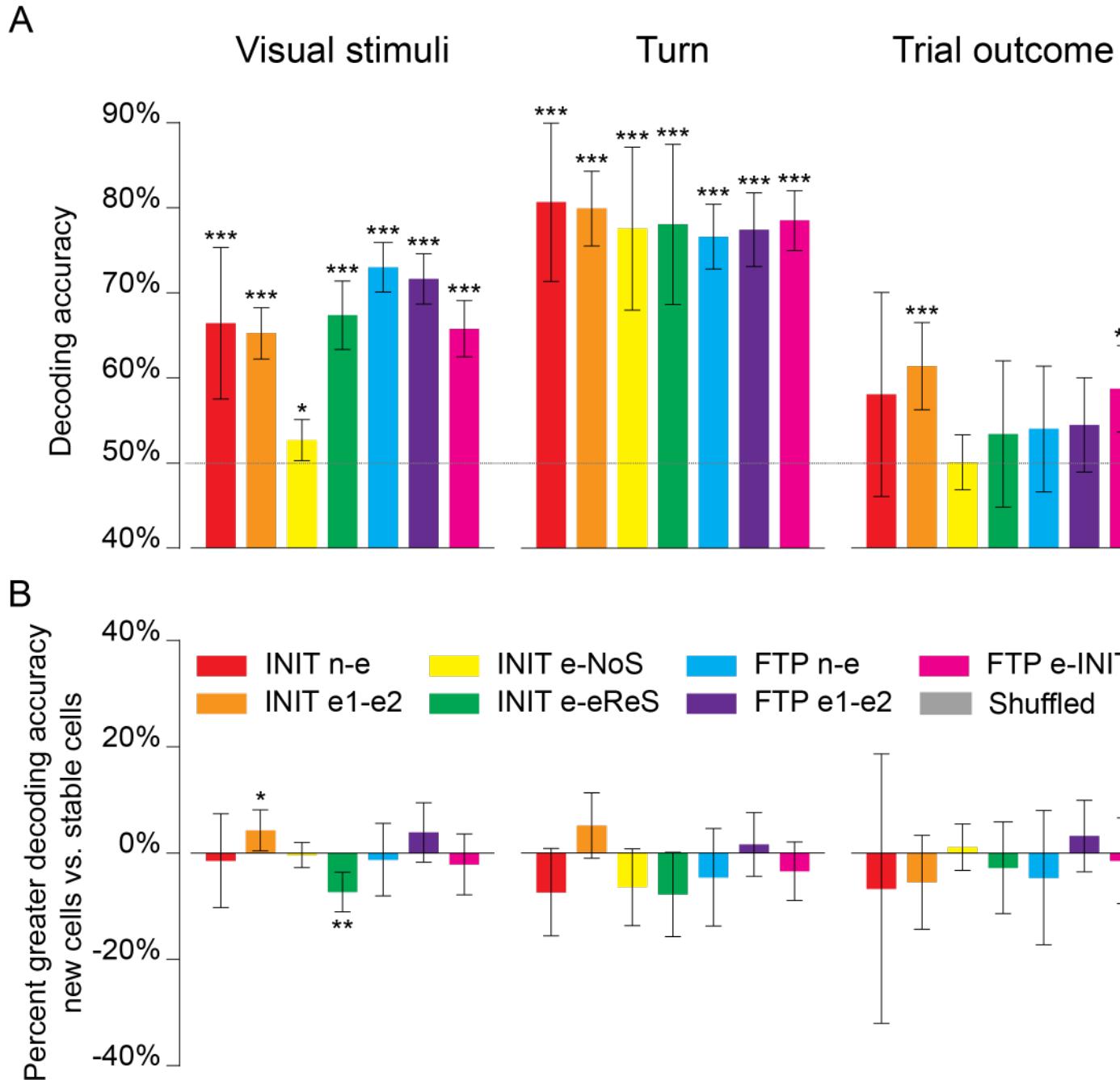


Figure 35. Cell matched decoding across paired sessions. (A) Decoding accuracy of the second session in the pairing, using all cells in each session. Significance tested against shuffled. (B) Decoding accuracy difference between using only newly active cells vs. only old active, stable cells for the second session in the pairing. The first session in the pairing is used to determine the matched and unmatched cells. *: p < 0.05, **: p < 0.01, ***p < 0.001, paired t-test. Error bars represent 95% confidence intervals.

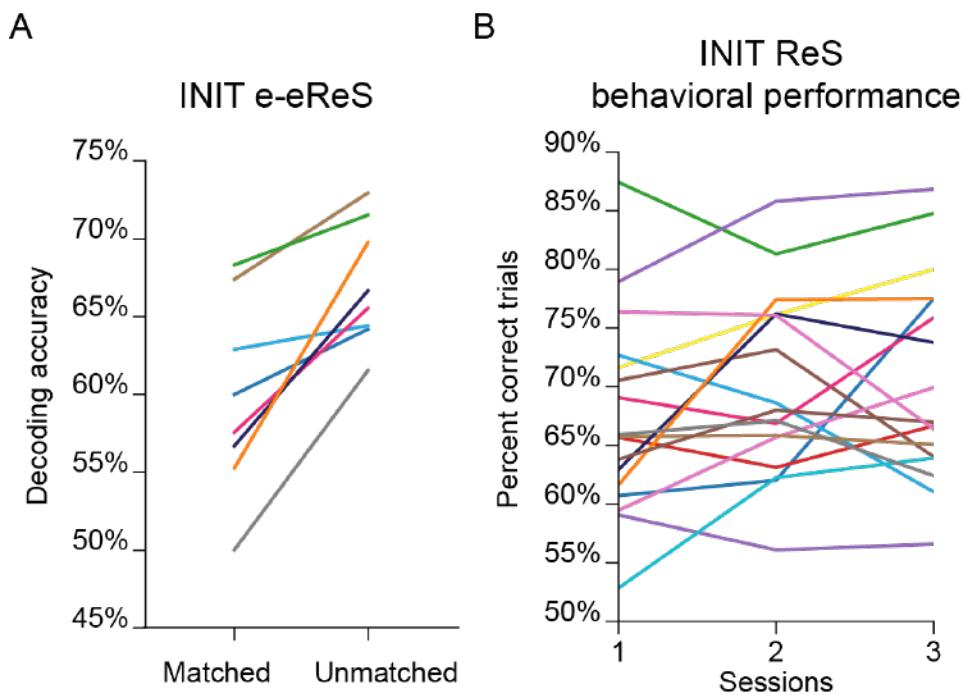


Figure 36. INIT e-ReS individual mice comparisons. (A) For each mouse, a pair of sessions is obtained. First is an expert INIT session. Second is also an expert INIT session, but where the visual stimuli is returned after several days of no visual stimuli being present during the behavior. The visual stimuli decoding accuracy for the second paired session is plotted using only stable active cells and using only newly active cells. (B) Behavioral performance is plotted for all mice during return of stimuli stage (INIT ReS).

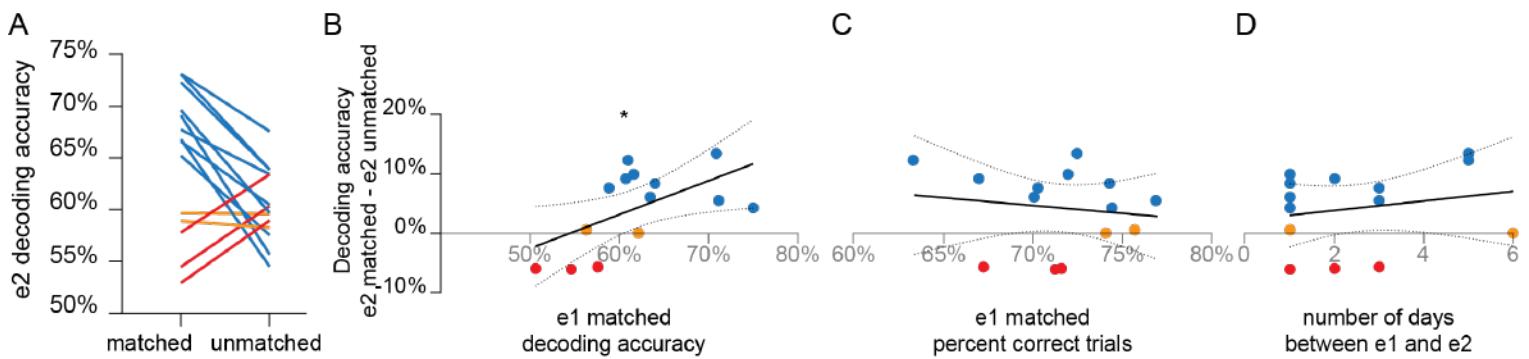
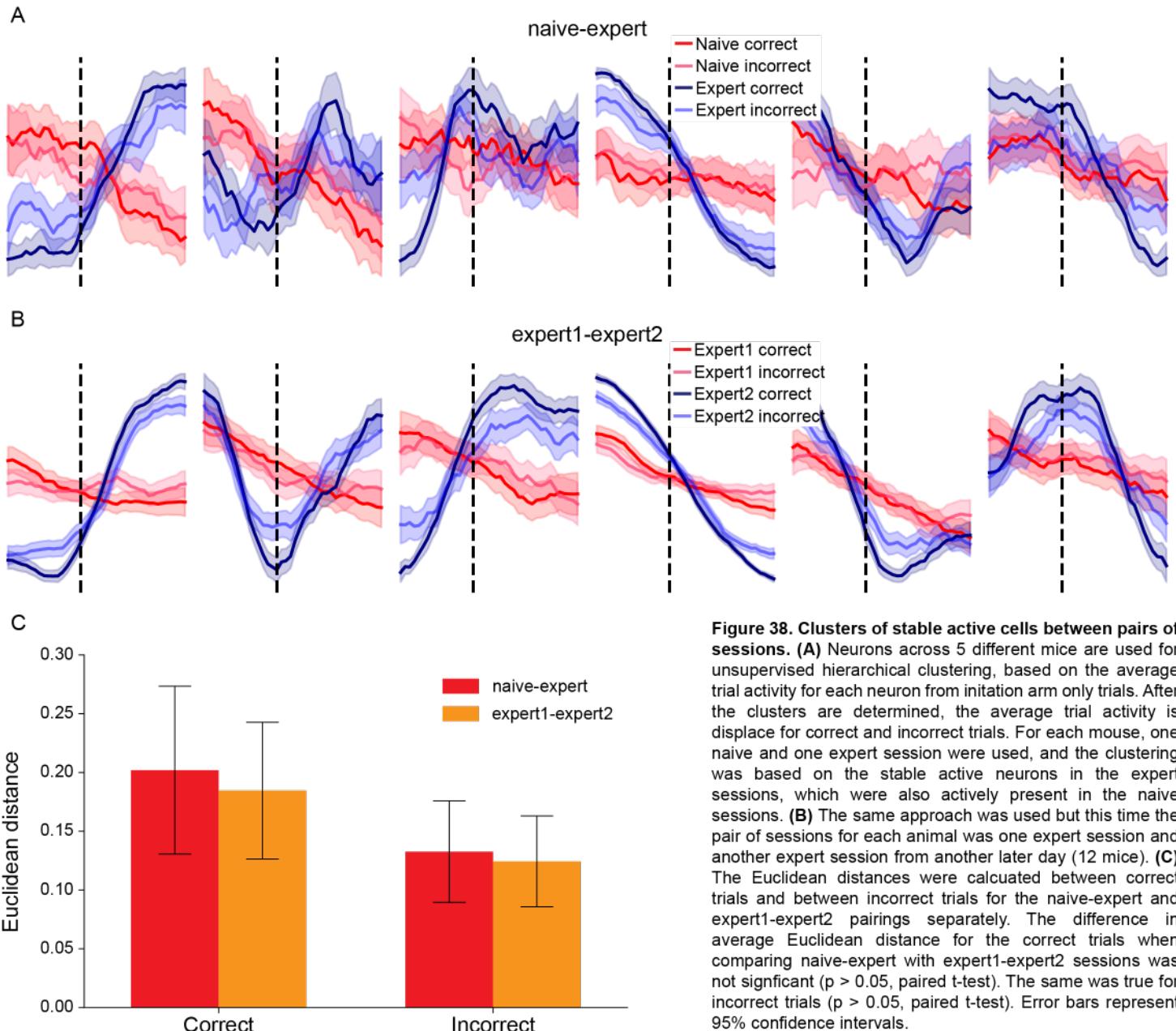


Figure 37. INIT e1-e2 individual mice comparisons. (A) Expert to expert INIT sessions, where the visual stimuli decoding accuracy of the second session of each pairing is tested using only matched or unmatched cells with the first session from the pair. Each pair corresponds to a different mouse. (B) Decoding accuracy difference between using only unmatched vs. only matched cells for the second paired session plotted against the decoding accuracy of the first paired session using matched cells only ($r^2 = 0.37$, $p < 0.05^*$). (C) Decoding accuracy difference between using only unmatched vs. only matched cells for the second paired session plotted against the decoding accuracy of the behavioral performance of the mouse in the first paired session ($r^2 = 0.02$, $p < 0.62$). (D) Decoding accuracy difference between using only unmatched vs. only matched cells for the second paired session plotted against the number of days between the two paired sessions ($r^2 = 0.04$, $p < 0.47$). All statistics were done using a simple linear regression with 95% confidence intervals.



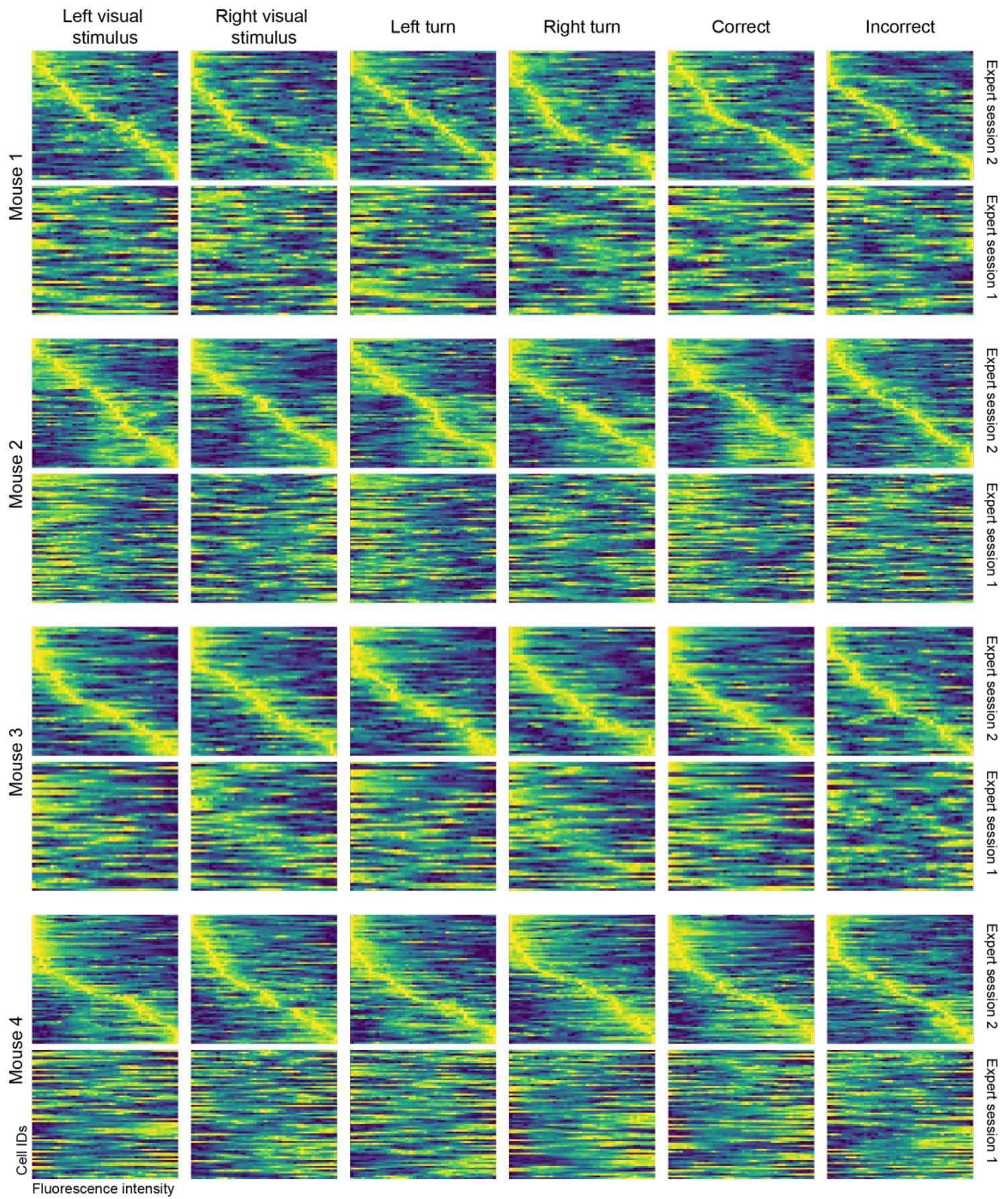


Figure 39. Heatmaps of stable active cell activity between expert sessions in individual mice. The same neurons that were active in an earlier expert session were then sorted by peak activity during a later expert session (top row), based on the average trial activity for each neuron within its respective trial category (left visual stimulus, right visual stimulus, left turn, right turn, correct, and incorrect). The order of cell IDs is then conserved and applied to the earlier expert session (bottom row). Data is used for 4 individual mice, where each pair of rows pertains to an expert and naive session from a single animal.