

Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior

Michael J Siniscalchi¹, Victoria Phoumthipphavong², Farhan Ali², Marc Lozano² & Alex C Kwan¹⁻³

The ability to shift between repetitive and goal-directed actions is a hallmark of cognitive control. Previous studies have reported that adaptive shifts in behavior are accompanied by changes of neural activity in frontal cortex. However, neural and behavioral adaptations can occur at multiple time scales, and their relationship remains poorly defined. Here we developed an adaptive sensorimotor decision-making task for head-fixed mice, requiring them to shift flexibly between multiple auditory-motor mappings. Two-photon calcium imaging of secondary motor cortex (M2) revealed different ensemble activity states for each mapping. When adapting to a conditional mapping, transitions in ensemble activity were abrupt and occurred before the recovery of behavioral performance. By contrast, gradual and delayed transitions accompanied shifts toward repetitive responding. These results demonstrate distinct ensemble signatures associated with the start versus end of sensory-guided behavior and suggest that M2 leads in engaging goal-directed response strategies that require sensorimotor associations.

Operant behaviors are structured around stimuli, actions and outcomes. Successful execution of a task requires selecting actions that are consistent with the contingencies between these task variables. Control of action selection in the brain should be both stable and flexible. On the one hand, stability allows a subject to sustain high performance to maximize reward. On the other hand, flexibility is essential for quickly adjusting behavior when a change in contingencies occurs. Striking the delicate balance between stability and flexibility is therefore a key requirement of adaptive decision-making. Moreover, a lack of balance between these opposing aspects of cognitive control is a hallmark of psychiatric disorders¹.

How do we know when to be stable or flexible in a changing environment? In tasks without explicit contextual cues, subjects may adjust their response strategy through reward feedback. Prior studies have observed task-dependent differences in neuronal firing rates and selectivity in multiple frontal cortical regions²⁻⁴. During periods of behavioral adjustment, evolution of cortical activity was found to be gradual and late, occurring on time courses that generally match or lag the improvement in task performance⁵⁻⁸. However, neurons in the frontal cortex exhibit substantial cell-to-cell variability in such time courses⁵. Population activity may therefore be more useful for capturing circuit dynamics⁹⁻¹¹. Using ensemble recordings, two studies examined reward-guided adaptations, and they found the corresponding changes in network activity to be surprisingly abrupt^{12,13}. Determining the functional significance of these findings, however, will require quantitative comparisons of ensemble activity transitions that differ in their dynamics. Transitions that are relatively gradual versus abrupt, or that differ in onset with respect

to behavioral changes, could reflect distinct underlying mechanisms for cognitive control.

To study adaptive sensorimotor decision-making in mice, we designed a head-fixed task that required animals to shift many times between three sets of stimulus-response contingencies. This task is a variant of arbitrary sensorimotor mapping, a classic model in which subjects are required to follow conditional rules^{14,15}, such as “for stimulus A, perform one action; for stimulus B, perform another action.” Once learned, the stimulus-response contingencies can then be switched, requiring the learning of novel mappings or retrieval of familiar associations. Associations are made by linking nonspatial stimuli or conditions to actions, and are therefore termed “arbitrary”¹⁶. A number of brain regions are involved in arbitrary sensorimotor mapping, including the frontal lobe, striatum, hippocampus and thalamus¹⁶. Within the frontal lobe, the dorsal premotor cortex has been implicated in the selection of motor programs based on antecedent conditions, as evidenced by the results of lesion studies¹⁷⁻¹⁹, electrophysiology^{5,6}, functional imaging^{20,21} and transcranial stimulation²² in humans and nonhuman primates.

Secondary motor cortex (M2) has been described as a potential rodent homolog of primate higher-order motor areas^{23,24}. Its location, adjacent to the medial prefrontal and primary motor regions, suggests that it may function as a cognitive-motor interface. A long line of research has linked the premotor cortex and neighboring regions to the generation of volitional movements²⁵⁻²⁷. Recent studies in rodents have also focused on the role of M2 in driving motor actions. Random-ratio lever-pressing was shown to become insensitive to reward devaluation in M2-lesioned mice, suggesting a role for M2 in goal-directed actions²⁸. Neural activity is

¹Interdepartmental Neuroscience Program, Yale University, New Haven, Connecticut, USA. ²Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut, USA. ³Department of Neuroscience, Yale University School of Medicine, New Haven, Connecticut, USA. Correspondence should be addressed to A.C.K. (alex.kwan@yale.edu).

Received 22 March; accepted 16 June; published online 11 July 2016; doi:10.1038/nn.4342

modulated before movement, reflecting involvement in action preparation and initiation^{29–31}. Moreover, M2 neurons encode not only current action, but also prior choice and outcome, indicating a broader role in decision-making²⁹. One early study showed that rats with lesions of medial frontal cortex—a broader region that includes M2—had deficits in a visual conditional motor task³².

To elucidate the relationship between frontal ensemble activity and adaptive behavior, we used two-photon calcium imaging to record from M2 neurons in behaving mice. We found distinct population activity patterns associated with each of the three sets of stimulus–response contingencies. Moreover, following a contingency switch, transitions between ensemble patterns occurred earlier and were more abrupt when animals were required to abort repetitive actions and use a conditional rule. In fact, this change in ensemble activity state could be detected after only a few error trials, preceding the more gradual recovery of behavioral performance. Our results uncover distinct neural transitions associated with different phases of voluntary behavior and identify a leading role for M2 in engaging actions that require the use of sensorimotor associations.

RESULTS

An adaptive decision-making task for head-fixed mice

We trained head-fixed mice to perform a task requiring flexible sensorimotor mapping. In each trial, fluid-restricted mice were presented with one of two randomized auditory stimuli, either logarithmic frequency-modulated sweeps from 5 to 15 kHz (‘upsweep’) or from 15 to 5 kHz (‘downsweep’), and had to respond with a lick to the left or right port (Fig. 1a and Supplementary Video 1). A correct response was rewarded with 2 μ L of water, while an incorrect response resulted in white noise. Trials were organized into blocks (Fig. 1b), each with a distinct set of stimulus–response contingencies: ‘sound-guided’ (upsweep–left; downsweep–right), ‘action-left’ (upsweep–left; downsweep–left), and ‘action-right’ (upsweep–right; downsweep–right).

When performance reached a criterion of 85% correct over 20 trials, a new block began with different contingencies. Sound and action blocks alternated, and no contextual cue was given to signal the block transition. Therefore, performance beyond the first block required flexible response selection and outcome monitoring. Mice were prepared for this task by initial training to an expert level on two-choice auditory discrimination; i.e., ~30 d on a task with only sound-guided trials. Here, we present data from mice with fewer than 6 sessions of experience in the adaptive decision-making task.

As expected, a switch in contingencies was associated with an immediate drop in correct response rate (Fig. 1c,d). Most incorrect responses were perseverative errors, indicating a failure to update response strategy for ~20 trials after the switch. We obtained concurrent calcium imaging and behavioral data during 9 sessions from 5 mice (Supplementary Table 1). On average, these mice performed 418 ± 49 trials per session, including 296 ± 38 rewarded trials and 9 ± 1 block switches (mean \pm s.e.m.; range: 6–19 switches; Supplementary Fig. 1a). To quantify motor output, we calculated the mean lick rates and the time of first lick for different trial types. Overall, licks were tightly locked to the time of auditory cue during correct trials (Fig. 1e). For congruent trials (in which stimulus–response contingencies match), lick rates were indistinguishable across sound and action blocks. For incongruent trials (for example, left action for upsweep during sound block versus downsweep during action-left block), there was a noticeable difference in mean lick rates and an increased latency to first lick (Supplementary Fig. 2). Nevertheless, the major determinant for the shape of the lick distribution was response direction: i.e., whether the animal chose left or right (Fig. 1e). Additionally, we used video tracking to monitor whisker and hindpaw positions, and found that their movements also depended mostly on response direction (Supplementary Fig. 3). Therefore, although tongue licks were the means for making operant responses in this head-fixed setup, mice performed more complex motor programs to indicate their choices.

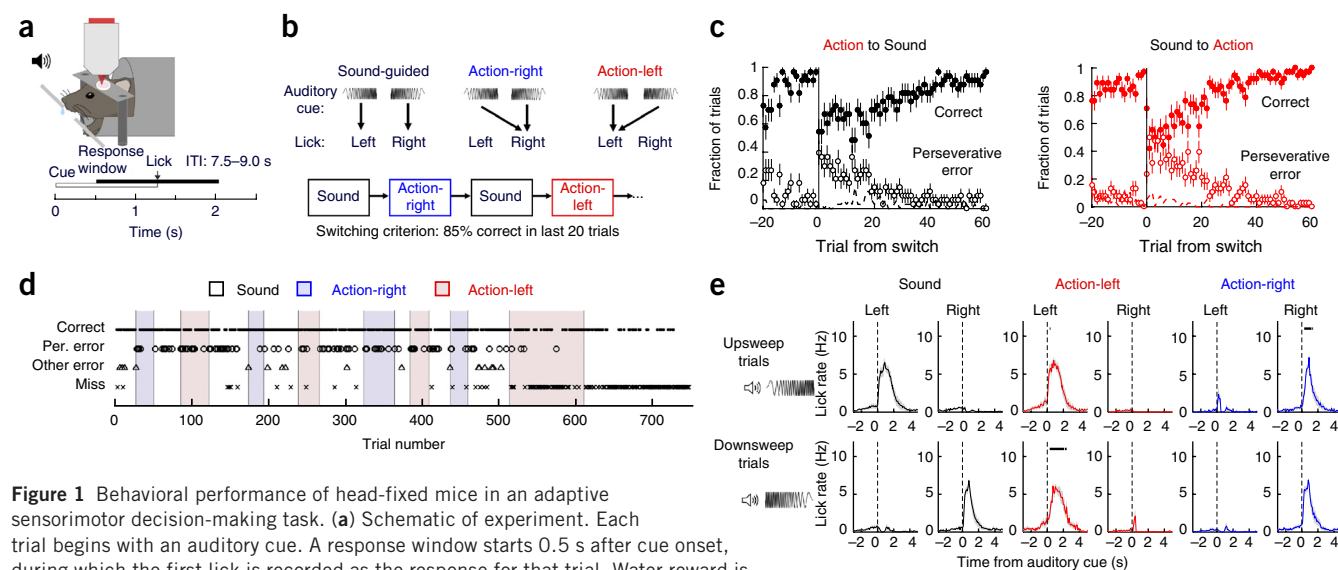


Figure 1 Behavioral performance of head-fixed mice in an adaptive sensorimotor decision-making task. **(a)** Schematic of experiment. Each trial begins with an auditory cue. A response window starts 0.5 s after cue onset, during which the first lick is recorded as the response for that trial. Water reward is delivered contingent on a correct response. ITI, intertrial interval. **(b)** Schematic of stimulus–response mappings and block design. **(c)** Behavioral performance surrounding a block switch from action to sound (left) or sound to action (right). Filled circles, hit rate. Open circles, perseverative error rate. Dotted line, other error rate. Mean \pm s.e.m. of 33 action to sound switches and 38 sound to action switches. **(d)** Performance from one example behavioral session. Trial outcomes: correct (filled circles), perseverative (Per.) error (open circles), other error (open triangles), or miss (cross). Vertical line, block switch. **(e)** Left and right lick rates for upsweep or downsweep sound cues during all correct sound-guided (black), action-left (red) and action-right (blue) trials. For each choice direction, lick rates in action trials were compared with those in sound trials in 0.1 s bins; black bars, significant differences ($P < 0.01$, paired t -test). Lines, mean; shading, \pm s.e.m. $n = 9$ sessions from 5 mice.

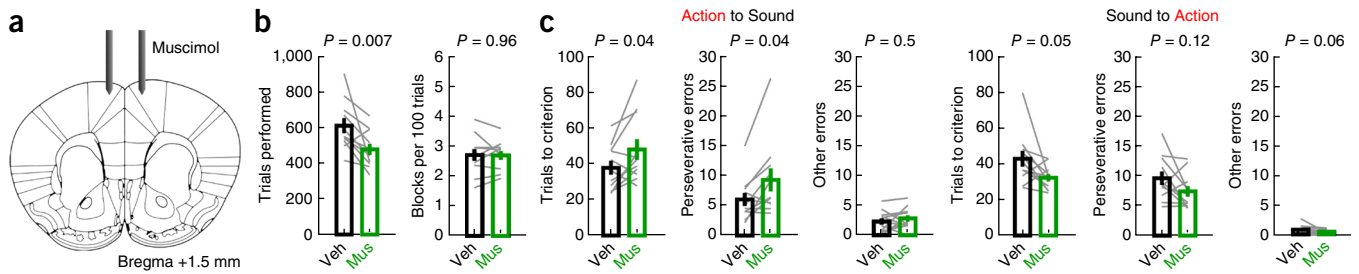


Figure 2 Bilateral inactivation of secondary motor cortex impairs adjustment to sound-guided trials. **(a)** Schematic of experiment. **(b)** Task performance after bilateral infusion of saline vehicle (Veh) or muscimol (Mus) into M2. **(c)** Effects of muscimol infusion on action-to-sound and sound-to-action block shifts. Gray lines, individual paired experiments. Bars, mean \pm s.e.m. Wilcoxon signed-rank test, action-to-sound: trials to criterion, $P = 0.042$, $W = 10$; perseverative errors, $P = 0.042$, $W = 10$; other errors, $P = 0.5$, $W = 24$. Sound-to-action: trials to criterion, $P = 0.054$, $W = 55$; perseverative errors, $P = 0.12$, $W = 51$; other errors, $P = 0.06$, $W = 46$. $n = 11$ mice.

Silencing M2 selectively impairs shift to sound-guided actions

To determine whether frontal cortical activity is necessary for adaptive decision-making in our task, we used the GABA_A receptor agonist muscimol to inactivate M2 bilaterally. Muscimol (5 mM, 46 nL per hemisphere) or saline vehicle was injected \sim 1 h before behavioral testing ($n = 11$ mice; **Fig. 2a** and **Supplementary Table 1**). We injected low-molecular-weight fluorescein to estimate the extent of the affected region, which included M2 and part of cingulate cortex (Cg1), but not other neighboring regions (**Supplementary Fig. 4**). Compared to controls, muscimol-injected mice performed fewer trials (**Fig. 2b**; saline: 608 ± 42 , muscimol: 476 ± 31 , mean \pm s.e.m.; $P = 0.007$, $W = 62$, Wilcoxon signed-rank test), although there was no difference in the number of switches per 100 trials (saline: 2.7 ± 0.2 , muscimol: 2.7 ± 0.1 , mean \pm s.e.m.; $P = 0.96$, $W = 34$). Notably, separate analyses of sound and action blocks revealed selective impairments in the animals' ability to engage sound-guided actions, evidenced by a marked (55%) increase in the number of perseverative errors per block (**Fig. 2c**; saline: 5.7 ± 1.1 , muscimol: 8.9 ± 1.9 , mean \pm s.e.m.; $P = 0.042$, $W = 10$, Wilcoxon signed-rank test), and a greater number of trials to reach criterion (saline: 38 ± 4 , muscimol: 48 ± 6 , mean \pm s.e.m.; $P = 0.042$, $W = 10$). Nevertheless, muscimol-injected mice eventually reached the criterion of $>85\%$ correct, indicating that the transition to a high level of performance was slowed, but not blocked, by M2 inactivation. Silencing had the opposite effect on shifts into action blocks, during which the mice required fewer trials to reach criterion, although this

effect fell short of statistical significance (saline: 43 ± 4 , muscimol: 32 ± 2 , mean \pm s.e.m.; $P = 0.054$, $W = 55$). Inactivation had no effect on the timing or rates of lick motor output (**Supplementary Fig. 5**). These results indicate a causal role for M2 in the flexible control of action selection. Additionally, the opposing effects of silencing are useful for understanding how the mouse performs the outlined task. One solution to the task would be to forget and relearn the relevant stimulus-response associations after each contingency change, as in a reversal task³³. This approach predicts symmetric changes in behavior following perturbations. An alternative approach would be to rely on these associations for sound-guided trials and then ignore them during action blocks to favor repeated selection of the same response. In this case, the mouse would perform the task by shifting the balance between conditional and nonconditional means of responding. The asymmetric deficits observed in our experiments are consistent with the second approach and implicate M2 in the breaking of repetitive actions and biasing choices based on learned associations.

Imaging task-related activity at cellular resolution in M2

To characterize neural activity, we injected adeno-associated viruses encoding GCaMP6s into layer 2/3 of M2 (AAV1-Syn-GCaMP6s-WPRE-SV40; **Fig. 3a**). GCaMP6s is a genetically encoded calcium indicator that exhibits a $\sim 25\%$ rise in fluorescence intensity per action potential in cortical pyramidal neurons³⁴. While mice performed the adaptive decision-making task, we used two-photon microscopy to

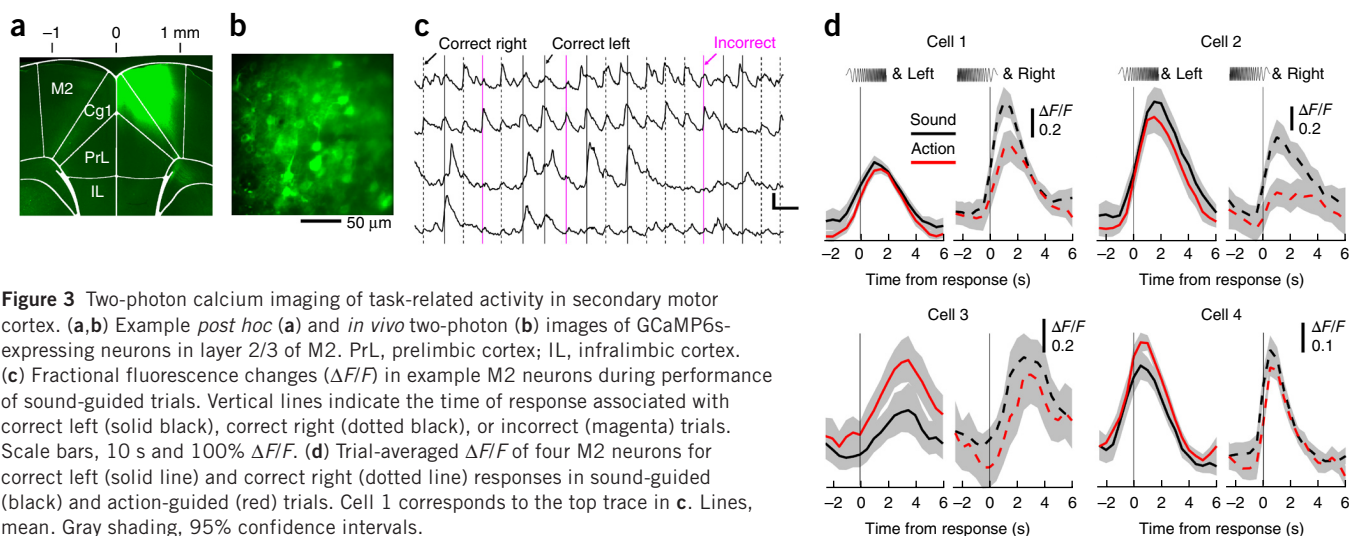


Figure 3 Two-photon calcium imaging of task-related activity in secondary motor cortex. **(a,b)** Example *post hoc* **(a)** and *in vivo* two-photon **(b)** images of GCaMP6s-expressing neurons in layer 2/3 of M2. PrL, prelimbic cortex; IL, infralimbic cortex. **(c)** Fractional fluorescence changes ($\Delta F/F$) in example M2 neurons during performance of sound-guided trials. Vertical lines indicate the time of response associated with correct left (solid black), correct right (dotted black), or incorrect (magenta) trials. Scale bars, 10 s and 100% $\Delta F/F$. **(d)** Trial-averaged $\Delta F/F$ of four M2 neurons for correct left (solid line) and correct right (dotted line) responses in sound-guided (black) and action-guided (red) trials. Cell 1 corresponds to the top trace in **c**. Lines, mean. Gray shading, 95% confidence intervals.

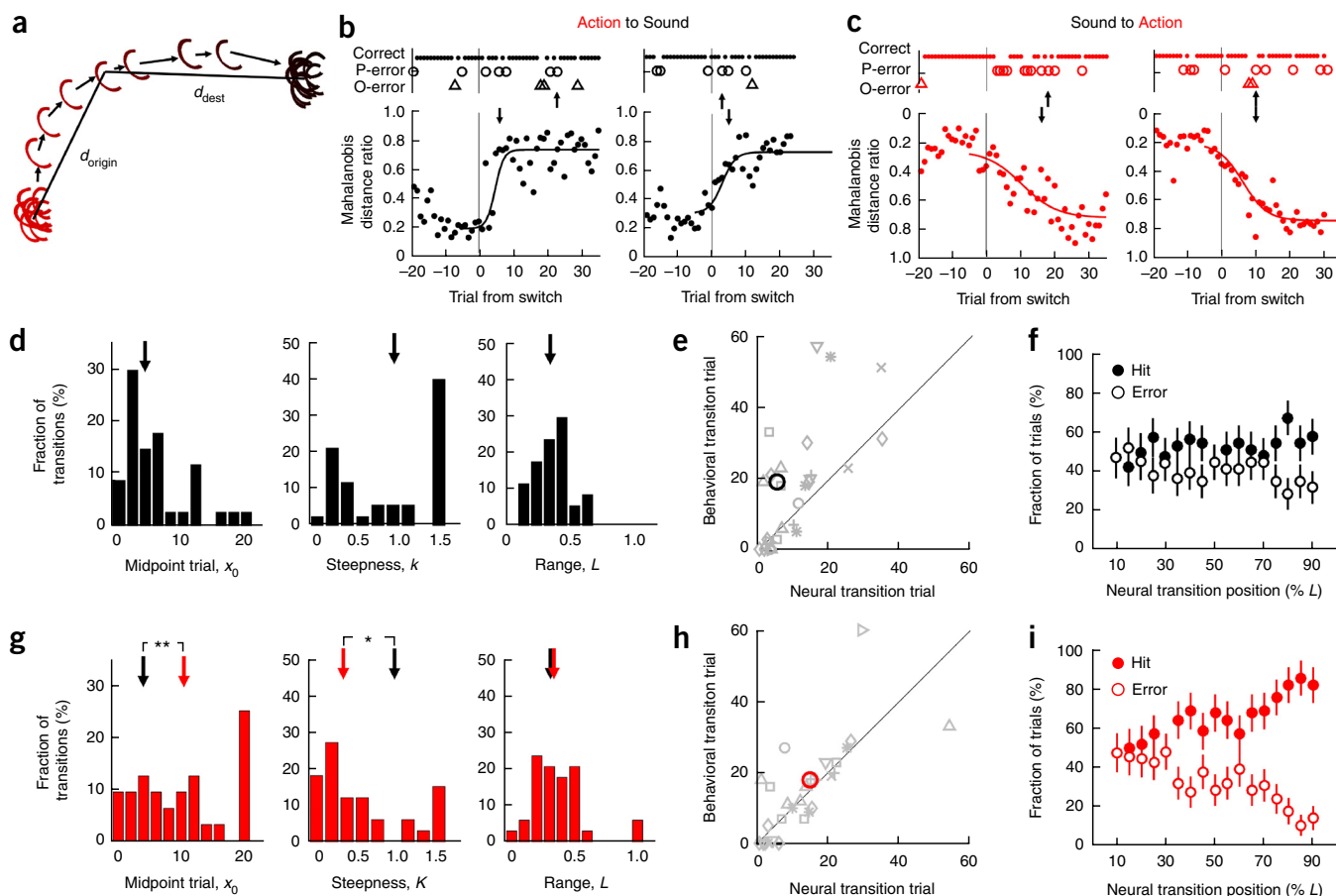


Figure 4 Transitions in ensemble activity occur earlier and are more abrupt following switch to sound-guided trials. (a) Schematic illustrating ensemble activity dynamics around a block switch. Each curved line represents a single-trial neural trajectory deduced from calcium imaging data. When contingencies switch, neural trajectories move within the representational space. Trial-by-trial location of ensemble activity patterns was determined by calculating a ratio of Mahalanobis distances, $d_{\text{origin}}/(d_{\text{origin}} + d_{\text{dest}})$, where d_{origin} and d_{dest} are the Mahalanobis distance from neural trajectory in the current trial to those of the 20 trials pre-switch for the last and current blocks, respectively. (b) Two switches from action to sound blocks. Top, trial outcomes: correct (filled circles), perseverative error (P-error, open circles), and other error (O-error, open triangles). Bottom, trial-by-trial location of ensemble activity patterns. Dots, Mahalanobis distance ratios for individual trials. Lines, fit to logistic functions. Up arrows, behavioral transition trials. Down arrows, neural transition trials. (c) Same as b for two switches from sound to action blocks. Vertical axis is inverted for presentation purposes. (d) Parameters extracted by fitting action-to-sound neural transitions with the logistic function. Arrows, median values. (e) Neural transition trials plotted against behavioral transition trials (Online Methods) for action-to-sound switches. Each symbol represents one block switch. Symbol shapes denote different sessions (Supplementary Fig. 6). Large circle, median value. (f) Mean hit and error rates at the behavioral trial corresponding to specific neural transition locations as estimated by the logistic fit for each action-to-sound transition. Circles, mean \pm s.e.m. (g–i) Same as d–f for switches from sound to action block; red arrows, median values; black arrows in g shown for comparison with c. Wilcoxon rank-sum test: for midpoint trial, x_0 , $P = 0.007$, $z = 2.70$; for steepness, k , $P = 0.03$, $z = -2.17$; $*P < 0.05$; $**P < 0.01$. Difference in range (L) was not significant (median L , sound: 0.36, action: 0.38; $P = 0.8$, $z = 0.28$). Rightmost bar of the histogram includes all instances above the range. $n = 33$ action-to-sound and 35 sound-to-action switches from 9 sessions from 5 mice.

record from 62 ± 6 cells per field of view (mean \pm s.e.m.; range, 26–83 cells; $n = 9$ sessions from 5 mice; Fig. 3b). Figure 3c shows four example M2 neurons with fluorescence transients ($\Delta F/F$) concurrent with responses during sound-guided trials. To examine how the use of conditional rules affects the activity of individual neurons, we averaged $\Delta F/F$ across correct trials for the congruent upswamp–left and downswamp–right conditions separately for sound and action blocks. Neural responses were diverse, even for neurons within the same field of view (Fig. 3d). During sound-guided trials, neurons could exhibit higher $\Delta F/F$ for specific associations—i.e., upswamp–left (cell 2) or downswamp–right (cells 1 and 3)—or have no preference (cell 4). The use of conditional rules modulated $\Delta F/F$ in some neurons (cells 1, 2 and 3) and in other cases had no effect (cell 4).

Neural transition is more rapid during shift to sound rule

The observed heterogeneity of neural responses opened the question of whether single-neuron activity in M2 reflects the components of an ensemble representation for specific task variables. If so, then population-level analyses might more effectively capture the content of such representations. Toward this end, we calculated population activity vectors from $\Delta F/F$ and used demixed principal component analysis (dPCA)^{35,36} to project the vectors in a reduced representational space (Online Methods). Plotting these vectors over time generates trajectories describing the time-dependent evolution of ensemble activity during behavior. To determine how the ensemble activity evolved around block switches on a trial-by-trial basis, we calculated the Mahalanobis distances between population activity vectors of each

trial and those of the 20 trials before the last or next block switch. Following a contingency switch, we found that the ensemble activity migrated away from the previous representational subspace toward a new subspace associated with the new rule (Fig. 4a). Comparisons of the transition dynamics following a switch into conditional versus nonconditional rules uncovered marked differences. Out of 33 action-to-sound and 38 sound-to-action transitions, 33 and 35 switches, respectively, could be fit with a logistic function to compare the onset and rate of shifts in population activity patterns (Fig. 4b,c). State transitions associated with the shift to sound-guided responses occurred after only several trials, much earlier than with shifts into repeated actions (Fig. 4d,g; sound: midpoint trial (x_0) = 4.0, action: x_0 = 10.4, median; P = 0.007, z = 2.70, Wilcoxon rank-sum test). Furthermore, breaking from repetitive to sound-guided responding involved transitions that were more abrupt (sound: steepness (k) = 1.02, action: k = 0.35, median; P = 0.03, z = -2.17, Wilcoxon rank-sum test). These differences in neural dynamics were not due to behavioral differences, because in this set of experiments trials to criterion were similar for the two rule types (sound: 39, action: 38, median; P = 0.9, z = 0.09, Wilcoxon rank-sum test; Supplementary Fig. 1a). Overall, these results suggest that ensemble activity patterns in M2 shift earlier and more steeply when animals are required to abort repetitive actions and engage conditional associations to perform sound-guided behavior.

To what extent must population activity resemble the final ensemble state in order to improve behavior? To address this question, we performed two analyses to compare the timing of neural and behavioral transitions. In the first analysis, we defined 'transition trials' for behavior (trials to criterion minus 20, the sliding window

for assessing criterion) and neural ensemble activity (Mahalanobis distance ratio equaling 75% L based on logistic fit; Online Methods). Block-by-block paired comparisons of neural and behavioral transition trials showed that ensemble activity in M2 shifted before the recovery of behavioral performance when adapting to conditional rules (Fig. 4e and Supplementary Fig. 6; P = 0.003, z = -2.96; Wilcoxon signed-rank test). By contrast, neural and behavioral changes occurred at around the same time for shifts to nonconditional responding (Fig. 4h; P = 0.19, z = 1.32; Wilcoxon signed-rank test). We should note, however, that the definitions used for transition trials were arbitrary. Therefore, we performed a second, less biased analysis in which we determined the mean performance at the behavioral trial corresponding to a series of different neural transition locations. Compared with shifts to action trials (Fig. 4i), transitions to sound-guided trials were associated with hit and error rates that diverged later (Fig. 4f), indicating that behavioral improvement occurred later along the time course of neural transitions. Taken together, these two analyses suggest that when shifting to sound-guided actions, neural ensemble transitions in M2 are nearly complete before behavioral performance improvement can be detected.

Distinct activity patterns accompany rule implementations

Our results indicated that rule shifts are associated with distinct transitions in network activity. This leads naturally to the question of what ensemble dynamics accompany successful rule implementation. We examined trajectories associated with correct responses in the 20 trials before the switch, when response strategies had stabilized (>85% correct by task design). Figure 5a shows the trajectories

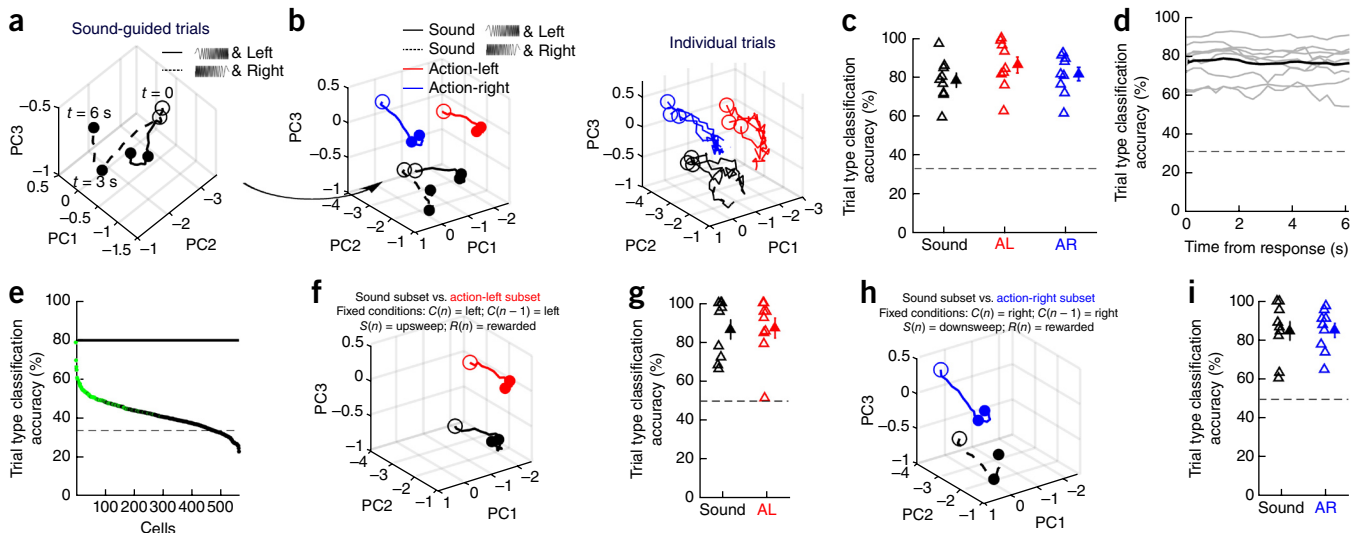


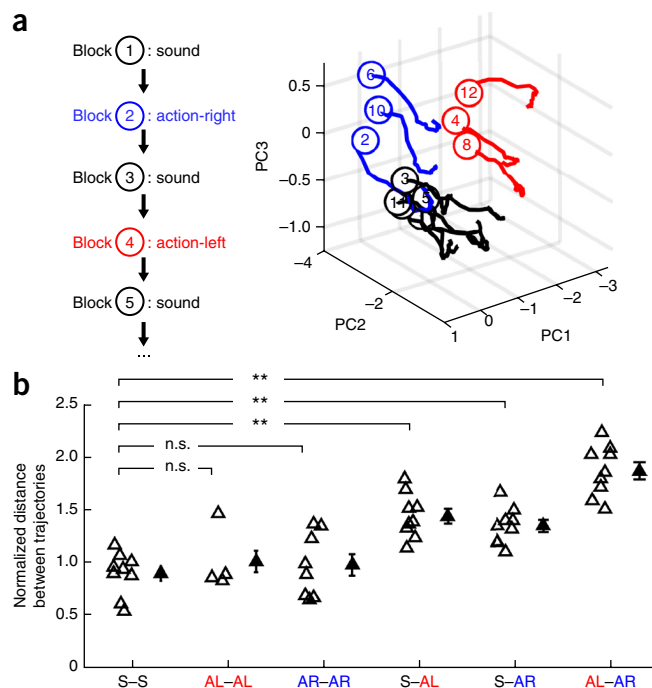
Figure 5 Multiple strategies are associated with distinct population activity patterns. (a) Neuronal circuit trajectories for 56 simultaneously imaged cells in one experiment. Trajectories determined from trial-averaged $\Delta F/F$ for 44 correct left (solid line) and 59 correct right (dotted line) responses in sound-guided trials. Open circles, time (t) of response. Filled circles, 3 and 6 s after response. PC, principal component. (b) Same axes as a, with additional trajectories from 51 correct action-left (red) and 51 correct action-right (blue) trials. Left, trajectories calculated using trial-averaged $\Delta F/F$. Right, 3 representative single-trial trajectories from each trial type. (c) Median accuracy of decoding trial type from individual population activity vectors. S, sound; AL, action-left; AR, action-right. Open triangles, individual experiments; filled triangles, mean \pm s.e.m.; dotted line, chance-level accuracy. (d) Temporal dependence of ensemble decoding accuracy. Gray, individual experiments. Black, mean. Dotted line, chance-level accuracy. (e) Median accuracy of decoding trial type from fluorescence transients of single neurons. Green circles, trial-type-selective cells; i.e., 95% percentile confidence intervals are above chance level of 33.3%. Black circles, other cells. Black line, mean decoding accuracy using ensemble activity. Dotted line, chance-level accuracy. (f,g) Same as b,c for trials in which stimulus was upswing, choice was left, outcome was reward for the current trial and choice was left for the prior trial. Trial type could be decoded well above chance (sound: $87 \pm 5\%$, action-left: $88 \pm 5\%$; versus chance level of 50%, $P = 7 \times 10^{-5}$, 7×10^{-5} , $t_8 = 7.50$, 7.52 , one-sample t -test). (h,i) Same as b,c for trials in which stimulus was downswing, choice was right, outcome was reward for the current trial and choice was right for the prior trial. Trial type could be decoded well above chance (sound: $85 \pm 5\%$, action-right: $86 \pm 4\%$; vs. chance level of 50%, $P = 7 \times 10^{-5}$, 8×10^{-6} , $t_8 = 7.54$, 10.01 , one-sample t -test). $n = 9$ sessions from 5 mice.

Figure 6 M2 ensembles revisit previous activity patterns upon re-exposure to corresponding trial type. **(a)** Neural circuit trajectories calculated from trial-averaged $\Delta F/F$ for each trial block during one behavioral session. Open circles, time of response. Black, sound-guided. Blue, action-right. Red, action-left. **(b)** Normalized distance between neural circuit trajectories from different trial types across all experiments (Online Methods) for sound (S), action-left (AL) and action-right (AR). Open triangles, median distances from individual experiments. Solid triangles, mean \pm s.e.m. Wilcoxon signed-rank test: S–S vs. AL–AL: $P = 0.6$, $W = 3$; S–S vs. AR–AR: $P = 0.5$, $W = 13$; S–S vs. S–AL: $P = 0.004$, $W = 0$; S–S vs. S–AR: $P = 0.004$, $W = 0$; S–S vs. AL–AR: $P = 0.004$, $W = 0$; $**P < 0.01$. $n = 9$ sessions from 5 mice, except for AL–AL ($n = 4$) and AR–AR ($n = 8$) because mice did not perform enough switches to experience the same block type again in some sessions. $**P < 0.01$; n.s., not significant.

of a 56-cell ensemble for left and right responses during sound-guided trials. The trajectories were initially indistinguishable and then diverged sharply after the animal made a response. Expanding this analysis to include action blocks revealed population activity patterns that occupied additional, distinct subspaces within the same representational space (Fig. 5b). To quantify rule representations present in the population code, we asked how accurately block type could be predicted from individual population activity vectors. For each session, we constructed a classifier based on linear discriminant analysis (Online Methods). Testing each classifier with fivefold cross-validation revealed that in all cases trial type could be decoded well above chance (Fig. 5c; sound: $78 \pm 3\%$, action-left: $86 \pm 4\%$, action-right: $82 \pm 3\%$; versus chance level of 33%, $P = 1 \times 10^{-6}$, 1×10^{-6} , 5×10^{-7} ; $t_8 = 13.2$, 12.8, 14.3; one-sample t -test; $n = 9$ sessions). Repetition of this analysis using a moving window yielded high decoding accuracy at all times during a trial (Fig. 5d), consistent with a global shift in engagement of the network rather than a simple change in the processing of cue, action or outcome related signals. Next we asked whether accuracy of the ensemble classifier could have been driven by a few highly rule-selective cells. When classifiers were trained on $\Delta F/F$ of individual cells, we found that 27% of the cells could be used to decode block types at rates above chance; however, accuracies fell along a continuum and at levels below the accuracy of the ensemble (Fig. 5e). To ensure that the differences in trajectories and decoding accuracies were not due to simple sensory or motor parameters, we computed trajectories with matched stimulus, prior choice, current choice and outcome conditions. Analyses of these congruent trials, which differed only by rule, yielded similar results (Fig. 5f–i and Supplementary Fig. 7a–d). Taken together, these results indicate that the behavioral implementation of specific conditional and nonconditional rules is associated with distinct network activity patterns in M2, such that population activity from any time during behavior can be used to decode task contingencies with high accuracy.

Activity patterns toggle between rule-related configurations

When animals solve trials with the same contingencies a second time, do M2 ensembles revisit similar activity patterns or does population activity migrate to a previously uncharted region of state space? Our task was well suited to address this question because blocks of the same trial type were presented multiple times within the same behavioral session. Figure 6a shows an example set of neural circuit trajectories for the first 12 trial blocks within one behavioral session, in which trajectories could be clearly grouped by block type and not by their temporal order. To quantify the representational similarity of ensemble dynamics on a block-by-block basis, we calculated the



mean Euclidean distances between all possible pairwise comparisons of trajectories within an experiment (Online Methods). We found that neural circuit trajectories from blocks of the same type had a relatively small distance of separation and were similarly compact (Fig. 6b and Supplementary Fig. 7e,f; for sound (S), action-left (AL) and action-right (AR): S–S versus AL–AL: $P = 0.6$, $W = 3$; S–S versus AR–AR, $P = 0.5$, $W = 13$; Wilcoxon signed-rank test). By contrast, trajectories from different block types were represented by markedly different ensemble activity (S–S versus S–AL, $P = 0.004$, $W = 0$; S–S versus S–AR, $P = 0.004$, $W = 0$; S–S versus AL–AR, $P = 0.004$, $W = 0$; corrected $\alpha = 0.01$, Wilcoxon signed-rank test with Bonferroni correction). These results indicate that, during adaptive decision-making, M2 toggles between distinct functional configurations as the animal repeatedly engages corresponding changes in task demands.

Comparison of task-related neural dynamics in M2, ALM and V1

Next we sought to determine whether the observed neural dynamics are specific to M2 or may be found in other brain regions also. For this purpose, we imaged neural ensembles in layer 2/3 of anterior lateral motor cortex (ALM; 65 ± 6 cells per field of view, mean \pm s.e.m.; 8 sessions from 4 mice; Supplementary Fig. 1b) and primary visual cortex (V1; 57 ± 7 cells per field of view; 4 sessions from 2 mice; Supplementary Fig. 1c) to compare with the data from M2 (62 ± 6 cells per field of view; 9 sessions from 5 mice). ALM has been implicated in motor planning and execution^{37,38}; however, it is ~ 1.5 mm distant from M2, and the relationship between the two frontal cortical regions is not understood. V1 was chosen as a control region because the task was performed in the dark and involved no visual stimulus. Multiple linear regression analysis showed that M2 neurons robustly encoded not only the choice of the current trial, but also the choices of the two prior trials (Fig. 7a). By contrast, a higher proportion of cells in ALM encoded the current choice; however, the signals decayed faster, resulting in weaker encoding of prior choices (Fig. 7b). Activity of M2 and ALM neurons could prefer either the ipsilateral or contralateral direction (Supplementary Fig. 8a,b), consistent with prior studies^{30,38}. Unexpectedly, choice signals were also

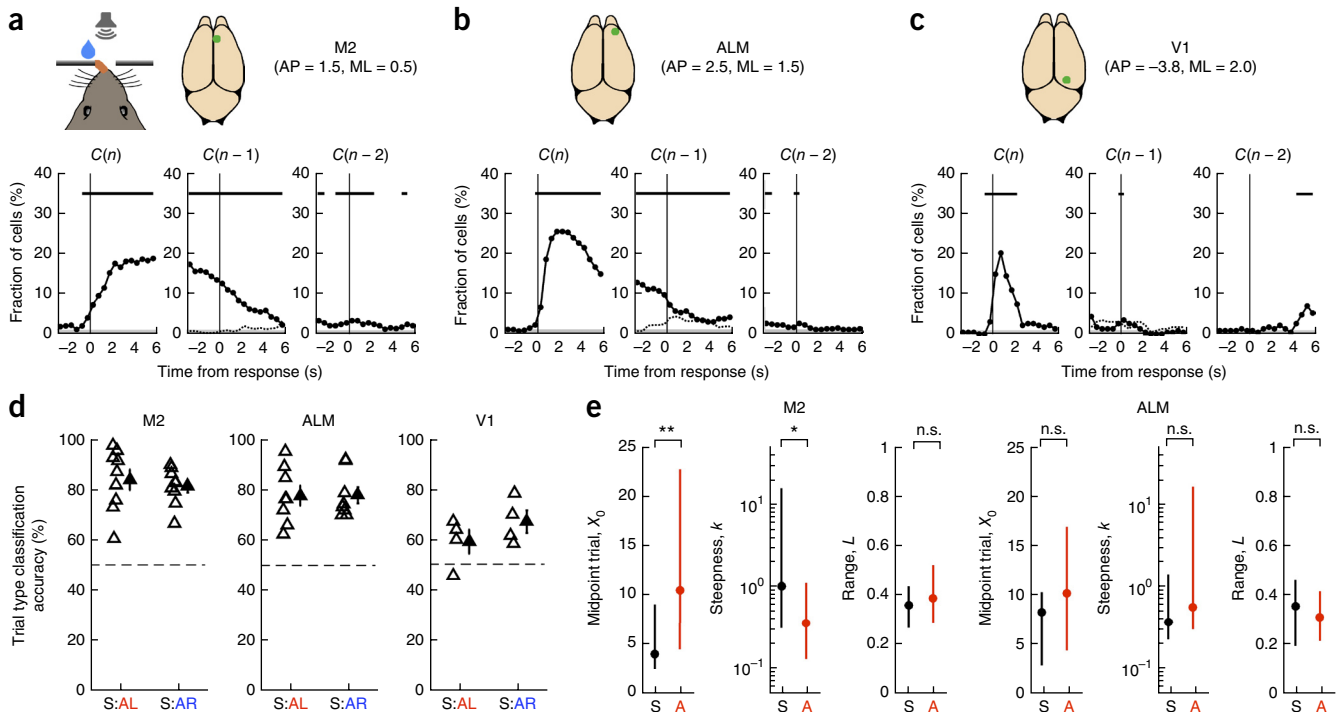


Figure 7 Comparison between neural activity patterns in M2, ALM, and V1 during flexible sensorimotor behavior. **(a)** Multiple linear regression analysis was used to evaluate the fraction of 562 M2 neurons encoding choice signals as a function of time. Regression was performed with a moving window (duration = 0.5 s, step = 0.5 s) to test for significance with $\alpha = 0.01$ on all sound-guided trials where current and prior outcomes were hits (i.e., $R(n) = 1$ and $R(n-1) = 1$). Black bars, significant fractions ($P < 0.01$, binomial test). Gray shading, significance level of 0.01. Dotted line in the middle panel, fraction of cells significant for the interaction term $C(n) \times C(n-1)$. $n = 9$ sessions from 5 mice. **(b)** Same as **a** for 518 ALM neurons. $n = 8$ sessions from 4 mice. **(c)** Same as **a** for 227 V1 neurons. $n = 4$ sessions from 2 mice. **(d)** Median accuracy of decoding trial type from individual population activity vectors restricted to matched trials, comparing M2, ALM and V1 ensembles. For S:AL, sound and action-left trial types were decoded from trials where stimulus was upsweep, choice was left, outcome was reward for the current trial and choice was left for the prior trial. For S:AR, sound and action-right trial types were decoded from trials where stimulus was downsweep, choice was right, outcome was reward for the current trial and choice was right for the prior trial. Trial type was decoded with high accuracy using ensemble activity from M2 (matched sound: action-left trials: $84 \pm 4\%$, $t_8 = 8.37$; $P = 3 \times 10^{-5}$; matched sound: action-right trials: $81 \pm 2\%$, $t_8 = 12.73$; $P = 1 \times 10^{-6}$; versus chance level of 50%, one-sample t -test). Open triangles, individual experiments. Filled triangles, mean \pm s.e.m. Dotted line, chance-level accuracy. **(e)** Neural transition parameters obtained by fitting action-to-sound (black) and sound-to-action (red) transitions with the logistic function, comparing M2 and ALM ensembles. Wilcoxon rank-sum test, sound vs. action, M2: midpoint trial (x_0), $P = 0.007$, $z = 2.70$; k , $P = 0.03$, $z = -2.17$; height (L), $P = 0.8$, $z = 0.28$; ALM: x_0 , $P = 0.11$, $z = 1.62$; k , $P = 0.4$, $z = 0.89$; L , $P = 0.51$, $z = -0.65$. * $P < 0.05$; ** $P < 0.01$; n.s., not significant. Filled circles, medians. Lines end at 25th and 75th percentiles.

observed in V1 (**Fig. 7c**). Choice selectivity in V1 was relatively weak, and $\Delta F/F$ was almost always higher when animals made an ipsilateral choice (**Supplementary Fig. 8c**). Because choice signals in V1 were transient and animals performed the task in the dark, we conjecture that the selectivity might relate to corollary discharge. To investigate ensemble activity, we employed the same dPCA and linear classifier analyses used for M2. We found that rule type could be decoded with high accuracy using ensemble activity from ALM (matched sound: action-left trials: $78 \pm 4\%$, $t_7 = 6.94$; $P = 2 \times 10^{-4}$; matched sound: action-right trials: $78 \pm 3\%$, $t_7 = 8.68$; $P = 5 \times 10^{-5}$; versus chance level of 50%, one-sample t -test; **Fig. 7d**), but at a much worse rate for V1 (matched sound: action-left trials: $58 \pm 5\%$, $t_3 = 1.88$; $P = 0.2$; matched sound: action-right trials: $67 \pm 4\%$, $t_3 = 3.76$; $P = 0.03$). Therefore, both ALM and M2 exhibited task-specific ensemble activity patterns. However, unlike what we found in M2, characterization of ensemble transitions in ALM did not reveal significant differences between switches to sound versus action blocks (sound: $x_0 = 8.2$, action: $x_0 = 10.2$, median; $z = 1.62$, $P = 0.11$; sound: $k = 0.37$, action: $k = 0.56$, median; $z = 0.89$, $P = 0.4$; Wilcoxon rank-sum test; **Fig. 7e**). There were also no detectable timing differences between neural and behavioral transitions in ALM (sound: $P = 0.5$, $z = 0.61$; action: $P = 0.13$,

$z = -1.50$; Wilcoxon signed-rank test; neural transition defined as 75% L). Taken together, these data indicate regionally specific ensemble dynamics associated with adaptive behavior.

DISCUSSION

The results support two insights regarding the function of higher-order motor cortex in adaptive choice behavior. First, fast and slow ensemble transitions are neural signatures for distinct phases of voluntary behavior. Comparison between transitions was possible because our task design allowed for multiple shifts between multiple contingencies within a single behavioral session. Second, the relative timing of neural and behavioral shifts, as well as the specific deficits following inactivation, highlighted a leading role for this region in the engagement of sensory cue-guided actions (**Supplementary Fig. 9**). This conclusion contrasts with previous studies of homologous or nearby prefrontal cortical regions, in which neural changes closely match or lag the time course of behavioral adaptation^{5,7,12}. A key difference is that prior studies have focused on the learning of novel sensorimotor mappings or new rules, whereas our task required animals to repeatedly disengage and re-engage the learned associations required for sound-guided trials. Likewise, although our task

shares important features with other assays for flexibility^{7,9,10,12,39–41}, there are also crucial differences (Online Methods).

We found that bilateral inactivation of M2 selectively impaired the shift into sound-guided actions. This observation is highly consistent with results of dorsal premotor lesions in primates, which disrupt both the learning of novel visuomotor associations and the engagement of previously learned mappings^{17–19}. Adaptation to action blocks was facilitated by M2 inactivation. This effect could result from a tendency to repeat the prior choice²⁹: if M2 normally biases animals toward sensory-cue-guided actions, then inactivation may remove an important brake on the nonconditional strategy. In our experiments, M2 inactivation slowed but did not preclude the eventual transition to high performance on sound-guided trials. This suggests that, at least for trained mice, two-choice auditory discrimination alone does not require M2 and may be subserved by other circuits⁴². Furthermore, we take the opposing effects of inactivation on shifts to sound-guided versus repeated actions as evidence that mice perform the task by balancing the use of conditional and nonconditional responses.

A key finding of this study concerns how specific parameters of ensemble activity transitions may relate to behavior. We found that ensemble transitions were more abrupt when animals needed to retrieve and begin using conditional associations. These fast transitions may be related to those observed in medial prefrontal cortex, which have been interpreted as neural correlates of insight¹² or as the abandonment of an inadequate internal model at the onset of exploration¹³. On what quantitative basis should transitions be classified as abrupt or gradual? We compared the steepness of these transitions directly to the slower transitions that accompanied adaptation to action blocks. Moreover, ensemble transitions occurred after only a few errors, whereas behavioral improvements took tens of sound-guided trials (Fig. 4e,f). The difference in neural and behavioral timing suggests that M2 neural activity had mostly adjusted while the animal was still systematically responding in a nonconditional manner. M2 may facilitate the engagement of sound-guided behavior by biasing the use of sensory information, suppressing repetitive actions, or both. By contrast, prior studies show that when an animal must acquire novel arbitrary associations, changes in cortical activity track behavioral improvements^{5,43} and lag the more rapid remapping in the striatum⁷. A major difference between these studies of fast learning and our study is that the auditory-motor associations were already well learned in our task.

We found that multiple rules were each associated with a distinct subset of population activity patterns. Such task-dependent changes in neural activity are reported in multiple frontal cortical regions across species^{2–4,12,44}. By asking the animal to shift repeatedly during a single session, we found that the network could return to a previously employed functional configuration to meet similar behavioral demands. This back-and-forth toggling of ensemble activity is reminiscent of the ensemble remapping observed in CA1 of the hippocampus during repeated exposure to spatial contexts^{45,46}. One study reported that changes in environmental context also cause network activity shifts in the rodent medial prefrontal cortex. However, the ensemble code is not identical upon re-exposure, potentially owing to a systematic drift over time⁴⁷. The divergent findings of repeatable versus drifting network states in the rodent frontal cortex could reflect regional differences or differences in how frontal areas encode cognitive versus environmental variables.

Several lines of evidence support the idea that the neural dynamics in M2 reflect changes in internal processes (for example, representation of task contingencies or motor planning and preparation) rather than differences in overt physical movements. First, three different

ensemble analyses with matched, congruent trial conditions indicated distinct neural dynamics in sound and action blocks (Fig. 5f–i and Supplementary Fig. 7), despite a lack of observable difference in motor output for the same sets of trials (Fig. 1e and Supplementary Fig. 2). Second, neural signals related to motor execution should be strongest at the time of response. Instead, we found that the rule-specific separation of population activity patterns was substantially above chance at all times across a trial (Fig. 5d). Third, perhaps the strongest evidence is that muscimol inactivation of M2 had no detectable effect on motor output (Supplementary Fig. 5), while clearly affecting behavioral flexibility.

What is the purpose of functional reconfiguration during adaptive decision-making? Ensemble activity patterns within multiple network subspaces reflect the diversity of neural representations in M2. Recent studies indicate functional roles for rodent M2 long-range projections to sensory cortices^{48,49} and dorsal striatum⁵⁰. Appropriate shifts in neural representations could allow M2 to exert differential top-down control in a task-dependent manner. Further study regarding the downstream impacts of frontal network transitions may yield insights into neuropsychiatric disorders in which cognitive flexibility is impaired. Plausibly, the cognitive rigidity characteristic of disorders such as schizophrenia could result from an inability of frontal cortical networks to shift or maintain stable ensemble states.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank D. Lee, M. Picciotto, and J. Taylor for discussions, and C. Posner and R. Hannibal for assistance with behavioral training. This work was supported by National Institute of Aging center grant P50AG047270 (A.C.K.), National Institute of Mental Health grant R21MH110712 (A.C.K.), NARSAD Young Investigator Award (A.C.K.), National Institutes of Health training grant T32NS041228 (M.J.S.), National Science Foundation Graduate Research Fellowship DGE-1122492 (M.J.S.), and a Brown-Coxe Postdoctoral Fellowship (F.A.).

AUTHOR CONTRIBUTIONS

M.J.S. and A.C.K. conceived the project. M.J.S. performed all experiments. V.P. assisted with mouse surgery and inactivation experiments. F.A. and M.L. assisted with behavioral training and histology. M.J.S. and A.C.K. analyzed the data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

- Griffiths, K.R., Morris, R.W. & Balleine, B.W. Translational studies of goal-directed action as a framework for classifying deficits across psychiatric disorders. *Front. Syst. Neurosci.* **8**, 101 (2014).
- Asaad, W.F., Rainer, G. & Miller, E.K. Task-specific neural activity in the primate prefrontal cortex. *J. Neurophysiol.* **84**, 451–459 (2000).
- Rich, E.L. & Shapiro, M. Rat prefrontal cortical neurons selectively code strategy switches. *J. Neurosci.* **29**, 7208–7219 (2009).
- Rodgers, C.C. & DeWeese, M.R. Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. *Neuron* **82**, 1157–1170 (2014).
- Mitz, A.R., Godschalk, M. & Wise, S.P. Learning-dependent neuronal activity in the premotor cortex: activity during the acquisition of conditional motor associations. *J. Neurosci.* **11**, 1855–1872 (1991).
- Chen, L.L. & Wise, S.P. Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. *J. Neurophysiol.* **73**, 1101–1121 (1995).

7. Pasupathy, A. & Miller, E.K. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**, 873–876 (2005).
8. Antzoulatos, E.G. & Miller, E.K. Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron* **71**, 243–249 (2011).
9. Mante, V., Sussillo, D., Shenoy, K.V. & Newsome, W.T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
10. Stokes, M.G. *et al.* Dynamic coding for cognitive control in prefrontal cortex. *Neuron* **78**, 364–375 (2013).
11. Wilson, R.C., Takahashi, Y.K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
12. Durstewitz, D., Vitoz, N.M., Floresco, S.B. & Seamans, J.K. Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* **66**, 438–448 (2010).
13. Karlsson, M.P., Tervo, D.G.R. & Karpova, A.Y. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* **338**, 135–139 (2012).
14. Bunge, S.A. *et al.* Neural circuitry underlying rule use in humans and nonhuman primates. *J. Neurosci.* **25**, 10347–10350 (2005).
15. White, I.M. & Wise, S.P. Rule-dependent neuronal activity in the prefrontal cortex. *Exp. Brain Res.* **126**, 315–335 (1999).
16. Wise, S.P. & Murray, E.A. Arbitrary associations between antecedents and actions. *Trends Neurosci.* **23**, 271–276 (2000).
17. Petrides, M. Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* **23**, 601–614 (1985).
18. Halsband, U. & Passingham, R.E. Premotor cortex and the conditions for movement in monkeys (*Macaca fascicularis*). *Behav. Brain Res.* **18**, 269–277 (1985).
19. Nixon, P.D., McDonald, K.R., Gough, P.M., Alexander, I.H. & Passingham, R.E. Cortico-basal ganglia pathways are essential for the recall of well-established visuomotor associations. *Eur. J. Neurosci.* **20**, 3165–3178 (2004).
20. Toni, I., Ramnani, N., Josephs, O., Ashburner, J. & Passingham, R.E. Learning arbitrary visuomotor associations: temporal dynamic of brain activity. *Neuroimage* **14**, 1048–1057 (2001).
21. Boettiger, C.A. & D'Esposito, M. Frontal networks for learning and executing arbitrary stimulus-response associations. *J. Neurosci.* **25**, 2723–2732 (2005).
22. Rushworth, M.F.S., Hadland, K.A., Paus, T. & Sipila, P.K. Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *J. Neurophysiol.* **87**, 2577–2592 (2002).
23. Murray, E.A., Bussey, T.J. & Wise, S.P. Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Exp. Brain Res.* **133**, 114–129 (2000).
24. Preuss, T.M. Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *J. Cogn. Neurosci.* **7**, 1–24 (1995).
25. Nachev, P., Kennard, C. & Husain, M. Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* **9**, 856–869 (2008).
26. Schall, J.D., Stuphorn, V. & Brown, J.W. Monitoring and control of action by the frontal lobes. *Neuron* **36**, 309–322 (2002).
27. Isoda, M. & Hikosaka, O. Switching from automatic to controlled action by monkey medial frontal cortex. *Nat. Neurosci.* **10**, 240–248 (2007).
28. Gremel, C.M. & Costa, R.M. Premotor cortex is critical for goal-directed actions. *Front. Comput. Neurosci.* **7**, 110 (2013).
29. Sul, J.H., Jo, S., Lee, D. & Jung, M.W. Role of rodent secondary motor cortex in value-based action selection. *Nat. Neurosci.* **14**, 1202–1208 (2011).
30. Erlich, J.C., Bialek, M. & Brody, C.D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72**, 330–343 (2011).
31. Murakami, M., Vicente, M.I., Costa, G.M. & Mainen, Z.F. Neural antecedents of self-initiated actions in secondary motor cortex. *Nat. Neurosci.* **17**, 1574–1582 (2014).
32. Passingham, R.E., Myers, C., Rawlins, N., Lightfoot, V. & Fearn, S. Premotor cortex in the rat. *Behav. Neurosci.* **102**, 101–109 (1988).
33. Fusi, S., Asaad, W.F., Miller, E.K. & Wang, X.-J. A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron* **54**, 319–333 (2007).
34. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
35. Machens, C.K., Romo, R. & Brody, C.D. Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).
36. Brendel, W., Romo, R. & Machens, C.K. Demixed principal component analysis. in *Advances in Neural Information Processing Systems 24* (eds. Shawe-Taylor J., Zemel, R.S., Bartlett, P.L., Pereira, F. & Weinberger, K.Q.) <http://papers.nips.cc/paper/4215-demixed-principal-component-analysis> (2011).
37. Guo, Z.V. *et al.* Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**, 179–194 (2014).
38. Li, N., Chen, T.-W., Guo, Z.V., Gerfen, C.R. & Svoboda, K. A motor cortex circuit for motor planning and movement. *Nature* **519**, 51–56 (2015).
39. Duan, C.A., Erlich, J.C. & Brody, C.D. Requirement of prefrontal and midbrain regions for rapid executive control of behavior in the rat. *Neuron* **86**, 1491–1503 (2015).
40. Asaad, W.F., Rainer, G. & Miller, E.K. Neural activity in the primate prefrontal cortex during associative learning. *Neuron* **21**, 1399–1407 (1998).
41. Darrach, J.M., Stefani, M.R. & Moghaddam, B. Interaction of N-methyl-D-aspartate and group 5 metabotropic glutamate receptors on behavioral flexibility using a novel operant set-shift paradigm. *Behav. Pharmacol.* **19**, 225–234 (2008).
42. Znamenskiy, P. & Zador, A.M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013).
43. Brasted, P.J. & Wise, S.P. Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur. J. Neurosci.* **19**, 721–740 (2004).
44. Wallis, J.D., Anderson, K.C. & Miller, E.K. Single neurons in prefrontal cortex encode abstract rules. *Nature* **411**, 953–956 (2001).
45. Wills, T.J., Lever, C., Cacucci, F., Burgess, N. & O'Keefe, J. Attractor dynamics in the hippocampal representation of the local environment. *Science* **308**, 873–876 (2005).
46. Leutgeb, S. *et al.* Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science* **309**, 619–623 (2005).
47. Hyman, J.M., Ma, L., Balaguer-Ballester, E., Durstewitz, D. & Seamans, J.K. Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proc. Natl. Acad. Sci. USA* **109**, 5086–5091 (2012).
48. Schneider, D.M., Nelson, A. & Mooney, R. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* **513**, 189–194 (2014).
49. Manita, S. *et al.* A top-down cortical circuit for accurate sensory perception. *Neuron* **86**, 1304–1316 (2015).
50. Rothwell, P.E. *et al.* Input- and output-specific regulation of serial order performance by corticostriatal circuits. *Neuron* **88**, 345–356 (2015).

ONLINE METHODS

Animals. We used adult male mice with C57BL/6J genetic background. Mice were housed in groups of 3–5, in 12-h/12-h light-dark cycle (lights off at 19:00), and most experiments were performed in late afternoons and evenings (16:00–midnight). At the start of experiments, mice were P51–117. No statistical tests were used to predetermine sample sizes, but sample sizes for this study are similar to those generally employed in the field. All experimental procedures were approved by the Institutional Animal Care and Use Committee, Yale University.

Surgery. Mice underwent two surgeries. For each surgery, the mouse was anesthetized with 2% isoflurane in oxygen during induction, which was then lowered to 1–1.5% for the remainder of the surgery. The mouse was placed over a water-circulating heating pad (TP-700, Gaymar Stryker) in a stereotaxic frame (David Kopf Instruments). Before the operation, the mouse was injected with carprofen (5 mg/kg, s.c.; #024751, Butler Animal Health) and dexamethasone (3 mg/kg, s.c.; Dexaject SP, #002459, Henry Schein Animal Health). The mouse was injected with carprofen immediately after surgery (5 mg/kg, s.c.) and each day for the following 3 d (5 mg/kg, s.c.). For the first surgery, an incision was made to expose the skull. Based on stereotaxic coordinates, the center location of the mouse secondary motor cortex (M2; AP = 1.5 mm and ML = 0.5 mm relative to bregma) was marked in the right hemisphere. In other experiments, we targeted the anterior-lateral motor cortex³⁷ (ALM; AP = 2.5 mm, ML = 1.5 mm) or the primary visual cortex (V1; AP = –3.8 mm, ML = 2 mm) on the right hemisphere. A stainless steel head plate (eMachineshop) was affixed to the skull with Metabond (C&B, Parkell, Inc.), and a thin layer of clear Metabond was then applied to cover the entire skull. Mice were given at least 1 week to recover before behavioral training. Head-plate-implanted mice were then trained on behavioral tasks (see below). Once a mouse reached a performance criterion of >90% correct rate on three consecutive days and was ready for imaging experiments, a second surgery was performed under anesthesia. Using a dental drill, a 3-mm-diameter craniotomy was made at the targeted location, which had been marked previously and remained visible through the Metabond. Dura was left intact and was irrigated with artificial cerebrospinal fluid (ACSF, in mM: 5 KCl, 5 HEPES, 135 NaCl, 1 MgCl₂, 1.8 CaCl₂; pH 7.3). Using a glass micropipette attached to a microinjection system (Nanoject II, Drummond), 32–46 nL of AAV1-Syn-GCaMP6s-WPRE-SV40 (5 × 10¹³ titer; UPenn Vector Core) was injected at a depth of 400 μm below dura at each of four locations, vertices of a 200-μm-wide square centered at the targeted cortical region. The glass micropipette was left in place for 5 min after injection to reduce backflow. A drop of warmed agar (1.2% in ACSF, Type III-A, High EEO, A9793, Sigma-Aldrich) was then applied to the cortical surface. A two-layer glass window was fabricated by first etching out a 2-mm-diameter circle from #0 thickness glass cover slip, then bonding with UV-activated polymer (61, Norland Optical Adhesive) to a #1 thickness, 3-mm-diameter round glass cover slip (64-0720 CS-3R, Warner Instruments). This glass window was then placed against the cortical surface. While applying light pressure, cyanoacrylate glue was added to the rim to attach the glass to the skull and Metabond. Mice were again given at least 1 week to recover before resuming behavioral training. Imaging experiments began when behavioral performance criterion was reached. Eight of 11 mice went through this procedure involving two surgeries. For the remaining three mice, the head plate implant, viral injection and window implant procedures were performed in the same surgery before behavioral training.

Behavioral setup. For head-fixed mouse behavior, we used a training apparatus that has two lick ports, thus enabling two alternative choices. The use of two lick ports was inspired by another study³⁷. Two metal screws were used to affix the head plate of the mouse onto a stainless steel mount. The mouse was then restrained inside an acrylic tube, which restricted gross body movements but allowed postural adjustments. The lick ports were fabricated from stainless steel 20-gauge needles, which were positioned at 90° and 270° with respect to the mouse's head orientation, and held in place by a 3D-printed plastic part mounted on a micromanipulator for fine positional adjustment. Water was delivered at the ports by gravity feed and the liquid volume was controlled by pneumatic valves (EV-2-24, Clippard) calibrated with an intravenous dripper to deliver ~2 μL per pulse. A battery-operated touch detector circuit signaled when the mouse's tongue contacted a lick port. Auditory stimuli were played through computer speakers placed directly in front of the animal. The intensity of the

auditory stimuli was calibrated to ~85 dB peak amplitude. Water delivery, lick detection and sound presentation were connected to a desktop computer via a data acquisition board (USB-201, Measurement Computing). Presentation software (Neurobehavioral Systems) controlled the entire behavioral system. An infrared webcam was used to monitor the animal while in the rig. Behavioral training was performed inside the closed compartment of an audio-visual cart that was dark and soundproofed with acoustic foams (5692T49, McMaster-Carr). For imaging, mice were tested using a replica of the behavioral training setup under a two-photon microscope.

Adaptive decision-making task. To motivate participation in the task, water consumption was restricted to behavioral sessions. Mice were trained for 1 session per day, 6 d per week. On the non-training day, water was provided *ad libitum* in the home cage for 15 min. The mice were trained through four phases to shape their behavior. Phase one (~2 d): mice were habituated to head fixation in the behavior box and trained to lick either one of the two ports for water reward. Mice were advanced to the next phase when they made >100 responses in a session. Phase two (~2 d): mice were trained to sample both ports. Here mice were required to lick the left port to obtain water rewards three times, followed by the right port for the next three rewards, and so on. Mice were advanced to the next phase when they made >100 correct responses in a session. Phase three (>15 d): animals underwent training for two-choice auditory discrimination. One of two auditory cues was presented to begin each trial: a 2-s-long train of 0.5-s-long logarithmic frequency modulated sweeps from 5 to 15 kHz ('upsweep') or from 15 to 5 kHz ('downsweep'). The stimuli were interleaved randomly from trial to trial. At 0.5 s following the onset of the auditory cue, a response window opened, lasting for a duration of 1.5 s. The animal's first lick within this response window was registered as its response for the trial. All other licks were logged but had no consequences. Once a response was recorded, playback of the auditory cue was terminated. A correct response, i.e., a left lick for upsweep or a right lick for downsweep, resulted in immediate delivery of 2 μL of water from the corresponding port. The next trial would begin 7 s following response. Incorrect responses resulted in 2 s of white noise presentation, with the next trial beginning 5 s later. Each trial had a total duration within a range from 7.5 to 9 s. Animals were allowed to perform trials until satiated (20 consecutive misses), typically after ~60 min. Training continued daily until a correct rate of >90% was attained for 3 consecutive days. For imaging experiments, mice were then trained under the two-photon microscope (with laser turned off) for habituation to the recording setup. All the mice were able to discriminate at >90% correct rate after 1–3 d of re-training. Finally, mice were tested on the adaptive decision-making task. The task always began with a sound block (S) indistinguishable from the two-choice auditory discrimination task. However, once the mouse reached a performance criterion of >85% correct for the last 20 trials, the stimulus-response-outcome contingencies changed from sound- to action-guided trials. In action-guided trials, task structure was identical to sound-guided trials. However, the correct response became fixed to one response direction, for example, always left, regardless of the stimulus identity. No cue signaled the change in contingencies. When the mouse reached performance criterion again, another block switch occurred. A sound block was always followed by an action block and vice versa. The second block was randomly chosen for each experiment to be action-left (AL) or action-right (AR). However, once the first action block was chosen, the block sequence became fixed for the remainder of the session. Therefore, the sequence of blocks could be one of two possibilities: (S-AL-S-AR-S-AL-S-AR-...) or (S-AR-S-AL-S-AR-S-AL-...). Each session was terminated after 20 consecutive misses (trials with no response). Mice typically performed the adaptive decision-making task for 60–90 min. Following each adaptive decision-making test, mice resumed daily two-choice auditory discrimination until the next recording session, up to a maximum of 7 adaptive decision-making tests.

Our behavioral paradigm consists of blocks of trials that require the animal to shift between conditional and non-conditional approaches to action selection. In principle, mice may solve this task by ignoring sensory information completely during action blocks. However, the temporally structured lick rates during action blocks (Fig. 1e) strongly suggest use of the stimulus for gating lick responses. Our task has similarities with other paradigms that test behavioral flexibility, but there are also crucial differences. In contrast to paradigms that use a contextual cue to instruct rapid executive control on a trial-by-trial basis^{9,10,39}, animals in

our experiment adapted on a time scale of tens of trials (Fig. 1c). This relatively slow rate of adaptation is akin to learning during arbitrary visuomotor mapping, where the animal's basis for action selection is updated gradually based on reward feedback^{7,40}. Our task also differs from other strategy- or set-shifting tasks for rodents^{12,41} because nonspatial stimuli were used to probe arbitrary sensorimotor associations that do not conform to classical definitions of exemplars or sets. Furthermore, analysis of the types of errors made during training suggests that mice perform two-choice auditory discrimination in part by suppressing a prepotent tendency to repeat a rewarded choice. Action trials could thus be considered a natural strategy to the animal, whereas sound-guided trials require weeks of training to achieve high performance. Therefore, one caveat for our task is that animals are likely to have different degrees of learned and intrinsic familiarity for sound versus action trials.

Two-photon calcium imaging. The two-photon microscope (Movable Objective Microscope, Sutter Instrument) was controlled using ScanImage software⁵¹. The excitation source was a Ti:Sapphire femtosecond laser (Chameleon Ultra II, Coherent). Excitation intensity was controlled by a Pockels cell (350-80-LA-02, Conoptics) and focused onto the sample with a 20×, N.A. 0.95 water immersion objective (Olympus). The time-averaged excitation laser intensity was 90–100 mW after the objective. To image fluorescence transients from GCaMP6s-expressing neurons, excitation wavelength was set at 920 nm, and emission was collected from 475–550 nm with a GaAsP photomultiplier tube. Time-lapse images were acquired at a resolution of 256 × 256 pixels and a frame rate of 3.62 Hz using bidirectional scanning. To synchronize behavior with imaging, a TTL pulse was sent at the beginning of each trial from the data acquisition board of the behavioral system to the imaging system to act as an external trigger for initiating image acquisition.

Inactivation. Mice were implanted with a head plate. The locations of M2 were marked on both hemispheres (AP = 1.5 mm, ML = 0.5 mm), and then covered with a thin layer of clear Metabond. Mice were then trained as described above, in preparation for the adaptive decision-making test. Craniotomies were performed at the marked locations. Using a glass micropipette attached to a microinjection system (Nanoject II, Drummond), ACSE, with or without muscimol (5 mM, 46 nL per hemisphere; cat. #195336, MP Biomedical), was injected at a depth of 400 μm into M2 of both hemispheres. Behavioral testing began 1–3 h following injection. The same mice were tested after saline and muscimol treatments on consecutive days in a counter-balanced design, with no blinding. The mice were randomized to receive either saline or muscimol first in an alternating manner depending on the order in which they reached the behavioral performance criterion. Twelve mice were allocated for this experiment; however, one was excluded due to equipment malfunction during testing.

Histology. Following experiments, mice were transcardially perfused with chilled formaldehyde solution (4% in phosphate-buffered saline). The brains were sectioned with a vibratome and imaged with an inverted wide-field fluorescence microscope.

Analysis of behavioral data. Timestamps of stimulus presentation, licks and water delivery were logged in a text file by Presentation software (Neurobehavioral Systems, Inc.). Scripts were written in Matlab to parse the log files. For the adaptive decision-making task, a perseverative error was defined as an incorrect response that would have been correct according to the last trial block's contingencies. For example, during an action-left block, the stimulus–response pairings of upsweep–left lick and downsweep–left lick would be 'correct'. Downsweep–right lick would be a 'perseverative error', because this stimulus–response pairing would have been correct in the preceding sound-guided block. The remaining possible stimulus–response pairing, upsweep–right lick, would be classified as an 'other error'. The number of trials performed included all correct and error trials, but excluded the miss trials when the mouse failed to lick within the response window. Miss trials typically occurred near the end of the session when the mouse was satiated. The number of trials to criterion was defined as the number of trials performed in a certain trial block before reaching a performance criterion of 85% correct for the last 20 trials. Therefore, the minimum value of this quantity is 20. Mean trials to criterion for each session was calculated excluding the first sound block, because contingency switches had not yet begun. Mean blocks per

100 trials, mean perseverative errors per block and mean other errors per block were calculated excluding the last block (i.e., trials after the last block switch). For analysis, we often compared conditions before and after block switches, which were defined as the 20 trials before or following a block switch. The first lick time was defined as the time of the first lick after sound onset for each trial, even if this occurred before the start of the response window. The first lick time is thus a sum of the reaction time and movement time. For this measurement, we excluded trials in which the mouse licked within 0.5 s before cue onset, in which case the first lick may represent the continuation of a spontaneous lick bout rather than a reaction to the stimulus.

Analysis of imaging data. Time-lapse fluorescence images were corrected for x - y motion using the TurboReg plug-in for ImageJ (NIH). We wrote a GUI in Matlab to select cell bodies as regions of interest (ROIs). Values of pixels within an ROI were averaged to generate $F_C(t)$. For each cell, we estimated the neuropil signal by drawing a doughnut⁵², by approximating the ROI area as a circle to estimate a radius r , then creating an annulus-shaped neuropil area with inner and outer diameters of $2r$ and $3r$. This neuropil area excluded pixels if they were part of the ROI of another cell body. Values of pixels within the annulus-shaped neuropil area were averaged to generate $F_N(t)$. To subtract the neuropil signal, we calculated $F(t) = F_C(t) - \alpha F_N(t)$, where α is a correction factor ranging from 0.2 to 0.6. The value of α was calibrated for each experiment to avoid overcorrection, by making sure that $F(t) > 0$ for each cell. For each ROI, the fractional change in fluorescence, $\Delta F/F(t)$, was calculated as

$$\frac{\Delta F}{F}(t) = \frac{F(t) - F_0(t)}{F_0(t)}$$

where $F_0(t)$ is the baseline fluorescence as a function of time. To estimate baseline, we first obtained $F_{\text{image}}(t)$, the mean pixel intensity for the entire 256 pixel × 256 pixel field of view as a function of time. $F_0(t)$ was then calculated as

$$F_0(t) = F^* \times \frac{F_{\text{image}}(t)}{F_{\text{image}}^*}$$

where $F_{\text{image}}(t)$ is the 10th percentile of $F_{\text{image}}(t)$ within a sliding window of 10 min duration. F^* and F_{image}^* are the 10th percentile of $F(t)$ and $F_{\text{image}}(t)$ within the first 10 min of the session, respectively. We verified that $F_{\text{image}}(t)/F_{\text{image}}^*$ does not vary with specific choices or rule blocks, and thus serves the purpose of compensating for slow, full-field signal drifts due to non-physiological sources. We repeated the ensemble analyses with two other methods for calculating baseline: one, estimating $F_0(t)$ using the 10th percentile of $F(t)$, on a per-cell basis, with a moving window of 10 min duration; and two, estimating $F_0(t)$ using the 10th percentile of $F(t)$ from the entire session, i.e., without a moving window. These different ways to estimate baseline led to qualitatively similar results for all the ensemble analyses.

Analysis of task-related activity and choice encoding. To calculate trial-averaged fluorescence transients, we created time bins that were 0.5 s wide and assigned each $\Delta F/F(t)$ value at a particular time t to the corresponding time bin relative to the animal's response. The binned $\Delta F/F(t)$ values were averaged to obtain trial-averaged $\Delta F/F$. To estimate uncertainty of the trial-averaged $\Delta F/F$, a bootstrap analysis was performed by drawing fluorescence transients per trial, with replacement, up to the same number used to construct the trial average. The median and 95% confidence intervals of trial-averaged $\Delta F/F$ were estimated from 1,000 iterations of this bootstrap analysis. To quantify choice encoding, we performed multiple linear regression analysis on the $\Delta F/F(t)$ of each cell using the following equation:

$$\frac{\Delta F}{F}(t) = a_0 + a_1 C(n) + a_2 C(n-1) + a_3 C(n)C(n-1) + a_4 C(n-2) + \epsilon(t)$$

where $C(n)$ was the choice of current trial, $C(n-1)$ was the choice of prior trial, $C(n-2)$ was the choice two trials ago, $\epsilon(t)$ was the error term and a values were regression coefficients. We coded a choice of left as 1 and right as -1.

We used a nonoverlapping 0.5-s-long moving window with step size of 0.5 s. A cell was deemed to encode one of the choice parameters or interaction if $P < 0.01$ for the corresponding regression coefficient. To avoid confounds from rule and reward signals, we analyzed only sound-guided trials in which $R(n) = 1$ (outcome of current trial = reward) and $R(n - 1) = 1$ (outcome of prior trial = reward). We did not analyze action trials because parameters such as $C(n)$ and $C(n - 1)$ were highly correlated by virtue of the task structure, obviating a simple interpretation of the analysis.

Analysis of neural circuit trajectories. Scripts for the ensemble analysis were written in Matlab and are available upon request. For state-space analysis, we used demixed principal component analysis³⁶ (dPCA). To prepare the imaging data for dPCA, $\Delta F/F(t)$ for each cell for each trial was aligned in time, from 0 to 6 s from the time of the response in that trial. We tried numerous other time windows and found similar results. This alignment led to an array with dimensions = cells \times time \times trials. Using this array, we averaged across 4 trial types: $C(n) = 1$, $R(n) = 1$, pre-switch sound trials; $C(n) = -1$, $R(n) = 1$, pre-switch sound trials; $C(n) = 1$, $R(n) = 1$, pre-switch action trials; $C(n) = -1$, $R(n) = 1$, pre-switch action trials. This trial-averaged array (cells \times time \times 4) was input into the dPCA algorithm³⁶ to demix time- and task-dependent variances and obtain principal components (PCs). To calculate neuronal circuit trajectories, single-trial or trial-averaged $\Delta F/F$ were projected onto the first three PCs. To characterize similarities between the neuronal circuit trajectories across blocks, we calculated the neuronal circuit trajectory for each block by using the trial-averaged fluorescence across the 20 trials pre-switch. The similarity between a pair of trajectories was quantified by calculating the mean of the Euclidean distances between the trajectories at matching time points in state space. To compare between different experiments, this distance was normalized for each experiment: the Euclidean distances were divided by the spread of all population vectors, calculated as the root mean square of distances between all population vectors and the centroid of the vectors. To quantify how the neuronal circuit trajectories evolve on a trial-to-trial basis, we used the Mahalanobis distance, which is a measure of distance between one point and another collection of points. We defined the origin as the 20 trials preceding a block switch, and the destination as the 20 trials preceding the next block switch. We were interested in the relative separation between the origin, an individual trial that occurred in between, and the destination. Therefore, for each time point of a trial, we calculated Mahalanobis distances, $d_{\text{origin}}(t)$ and $d_{\text{dest}}(t)$, from the individual trial (1 three-dimensional value) to the origin and destination respectively (20 three-dimensional values). The $d_{\text{origin}}(n)$ and $d_{\text{dest}}(n)$ for each individual trial is the median of $d_{\text{origin}}(t)$ and $d_{\text{dest}}(t)$ of the ~ 30 time points within a trial. To estimate the location of an individual trial relative to the origin and destination, we calculated the ratio of Mahalanobis distances, $d_{\text{origin}}(n)/(d_{\text{origin}}(n) + d_{\text{dest}}(n))$. For the Mahalanobis distance ratios, which are a function of trial number from switch, we fitted with a logistic function,

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} + L_{\text{min}}$$

where x_0 is the midpoint trial, k is the steepness, L is the range and L_{min} is the minimum value. The parameter L_{min} is not fitted, but rather is estimated for each transition by calculating the mean of Mahalanobis distance ratio using trials -5 to -1 from switch. We fitted every neural ensemble transition using this method, but excluded those in which the midpoint trial $x_0 < -5$ or $x_0 > 200$, indicating a poor fit. Based on this criterion, we excluded none (0/33) of the action-to-sound shifts and 8% (3/38) of the sound-to-action shifts in our analysis of M2 neural ensembles. For analysis of the ALM data set, we excluded 8% (2/26) of the action-to-sound shifts and 3% (1/32) of the sound-to-action shifts. When comparing behavioral and neural transitions, we defined 'behavioral transition trial' as the trials to criterion (85% correct for 20 trials) minus 20, i.e. the first of the sequence of 20 trials leading to block switch. The 'neural transition trial' was defined as the trial when the first term of the logistic fit of

Mahalanobis distance ratios reached a value of 75% L . That is, the trial x that satisfies this equation:

$$0.75L = \frac{L}{1 + e^{-k(x-x_0)}}$$

This definition is arbitrary; it is unknown how much the population activity pattern must resemble the final pre-switch ensemble state in order to qualify as a transition. Therefore, in another analysis we first fitted each neural transition with the logistic function and identified the behavioral trial corresponding to each 5% L step of neural transition from 10 to 90% L . We then calculated the mean hit and error rates at those corresponding behavioral trials, thus plotting the relationship between behavioral performance and neural transition without explicitly defining a transition trial.

Decoding population activity. To determine how well ensemble dynamics could be used to predict trial type, we first selected those imaging frames that occurred between 0 to 6 s from time of response out of the frame-by-frame imaging data (i.e., $\Delta F/F(t)$). We then projected these $\Delta F/F(t)$ onto the PCs deduced from dPCA to obtain population activity vectors. This procedure reduced the dimensionality of our data from (frames \times cells) to (frames \times 3). Each population activity vector in this analysis came from one of three possible trial types: $R(n) = 1$, sound trials before the switch; $R(n) = 1$, action-left trials before the switch; $R(n) = 1$, action-right trials before the switch; other trial types were not considered for the decoding analysis. Using a randomly chosen fraction (80%) of the population activity vectors, we constructed a classifier based on linear discriminant analysis, using Mahalanobis distances with stratified covariance estimates (the "classify" function in Matlab with "Mahalanobis" option). We then tested the performance of this classifier on the remaining 20% of the population activity vectors, comparing the classification results with actual trial types. This fivefold cross-validation process was repeated 1,000 times to obtain a median estimate of classifier accuracy. To investigate decoding accuracy across time, the timing information of each population activity vector relative to the time of response in each trial was retained. We then ran a separate decoding analysis on the population activity vectors measured during each time period, using a non-overlapping sliding window with duration of 0.28 s and step size of 0.28 s. This window duration is the inverse of frame rate, which was 3.6 Hz. To decode from single-cell activity, $\Delta F/F(t)$ of each cell was used instead of population activity vectors as inputs to construct the classifier.

Statistics. Statistical tests were performed in Matlab and are indicated in the main text or figure legends. Unless otherwise noted, a Wilcoxon signed-rank test was used for all paired comparisons. For two-sample, unpaired comparisons, a Wilcoxon rank-sum test was used. Paired t -tests were used for bin-wise analysis of lick rates. For quantification of choice signals as a function of time, multiple linear regression was first performed as detailed above; a binomial test was then applied to the proportion of cells significantly encoding choice within each time-bin. For ensemble decoding analyses, mean classification accuracy was tested against chance level using a one-sample t -test. For t -tests, the sampling distribution of the mean was assumed to be normal, but this was not formally tested. All t -tests were two-tailed. Statistics are available in the **Supplementary Methods Checklist**.

Code availability. The custom Matlab code used for this study is available upon request.

Data availability. The data that support the findings of this study are available from the corresponding author upon request.

51. Pologruto, T.A., Sabatini, B.L. & Svoboda, K. ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* **2**, 13 (2003).
52. Peron, S.P., Freeman, J., Iyer, V., Guo, C. & Svoboda, K. A cellular resolution map of barrel cortex activity during tactile behavior. *Neuron* **86**, 783–799 (2015).