

Network state changes in sensory thalamus represent learned outcomes.

Masashi Hasegawa^{1,2}, Ziyang Huang¹, Jan Gründemann^{1,2*}

Affiliation

¹ Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE),
Neural Circuit Computations; Bonn, Germany

² University of Basel, Department of Biomedicine; Basel, Switzerland

*Corresponding author. Email: jan.grundemann@dzne.de

Abstract

Thalamic brain areas play an important role in adaptive behaviors. Nevertheless, the population dynamics of thalamic relays during learning across sensory modalities remain mostly unknown. Using a cross-modal sensory reversal learning paradigm combined with deep brain two-photon calcium imaging of large populations of auditory thalamus (MGB) neurons, we identified that MGB neurons are biased towards reward predictors independent of modality. Additionally, functional classes of MGB neurons aligned with distinct task periods and behavioral outcomes, both dependent and independent of sensory modality. During non-sensory delay periods, MGB ensembles developed coherent neuronal representation as well as distinct co-activity network states reflecting predicted task outcome. These results demonstrate flexible cross-modal ensemble coding in auditory thalamus during adaptive learning and highlight its importance in brain-wide cross-modal computations during complex behavior.

Summary

Deep brain imaging reveals flexible network states of sensory thalamus predicting task outcome in mice.

Main

Relaying sensory information from subcortical areas to sensory cortices is a fundamental function of thalamus^{1,2}, yet we are only starting to understand the computational role of individual sensory thalamic nuclei in flexible encoding during adaptive behaviors³⁻⁷. Auditory thalamus (medial geniculate body, MGB) flexibly encodes sensory information across associative learning⁸⁻¹², and populations of individual MGB neurons exhibit diverse single cell response adaptations to conditioned tone stimuli in a behavioral state-dependent manner during fear learning⁵. In addition to learning-related auditory response plasticity, MGB neurons process cross-modal sensory inputs in a complex manner^{10,13}. For example, visual stimuli enhance MGB tone responses non-linearly¹⁰, while tactile stimuli affect MGB tone responses bidirectionally¹³, indicating that, similar to cortical and collicular brain areas, MGB processes cross-modal sensory inputs in addition to its auditory relay function¹⁴⁻¹⁷. These data suggest that MGB is an active computational unit which processes complex information across sensory modalities upon adaptive behaviors. Nevertheless, it remains unknown how large-scale neuronal dynamics in auditory thalamus represent sensory stimuli of different modalities that change their assigned value and expected outcome during flexible learning.

Results

Auditory thalamus neurons display various response patterns to cross-modal sensory stimuli in reward associative learning.

To address this question, we developed a cross-modal sensory Go/Nogo reversal learning task in mice to test for neuronal population dynamics during sensory learning and for cognitive flexibility upon changing reward contingencies. Mice (N = 8 animals) were trained to associate counterbalanced auditory (12 kHz tone) and visual (rightward drifting grating) stimuli as Go (reward predictor) or Nogo (not rewarded) cues (Fig. 1a-c). Once mice learned the task at expert level (*Initial learning*), the reward contingency was reversed and the previous Go cue became non-predictive of the reward, while the previous Nogo cue turned into a reward-predictor (*Reversal learning*). Animals learned the initial rule within approximately 10 days (Fig. 1d-f and fig. S1). Upon reversal learning, task performance dropped initially, yet mice learned the new stimulus-reward rule again with a similar learning rate (N = 8 mice, Fig. 1d-g and fig. S1). These data indicate that mice flexibly associate sensory stimuli with reward outcome across initial and reversal learning using similar learning strategies.

To track the activity patterns of MGB neurons during learning, we performed longitudinal *in vivo* two-photon calcium imaging of large populations of individual MGB relay neurons through a gradient refractive index (GRIN) lens across all stages of the learning paradigm (Fig. 1h, Movie S1). During naive and expert phases of initial and reversal learning, MGB neurons exhibited a large variety of distinct response patterns to the auditory tone as well as the visual stimulus and combinations thereof (Fig. 1h-j, fig. S2), indicating that subsets of MGB neurons are inherently responsive to cross-modal sensory stimuli. Nevertheless, the learning-induced reward association of the Go stimulus during initial or reversal learning altered the proportion of stimulus-responsive MGB neurons towards the Go stimulus regardless of the sensory modality (Fig. 1k, fig. S2). These results demonstrate that auditory thalamus processes cross-modal sensory information during discriminatory reward learning and that MGB responsiveness is dynamically biased towards reward predictors independent of stimulus modality.

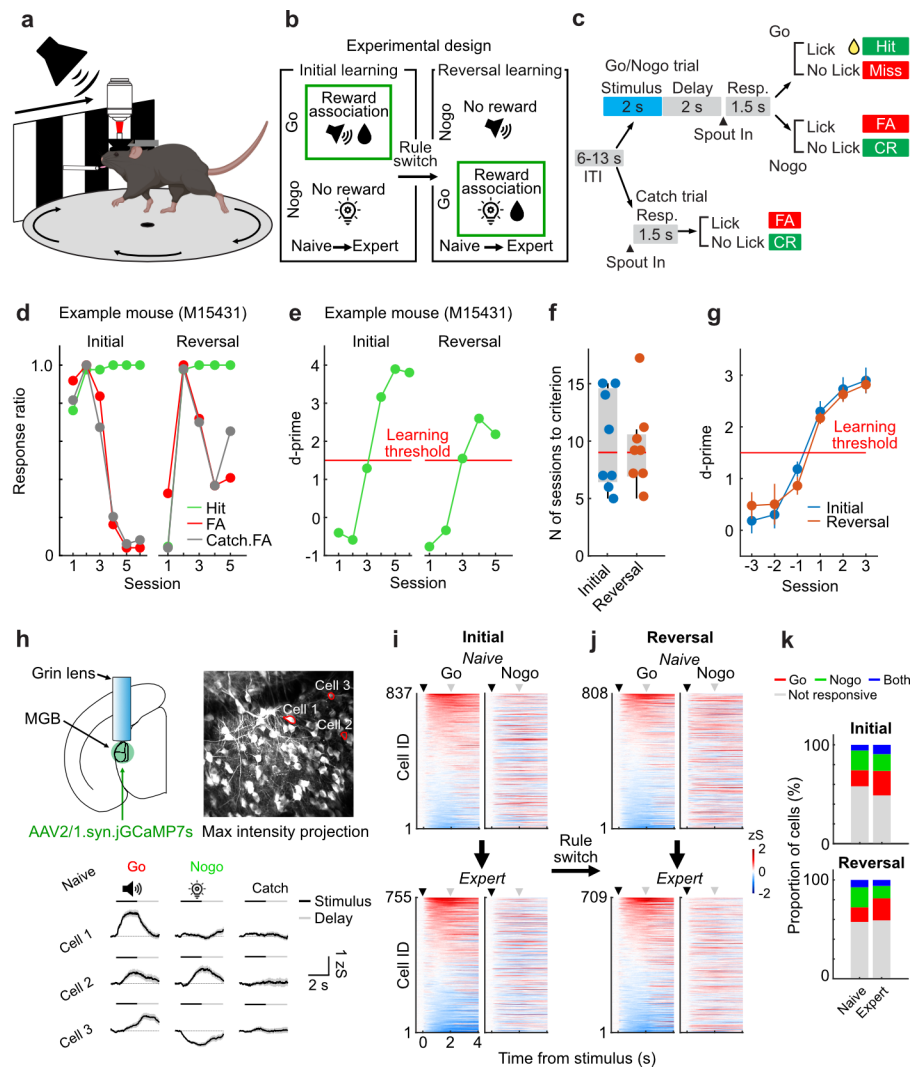


Fig. 1. Auditory thalamus neurons exhibit diverse responses and dynamic bias towards reward predictors upon cross-modal reversal learning.

a, Schematic of the behavioral setup. **b**, Experimental design of the sensory Go/Nogo reversal learning paradigm. **c**, Trial structure. **d**, Task performance of one representative mouse. **e**, Example d-prime transition of the same mouse shown in (**d**). Learning threshold: d-prime > 1.5. **f**, Number of sessions to reach learning criterion (N = 8 mice). Stimulus-reward association sequence (auditory to visual vs. visual to auditory) is counterbalanced across mice (N = 4 for each group). **g**, Transitions of d-prime around the learning threshold are similar between initial and reversal learning (mean ± SEM, N = 8 mice). **h**, Top left: Schematic of GRIN lens implantation. Top right: Example two-photon image (max intensity projection) of MGB. The contrast is enhanced for improved visualization. Bottom: Ca²⁺ traces from three example neurons in Go, Nogo and Catch trials (mean ± SEM). **i**, Mean individual cell activities in Go and Nogo trials during the initial learning (single session for Naive and Expert, N = 8 mice). Cells were sorted by the mean response amplitude during the stimulus presentation in Go trials. Cell IDs are matched across Go and Nogo trials in each learning phase. Black and gray triangles represent the stimulus and delay period onset, respectively. **j**, Mean individual cell activities in Go and Nogo trials during the reversal learning (single session for Naive and Expert, N = 8 mice). **k**, Proportion of stimulus-responsive cells from the data shown in **i** and **j**. Top: The transition of the proportion of stimulus-responsive cells from naive to expert phases in the initial learning. The proportions of the stimulus-responsive cells were altered from Naive to Expert phases in both initial and reversal learning (both $p < 0.01$, Chi-square). Parts of panel 1a were created with BioRender.com.

Functional neuronal subgroups predict task-outcome in MGB.

Next, we separated MGB neurons into functional subgroups by using a k-means-based cluster analysis approach. MGB neurons exhibited distinct stable, learning-enhanced as well as learning-inhibited responses to the reward-predicting Go stimuli (Fig. 2a,b). These learning-related functional clusters in MGB emerged regardless of the modality of the Go stimulus (Fig. 2a,b, right). In addition to stimulus-driven responses, subsets of MGB neurons developed ramping activity during the non-sensory, pre-reward delay period of Go trials (Ramp-up or

Ramp-down clusters), during which the animal has to retain the stimulus type (Go vs. Nogo) and prepare for the action. This ramping activity was specific to the Go stimulus, given that only a small proportion of neurons exhibited ramping activity during Nogo delay periods (fig. S3d,e).

We then compared the activity patterns of the same neurons across expert states in initial and reversal learning when the animals are successfully performing the task, and found task-specific MGB neurons that are active during distinct trial epochs irrespective of the sensory modality (Fig. 2c, left). In contrast, a second modality-specific group of MGB neurons was modulated by the trial epoch in a sensory modality-specific manner (Fig. 2c, middle). These results demonstrate that reward learning drives heterogeneous neuronal plasticity in MGB that flexibly reflects task features and reward outcomes in a modality-driven as well as modality-independent manner.

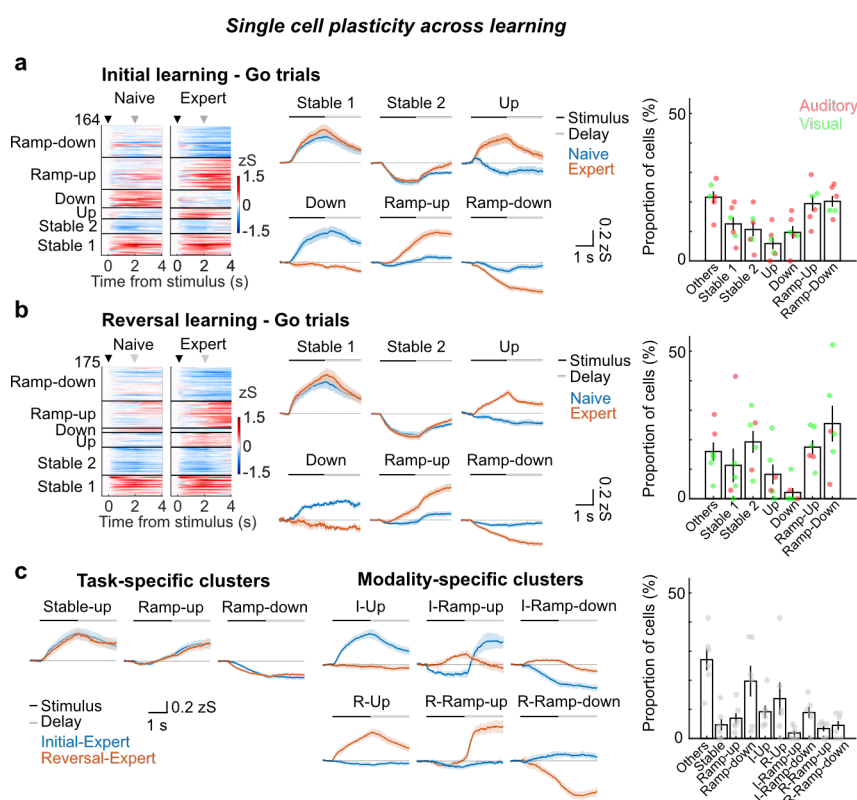


Fig. 2. Reward learning induces heterogeneous single cell plasticity in a task- and modality-specific fashion.

a, Functional subgroups of MGB neurons in initial learning. Left: Heatmaps of single cell activities before and after learning in initial learning. Cells were clustered into functional subgroups depending on their activity patterns ($n = 164$ cells from 6 mice). Middle: Average calcium traces (mean \pm SEM) of the functional subgroups shown on the left side. Right: Proportion of cells in each cluster. Each dot represents the data from the individual mouse. Red and green dots represent the type of stimulus-reward association. **b**, Learning-related functional subgroups in reversal learning ($N = 6$ mice, $n = 175$ cells). Figure structures are the same as (a). **c**, Functional subgroups exhibiting task and modality-specific plasticity in the two expert phases in initial and reversal learning. Left: Average calcium traces (mean \pm SEM) of the task-specific functional subgroups. Middle: Average calcium traces (mean \pm SEM) of the modality-specific functional subgroups. Right: Proportion of cells in each cluster.

Functional plasticity in MGB is not exclusively driven by behavioral variables.

During the cross-modal Go/Nogo paradigm, mice adapted their behavior flexibly and reversibly. Behavioral variables such as body movement, licking^{9,18,19} or arousal level²⁰ affect neural activity in a brain-wide manner. Depending on the trial epoch, approximately 5.7-38.6 % of MGB neurons exhibited correlations of Ca^{2+} activity with behavioral variables (fig. S4a,b, threshold: $r = 0.2$). However, the behavioral modulation of Ca^{2+} activity was not

systematically changed across learning (fig. S4c,d). In addition, we found that the strength of the correlation of Ca^{2+} activity and locomotion or pupil size in neurons that were tracked across the behavioral paradigm was unchanged or even decreased after learning and upon behavioral adaptation (fig. S4e,f). In addition to locomotion and arousal, the number of licks during the pre-reward delay period increased after learning (fig. S4g,h). Nevertheless, anticipatory licking did not systematically modulate MGB activity during expert phases (fig. S4i,j) and did not correlate with the proportion of ramping cells during the delay period (fig. S4k). Taken together, these results indicate that subsets of individual MGB neurons correlate with behavioral variables such as locomotion, pupil size and licking, yet changes in behavior are not the main driver of functional plasticity and changes in learning-related activity in MGB.

Learning induces coherent neuronal population representations during reward-preceding periods.

In addition to the classification of individual stable and plastic MGB neurons, we next asked if and how the neural population representation of MGB changes upon learning. Throughout an individual session as well as across learning, the trial-by-trial population vector correlation (PVC, Pearson's r)^{21–23} of the stimulus period (Go and Nogo trials) remained stable (Fig. 3a,b, figs. S5-6). In contrast, the PVC specifically increased during the reward-preceding delay period of Go trials once mice learned the stimulus-reward contingency in initial learning and flexibly adapted to the previous non-rewarded delay period after reversal learning (Fig. 3c,d, figs. S5-6). Thus, upon associative learning, the neuronal representation of the MGB population becomes more similar during the reward-preceding delay period in a sensory modality-independent fashion (figs. S5-6). The PVC increase was specific to the outcome and not the action given that it occurred only during the delay period of Hit, but not False Alarm trials, where mice incorrectly licked to a Nogo stimulus. (Fig. 3e,f, and figs. S7-8). Furthermore, the delay period PVC increase was present in trials with and without anticipatory licking (Fig. 3g) indicating that enhanced PVCs during the delay period are not driven by the preparatory behavior of the animal or lick impulsivity. Removing delay cells (i.e., Ramp-up and Ramp-down cells, see Fig. 2a,b; see fig. S9 for the removal of unspecific subclusters) did not affect the increase in PVC (Fig. 3h), indicating that the change in population representation during the delay period is driven by the total population of MGB neurons.

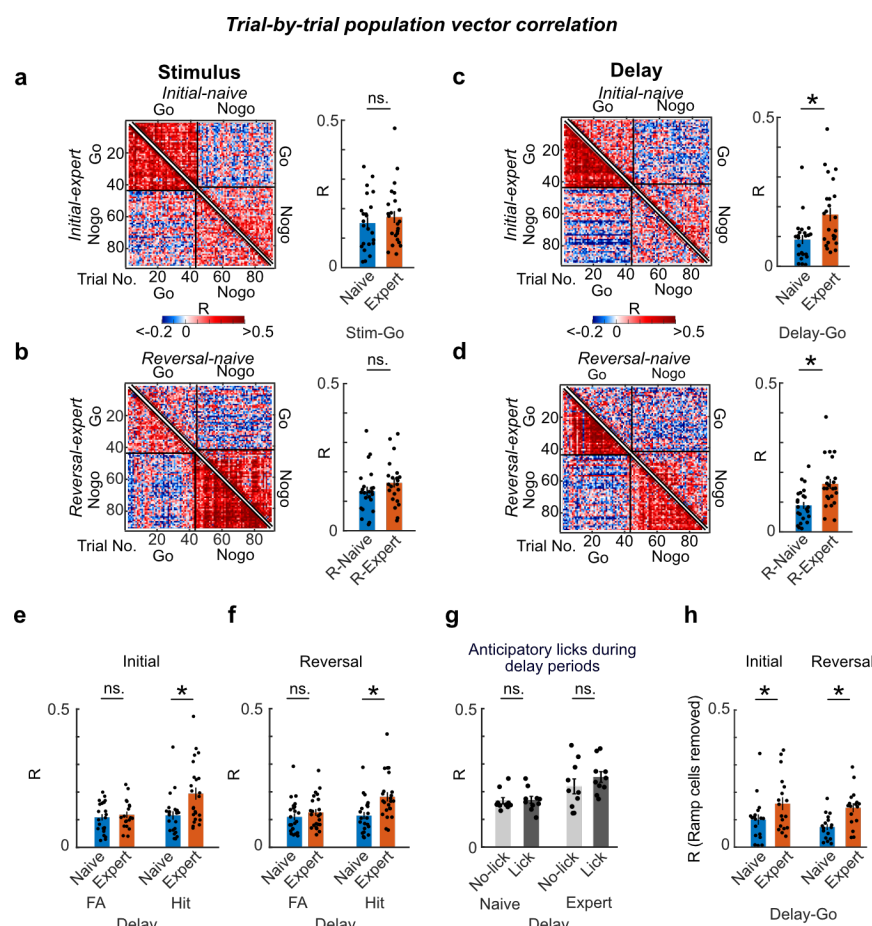


Fig. 3. Learning-induced remodeling of neural population level representations of task features in MGB.

a-d, Representative single session trial-by-trial population vector correlation matrices of stimulus (a, b) and delay (c, d) periods from one example mouse. Bar charts show the means of Go trial PV correlation between naive and expert phases in initial (a, c) and reversal learning (b, d) from all mice (N = 8 mice, 2-3 sessions / mouse, see figs. S5-8). **e,f**, Mean trial-by-trial population vectors correlation of the delay period of Hit and False Alarm (FA) trials in different learning stages (N = 8 mice, 1-3 sessions / mouse, see figs. S7, 8). **g**, Mean trial-by-trial correlation in the delay period in Hit trials with or without anticipatory licking in the initial learning (N = 8 mice, 1-3 sessions / mouse, see figs. S5, 6). Due to the small number of the Hit trials with licking in reversal learning, we excluded the reversal learning data for this analysis. **h**, Mean trial-by-trial correlation in the delay period after removing all ramping cells found in Fig. 2a,b in all mice (N = 6 mice, 2-3 sessions / mouse, see fig. S9). Each dot is the mean R value in individual sessions in each training stage for all mice. Statistical tests: Linear mix model (a-f, h, *p < 0.01, details in Table S1). Wilcoxon rank sum test (g).

Associative reward learning changes the co-activity network structure in MGB.

Next, we analyzed how the co-activity network structure changes across learning in MGB by computing the weighted undirected graphs of MGB population activity^{24,25}. Here, nodes in the co-activity network represent individual MGB neurons and edges the positive pairwise Pearson's correlation coefficients of neural activity between all neuron pairs (Fig. 4a). During the delay periods of the Go trials, the hubness (average of the activity correlations between all cell pairs) increased from the naive to the expert phases, while it was unaffected in the Nogo trials (Fig. 4b). Furthermore, the mean shortest path length between any two neurons became shorter in the expert phase in the Go trials, which is consistent with a global increase of hubness (Fig. 4b). Local co-activity structures represented by the cluster coefficient of triad did not increase from the naive to expert phases (fig. S10a), indicating that the observed changes are a global event in the total MGB population. Upon reversal learning, the co-activity structure during the delay period was actively remodeled, both globally (hubness and path length) and locally (cluster coefficient) (Fig. 4b, fig. S10a), indicating that MGB co-activity network

structure is dynamic during changing stimulus-reward contingencies. Furthermore, removing the cluster of delay ramping neurons (or a similar number of random neurons) from the analysis did not affect the changes in the co-activity network structure (fig. S10b,c). These results indicated that the general MGB population mediates the remodeling of the co-activity network. Anti-correlation networks (negative Pearson's r) remained unchanged in all learning stages (fig. S11). In summary, global co-activity network structures in MGB changed dynamically during the reward-preceding delay period and flexibly re-adjusted after learning in a modality-unspecific fashion. Changes in co-activity structures during the stimulus period were less consistent and limited to the reversal learning period (fig. S12). Altogether, these results demonstrate that the functional co-activity structure in MGB can be rapidly re-organized through associative reward learning and learning rule switches, which could help to support the cognitive processing of task-relevant information^{24,25}.

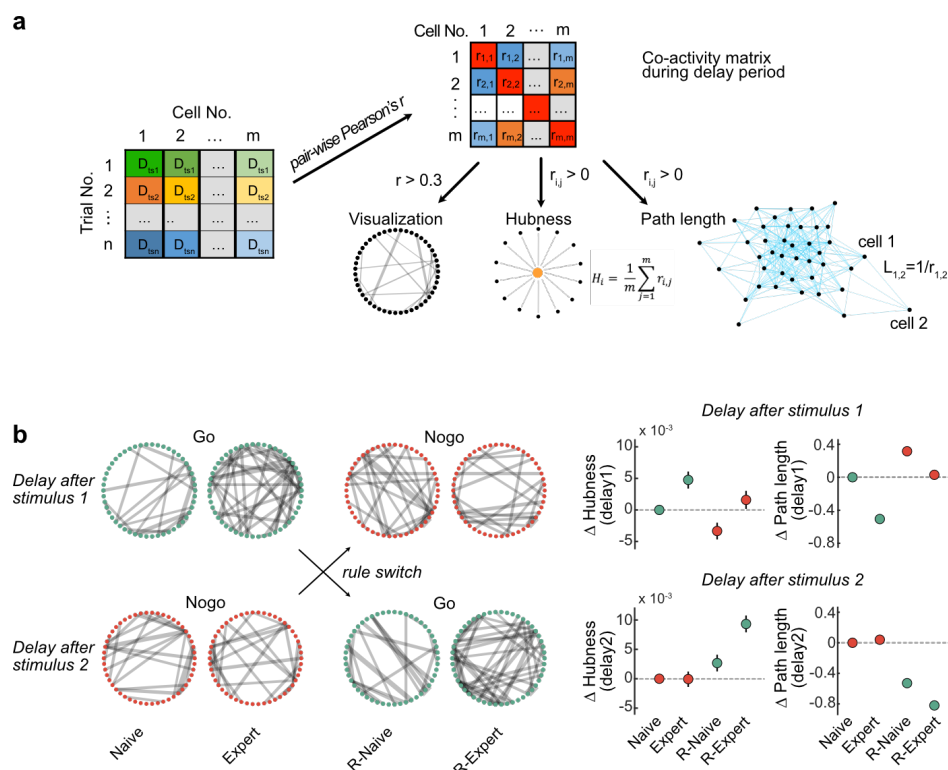


Fig. 4. Learning re-organizes MGB co-activity structure.

a, Co-activity matrix construction and network analysis. The co-activity matrix represents the cell-by-cell pair-wise Pearson's correlation (r) from all concatenated time series vectors of the delay periods (D_{ts}) from all trials (see left matrix for construction of time series vector). The co-activity strength is visualized by line thickness in a circular plot for each cell pair ($r > 0.3$, dots indicate individual neurons). Global MGB co-activity strength is quantified by hubness and shortest path length. D_{tsn} : 2 s time series during delay period in trial n . $r_{i,j}$: Pearson's correlation of calcium activity between cell i and j . H_i : hubness of cell i to all its neighbors. $L_{1,2}$: path length between cell 1 and 2.

b, Left: Representative circular plots show an example MGB co-activity network structure at different learning stages of Go and Nogo trials for one representative mouse before and after learning the rule switch. Right: Changes of hubness and path length across all neuron pairs in each learning stage which are baselined to the values in the naive Go (Stimulus 1) or Nogo (Stimulus 2) condition ($N = 210$ neurons from 6 mice, 3000 bootstraps). Error bars: 95 % confidence interval of mean.

Discussion

Recent work has shed light on the role of sensory thalamus in adaptive behaviors^{3,4}, while the neural dynamics of sensory thalamus in flexible learning of complex tasks across multiple sensory modalities remained elusive. Here, we combined a cross-modal (auditory and visual) reversal learning paradigm with longitudinal deep-brain two-photon calcium imaging in medial

geniculate body (MGB) and demonstrated that sensory thalamus dynamically encodes sensory as well as task-related information, adapting its neural responses to varying task rules.

We find that auditory thalamus exhibits adaptive processing of auditory and visual information (Fig. 1h-j and figs. S13, 14) similar to sensory cortices^{15,16}. Responses of MGB neurons to auditory and visual stimuli are plastic and modulated upon cross-modal stimulus-reward learning (Fig. 1k), indicating that sensory encoding in MGB is dynamic upon cognitively demanding tasks irrespective of the sensory modality (fig. S2). A subset of MGB neurons developed ramping activity during the reward-preceding delay period (Fig. 2a-c), during which the animal has to hold the predictive information of the stimulus about the trial outcome, which could reflect reward anticipation¹¹ or short-term memory to guide the next action²⁶, similar to cortical and thalamic areas of the mouse motor system²⁷. Reciprocal loops between motor thalamus and cortex have been shown to be necessary to maintain ramping activities in either region during delay periods²⁷, suggesting that, similar to the motor system, auditory thalamus could cooperatively maintain delay ramping activity together with auditory and other cortices.

On the population level, learning can either increase^{28,29} or decrease^{30,31} trial-by-trial correlation, whilst other cognitive factors such as attention^{32,33} exhibit a bi-directional influence. In our reversal learning task, the neuronal representation of sensory stimuli in auditory thalamus was stable across learning, regardless of whether they were predictive of a reward (Go cue) or not (Nogo cue). These findings are similar to previous observations for conditioned stimulus presentations in aversive auditory fear learning⁵. In contrast to the stimulus period, the MGB population response (trial-by-trial correlation) became more coherent across trials during the non-sensory pre-reward delay period, once mice learned the task-reward rule, irrespective of the sensory modality (Fig. 3a-d, figs. S5, 6). This trial-by-trial population vector correlation increase during the delay period was specific to Hit trials, but not False Alarm trials (Fig. 3e, f, figs. S6, 7), indicating that the re-organization of the population activity during the delay period is dependent on the predicted task outcome and not the general reward preparation and consumption movement per se (e.g., licking, Fig. 3g). Recently, trial-by-trial correlations were suggested to account for the optimization of information communication between upstream sensory information encoding and downstream read-out guiding behaviors in different population representation subspaces³⁴⁻³⁶. Together, our findings link changes in population dynamics of auditory thalamus to a wider brain network that is implicated in functions of task representation³⁷, short term memory^{27,38}, outcome-dependent actions³⁹ and reward anticipation¹¹.

Co-activity network structure during the reward-preceding delay period was remodeled across associative learning regardless of the sensory modality, while it remained by-and-large stable during the stimulus period (Fig. 4b, figs. S10, 12). The distinct changes of co-activity network states between the stimulus and delay periods may result in differential information transfer to downstream areas to prime information integration and guide behavioral outputs in Go or Nogo trials^{7,40,41}. While the global co-activity network changed upon learning (i.e., path length and hubness), the local co-activity network (i.e., cluster coefficient) did not systematically change during the delay period, and ramping cells did not contribute to the enhancement of the co-activity network (fig. S10). This suggests that the activity of the total MGB population supports the functional network structure and not individual functional subgroups of neurons. Upon reversal learning, the co-activity structures were dynamically remodeled not only during the delay but also during the stimulus periods (fig. S12), which indicates that reversal learning might depend on a distinct learning rule, resulting in the generalization of the Go stimulus from

a sensory representation towards a predictive state representation that could be flexibly assigned to future learning rule updates. It remains to be tested if and how these changes are mediated and whether they are generated locally or if they are co-dependent on external activity, e.g., primary or associative cortical brain areas⁴².

What are the neural mechanisms that guide the functional plasticity in MGB from naive to expert phases across reward associative learning? Given the sparse local recurrent connectivity within mouse sensory thalamus^{1,43,44} and negligible proportions of local inhibitory neurons^{5,45}, local microcircuit plasticity is unlikely to drive changes in response patterns. Plausible scenarios could include changes of synaptic plasticity, neuromodulation or adaptations of long range bottom-up excitatory^{46,47} and/or inhibitory^{3,4} inputs. In addition, corticothalamic feedback^{2,48} could stabilize changes in MGB population activity upon learning, either directly or di-synaptically via the thalamic reticular nucleus (TRN). Future studies will need to test how distinct circuit elements and their combination can affect functional population level plasticity in auditory thalamus.

While reward learning biased the sensory responses of MGB neurons towards reward-associated Go stimuli during the task regardless of sensory modality (Fig. 1k, fig. S2), the sensory responses in off-task non-rewarded mapping sessions before and after learning (i.e., passive measurements of tuning to multimodal sensory stimuli) were not biased to the on-task Go stimuli (figs. S13, 14). Specific multisensory enhancement to the Go stimulus, which was previously observed during appetitive learning in MGB¹⁰, did not take place during off-task, non-rewarded sessions (figs. S15, 16). This could be due to a fast devaluation of reward-conditioned stimuli¹¹. Finally, while uni and multisensory responses could be altered on the single cell level (figs. S13, 14), sensory stimuli could be reliably decoded from the MGB population activity across learning (fig. S17). The complementary mechanisms of single cell plasticity and population-level stability of sensory coding could be crucial to allow for dynamic neuronal representations upon learning, while ensuring stable representations of the environment^{5,49}.

Altogether, our study reveals that auditory thalamus displays flexible adaptations of single cell responses and co-activity network states that align not only with sensory but also task-period and outcome-relevant information, which change bi-directionally upon reversal learning, highlighting the role of sensory thalamus in complex neural computations for adaptive behaviors.

References

1. Jones, E. G. *The Thalamus*. (Cambridge University Press, 2007).
2. Guillery, R. W. & Sherman, S. M. Thalamic Relay Functions and Their Role in Corticocortical Communication: Generalizations from the Visual System. *Neuron* **33**, 163–175 (2002).
3. Wimmer, R. D. *et al.* Thalamic control of sensory selection in divided attention. *Nature* **526**, 705–709 (2015).
4. Nakajima, M., Schmitt, L. I. & Halassa, M. M. Prefrontal Cortex Regulates Sensory Filtering through a Basal Ganglia-to-Thalamus Pathway. *Neuron* **103**, 445–458.e10 (2019).
5. Taylor, J. A. *et al.* Single cell plasticity and population coding stability in auditory thalamus upon associative learning. *Nat. Commun.* **12**, 1–14 (2021).
6. Fratzl, A. & Hofer, S. B. The caudal prethalamus: Inhibitory switchboard for behavioral control? *Neuron* **110**, 2728–2742 (2022).

7. Rikhye, R. V., Wimmer, R. D. & Halassa, M. M. Toward an Integrative Theory of Thalamic Function. *Annu. Rev. Neurosci.* **41**, 163–183 (2018).
8. Weinberger, N. M. The medial geniculate, not the amygdala, as the root of auditory fear conditioning. *Hear. Res.* **274**, 61–74 (2011).
9. Gilad, A., Maor, I. & Mizrahi, A. Learning-related population dynamics in the auditory thalamus. *Elife* **9**, (2020).
10. Komura, Y., Tamura, R., Uwano, T., Nishijo, H. & Ono, T. Auditory thalamus integrates visual inputs into behavioral gains. *Nat. Neurosci.* **8**, 1203–1209 (2005).
11. Komura, Y. *et al.* Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature* **412**, 546–549 (2001).
12. Barsy, B. *et al.* Associative and plastic thalamic signaling to the lateral amygdala controls fear behavior. *Nat. Neurosci.* **23**, 625–637 (2020).
13. Lohse, M., Dahmen, J. C., Bajo, V. M. & King, A. J. Subcortical circuits mediate communication between primary sensory cortical areas in mice. *Nat. Commun.* **12**, 1–14 (2021).
14. Meredith, M. A. & Stein, B. E. Interactions among converging sensory inputs in the superior colliculus. *Science* **221**, 389–391 (1983).
15. Knöpfel, T. *et al.* Audio-visual experience strengthens multisensory assemblies in adult mouse visual cortex. *Nat. Commun.* **10**, 5684 (2019).
16. Morrill, R. J. & Hasenstaub, A. R. Visual Information Present in Infragranular Layers of Mouse Auditory Cortex. *J. Neurosci.* **38**, 2854–2862 (2018).
17. Gruters, K. G. & Groh, J. M. Sounds and beyond: multisensory and other non-auditory signals in the inferior colliculus. *Front. Neural Circuits* **6**, 96 (2012).
18. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* **22**, 1677–1686 (2019).
19. Steinmetz, N. A., Zatzka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).
20. McGinley, M. J., David, S. V. & McCormick, D. A. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* **87**, 179–192 (2015).
21. Kriegeskorte, N. & Kievit, R. A. Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17**, 401–412 (2013).
22. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
23. Kohn, A., Coen-Cagli, R., Kanitscheider, I. & Pouget, A. Correlations and Neuronal Population Information. *Annu. Rev. Neurosci.* **39**, 237–256 (2016).
24. Bassett, D. S. & Sporns, O. Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).
25. Gava, G. P. *et al.* Integrating new memories into the hippocampal network activity space. *Nat. Neurosci.* **24**, 326–330 (2021).
26. Svoboda, K. & Li, N. Neural mechanisms of movement planning: motor cortex and beyond. *Curr. Opin. Neurobiol.* **49**, 33–41 (2018).
27. Guo, Z. V. *et al.* Maintenance of persistent activity in a frontal thalamocortical loop. *Nature* **545**, 181–186 (2017).
28. Peters, A. J., Chen, S. X. & Komiyama, T. Emergence of reproducible spatiotemporal activity during motor learning. *Nature* **510**, 263–267 (2014).
29. Komiyama, T. *et al.* Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* **464**, 1182–1186 (2010).
30. Gu, Y. *et al.* Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* **71**, 750–761 (2011).
31. Jeanne, J. M., Sharpee, T. O. & Gentner, T. Q. Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* **78**, 352–363 (2013).

32. Ni, A. M., Ruff, D. A., Alberts, J. J., Symmonds, J. & Cohen, M. R. Learning and attention reveal a general relationship between population activity and behavior. *Science* **359**, 463–465 (2018).
33. Ruff, D. A. & Cohen, M. R. Attention can either increase or decrease spike count correlations in visual cortex. *Nat. Neurosci.* **17**, 1591–1597 (2014).
34. Panzeri, S., Moroni, M., Sifaai, H. & Harvey, C. D. The structures and functions of correlations in neural population codes. *Nat. Rev. Neurosci.* **23**, 551–567 (2022).
35. Ruff, D. A., Ni, A. M. & Cohen, M. R. Cognition as a Window into Neuronal Population Space. *Annu. Rev. Neurosci.* **41**, 77–97 (2018).
36. Valente, M. *et al.* Correlations enhance the behavioral readout of neural population activity in association cortex. *Nat. Neurosci.* **24**, 975–986 (2021).
37. Samborska, V., Butler, J. L., Walton, M. E., Behrens, T. E. J. & Akam, T. Complementary task representations in hippocampus and prefrontal cortex for generalizing the structure of problems. *Nat. Neurosci.* **25**, 1314–1326 (2022).
38. Morcos, A. S. & Harvey, C. D. History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat. Neurosci.* **19**, 1672–1681 (2016).
39. Courtin, J. *et al.* A neuronal mechanism for motivational control of behavior. *Science* **375**, eabg7277 (2022).
40. Guest, J. M., Bast, A., Narayanan, R. T. & Oberlaender, M. Thalamus gates active dendritic computations in cortex during sensory processing. *bioRxiv* 2021.10.21.465325 (2021) doi:10.1101/2021.10.21.465325.
41. Avena-Koenigsberger, A., Misic, B. & Sporns, O. Communication dynamics in complex brain networks. *Nat. Rev. Neurosci.* **19**, 17–33 (2017).
42. Banerjee, A. *et al.* Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature* **585**, 245–250 (2020).
43. Halassa, M. M. & Kastner, S. Thalamic functions in distributed cognitive control. *Nat. Neurosci.* **20**, 1669–1679 (2017).
44. Mukherjee, A. *et al.* Variation of connectivity across exemplar sensory and associative thalamocortical loops in the mouse. *Elife* **9**, e62554 (2020).
45. Jager, P. *et al.* Dual midbrain and forebrain origins of thalamic inhibitory interneurons. *Elife* **10**, e59272 (2021).
46. Metzger, R. R., Greene, N. T., Porter, K. K. & Groh, J. M. Effects of Reward and Behavioral Context on Neural Activity in the Primate Inferior Colliculus. *J. Neurosci.* **26**, 7468–7476 (2006).
47. Lee, T.-Y., Weissenberger, Y., King, A. J. & Dahmen, J. C. Midbrain encodes sound detection behavior without auditory cortex. *bioRxiv* 2023.06.07.544013 (2023) doi:10.1101/2023.06.07.544013.
48. Williamson, R. S., Hancock, K. E., Shinn-Cunningham, B. G. & Polley, D. B. Locomotion and Task Demands Differentially Modulate Thalamic Audiovisual Processing during Active Search. *Curr. Biol.* **25**, 1885–1891 (2015).
49. Betzel, R. F., Wood, K. C., Angeloni, C., Neimark Geffen, M. & Bassett, D. S. Stability of spontaneous, correlated activity in mouse auditory cortex. *PLoS Comput. Biol.* **15**, e1007360 (2019).

Acknowledgments:

We thank the HHMI Janelia GENIE Project for making jGCaMP7 available; Raymond Strittmatter, Patrick Schlenker, Robert Häring and Simon Saner for mechanical workshop and electrical engineering support; all members of Gründemann laboratory for their general support and feedback on the manuscript.

Funding:

Swiss National Science Foundation (SNSF professorship PP00P3_170672, J.G.)
 European Research Council Starting Grant 803870 (J.G.)
 The Forschungsfonds Nachwuchsforschende of the University of Basel (M.H.)
 Forschungspreis der Schweizerischen Hirnliga (J.G.)
 The Department of Biomedicine at the University of Basel (J.G.)
 Deutsche Forschungsgemeinschaft (DFG), SFB 1089 (Teilprojekt, J.G.)
 Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Bonn, Germany (J.G.)

Author contributions:

Conceptualization: MH, JG
 Methodology: MH, ZH, JG
 Investigation: MH, ZH, JG
 Visualization: ZH, MH, JG
 Funding acquisition: JG, MH
 Project administration: JG, MH
 Supervision: JG
 Writing – original draft, review & editing: MH, ZH, JG

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: Data to generate figures and supplementary figures in this paper can be accessed in Data S1 and S2, respectively. Statistical results are shown in Table S1. Custom-written code used to generate figures and supplementary figures are available at https://gin.g-node.org/GrundemannLab/ThalamicNetworkStates_Code. Data is available at https://gin.g-node.org/GrundemannLab/ThalamicNetworkStates_Data.

Supplementary Materials

Materials and Methods
 References (50–60)
 Figs. S1 to S17
 Table S1
 Movie S1
 Data S1 to S2

Methods

Animals: All experiments were performed in accordance with the institutional guidelines of University of Basel and were approved by the Cantonal Veterinary Office of Basel-Stadt, Switzerland. Six to 10-week-old (at the start of the experiment) male C57BL/6J mice were used throughout this study. Animals were housed on a 12-hour light / dark cycle at an ambient temperature (22 ° C) and humidity (55 %) and had free access to food and water until the initiation of the behavioral experiment. Throughout the behavioral experiment, animals were placed under food restriction and their body weights were maintained at 85-90% of their free-feeding weights. Well-being was monitored daily through the entire experimental period. No statistical methods were applied to pre-determine the sample size for each experiment. The investigator was not blinded for surgery, behavioral and imaging experiments or data analysis.

Surgical procedures: Buprenorphine (0.1 mg/kg) was subcutaneously injected approximately 30 min before surgery for analgesia. Then, mice were anesthetized with isoflurane (1.5 - 2.0% maintenance) through the oxygen-enriched air (95 %, 1-3 l/min, Oxymat III, Weinmann). Anesthesia level was monitored via breathing rates and foot and tail reflexes before and during surgery. Mice were placed in a stereotaxic apparatus (Model 1900, Kopf Instruments), and their body temperature was maintained through a heating pad (Rodent warmer, 53800M, Stoelting). Their eyes were covered by an eye protective cream (Bepanthen Augen und Nasensalbe, Bayer). The mixture of Lidocaine (10 mg/kg) and Ropivacaine (3 mg/kg) was injected under the skin over the skull for local anesthesia. Stereotaxic viral injections were performed as previously described⁵. Briefly, a small craniotomy was performed above the medial geniculate body (MGB, AP: -3.28, ML: -1.9, DV: -3.0 mm) by using a stereotaxic drill (Model 1911, Kopf) with a burr drill bit (105-0135-225, Kyocera). A pulled glass pipette (2-000-001, Drummond Scientific) filled with AAV vector was slowly lowered into the brain with the help of a micropositioner (Model 2650, Kopf). AAV2/1.syn.jGCaMP7s⁵⁰ (Addgene, 104487-AAV1, ca 500nl, diluted by sterile PBS, 1-2x) was injected into MGB with a pressure ejection system (Picospritzer, Parker). One to two weeks after viral injection, a gradient refractive index (GRIN) lens (0.6 mm diameter, 7.3 mm length or 1.0 mm diameter, 4 mm length, Inscopix) was implanted during the second surgery (anesthesia and analgesia, see above). 0.6 mm lenses were implanted as previously described⁵. Briefly, a 0.8 mm diameter craniotomy was performed above MGB (drill: 105-0709.400, Kyocera) and a small track was cut with a 0.7 mm sterile needle. Next, the GRIN lens was slowly advanced into the brain using the Micropositioner (Model 2650). For the implantation of 1.0 mm diameter lenses, a 1.2 - 1.3 mm craniotomy was performed above MGB using a hand drill (503599, World Precision Instrument) with a burr drill bit (200 µm diameter, C1.104.002, Bösch Dental). Tissue above MGB was slowly aspirated through a sterile blunt needle (27G, Endo irrigation cannula) connected to a suction system. Sterilized phosphate buffer saline (PBS) was used to irrigate the brain until the bleeding stopped around the aspirated site. Next, the 1.0 mm lens was slowly advanced into the brain with the micropositioner. Both, 0.6 and 1.0 mm lenses were fixed to the skull with light curable glue (Loctite 4305, Henkel). A custom-made head bar was attached to the skull next to the GRIN lens, and the skull was sealed with Scotchbond (3M), Vetbond (3M) and dental acrylic (Paladur, Kulzer, Orth Jet, Lang Dental and/or C&B Super-Bond, Sun Medical). Meloxicam (5 mg/kg) was injected subcutaneously after the surgery for post-operative analgesia.

Behavioral apparatus: The behavioral apparatus was housed in a light shield chamber under a custom-built two-photon microscope (Independent NeuroScience Services (INSS), UK).

Mice were head-fixed by a custom-designed holding system and placed on a running-wheel connected to a rotary encoder (E6A2-CS3E, Omron) to measure locomotion. Auditory and visual stimuli were presented with a speaker (ES1, Tucker Davis Technologies, placed at upper-right, 10 cm from the mouse head) and a 7-inch screen (Adafruit 1667, placed 10 cm from the right side of the mouse face at a parallel angle), respectively. The screen system was modified to minimize the light exposure to photomultiplier tubes (PMTs) during a visual stimulus presentation in the two-photon imaging^{51,52}. During the experiment, the gray background was continuously presented from the screen. A lick spout was mounted on the custom-built retractable stage controlled by Trinamic motion control language (TMCL). A reward (soy milk) was delivered by a custom-designed peristaltic pump system by using a micropump (mp6, Bartels Mikrotechnik). A licking of a reward spout was detected by a lick detector modified from the detector described in a previous study⁵³. The experiments were controlled by a custom-written program in MATLAB (MathWorks) with NI USB-6008 (National Instruments) and RZ 6 (Tucker Davis Technologies), and the timing of TTL input/output of behavioral events were recorded by RZ 6 at 50 kHz sampling rate.

***In vivo* two-photon calcium imaging:** *In vivo* two-photon calcium imaging was performed using a custom-built two-photon microscope (INSS, UK). The microscope was equipped with a resonant scanning system and a pulsed Ti:sapphire laser ($\lambda = 940$ nm, Chameleon Vision S, Coherent). A motorized three-axis system (Zaber motor) connected with a microscope head (Z-direction) and breadboard under the behavioral apparatus (X-Y directions) enabled locating an objective lens above the GRIN lens. The microscope system was controlled by ScanImage software (Vidrio Technologies). Green and red fluorescent photons were collected with an objective lens (x16, 0.80 NA, Nikon). Photons were separated by a dichroic mirror (T565lpxr, long pass, Chroma) and barrier filters (green: ET510/80m, red: ER630/75m), and measured by PMTs (PMT2101, Thorlab). The imaging frame was 512 x 512 pixels, and the frame rate was approximately 30 Hz. Fields of views (FOVs) of two-photon images were approximately 330 μm x 330 μm (at 3.0x zoom, 1.0 mm diameter GRIN lens, N = 6 mice) or approximately 400 μm x 400 μm (at 2.5x zoom, 0.6 mm diameter GRIN lens, N = 2 mice). Note that GRIN lens FOVs do not correspond to the actual size of the imaged brain areas due to the spatially non-uniform optical distortion inherent to GRIN lenses⁵⁴.

Two-photon Image Processing: Two-photon images were processed using Suite2P⁵⁵. The images were motion-corrected, and regions of interest (ROIs) were automatically generated. Next, experimenters curated ROIs and sorted them as neurons or not. A small portion of ROIs were drawn using a manual ROI drawing function of Suite2p. A neuropil signal was also calculated for each ROI by Suite2P. A correction coefficient (0.7) was multiplied with the neuropil signal, and each ROI signal was subtracted from this value and handled as Ca^{2+} signal of individual neurons. To extract the same ROI signals through the behavioral and sensory mapping sessions, we employed ROIMatchpub (<https://github.com/ransona/ROIMatchPub>). For behavioral sessions, we selected 2-3 sessions per learning phase (initial naive, initial expert, reversal naive and reversal expert) for each mouse. The naive phase consisted of the first 2-3 behavioral sessions, and the expert phase consisted of the 2-3 sessions in which d-prime (d' , discriminability index)⁵⁶ reached over 1.5. Across these selected sessions, we tracked the same ROIs wherever possible. The ROI tracking procedure was done separately for behavioral and sensory mapping sessions. The same ROI signals, or matched cell data, were used for the following analyses: k-means clustering of the Go/Nogo task (Fig. 2a-c, fig. S3), co-activity network structure analysis (Fig. 4 and figs. S9-12), correlation analysis for behavioral variables and Ca^{2+} signals (fig. S4a-f) as well as the sensory mapping analysis (figs. S13-17) unless stated otherwise. For the trial-by-trial population vector correlation analysis, co-activity

network structure analysis as well as decoder training (see below), Ca^{2+} signals of individual neurons were detrended and lowpass filtered to 5 Hz with a Butterworth filter.

Sensory mapping: Before the start of the sensory mapping sessions, mice were head-fixed under the custom-built two-photon microscope and habituated to the behavioral apparatus and environment for minimum 3 days. Each session started with a 1 min habituation period. Auditory (4, 8, 12, 16 or 20 kHz pure tones at 75 dB, 2 s), visual (Upward, downward, rightward or leftward sine wave drifting gratings, 100% contrast, 2 Hz, 0.05 cycle per degree, 2 s) or multisensory stimuli (combination of auditory and visual stimuli, e.g., 4 kHz pure tone with upward drifting grating, 2 s) were presented. The sensory stimuli were presented in a pseudo-random order. Each auditory, visual and the multisensory stimulus was presented 8 times (e.g., 4 kHz pure tone, rightward grating, and the combination of them were presented for 8 times), and a total of 240 trials (8 trials x 30 stimulus types of uni and multisensory stimuli) were performed in one session per day. Inter-trial interval (ITI) was 6-9 s. For one mouse, each stimulus was presented 5 times and a total of 150 trials / session were performed. Sensory mapping was performed for 2-3 consecutive sessions before and after the sensory Go/Nogo reversal learning paradigm.

Behavioral training in a sensory Go/Nogo reversal learning task: Following sensory mapping sessions, mice were pre-trained to lick a spout to receive a liquid reward (soy milk) for 1-2 days under head-fixation. Next, the animals were trained to perform a sensory Go/Nogo task, which consisted of Go trials (30%), Nogo trials (35%) and catch trials (35%) (total number of trials: 140 per session). At the beginning of each trial, a 6-13 s ITI was initiated. Then, either an auditory stimulus (Go cue, 12 kHz pure tone, 75 dB, 2 s), a visual stimulus (Nogo cue, rightward drifting grating, 2s) or no stimulus (2 s blank period, catch trial) was presented. The sensory stimulus or blank period was followed by a delay period (2 s) without any sensory stimulus. At the end of the delay period, a response window (1.5 s) was initiated and a retractable lick spout moved forward to the mouse. Go trials required mice to lick the spout (Hit) to obtain a reward (8-10 μl soya milk), otherwise the trial was considered as an error (Miss). In Nogo trials, mice were required to withhold the lick response (Correct Rejection, CR). If the mice licked the spout in the Nogo trial, the trial was considered as an error (False Alarm, FA). In catch trials mice were required to withhold the lick response (catch-CR). If mice licked the spout, the trial was considered as an error (catch-FA). If mice accidentally touched the spout too early (within 200 ms after the response window onset) (e.g., due to grooming), the contact was not considered as a response. Catch trials were introduced to ensure that mice identify the Go cue as a reward predictor. Task performances for Nogo and catch trials developed similarly (Fig. 1d,e, and fig. S1). Thus, the task performances in those trials were combined to calculate the task performance index, d-prime (d') as follows: $d' = Z(\text{Hit ratio}) - Z(\text{False alarm ratio})$. Hit and False alarm ratios were calculated as follows: Hit ratio = the number of Hit trials/(the number of Hit trials + the number of Miss trials); False alarm ratio = the number of false alarm trials/(the number of false alarm trials + the number of correction rejection trials). If the ratio reached 1.0, the ratio was adjusted to calculate d' ⁵⁶. Z is the inverse cumulative distribution function, and Z (Hit ratio) and Z (False alarm ratio) were calculated by using the qnorm function (<https://de.mathworks.com/matlabcentral/fileexchange/48978-qnorm-matprobabilities-dblmean-matsigmas-booluseapproximation>). Once d' reached values above 1.5 for three consecutive days, mice were considered experts and a reversal learning paradigm was initiated. Upon reversal learning, the stimulus-reward contingency was switched. In the group in which the auditory stimulus was presented as a Go cue, the visual stimulus was now presented as a Go cue and rewarded and vice versa for animals in which the visual stimulus was initially presented as a Go cue. Once d' reached values above 1.5 for three

consecutive days, mice were considered reversal experts. The order of the stimulus-reward contingency was counterbalanced between mice, i.e., four mice were trained to the auditory stimulus as a Go cue first and four mice were trained to the visual stimulus as a Go cue first. Specific cases of data exclusion: Mice were trained for a maximum of 28 sessions. Due to slower learning, one out of eight mice reached only two reversal expert sessions before the training had to be terminated. Furthermore, in two out of eight mice the training had to be briefly suspended due to technical issues, which did not have a major impact on task performance after re-initiation of the paradigm. In one mouse, an imaging artifact gradually appeared from the lateral edges of the two-photon image in one session. Data from the later stage of this session was excluded from analysis. In one mouse, the data recording of one session was aborted in the middle of the session due to the malfunction of the acquisition system. Since the number of trials reached more than half of the session, the data obtained in this session was included for the data analyses.

Video recording: Mouse behavior was monitored during the experiment with a CMOS camera (a2A2590-60umPRO, Basler) equipped with a CCTV lens (Moritex ML-M1616UR) and a band-pass filter (DB850, Midwest) (N = 7 mice) or a Raspberry Pi Camera Module 2 with a shortpass filter (FES850, Thorlabs) controlled by Raspberry Pi 3 model B+ (N = 1 mouse). Either camera system was located on the left side of the mouse together with a custom-made infrared LED system (830 nm). In the CMOS camera system, each frame acquisition was synchronized with the acquisition of a two-photon image (ca. 30 Hz) through ScanImage. In the raspberry Pi camera system, each frame was acquired at ca. 30 Hz in free-run mode and image acquisition was monitored by internally-generated TTLs. During offline analysis, the tongue and pupil were detected and tracked by animal pose estimation using DeepLabCut (DLC)⁵⁷. The experimenter manually labeled the tongue and the eight points on the edges of the pupil (top, top-right, right, bottom-right, bottom, bottom-left, left, top-left) for each mouse to train the model. Videography based licking-behavior was detected if the tongue was tracked by DLC for at least two consecutive frames. The pupil area was calculated by using a circle fit function⁵⁸ (fitcircle, https://mathworks.com/matlabcentral/fileexchange/15060-fitcircle-m?s_cid=ME_prod_FX) from available data points at each frame. Pupil area data was smoothed by using a hampel filter (MATLAB built-in function, number of neighbors, 10; number of standard deviations, 1.0).

Mean individual cell activity during Go and Nogo trials: Heatmaps were used to visualize mean activities of individual MGB neurons pooled from all mice across learning (N = 8 mice, Fig. 1i, j). For the naive phase, the data of the first training session in both initial and reversal learning were used to show the unconditioned responses to sensory stimuli. For the expert phase, the data from the session with the highest d' in both initial and reversal learning were used to show well-conditioned responses to the sensory stimuli. The calcium data was baselined to the mean during 0.5 s before stimulus presentation in each trial. The data of individual calcium traces were averaged across Go and Nogo trials separately (maximum 42 trials for Go trials, and maximum 49 trials for Nogo trials in a single session). Cell IDs were sorted according to the amplitude of the mean sensory response during the stimulus presentation (2 s) in the Go trials.

Proportion of stimulus-responsive cells in the sensory Go/Nogo task: Stimulus-responsiveness was determined through a two-step procedure. First, we performed a signed-rank test to examine if the sensory response of each cell was significantly different from zero. In each cell, the sensory response during the stimulus presentation (2 s) was averaged in Go and Nogo trials. Then, the means of the sensory responses pooled across Go or Nogo trials

were used for a signed-rank test of each cell. Cells with statistical significance in the signed-rank test were selected for response thresholding (threshold: median z-score $> \pm 0.2$). In addition, for auditory stimulus trials, the sensory response during the stimulus onset (0.3 s) was averaged and analyzed in the same manner as described above to catch fast-adapting MGB neurons. If a neuron was classified as sensory responsive during the whole 2 s stimulus presentation period or as onset responsive, the neuron was included as auditory responsive. Neurons were then classified as Go, Nogo or ‘both’ responsive cells depending on the trial type. Neurons that did not pass the detection threshold were classified as non-responsive cells.

K-means cluster analysis: K-means cluster analysis was performed to sort individual neurons into functional subgroups. The calcium traces of matched individual neurons ($N = 6$ mice, $n = 210$ cells) were averaged across Go or Nogo trials for 2-3 sessions in each training phase (naive, expert, reversal naive and reversal expert). The mean Go or Nogo responses of each neuron were concatenated between naive and expert phases in both initial and reversal learning (time-series concatenation). The time-series for each cell was composed of the concatenated stimulus and delay periods of naive and expert training phases, while initial and reversal learning were treated as independent observations. This time series underwent principal component analysis followed by k-means clustering (cosine distance) to sort the individual neurons into functional clusters. Thirty clusters were generated, and clusters showing similar activity patterns were merged manually⁵. After generating the merged clusters, cell IDs in each merged cluster were separated to the initial and reversal learning data. To track the activity patterns of the MGB neurons between the two expert phases in initial and reversal learning, we performed k-means cluster analysis by using the data of the two expert phases in the initial and reversal learning (Fig. 2c, fig. S3a-c). After the preprocessing described above, the data of the two expert phases were concatenated between initial and reversal learning, then k-means clustering was performed.

Correlation analysis between behavioral variables and neuronal activity: Correlations between behavioral variables (locomotion, pupil size) and single cell Ca^{2+} traces were calculated as the Pearson’s correlation coefficient in Hit trials across learning (fig. S4a-f). The first session in the naive phase and the session with the highest d-prime in the expert phase were selected in both initial and reversal learning. Pupil, locomotion and calcium data of day-matched cells was down-sampled to 5 Hz. Cells that exhibited correlations between the behavioral variable and the calcium data of $|r| > 0.2$ and $p < 0.05$ were considered as significantly correlated. R-value distributions were compared between the naive and expert phase in both initial and reversal learning by a Kolmogorov-Smirnov test (ks-test2, MATLAB). To measure the change of correlation across learning, the difference of $|r|$ values between naive and expert phases in matched cells was measured for stimulus, delay and ITI periods (signed-rank test). The pupil data was analyzed in the same manner except that the pupil data was smoothed by a Hampel filter (see above) and z-scored (fig. S4b,d,f).

Anticipatory licking analysis: Anticipatory licking was quantified during the delay period. To visualize how the number of anticipatory licks change across learning, the proportion of trials with anticipatory licking was plotted as a function of the number of anticipatory licks per trial. The number of trials with or without anticipatory licking was pooled from all mice in each training phase, and this pooled data was used for chi-square testing (fig. S4g). The probability of anticipatory licks per time bin was calculated across mice (fig. S4h). For visualization purposes, lick probability was smoothed by a moving average (5 frames, movmean, MATLAB). To examine how anticipatory licking influenced MGB activity during the delay period, Ca^{2+} traces of Hit trials with and without anticipatory licking were separated and baselined to the mean Ca^{2+} fluorescence 150 ms before the delay period onset (fig. S4i, same baseline duration for fig. S4j).

Trial-by-trial population vector correlation: Before the construction of the population vector (PV), calcium data for each cell was baselined to the mean of the 2 s pre-stimulus period on a trial-by-trial basis. Next, the mean calcium responses during the stimulus and delay periods were calculated and the population vector constructed. Trial-by-trial population vector correlations (PVC, Pearson's) were calculated across the stimulus and delay periods of all trials and plotted as a N-by-N correlation matrix per session (N = number of trials). Mean correlation values from all trial pairs in each session were used for summary statistics. A linear mixed model was used to compare the difference of R values (fixed effect: naive vs expert, random effect: mouse ID). For a fair comparison between trials with and without anticipatory licking, only sessions with more than 6 trials with and without anticipatory licking each were included in the analysis (Fig. 3g). Reversal learning was excluded from the analysis of anticipatory licking due to low trial numbers. To examine the contribution of ramping cells to the PVC during the Go delay period ramp-up and ramp-down cells were removed from the PV, and PVC values of the Go delay period were re-calculated. The same number of cells as ramp-up and ramp-down cells were randomly removed to generate the shuffle dataset (nShuffle = 30, fig. S9).

Weighted graph-based analysis of MGB co-activity network structure: Undirected weighted graph that represents the relationship of calcium activity between MGB neurons were computed based on matched cell data (n = 210 neurons). In the graph, nodes and edges stand for cell identities and the pairwise Pearson's correlation coefficients between two neuron pairs, respectively. The calcium activity data was baselined to the mean of the 2 s pre-stimulus period of each trial. The pair-wise correlation coefficient of the time-series data (2 s) during the stimulus and delay periods was calculated between all cell pairs in Go and Nogo trials. Each graph with the number of nodes M, which are identical to the number of matched cells is represented by its adjacency M-by-M symmetric matrix C where each element r_{ij} is the Pearson's correlation coefficients ($-1 \leq r \leq 1$) between two nodes (neurons) i and j. Positive and negative edges were analyzed separately (fig. S11). The strength of the global co-activity network structure among MGB neurons was quantified by hubness, i.e., the mean of the pair-wise activity correlation of each node to all the others:

$$H_i = \frac{1}{m} \sum_{j=1}^m r_{i,j}$$

We estimated the global communication efficiency between any two nodes by measuring the geodesic (shortest) path length in the graph. First, a weighted graph with the length between two nodes, i and j as $L_{ij}=1/r_{ij}$ was created. Next, the shortest path length between any two nodes was computed using Dijkstra's algorithm among 3903 cell pairs. The local co-activity network structures were quantified by the cluster coefficient of the triad. The cluster coefficient of the triad was calculated among 50376 triads⁵⁹. w_{ij} is the edge weight between any two nodes. \hat{w}_{ij} is the normalized edge weight by the maximum weight in the adjacency matrix between cell i and j:

$$C_i = \frac{1}{k_i(k_i - 1)} \sum_{j,k} (\hat{w}_{ij} \hat{w}_{ik} \hat{w}_{jk})^{1/3}$$

$$\hat{w}_{ij} = w_{ij}/\max(w)$$

The same number of cells as ramp-up and ramp-down cells were randomly removed to generate the shuffle dataset ($n_{\text{Shuffle}} = 30$, fig. S10). Network structure was recalculated using the shuffle dataset. Functional connectivity parameters were calculated using the BrainConnectivity toolbox (<https://github.com/jblocher/matlab-network-utilities/tree/master/BrainConnectivity>). Plotting and confidence interval calculation were adapted from DABEST toolbox (<https://github.com/ACCLAB/DABEST-Matlab>).

General sensory responsiveness: To quantify the general sensory responsiveness of MGB neurons before and after the Go/Nogo learning in sensory mapping session (fig. S13), the calcium data during the stimulus presentation (2 s) was split in half (0 - 1 s and 1 - 2 s). The mean of the calcium data during these two periods was handled as individual data points. Next, it was tested whether individual MGB neurons were responsive to auditory, visual and/or multi-sensory stimulus through the same two-step procedures as described above. In multi-sensory trials, the data from all visual grating directions were averaged, and the sensory response to auditory frequencies were tested. In fig. S13c, cells were defined as excited/inhibited if they were responsive to any auditory frequency or grating direction. In fig. S13d, the response amplitude of the cells that were responsive to at least one of the sensory stimuli in pre-learning were compared to the response amplitude of the same sensory stimuli in post-learning sessions. In fig. S13e, change of peak tuning frequency or grating direction was calculated if a cell was responsive to any frequency/direction in both pre- and post-learning session. In fig. S13g and h, the time series of the averaged population response and the response amplitude of the cells with the significant response to the reward-associated stimulus in Go/Nogo training (12 kHz, rightward drifting grating, 12 kHz with rightward drifting grating) were compared between the pre- and post-learning session. In fig. S13i, the change of peak tuning frequency or grating direction was calculated if a cell was responsive to the conditioned stimulus (12 kHz, rightward-drifting grating, 12 kHz with rightward drifting grating) both in pre- and post-learning session.

K-means clustering of tuning curves: To sort MGB neurons into groups with similar frequency tuning patterns, k-mean clustering ('correlation' distance) was performed with 20 features (5 frequencies, multi-/uni-sensory and pre/post training) (fig. S15). Sensory responses to the auditory (4, 8, 12, 16 and 20 kHz) and multi-sensory stimuli (4, 8, 12, 16 and 20 kHz with drifting gratings) were averaged across trials and sessions (1-3 sessions). Visual stimulus feature (i.e., grating directions) was collapsed in multi-sensory trials. The calcium data in the response period was baselined to 0.5 s before stimulus presentation. The mean of the calcium activity during the stimulus presentation (2 s) was used to generate frequency tuning curves. F : mean calcium response during stimulus presentation period. m : cell number (pooled across 6 mice). f : auditory frequency. Clustering matrix:

$$\begin{pmatrix} pre_{multi}F_{f1}^1 & \dots & pre_{multi}F_{f5}^1 & post_{multi}F_{f1}^1 & \dots & post_{multi}F_{f5}^1 & pre_{uni}F_{f1}^1 & \dots & pre_{uni}F_{f5}^1 & post_{uni}F_{f1}^1 & \dots & post_{uni}F_{f5}^1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ pre_{multi}F_{f1}^m & \dots & pre_{multi}F_{f5}^m & post_{multi}F_{f1}^m & \dots & post_{multi}F_{f5}^m & pre_{uni}F_{f1}^m & \dots & pre_{uni}F_{f5}^m & post_{uni}F_{f1}^m & \dots & post_{uni}F_{f5}^m \end{pmatrix}$$

$$m = 1, 2, \dots, 233$$

$$f1 \dots f5: 4, 8, 12, 16, 20 \text{ kHz}$$

To examine if frequency tuning curves were stable or plastic, similarity of the frequency tuning curves between pre- and post-learning sessions for each cell was quantified by Pearson's

correlation (R) in auditory and multi-sensory trials separately. The distribution of Pearson's R pooled across all cells was compared between auditory and multi-sensory trials by ks-test (fig. S15c) to examine in which trial the frequency tuning curve would be more plastic or stable.

Multi-sensory index: The multi-sensory index was calculated by the division of the average response in multi-sensory trials (AV) to the sum of average response in uni-sensory trials (A: auditory trials, V: visual trials). Cells with a non-significant response in both AV and A trials were excluded from multi-sensory index calculation. Cells with different calcium response sign in multi-sensory response and sum of uni-sensory response were excluded ($A_f V \cdot (A_f + V) < 0$). Absolute value of sensory response was taken to demonstrate the principle of multi-sensory integration (linear or non-linear). Visual grating directions were averaged in multi-sensory response as well as uni-sensory response. First order exponential fitting was conducted to reveal the trend of data distribution for each auditory frequency (fig. S15d). Multi-sensory indexes of all individual cells and all auditory frequencies were compared regardless of response sign and amplitude (fig. S16). Two-dimensional two-sample ks-test was performed to test if the distribution of the multi-sensory index was comparable between pre- and post-learning (See, Table S1).

$$\text{Multisensory index}_f = \frac{|A_f V|}{|A_f + V|} \quad f = 4, 8, 12, 16, 20 \text{ kHz}$$

Single cell correlation analysis across multisensory mapping sessions: Single cell averaged peri-stimulus time histograms (PSTH) of uni- and multisensory trials for stimulus features (auditory frequency, grating direction) in pre- and post-learning were sorted by stimulus amplitude (0 – 2 s) and plotted as heat maps (fig. S14). Next, the Pearson's correlation between the response time series of individual neurons before and after learning was calculated and average across all cells (fig. S14)⁶⁰.

Decoder analyses: Linear support vector machine (SVM) decoders were trained on the stimulus period (normalized by 1 s pre-stimulus baseline) for all tracked neurons in each multi-sensory mapping session for each animal. Time series data were down-sampled to 3 Hz to avoid overfitting. The decoders trained on individual sessions were tested in a pair-wise manner for all sessions (fig. S17a). For modality decoding (fig. S17b), decoders were trained on 70 % of the data and tested on a 30 % hold-out test-set. Same numbers of trials from each modality were selected for training and testing in 50 iterations to avoid a biased accuracy measurement due to unbalanced trial number in each modality (50 % multi-sensory, 16.7 % auditory, 13.3 % visual). For auditory frequency decoding (fig. S17c), trial sub-sampling was not necessary since each auditory frequency was presented for a same number of trials. Uni-sensory and multi-sensory trials were combined in one dataset. Visual stimulus feature was collapsed in multi-sensory trials leaving only frequency labels (4, 8, 12, 16 and 20 kHz). Train/test ratio remained 70/30 (140 training trials, 60 test trials). Simple accuracy (modality: $x / 27$ trials, frequency: $x / 60$ trials, x = number of correct decoded trials) in all iterations were averaged yielding pair-wise session decoding accuracy. Shuffle data ($n = 100$) were generated by circular permutation of the down-sampled time series features for each train/test iteration ($n = 50$) from which mean accuracy values of shuffle iterations were taken to depict chance decoding accuracy and further averaged to yield pair-wise session decoding accuracy (fig. S17b,c).

Histology: After training, mice were transcardially perfused with phosphate buffer saline (PBS) followed by ca. 40 ml 4 % paraformaldehyde (PFA) in PBS. Immediately after perfusion, brains were removed and post-fixed in 4 % PFA overnight at 4 °C. Then, brains were stored in PBS at 4 °C until dissection. 150 µm coronal slices were prepared using a vibratome (Campden Instruments) and immunostained for calretinin as an anatomical marker as described previously⁵. Briefly, after PBS washes, brain slices were immersed in blocking solution (10% normal horse serum, S-2000-20, Vector Laboratories) with 0.5 % Triton (T8787, Sigma-Aldrich) in PBS for 2 hours at room temperature. Next, slices were incubated in primary antibody (goat anti-calretinin, 1:1000, CG1, Swant) in carrier solution (1% normal horse serum with 0.5% Triton PBS) overnight at 4 °C. Slices were washed again in 0.5% Triton PBS and incubated for 2 h at room temperature or overnight at 4 °C in secondary antibody (donkey anti-goat 647, 1:1000, A-21447, ThermoFisher) in carrier solution. After final washes by PBS, slices were mounted on slides and cover slipped using 22 x 50 mm, 0.16 - 0.19 mm thick cover glass (FisherScientific). Images were acquired with a LSM700 confocal microscope (Zeiss), Axio Scan Slide Scanner (Zeiss) or Olympus BX63. Acquired images were post-processed with ImageJ (<https://imagej.nih.gov/ij/>) to locate the implantation site of the GRIN lenses.

Statistical methods: Statistical analysis was performed in MATLAB (MathWorks). Alpha level was set at 0.05 and Bonferroni correction was applied to statistical tests (see Table S1). Chi-square test, rank sum test, signed-rank test, linear mixed model (LMM), two-sample Kolmogorov-Smirnov test (ks-test2) and two-dimensional two-sample Kolmogorov-Smirnov test (<https://github.com/brian-lau/multdist>) were performed for datasets indicated in Table S1. Data are presented as mean ± SEM unless otherwise stated. Statistical results and p values are presented in Table S1.

References

50. Dana, H. *et al.* High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).
51. Leinweber, M. *et al.* Two-photon calcium imaging in mice navigating a virtual reality environment. *J. Vis. Exp.* e50885 (2014).
52. Reiff, D. F., Plett, J., Mank, M., Griesbeck, O. & Borst, A. Visualizing retinotopic half-wave rectified input to the motion detection circuitry of *Drosophila*. *Nat. Neurosci.* **13**, 973–978 (2010).
53. Marbach, F. & Zador, A. M. A self-initiated two-alternative forced choice paradigm for head-fixed mice. *bioRxiv* 073783 (2016) doi:10.1101/073783.
54. He, C. *et al.* Complex vectorial optics through gradient index lens cascades. *Nat. Commun.* **10**, 1–8 (2019).
55. Pachitariu, M. *et al.* Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv* 061507 (2017) doi:10.1101/061507.
56. Macmillan, N. A. & Creelman, C. D. Detection theory: A user's guide, 2nd ed. <https://psycnet.apa.org › recordhttps://psycnet.apa.org › record> **2**, 492 (2005).
57. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
58. Reimer, J. *et al.* Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron* **84**, 355–362 (2014).

59. Onnela, J.-P., Saramäki, J., Kertész, J. & Kaski, K. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **71**, 065103 (2005).
60. Peyrache, A., Benchenane, K., Khamassi, M., Wiener, S. I. & Battaglia, F. P. Principal component analysis of ensemble recordings reveals cell assemblies at high temporal resolution. *J. Comput. Neurosci.* **29**, 309–325 (2010).

Movie S1.

An example movie of Hit trials in naïve and expert phases. For visualization, the images are cropped and resized. The movie of the calcium imaging was also temporally smoothed.

Data S1.

Source data for Figures in main text (Figs. 1-4).

Data S2.

Source data for Supplementary Figures.

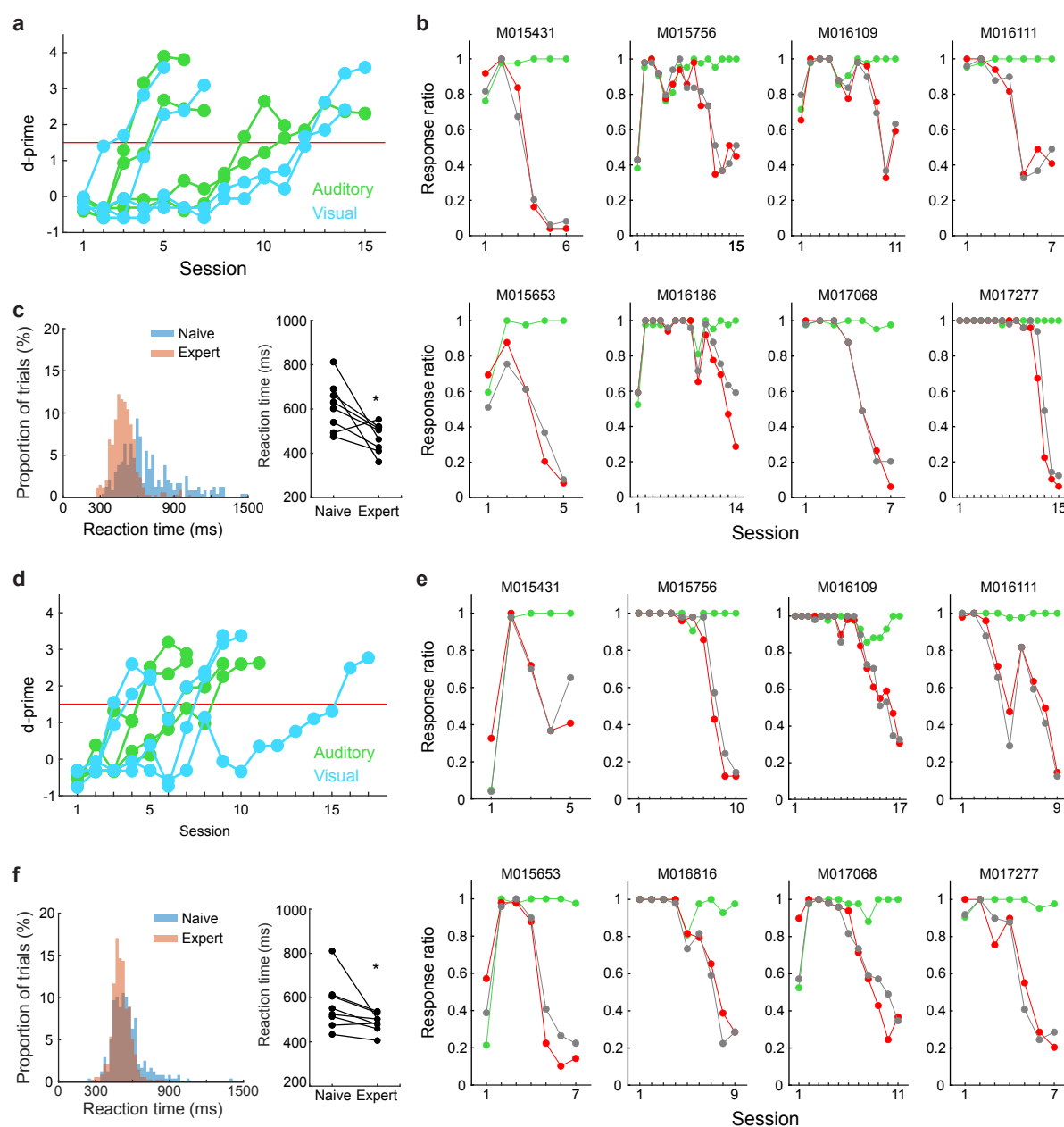


fig. S1 | Task performance of individual mice during the sensory Go/Nogo task. (a-c) The task performance of individual mice and Hit reaction time (RT) in the initial learning. (a) Transition of d-prime from all mice. (b) Transition of the task performance of individual mice. Top row: Auditory-reward group. Bottom row: Visual-reward group. (c) Left: Hit RT pooled across all mice in the naive and expert phases. Right: Transition of median Hit reaction time. Each dot represents the individual mouse data. RT in the expert phase was shorter than that in the naive phase (signed-rank test). The first behavioral session in the naive phase and the session with the highest d-prime in the expert phase were used for the RT analysis. (d-e) The task performance of individual mice and Hit RT in the reversal learning. The figure configurations are the same as a-c. Reaction time in the expert phase was shorter than that in the naive phase (signed-rank test). * in the figures represent the statistical significance. *, $p < 0.05$.

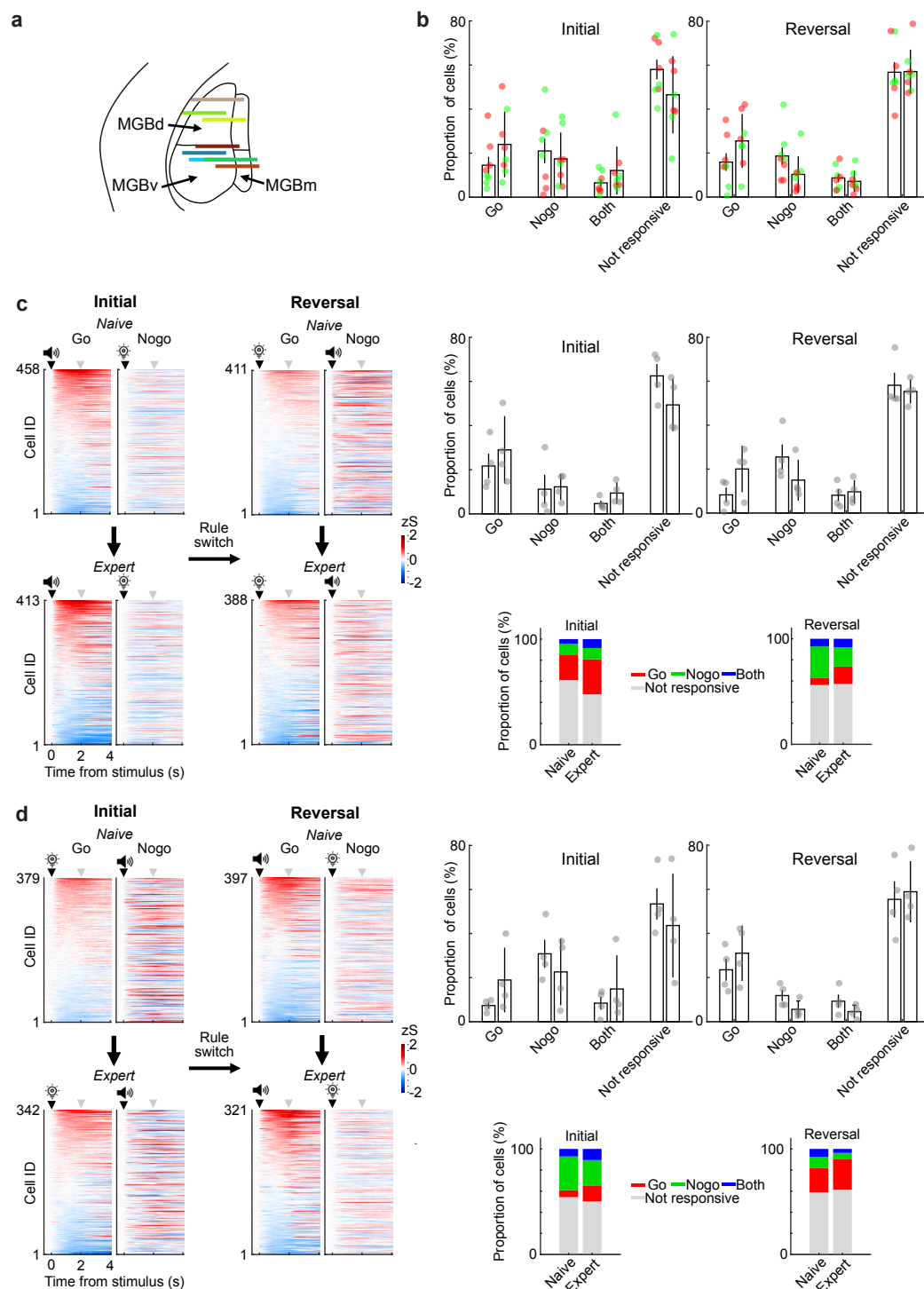


fig. S2 | Activity of individual cells and the proportion of stimulus-responsive cells in auditory and visual reward groups.

(a) GRIN lens front and putative field of view (FOV) from all mice (N = 8 mice). Each line represents the lens front and the putative FOVs of the two-photon image. For the 1.0 mm lens, the FOV is approximately 330 μ m x 330 μ m. For the 0.6 mm lens, the FOV is approximately 400 μ m x 400 μ m. Due to the distortion of the image through the GRIN lens, these FOV values do not necessarily reflect the actual size of the brain image. (b) The proportion of stimulus-responsive cells from all mice (N = 8 mice). (c) Mean individual cell responses and the proportion of stimulus-responsive cells from Auditory-reward (initial learning) → Visual reward (reversal learning) group. Left: Heatmaps of the mean individual cell activities in Go and Nogo trials of a single representative session during the initial learning (n = 458 cells in the naive phase, n = 411 cells in the expert phase cells from 4 mice) and reversal learning (n = 413 cells in the naive phase, n = 388 cells in the expert phase from 4 mice). Cells were sorted by the averaged response amplitude during the stimulus presentation in Go trials. Cell IDs are matched across Go and Nogo trials in each learning phase. Black and gray triangles represent the stimulus and delay period onset, respectively. Right top: The proportion of stimulus-responsive cells in initial and reversal learning. Right bottom: The stacked-bar charts of the stimulus-responsive cells from naive to expert phase in the initial and reversal learning. The proportions of stimulus responsive cells was altered from the naive to the expert phase across learning in both initial and reversal learning ($p < 0.01$, chi-square test, in both initial and reversal learning). (d) Mean individual cell responses and the proportion of stimulus-responsive cells from Visual-reward (initial learning) → Auditory reward (reversal learning) group. The figure configurations are the same as (c). The proportions of stimulus responsive cells were altered from the naive to the expert phase across learning in both initial and reversal learning ($p < 0.01$ for Initial and $p < 0.05$ reversal learning, after Bonferroni correction, chi-square test).

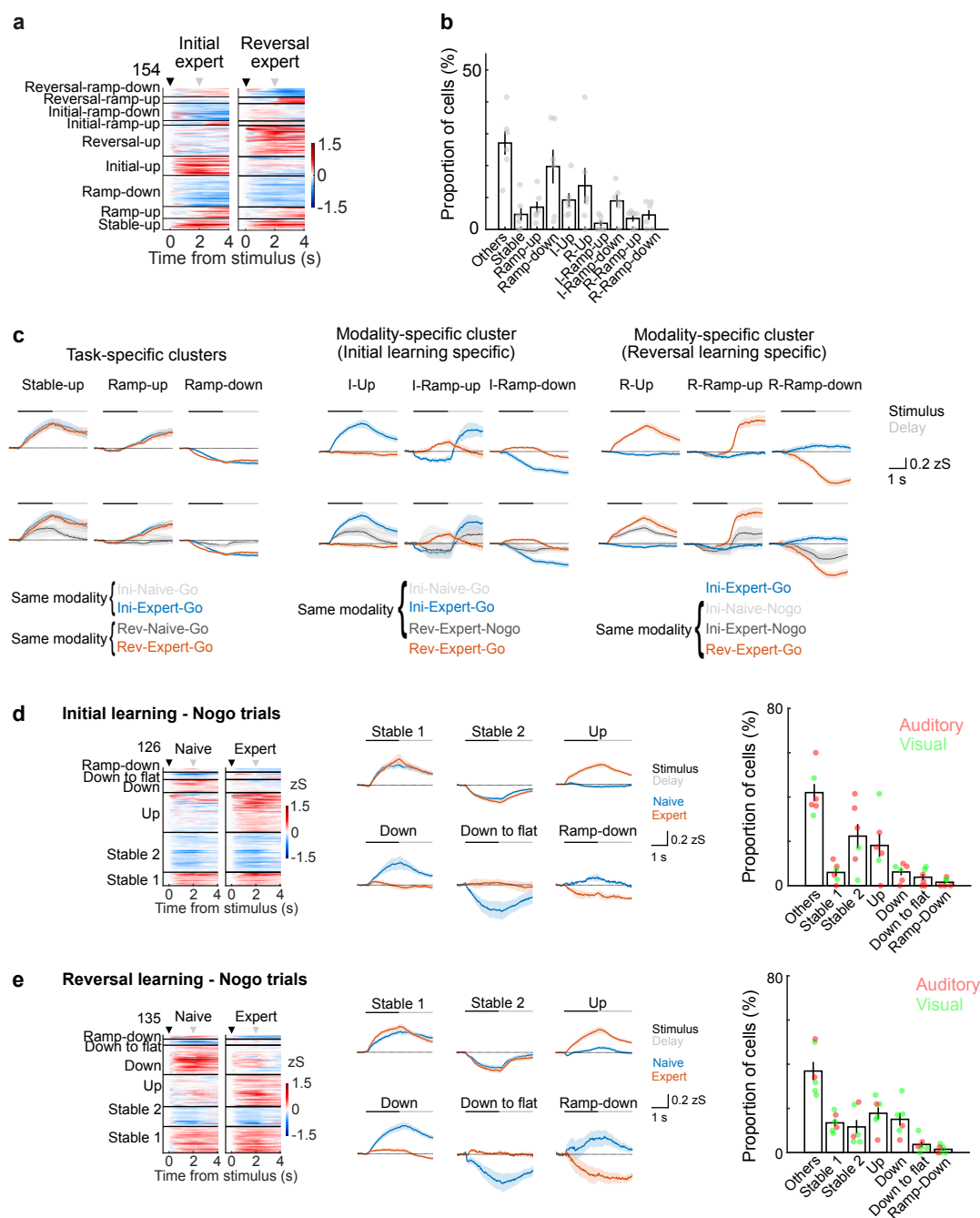


fig. S3 | Task-specific and modality-specific clusters in Go trials and functional clusters in Nogo trials.

(a-c) Functional subgroups with the task and modality-specific learning plasticity in the two expert phases in both initial and reversal learning. (a) Heatmaps of single cell activities in the two expert phases (N = 6 mice, n = 154 cells). (b) Proportion of cells in each functional cluster. Each dot represents the data from the individual mouse. (c) Average calcium traces (mean \pm SEM) of the functional subgroups shown in (a). Top row: Average calcium traces in the two expert phases. Left: Task-specific clusters. Middle: Modality-specific clusters showing the plasticity in the initial learning. Right: Modality-specific clusters showing the plasticity in the reversal learning. Bottom row: Average calcium traces of the two expert phases with additional traces of other conditions with the same sensory modality. The additional traces were overlaid to show how reward learning modulated the neuronal activity within the same sensory modality. The figure configuration from Left to Right is the same as the top row. (d) Functional subgroups of MGB neurons in Nogo trials in initial learning. Left: Heatmaps of single cell activities in the naive and expert phases in initial learning (N = 6 mice, n = 126 cells). Middle: Average calcium traces (mean \pm SEM) of the functional subgroups. Right: Proportion of cells in each cluster. Each dot represents the data from the individual mouse. Red and green dots represent the types of stimulus-reward condition. (e) Functional subgroups of MGB neurons in Nogo trials in reversal learning (N = 6 mice, n = 135 cells). The figure structures are the same as (d).

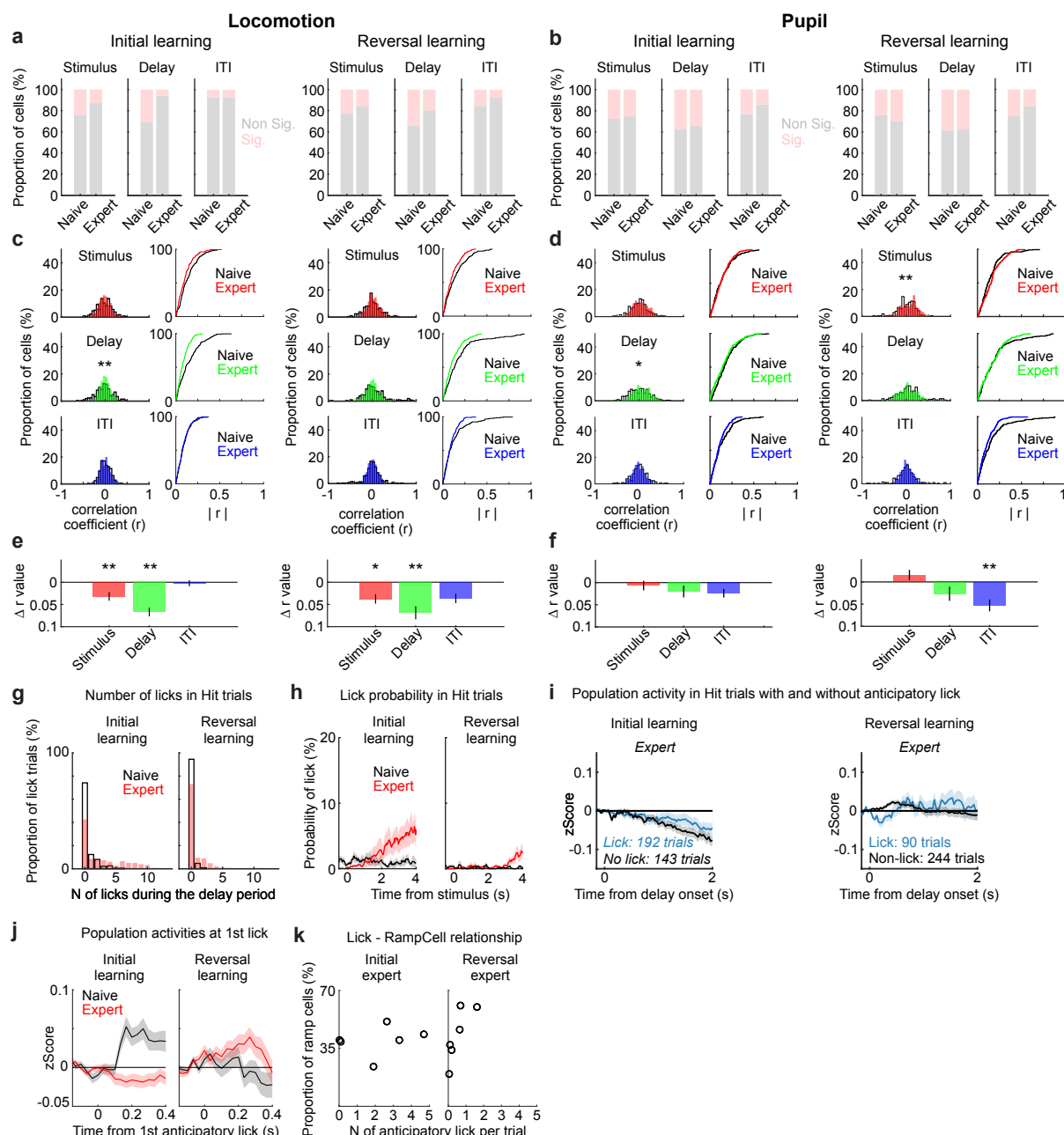


fig. S4 | Behavioral variables are not the main source of MGB plasticity. (a, c and e) Correlation analysis of Ca^{2+} trace and locomotion in Hit trials in initial and reversal learning ($N = 6$ mice, $n = 210$ cells). The longitudinally tracked cells were used for this analysis. (a) Left: The proportion of the significantly correlated cells in each time window (Stimulus, Delay and ITI periods) in the initial learning (Left) and reversal learning (Right). (c) The distributions of the locomotion - Ca^{2+} correlation between the naive (black) and expert phases (color) in the initial learning (Left) and reversal learning (Right) from all cells. Top: The probability density function during the stimulus period. Middle: Probability density function during the delay period. Bottom: Probability density function during the ITI period. ITI periods is 2 s before the stimulus presentation. (e) The change in the magnitude of correlation. The difference of the absolute r value for each MGB neuron between the naive and expert phases (Mean \pm SEM) was calculated. (b, d and f) Correlation analysis of Ca^{2+} trace and pupil size in Hit trials in the initial and reversal learning ($N = 6$ mice, $n = 210$ cells). Figure configurations are same as a, c and e for locomotion analysis. (g) The proportion of Hit trials with anticipatory licking during the delay period across all mice ($N = 8$ mice). The number of trials with anticipatory licking increased from the naive to expert phases in both initial and reversal learning ($p < 0.01$ for initial and reversal learning, after bonferroni correction, chi-square test). (h) The probability of anticipatory licking at each time bin (Mean \pm SEM) in Hit trials from all mice ($N = 8$ mice). The probability of anticipatory lick was calculated at each time bin across Hit trials for each mouse, then averaged across all mice. (i) Average population activities of MGB neurons (Mean \pm SEM) during the delay period in Hit trials with and without anticipatory licking across all mice ($N = 8$ mice). Left: Population activities in the initial expert (Lick trial = 192 trials and $n = 755$ cells, non-lick trial = 143 trials and $n = 755$ cells), Right: Population activities in the reversal expert (Lick trial = 90 trials and $n = 637$ cells, non-lick trial = 244 trials and $n = 709$ cells). (j) Average population activities of MGB neurons (Mean \pm SEM) aligned at the first anticipatory licking onset during the delay period. Left: Average population activities of MGB neurons in the initial learning (61 trial, $n = 749$ cells for naive, 192 trials, $n = 755$ cells for expert). Right: Average population activities in the reversal learning (13 trials, $n = 604$ cells for naive, 90 trials, $n = 637$ cells for expert). (k) The proportion of ramping cells (ramp-up and ramp-down cells) found in k-means analysis as a function of the mean of the number of anticipatory licking during the delay period per trial in Hit trials. * and ** in the figures represent the statistical significance after Bonferroni correction. *, $p < 0.05$, **, $p < 0.01$.

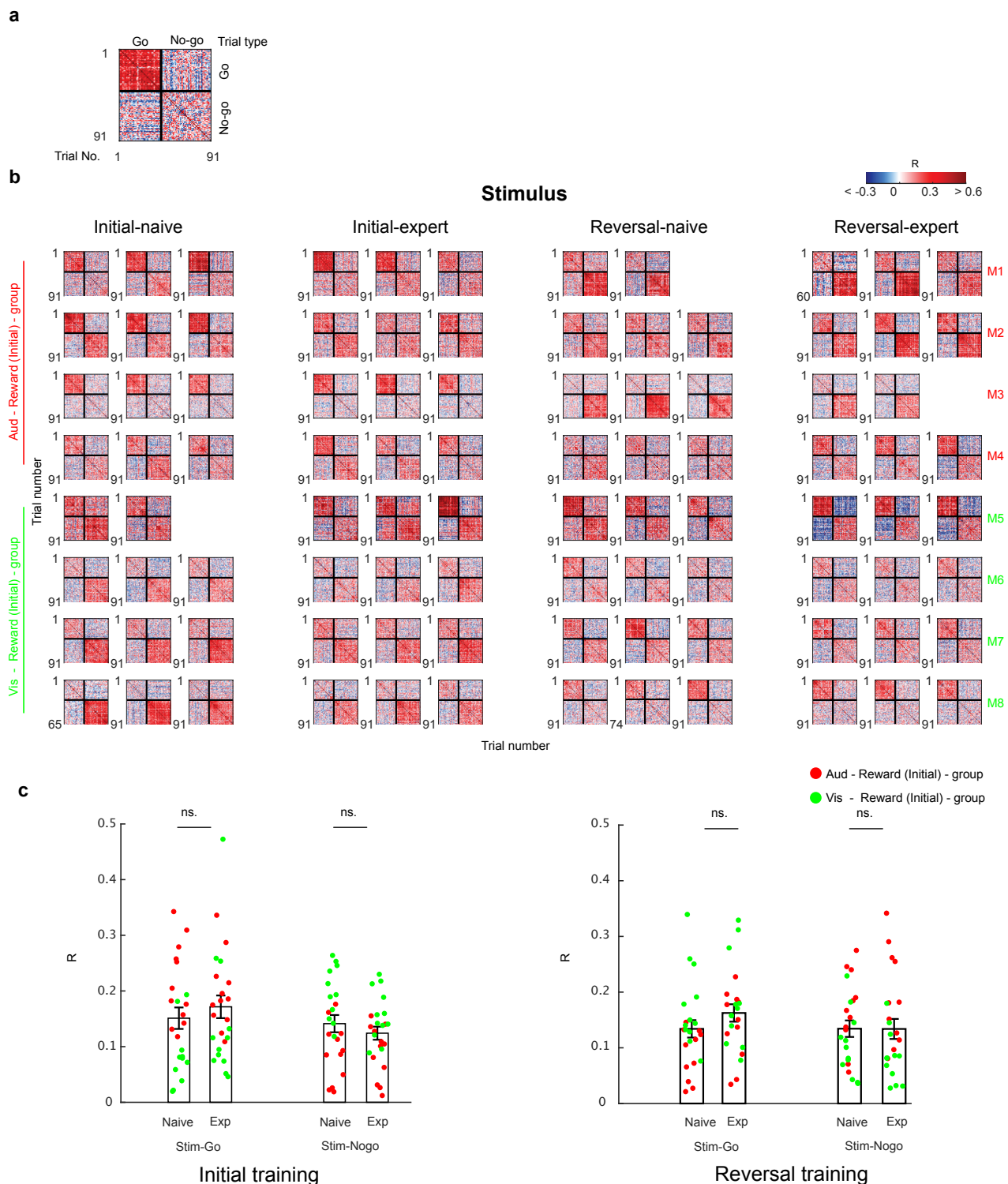
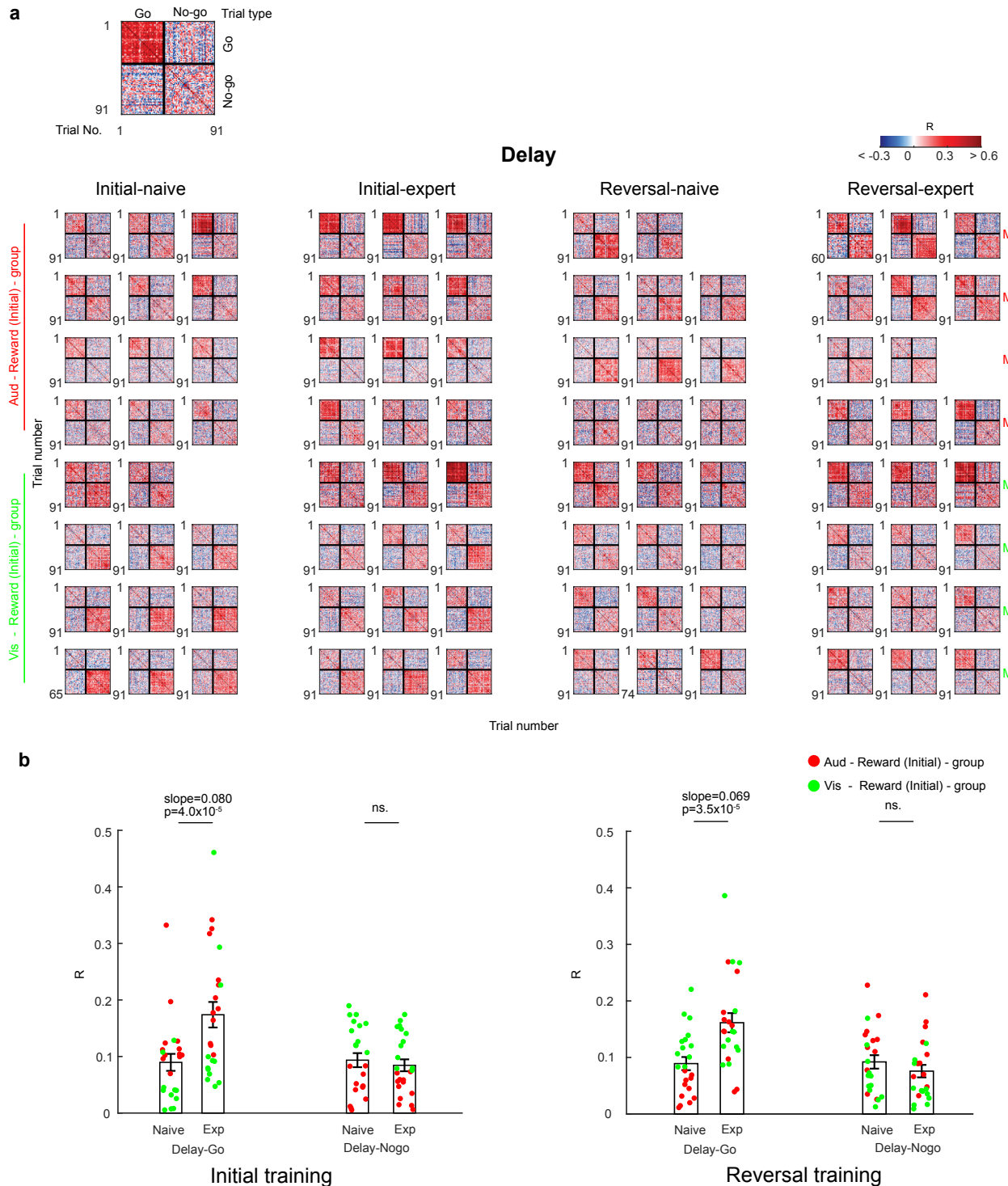


fig. S5 | Stable evolution of MGB activity patterns during the stimulus period in the Go/Nogo learning task. (a) Representative trial-by-trial correlation matrix within and between Go/Nogo trials in one session. (b) Stimulus activity correlation matrices from all mice in different learning stages (Naive/Expert/Reversal-naive/Reversal-expert). Each matrix represents pair-wise population vector correlation (PVC) for all trials in one learning session. (c) Mean correlation value (R) in each session shown in (b) was grouped into initial (left) and reversal training (right). PVC in stimulus period in Expert and Reversal-expert did not significantly differ from Naive or Reversal-naive condition respectively (details in Table S1). Visualization of counterbalanced stimulus-reward pairing by color of dots. Both groups showed consistent trend of activity pattern. Red: Mice trained on auditory-reward pairing in initial training session. Green: Mice trained on visual-reward pairing in initial training session.



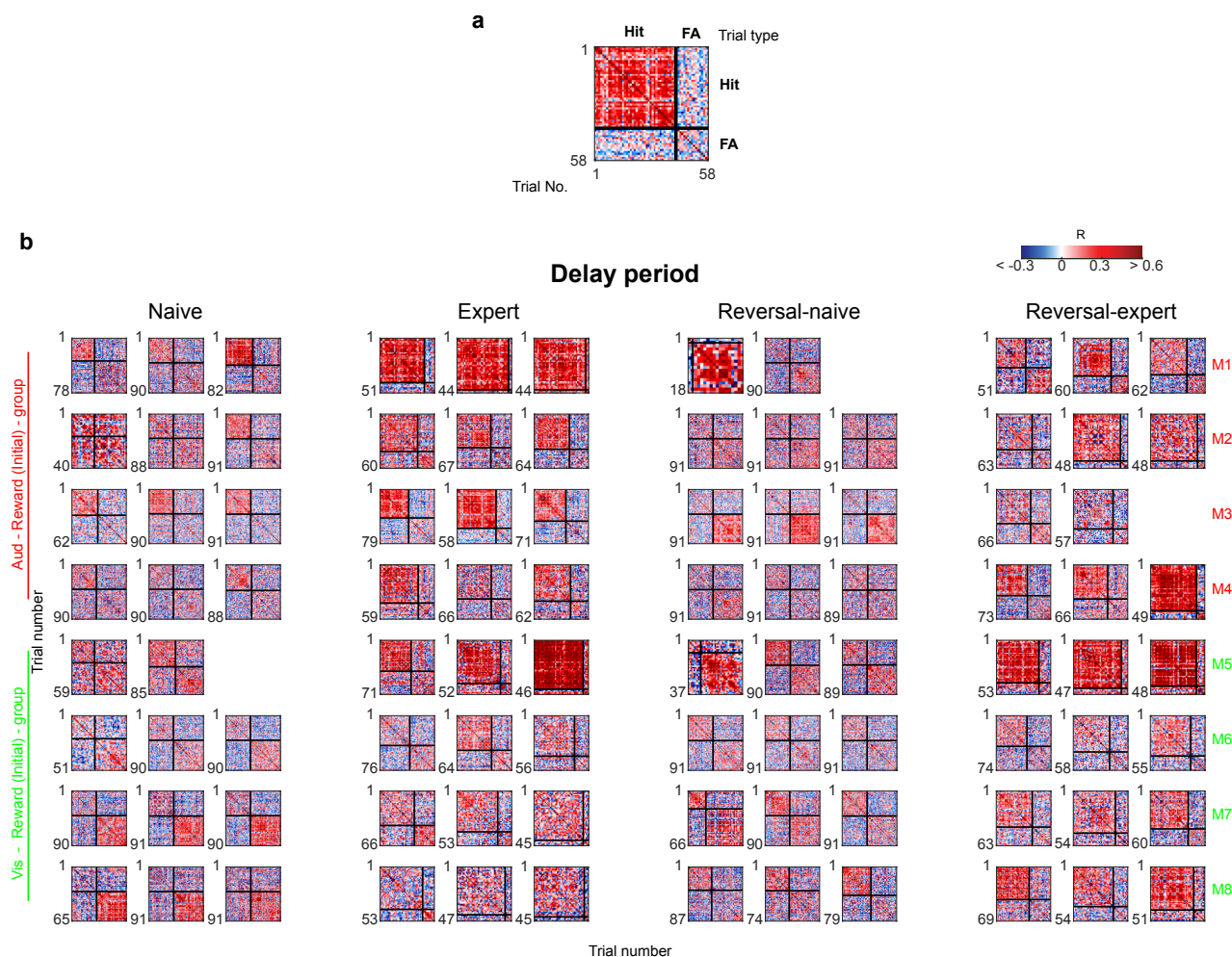


fig. S7 | Coherence of MGB population activity is increased during the delay period of Hit but not False Alarm (FA) trials. (a) Representative trial-by-trial PVC matrix within and between Hit and False Alarm trials of one session. (b) Delay activity PVC matrices from all mice in different learning stages (Naive/Expert/Reversal-naive/Reversal-expert). Each matrix represent pair-wise PVC for all trials in one learning session. Summary data were shown in Fig.3 e, f and Table S1.

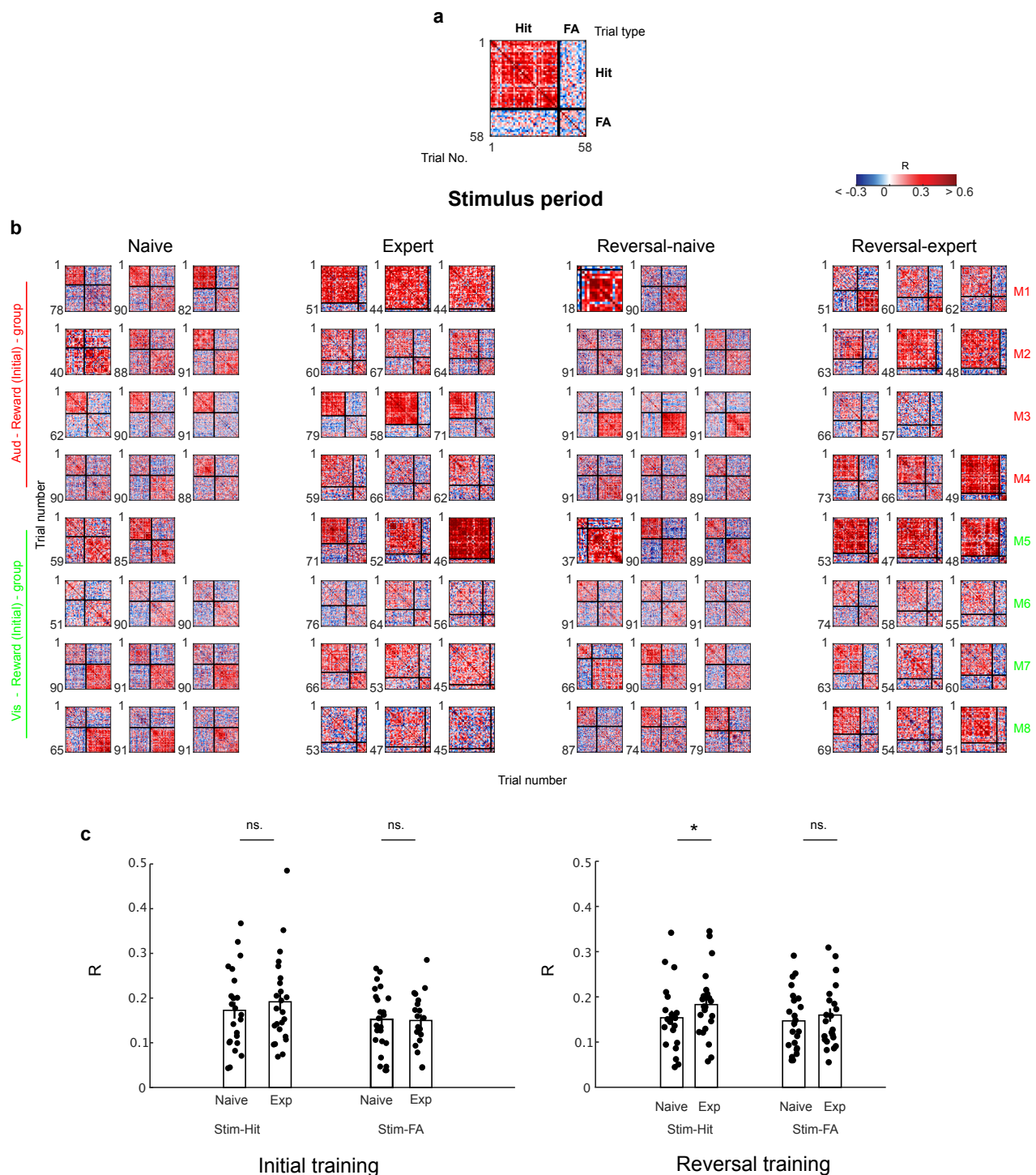


fig. S8 | Coherence of MGB population activity is not increased during the stimulus period of Hit and False Alarm (FA) trials, but remodeled after reversal learning. (a) Representative trial-by-trial PVC matrix within and between Hit and False Alarm trials of one session. (b) Stimulus PVC matrices from all mice in different learning stages (Naive/Expert/Reversal-naive/Reversal-expert). Each matrix represent pair-wise PVC for all trials in one learning session. (c) Summary data. (N = 8, details in Table S1).

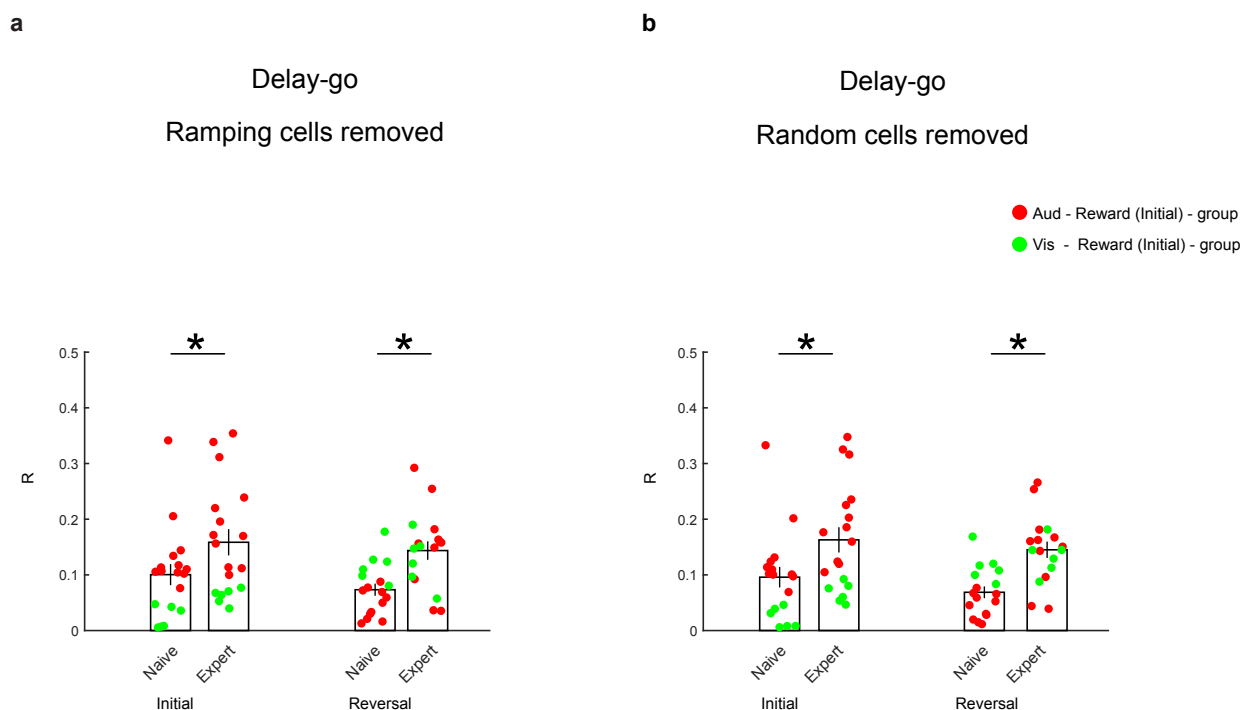


fig. S9 | Increase of population vector correlation (PVC) in MGB during learning depends on MGB population activity and is not driven by ramping-cells. (a) Mean trial-by-trial PVC (R) during the delay period of Go trials in different learning stages. Increase of PVC persisted even if ramping cells (ramp-up and ramp-down, Fig. 2a, b) were removed (* $p < 0.01$, details in Table S1). (b) Increase of PVC was dependent on the MGB population activity instead of subsets of neurons. Same number of cell as ramping cells (ramp-up + ramp-down) in each mouse were randomly removed before calculating PVC (number of random cells = 30, * $p < 0.01$, details in Table S1).

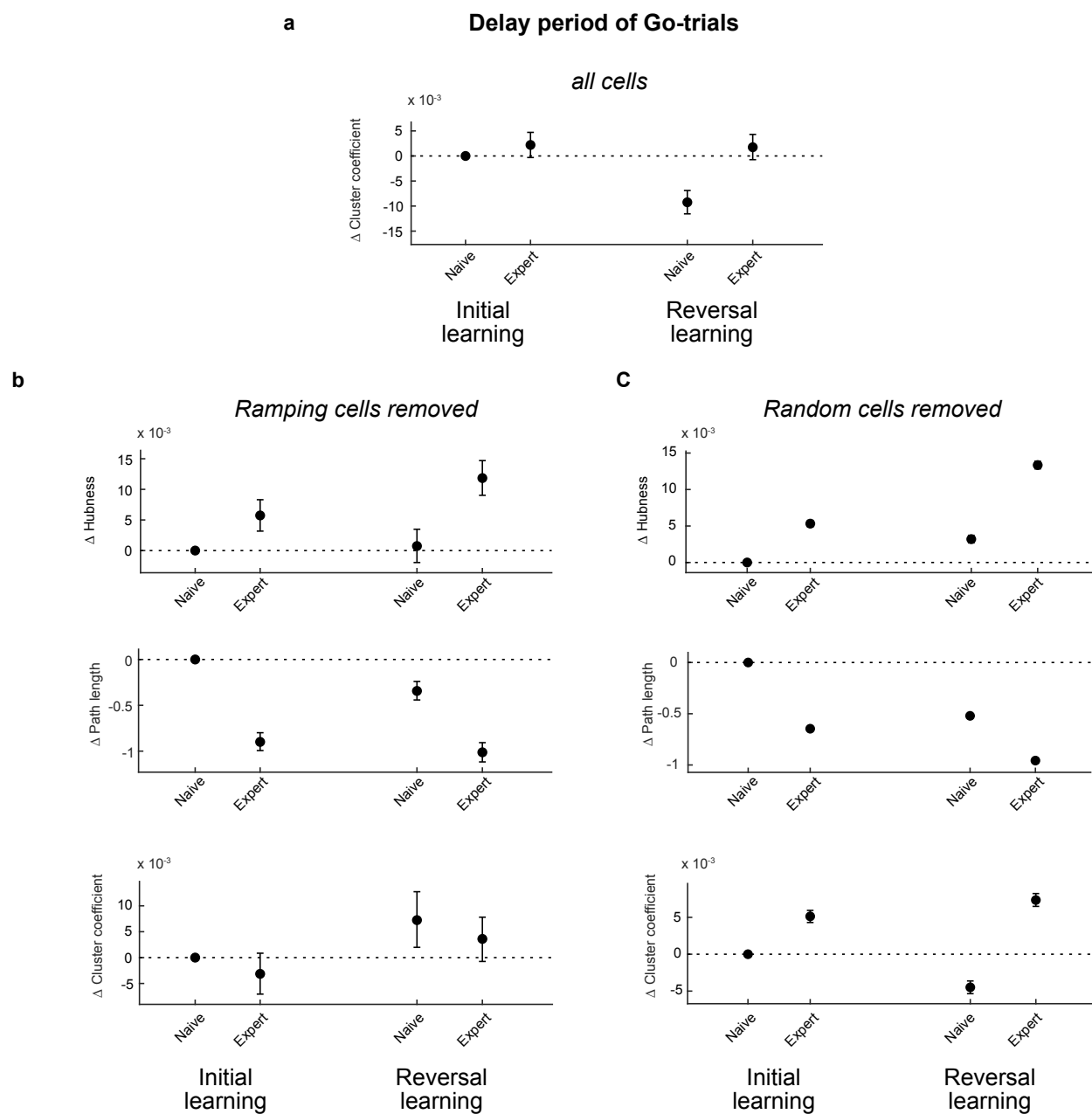


fig. S10 | MGB co-activity network structure does not depend on ramping cells. (a) Decrease of local connectivity was observed specifically in reversal-naive learning stage when reward contingency was changed, suggesting a re-organization of pre-configured connectivity that was established in Initial-expert stage. (b) Ramp-cells were not the sole contributors of connectivity re-organization. Removing ramp-cells (ramp-up and ramp-down, Fig. 2a, b) or random cells did not abolish the change of network architecture which was an indication of MGB general ensemble remodeling during reward-association learning task (detail in Methods).

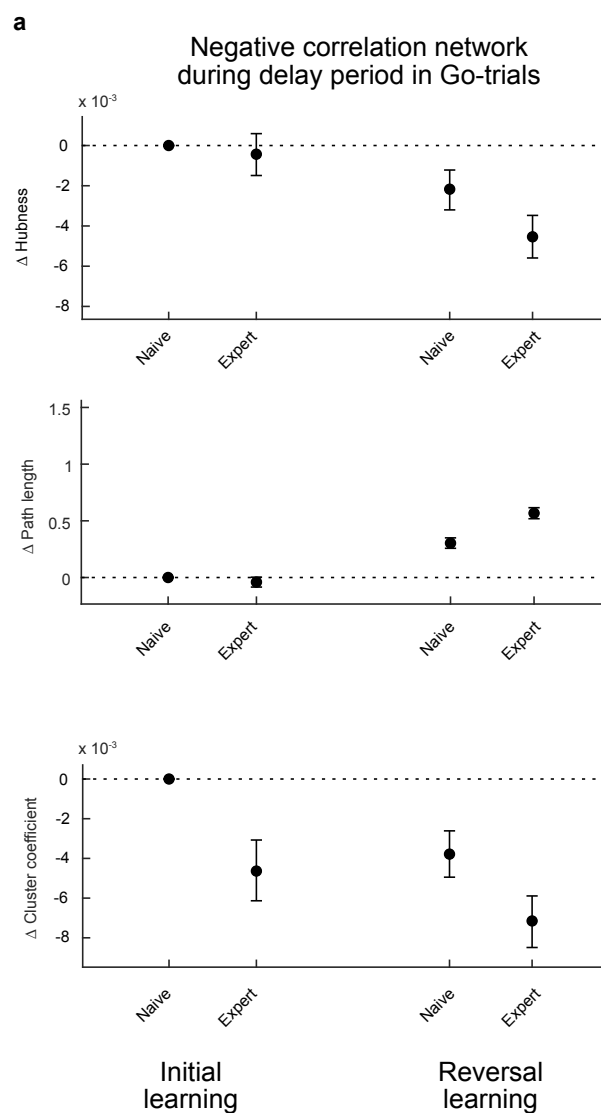


fig. S11 | Negative correlation co-activity network during delay period in Go-trials. (a) There was no prominent change of global negative co-activity structure from Naive to Expert stage in Initial and Reversal learning phase which was in accordance to the increase of global positive co-activity network (detail in Methods). Local negative co-activity network became less connected from Naive to Expert stage suggesting an increase of connectivity during delay period when the mouse learned the task structure. Error bars: 95 % confidence interval of mean.

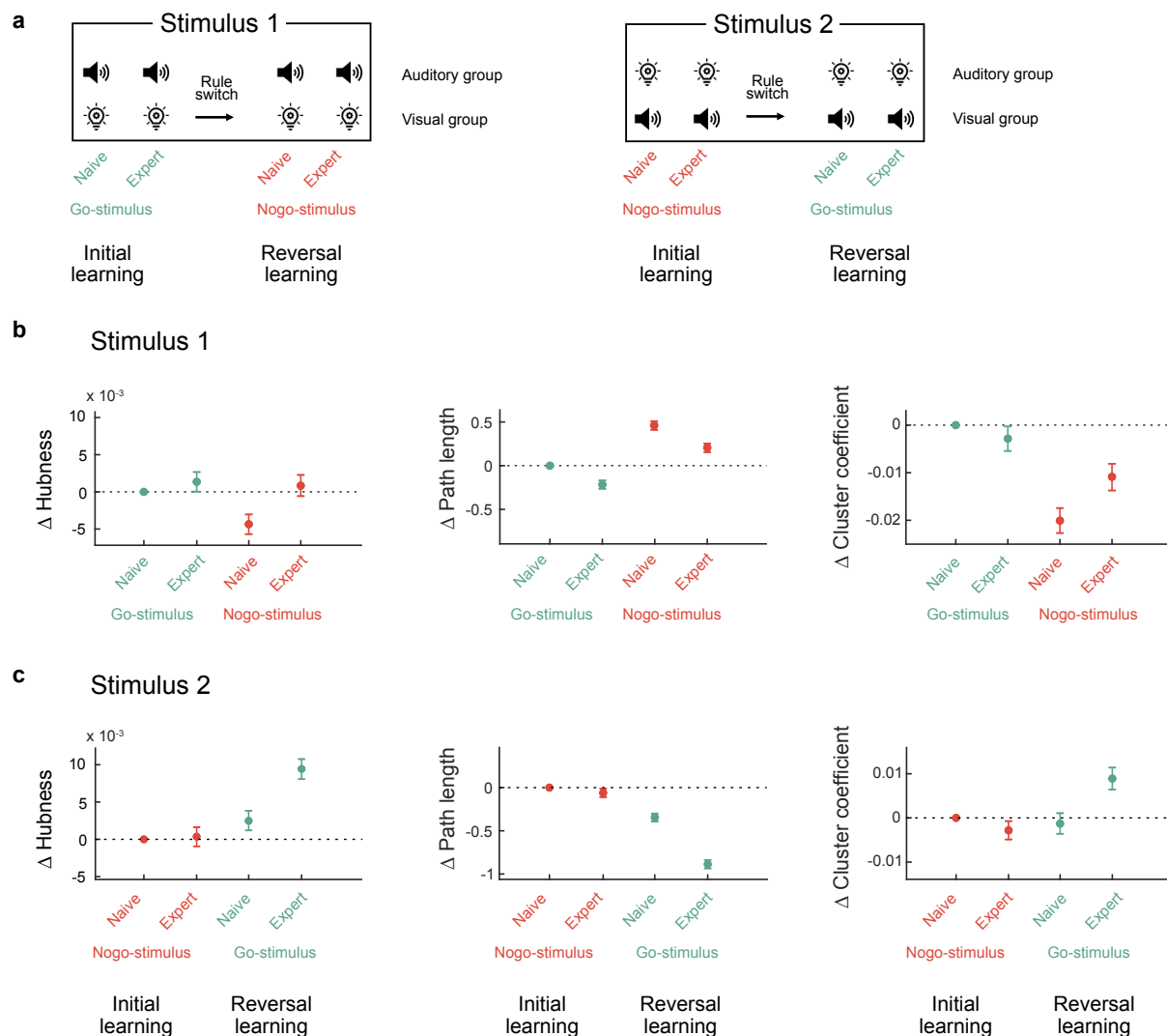


fig. S12 | MGB co-activity networks during the stimulus period. (a) Experimental design. (b and c) Change of hubness, path length and cluster coefficient across all neuron pairs in relation to the naive Go (or) Nogo condition (detail in Materials and Methods). There was no obvious change of global and local connectivity in initial training phase during stimulus presentation period for within group identical stimulus 1 and 2. Conversely in reversal learning, global and local connectivity of both Go and Nogo stimuli were strengthened. Neural representation of stimuli that were associated or dissociated with reward contingency were remodeled in the reversal naive phase. Error bars: 95 % confidence interval of mean.

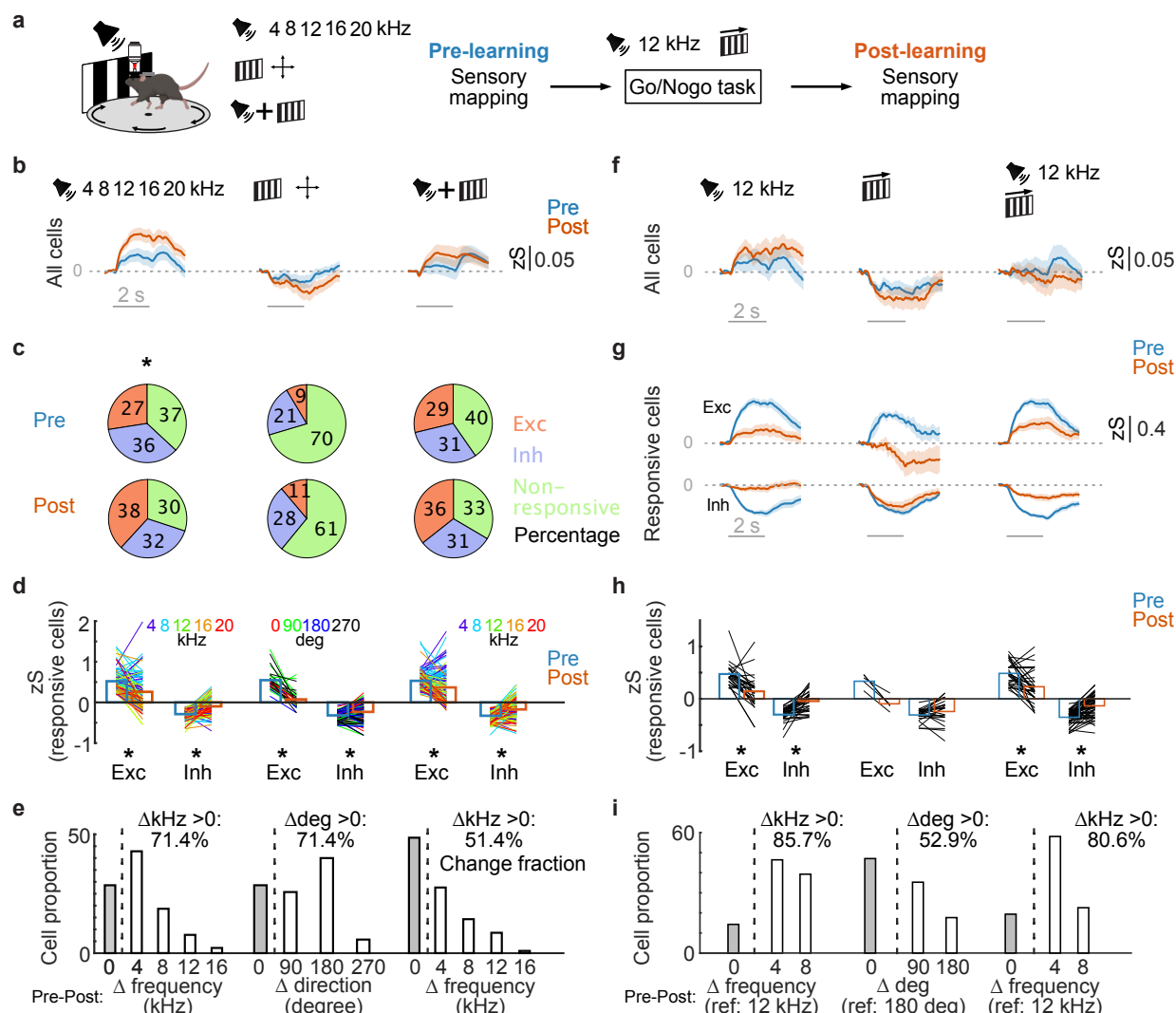


fig. S13 | Sensory coding in MGB is not biased to the reward-associated stimuli in non-rewarded sensory mapping sessions. (a) Schematics of sensory mapping. Various auditory and visual stimuli were presented as unisensory stimulus, and the combination of them were presented as multisensory stimulus. (b) Average population calcium activity in uni- and multi-sensory trials before and after learning ($n = 233$ cells, 6 mice). (c) Proportions of the excited and inhibited cells to auditory, visual and multi-sensory stimuli. (d) Transition of the sensory response amplitude from the cells showing the significant response at pre-learning stage were plotted. Each line represents sensory responses before and after learning from the same cell for a specific stimulus feature (tone frequency or grating direction). (e) Changes of the best frequency/direction tuning across learning. (f) Averaged population calcium activity to the reward-conditioned sensory stimuli (12 kHz tone, rightward grating, combination of them, $n = 233$ cells, 6 mice). (g) Average population responses of the excited or inhibited cells to the conditioned stimuli. (h) Transition of the sensory response amplitude to the conditioned stimuli. Each line represents the sensory response of the same cell before and after learning. (i) Changes of the best frequency/direction tuning referred to the conditioned stimuli. Statistical details in Table S1. Parts of panel S13a were created with BioRender.com.

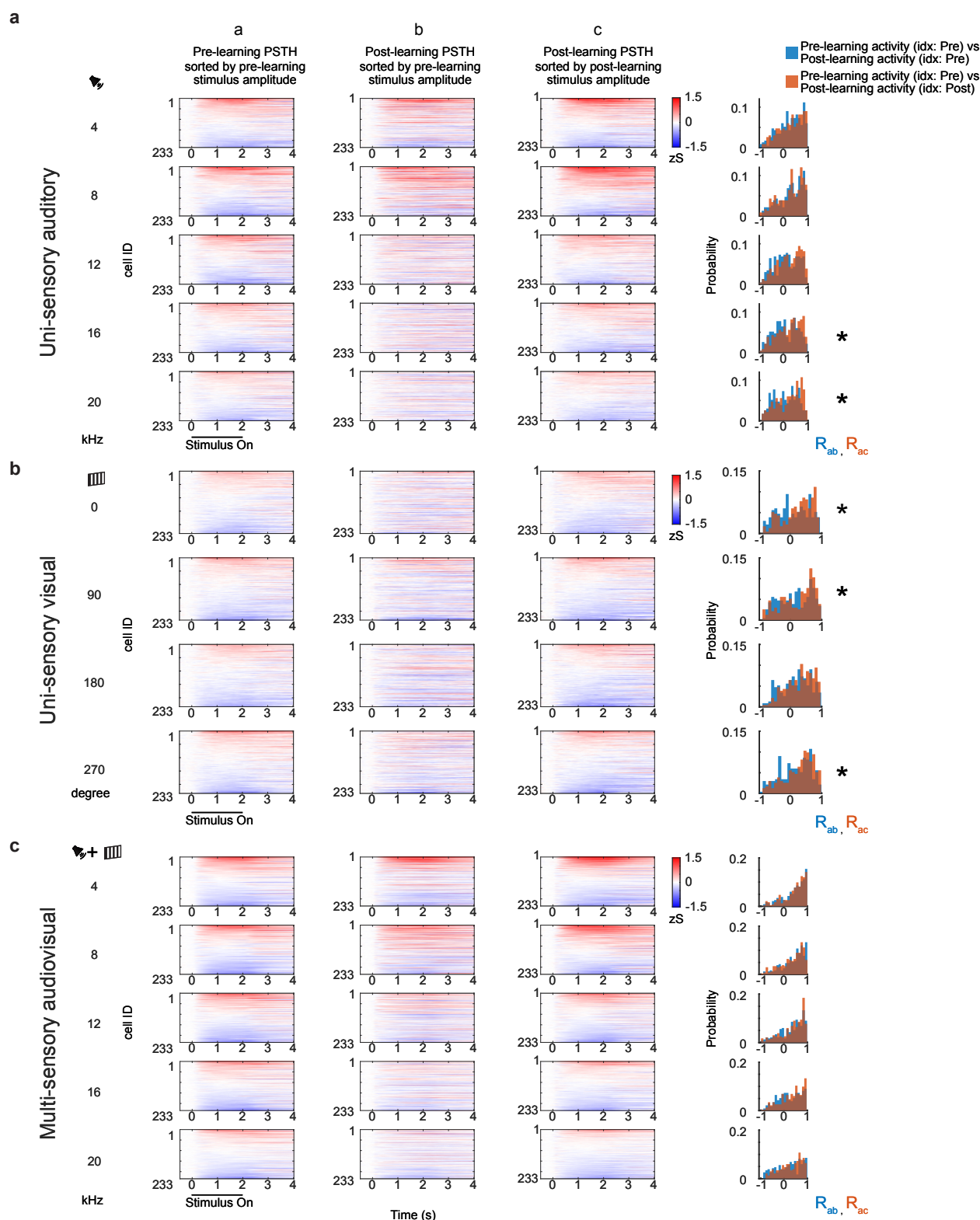


fig. S14 | MGB single cell responses across sensory mapping. (a) Heat-map shows average peri-stimulus time histogram (PSTH) of cell response in pre-learning for uni-sensory auditory stimulus sorted by pre-learning response amplitude during stimulus presentation (column a). Same sorting index was used to sort post-learning PSTH (column b). Reduced response amplitude with same sorting index indicated a remapping of sensory response to the same auditory frequency. Post-learning responses were not diminished after learning if cells were sorted with amplitude in post-learning session (column c). Correlation of cell response time series using two different indexes was compared to quantify the extent of remapping (R_{ab} , R_{ac} , 2-sample ks-test). (b, c) Same measurement for uni-sensory visual response and multi-sensory response.

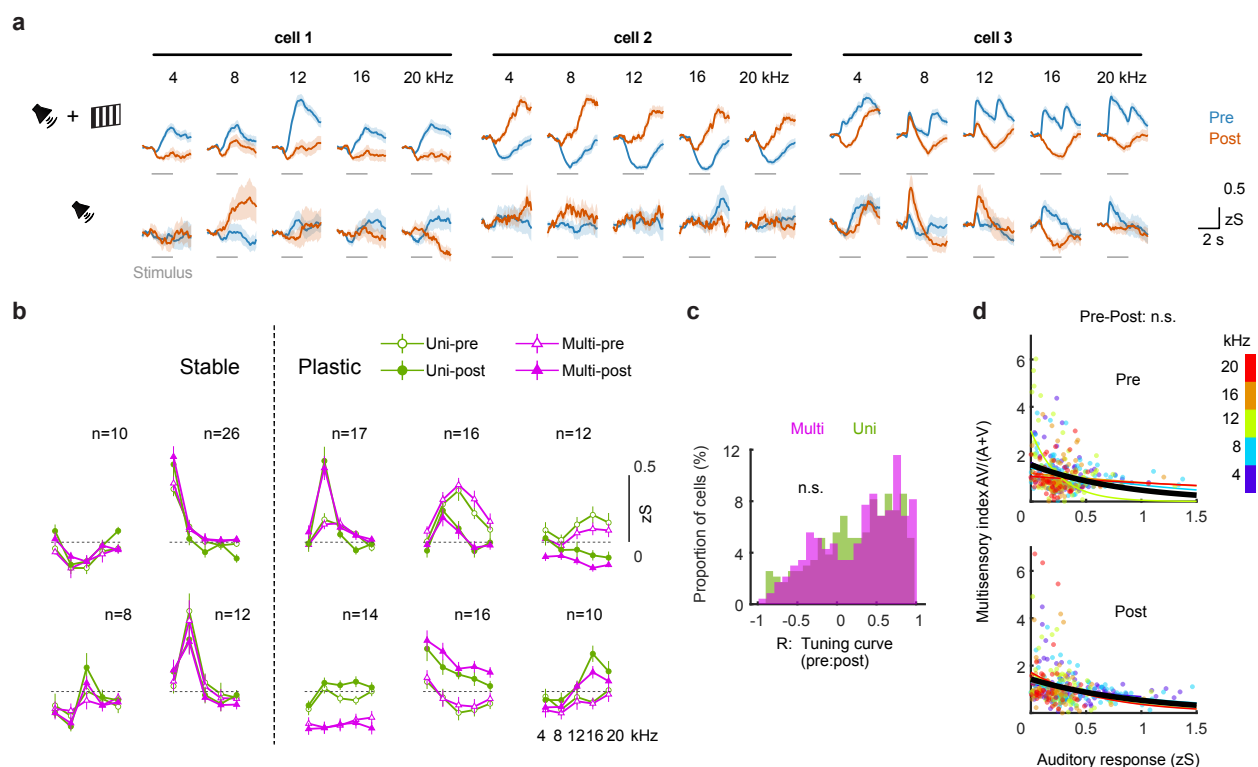


fig. S15 | Single cell tuning properties are not biased to the reward-associated stimuli in non-rewarded sensory mapping sessions. (a) Example sensory responses of individual MGB neurons to auditory and multisensory stimuli. Individual neurons showed stable or plastic sensory response patterns to auditory and multisensory stimuli. (b) Transition of tone-frequency tuning across learning. The neurons showing a similar changing pattern of frequency tuning were grouped by k-means clustering. (c) Distribution of tuning curve similarity of individual neurons across learning. Tuning curve similarities are comparable in auditory and multisensory trials. R is Pearson's correlation coefficient between pre- and post-learning tuning curve ($n = 233$ cells, 6 mice). (d) Multisensory index as a function of unisensory tone response. Two-dimension distribution of unisensory response (auditory) versus multisensory index was compared between pre-learning and post-learning conditions (5 frequencies were pooled, $p > 0.05$, details in Table S1).

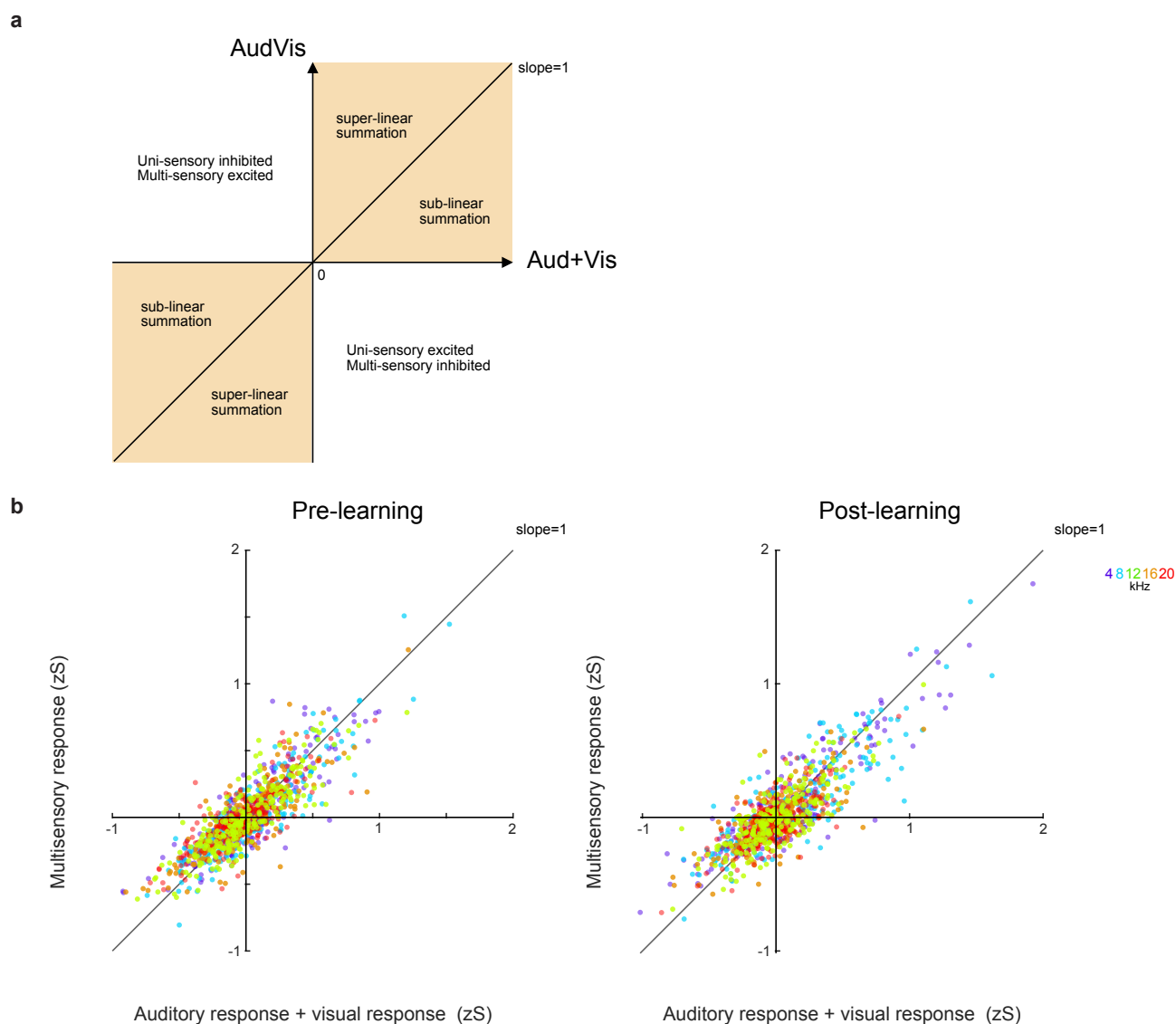


fig. S16 | Multisensory integration remains unchanged after cross-modal reward learning. (a) Multi-sensory integration principle is described in each quadrant. Data in 1st and 3rd quadrant were included in analysis in fig. S15d. (b) Average response of Multi-sensory trials (Audiovisual trials) for each cell and for each auditory frequency were plotted against the sum of average response of Uni-sensory trials (Auditory and Visual trials). Diagonal line represents unity between multisensory responses and the linear sum of individual unisensory responses. Response distribution was not different between pre-learning and post-learning sessions (2-dimensional 2-sample KS-test, details in Table S1). Each dot indicates a cell (repeat measurement form 1-3 sessions).

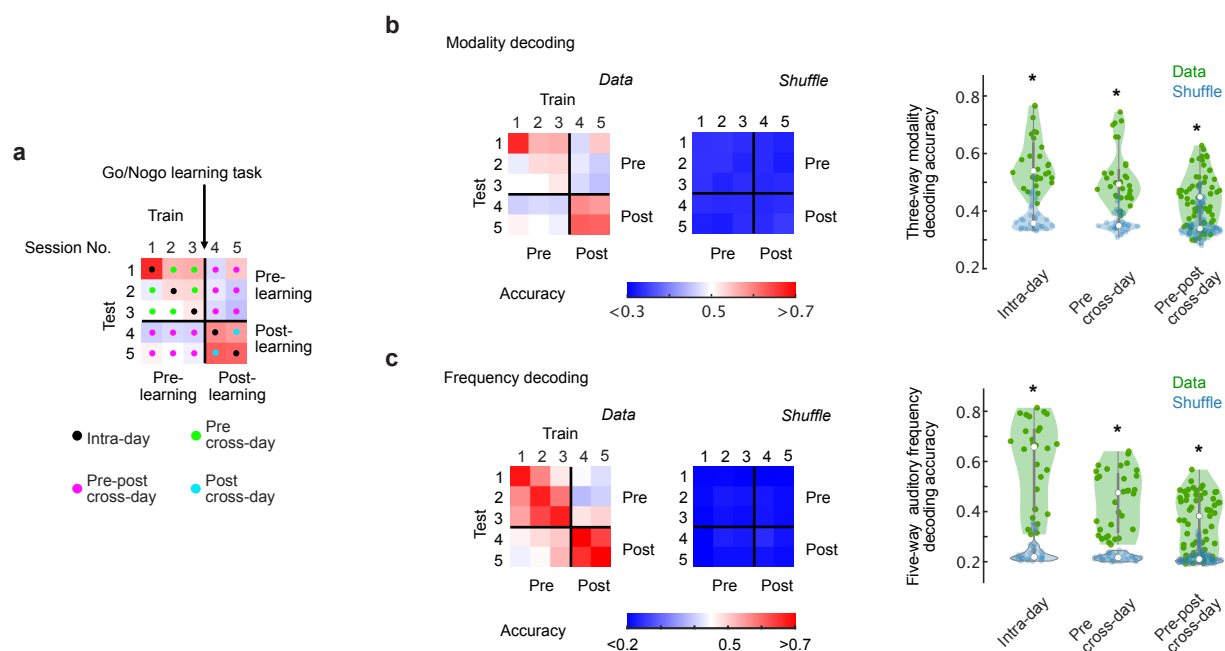


fig. S17 | The total MGB population reliably encodes sensory information. (a) Decoding design matrix. Linear-SVM decoder was trained and tested in a session-by-session manner. (b) Three-way modality decoding. Left: Matrix of modality decoding accuracy from one representative mouse. Middle: Matrix of modality decoding accuracy from the same mouse generated by the shuffled data. Right: Decoding accuracies grouped by intra-day, pre cross-day and pre-post cross-day conditions. Modality decoding accuracies were above chance ($*p < 0.01$). Each dot represents the one session data from one mouse. Shuffle data showed the chance level decoding accuracy. (c) Five-way tone-frequency decoding. The figure structures are the same as (b). Frequency decoding accuracies were above chance ($*p < 0.01$, details in Table S1).

Supplementary Tables

Table. S1. Summary of statistical tests

Fig	Method	Statistics value	df	p	bonferroni-corrected p	detail description	notes
Fig.1f	Signed-rank test	2.00E+01		8.44E-01	NA	N of session to reach criterion between initial and reversal learning	
Fig.1k, left	Chi-square test	2.95E+01		1.74E-06	3.48E-06	Initial learning, cell type ratio	2x bonferroni correction
Fig.1k, Right	Chi-square test	2.66E+01		7.22E-06	1.44E-05	Reversal learning, cell type ratio	2x bonferroni correction
Fig3. a-d	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	1.62E-02	45	3.18E-01	1.27E+00	Main-effect slope, Go,Stim	4x bonferroni correction
		8.03E-02	45	9.89E-06	3.96E-05	Main-effect slope, Go,Delay	4x bonferroni correction
	Linear Mix Model (fix:RevNai-RevExp, random: Mouse ID)	2.55E-02	44	2.53E-02	1.01E-01	Main-effect slope, Go,Stim	4x bonferroni correction
		6.88E-02	44	8.79E-06	3.52E-05	Main-effect slope, Go,Delay	4x bonferroni correction
Fig3. e,f	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	7.51E-02	45	3.49E-05	1.40E-04	Main-effect slope, Hit,Delay	4x bonferroni correction
		1.27E-02	40	9.72E-02	3.89E-01	Main-effect slope, FA, Delay	4x bonferroni correction
	Linear Mix Model (fix:RevNai-RevExp, random: Mouse ID)	6.36E-02	43	5.43E-05	2.17E-04	Main-effect slope, Hit,Delay	4x bonferroni correction
		1.47E-02	44	2.47E-01	9.88E-01	Main-effect slope, FA, Delay	4x bonferroni correction
Fig3. g	Wilcoxon Rank Sum Test	9.10E+01		3.08E-01	6.15E-01	Expert-Hit, Lick/Nolick	2x bonferroni correction
		9.50E+01		4.73E-01	9.45E-01	Naive-Hit, Lick/Nolick	PVC-Hit-Lick/Nolick
Fig3. h	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	5.81E-02	34	1.07E-03	2.14E-03	Main-effect slope, Go,Delay (Nai Exp)	2x bonferroni correction (removing ramp cells)
		6.60E-02	32	2.47E-04	4.95E-04	Main-effect slope, Go,Delay (RevNai RevExp)	
Fig	Method	Statistics value	df	p	bonferroni-corrected p	detail description	notes
fig. S1c	Signed-rank test	3.50E+01		1.56E-02	3.13E-02	Hit RT between naive and expert phase	2x bonferroni correction
fig. S1f	Signed-rank test	3.50E+01		1.56E-02	3.13E-02	Hit RT between naive and expert phase	2x bonferroni correction
fig. S2c	Chi-square	1.84E+01		3.57E-04	7.14E-04	Proportion of stimulus-responsive cell in Aud-reward 1st in Initial learning	2x bonferroni correction
fig. S2c	Chi-square	2.85E+01		2.80E-06	5.60E-06	Proportion of stimulus-responsive cell in Aud-reward 1st in Reversal learning	2x bonferroni correction
fig. S2d	Chi-square	1.75E+01		5.57E-04	1.11E-03	Proportion of stimulus-responsive cell in Vis-reward 1st in Initial learning	2x bonferroni correction
fig. S2d	Chi-square	1.20E+01		7.50E-03	1.50E-02	Proportion of stimulus-responsive cell in Vis-reward 1st in Reversal learning	2x bonferroni correction
fig. S4c, left	ks-test (2 sample)	1.24E-01		7.36E-02	2.21E-01	r value distributions between naive and expert phase during the stimulus period in Hit trials, locomotion	3x bonferroni correction
fig. S4c, left	ks-test (2 sample)	1.81E-01		1.73E-03	5.19E-03	r value distributions between naive and expert phase during the delay period in Hit trials, locomotion	3x bonferroni correction

fig. S4c, left	ks-test (2 sample)	6.19E-02	8.04E-01	2.41E+00	r value distributions between naïve and expert phase during the ITI period in Hit trials, locomotion	3x bonferroni correction	
fig. S4c, right	ks-test (2 sample)	1.33E-01	4.35E-02	1.30E-01	r value distributions between naïve and expert phase during the stimulus period in Hit trials, locomotion	3x bonferroni correction	
fig. S4c, right	ks-test (2 sample)	1.33E-01	4.35E-02	1.30E-01	r value distributions between naïve and expert phase during the delay period in Hit trials, locomotion	3x bonferroni correction	
fig. S4c, right	ks-test (2 sample)	7.14E-02	6.42E-01	1.93E+00	r value distributions between naïve and expert phase during the ITI period in Hit trials, locomotion	3x bonferroni correction	
fig. S4d, left	ks-test (2 sample)	1.05E-01	1.88E-01	5.64E-01	r value distributions between naïve and expert phase during the stimulus period in Hit trials	3x bonferroni correction	
fig. S4d, left	ks-test (2 sample)	1.57E-01	9.80E-03	2.94E-02	r value distributions between naïve and expert phase during the delay period in Hit trials	3x bonferroni correction	
fig. S4d, left	ks-test (2 sample)	1.00E-01	2.32E-01	6.95E-01	r value distributions between naïve and expert phase during the ITI period in Hit trials	3x bonferroni correction	
fig. S4d, right	ks-test (2 sample)	1.81E-01	1.73E-03	5.19E-03	r value distributions between naïve and expert phase during the stimulus period in Hit trials	3x bonferroni correction	
fig. S4d, right	ks-test (2 sample)	8.10E-02	4.80E-01	1.44E+00	r value distributions between naïve and expert phase during the delay period in Hit trials	3x bonferroni correction	
fig. S4d, right	ks-test (2 sample)	1.00E-01	2.32E-01	6.95E-01	r value distributions between naïve and expert phase during the ITI period in Hit trials	3x bonferroni correction	
fig. S4e, left	signed-rank	8.49E+03	3.32E-03	9.96E-03	Delta of absolute r value during the stimulus period	3x bonferroni correction	
fig. S4e, left	signed-rank	5.96E+03	1.00E-08	3.00E-08	Delta of absolute r value during the delay period	3x bonferroni correction	
fig. S4e, left	signed-rank	1.08E+04	7.46E-01	2.24E+00	Delta of absolute r value during the ITI period	3x bonferroni correction	
fig. S4e, right	signed-rank	8.50E+03	3.47E-03	1.04E-02	Delta of absolute r value during the stimulus period	3x bonferroni correction	
fig. S4e, right	signed-rank	7.69E+03	1.24E-04	3.72E-04	Delta of absolute r value during the delay period	3x bonferroni correction	
fig. S4e, right	signed-rank	9.31E+03	4.44E-02	1.33E-01	Delta of absolute r value during the ITI period	3x bonferroni correction	
fig. S4f, left	signed-rank	1.12E+04	9.18E-01	2.75E+00	Delta of absolute r value during the stimulus period	3x bonferroni correction	
fig. S4f, left	signed-rank	9.74E+03	1.29E-01	3.86E-01	Delta of absolute r value during the delay period	3x bonferroni correction	
fig. S4f, left	signed-rank	9.78E+03	1.42E-01	4.25E-01	Delta of absolute r value during the ITI period	3x bonferroni correction	
fig. S4f, right	signed-rank	1.30E+04	3.22E-02	9.66E-02	Delta of absolute r value during the stimulus period	3x bonferroni correction	
fig. S4f, right	signed-rank	1.08E+04	7.71E-01	2.31E+00	Delta of absolute r value during the delay period	3x bonferroni correction	
fig. S4f, right	signed-rank	8.16E+03	9.40E-04	2.82E-03	Delta of absolute r value during the ITI period	3x bonferroni correction	
fig. S4g, left	Chi-square	5.56E+01	9.09E-14	1.82E-13	Frequency of lick and non-lick trial between naïve and expert in initial learning	2x bonferroni correction	
fig. S4g, right	Chi-square	4.34E+01	4.36E-11	8.72E-11	Frequency of lick and non-lick trial between naïve and expert in reversal learning	2x bonferroni correction	
fig. S5	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	1.62E-02	45	3.18E-01	1.27E+00	Main-effect slope, Go,Stim	4x bonferroni correction

		-1.96E-02	45	4.98E-02	1.99E-01	Main-effect slope, NoGo, Stim	
	Linear Mix Model (fix:RevNai-RevExp, random: Mouse ID)	2.55E-02	44	2.53E-02	1.01E-01	Main-effect slope, Go,Stim	
		-4.29E-03	44	6.87E-01	2.75E+00	Main-effect slope, NoGo, Stim	
fig. S6	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	8.03E-02	45	9.89E-06	3.96E-05	Main-effect slope, Go,Delay	4x bonferroni correction
		-1.10E-02	45	1.07E-01	4.28E-01	Main-effect slope, NoGo, Delay	
	Linear Mix Model (fix:RevNai-RevExp, random: Mouse ID)	6.88E-02	44	8.79E-06	3.52E-05	Main-effect slope, Go,Delay	
		-1.85E-02	44	2.21E-02	8.85E-02	Main-effect slope, NoGo, Delay	
fig. S8b	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	1.52E-02	45	3.50E-01	1.40E+00	Main-effect slope, Hit,Stim	4x bonferroni correction
		7.73E-04	40	9.35E-01	3.74E+00	Main-effect slope, FA, Stim	
	Linear Mix Model (fix:RevNai-RevExp, random: Mouse ID)	2.86E-02	43	9.31E-03	3.73E-02	Main-effect slope, Hit,Stim	
		9.81E-03	44	3.61E-01	1.45E+00	Main-effect slope, FA, Stim	
fig. S9a	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	5.81E-02	34	1.07E-03	2.14E-03	Main-effect slope, Go,Delay (Nai Exp)	2x bonferroni correction (removing ramp cells)
		6.60E-02	32	2.47E-04	4.95E-04	Main-effect slope, Go,Delay (RevNai RevExp)	
fig. S9b	Linear Mix Model (fix:Nai-Exp, random: Mouse ID)	6.70E-02	34	1.69E-04	3.37E-04	Main-effect slope, Go,Delay (Nai Exp), random removal	2x bonferroni correction (randomly removing ramp cells)
		7.24E-02	32	2.27E-05	4.55E-05	Main-effect slope, Go,Delay (RevNai RevExp), random removal	
fig. S13c	Chi-square Test	6.24E+00		4.41E-02	NA	Uni-Aud, Pre:Post learning	Response cell fraction
		4.61E+00		9.98E-02	NA	Uni-Vis, Pre:Post learning	
		3.20E+00		2.02E-01	NA	Multi-AudVis, Pre:Post learning	
fig. S13d	Rank sum test	1.38E+04		7.62E-10	1.52E-09	Uni-Aud-Exc, Pre:Post learning	2x bonferroni correction
		1.50E+04		4.95E-20	9.90E-20	Uni-Aud-Inh, Pre:Post learning	2x bonferroni correction
		1.13E+03		7.32E-08	1.46E-07	Uni-Vis-Exc, Pre:Post learning	2x bonferroni correction
		6.51E+03		9.20E-04	1.84E-03	Uni-Vis-Inh, Pre:Post learning	2x bonferroni correction
		1.87E+04		2.55E-06	5.10E-06	Multi-AudVis-Exc, Pre:Post learning	2x bonferroni correction
		3.03E+04		1.44E-19	2.88E-19	Multi-AudVis-Inh, Pre:Post learning	2x bonferroni correction
fig. S13h	Rank sum test	9.31E+02		9.88E-06	1.98E-05	Uni-12kHz-Exc, Pre:Post learning	2x bonferroni correction
		7.15E+02		5.69E-07	1.14E-06	Uni-12kHz-Inh, Pre:Post learning	2x bonferroni correction n=4, sample too small for ranksum test
		NA		NA	NA	Uni-180deg-Exc, Pre:Post learning	2x bonferroni correction
		4.66E+02		1.04E-01	2.08E-01	Uni-180deg-Inh, Pre:Post learning	2x bonferroni correction
		1.15E+03		4.98E-04	9.96E-04	Multi-12kHz180deg-Exc, Pre:Post learning	2x bonferroni correction
		1.65E+03		8.01E-07	1.60E-06	Multi-12kHz180deg-Inh, Pre:Post learning	2x bonferroni correction
fig. S14	ks test (2-sample)	5.58E-02		8.51E-01	4.26E+00	D value, Uni-Aud: 4 kHz	R-Pre-learning
		4.29E-02		9.80E-01	4.90E+00	D value, Uni-Aud: 8 kHz	activity (idx:
		1.42E-01		1.67E-02	8.34E-02	D value, Uni-Aud: 12 kHz	Pre) vs

		1.63E-01	3.50E-03	1.75E-02	D value, Uni-Aud: 16 kHz	R-Post-learning
		1.55E-01	6.71E-03	3.36E-02	D value, Uni-Aud: 20 kHz	activity (idx:
		2.10E-01	5.21E-05	2.09E-04	D value, Uni-Vis: 0 deg	R-Pre-learning
		1.63E-01	3.50E-03	1.40E-02	D value, Uni-Vis: 90 deg	activity (idx:
		1.37E-01	2.22E-02	8.87E-02	D value, Uni-Vis: 180 deg	Pre) vs
		1.63E-01	3.50E-03	1.40E-02	D value, Uni-Vis: 270 deg	R-Post-learning
		6.01E-02	7.82E-01	3.91E+00	D value, Multi-Aud: 4 kHz	R-Pre-learning
		8.15E-02	4.06E-01	2.03E+00	D value, Multi-Aud: 8 kHz	activity (idx:
		8.58E-02	3.43E-01	1.71E+00	D value, Multi-Aud: 12 kHz	Pre) vs
		1.46E-01	1.24E-02	6.21E-02	D value, Multi-Aud: 16 kHz	R-Post-learning
		1.16E-01	8.11E-02	4.06E-01	D value, Multi-Aud: 20 kHz	activity (idx:
fig. S15c	ks test (2-sample)	7.73E-02	4.75E-01		R-Multi-sensory vs R-Uni-sensory	all cells were included
		5.62E-02	1.04E-01		Multi-sensory response distribution (2d, x: Aud+Vis, y: index), pre vs post	all frequencies were mixed
fig. S15d	2-dimension K-S test (2-sample)	9.57E-02	1.85E-01		Multi-sensory index distribution (2d, x:Aud, y: index), pre vs post	all frequencies were mixed
fig. S16	2-dimension K-S test (2-sample)	5.62E-02	1.04E-01		Multi-sensory response distribution (2d, x: Aud+Vis, y: index), pre vs post	all frequencies were mixed
fig. S17b	Wilcoxon Rank Sum Test	1.18E+03	3.66E-10	1.10E-09	Intra-day, data:shuffle, modality decoding	3x bonferroni correction
		1.54E+03	2.18E-11	6.54E-11	Pre-cross day, data:shuffle, modality decoding	3x bonferroni correction
		5.00E+03	2.15E-08	6.45E-08	PrePostCrossDay, data:shuffle, modality decoding	3x bonferroni correction
fig. S17c	Wilcoxon Rank Sum Test	1.18E+03	3.29E-10	9.88E-10	Intra-day, data:shuffle, frequency decoding	3x bonferroni correction
		1.55E+03	6.51E-12	1.95E-11	Pre-cross day, data:shuffle, frequency decoding	3x bonferroni correction
		5.27E+03	3.19E-12	9.58E-12	PrePostCrossDay, data:shuffle, frequency decoding	3x bonferroni correction