Article

Rapid emergence of latent knowledge in the sensory cortex drives learning

https://doi.org/10.1038/s41586-025-08730-8

Received: 8 May 2024

Accepted: 3 February 2025

Published online: 19 March 2025

Check for updates

Céline Drieu[1,2]✉, Ziyi Zhu[3], Ziyun Wang[1], Kylie Fuller[1], Aaron Wang[1], Sarah Elnozahy[1,4] & Kishore Kuchibhotla[1,2,3]✉

Rapid learning confers significant advantages on animals in ecological environments. Despite the need for speed, animals appear to only slowly learn to associate rewarded actions with predictive cues[1–4]. This slow learning is thought to be supported by gradual changes to cue representation in the sensory cortex[2,5]. However, evidence is growing that animals learn more rapidly than classical performance measures suggest[6,7], challenging the prevailing model of sensory cortical plasticity. Here we investigated the relationship between learning and sensory cortical representations. We trained mice on an auditory go/no-go task that dissociated the rapid acquisition of task contingencies (learning) from its slower expression (performance)[7]. Optogenetic silencing demonstrated that the auditory cortex drives both rapid learning and slower performance gains but becomes dispensable once mice achieve 'expert' performance. Instead of enhanced cue representations[8], two-photon calcium imaging of auditory cortical neurons throughout learning revealed two higher-order signals that were causal to learning and performance. A reward-prediction signal emerged rapidly within tens of trials, was present after action-related errors early in training, and faded in expert mice. Silencing at the time of this signal impaired rapid learning, suggesting that it serves an associative role. A distinct cell ensemble encoded and controlled licking suppression that drove slower performance improvements. These ensembles were spatially clustered but uncoupled from sensory representations, indicating higher-order functional segregation within auditory cortex. Our results reveal that the sensory cortex manifests higher-order computations that separably drive rapid learning and slower performance improvements, reshaping our understanding of the fundamental role of the sensory cortex.

Despite the value of rapid learning in ecological environments, most laboratory models of rodent learning show that linking sensory cues with reinforced actions is a slow, gradual process[1–4,9]. An alternative view suggests that animals, including humans, rapidly learn relationships between cues, actions and reinforcement[6], even if they continue to make ongoing performance errors[7,10]. Recent behavioural studies in rodents have begun to reconcile these views, arguing that latent task knowledge (that is, discriminative contingencies) can emerge rapidly even though behavioural performance appears to improve only gradually[7]. We aim to explore how these two dissociable behavioural processes—rapid acquisition of contingencies and slower performance improvements—are implemented in the brain.

An attractive brain region to consider is the sensory cortex, as it is thought to underlie instrumental learning by modulating the representation of sensory cues that drive behaviour. Indeed, the timescale of plasticity for these cue-related responses mirrors the slow and gradual improvements in behavioural performance[1,2,5,9]. This raises a fundamental challenge: if animals learn discriminative contingencies rapidly but sensory cortical cue representations change slowly[1,2,8], the causal model that links cue-related plasticity to learning becomes problematic. One possible solution is that the sensory cortex has a role beyond cue-related representational plasticity and directly represents higher-order signals that associate reinforced actions with predictive cues. Here we tested this hypothesis and focused on the auditory cortex (AC), where cue-related plasticity is thought to increase the salience of the conditioned stimuli to underlie audiomotor learning[1,5].

We trained head-fixed mice to lick to a target tone (S+) for water reward and to withhold licking to a foil tone (S−) to avoid a timeout (auditory go/no-go task; Fig. 1a). To isolate AC computations related to higher-order functions and unrelated to complex auditory processing or perceptual sharpening, we used three-quarter octave-spaced pure tones. Processing of these simple auditory cues does not require the AC, and the tones are easily discriminable at the perceptual level. To support this, two-photon imaging of AC excitatory neurons showed that stimulus identity could be accurately decoded from AC activity from the first training day with no subsequent improvement

[1]Department of Psychological and Brain Sciences, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD, USA. [2]Johns Hopkins Kavli Neuroscience Discovery Institute, Baltimore, MD, USA. [3]Department of Neuroscience, School of Medicine, Johns Hopkins University, Baltimore, MD, USA. [4]Present address: Sainsbury Wellcome Centre, London, UK. ✉e-mail: cdrieu1@jhu.edu; kkuchib1@jhu.edu

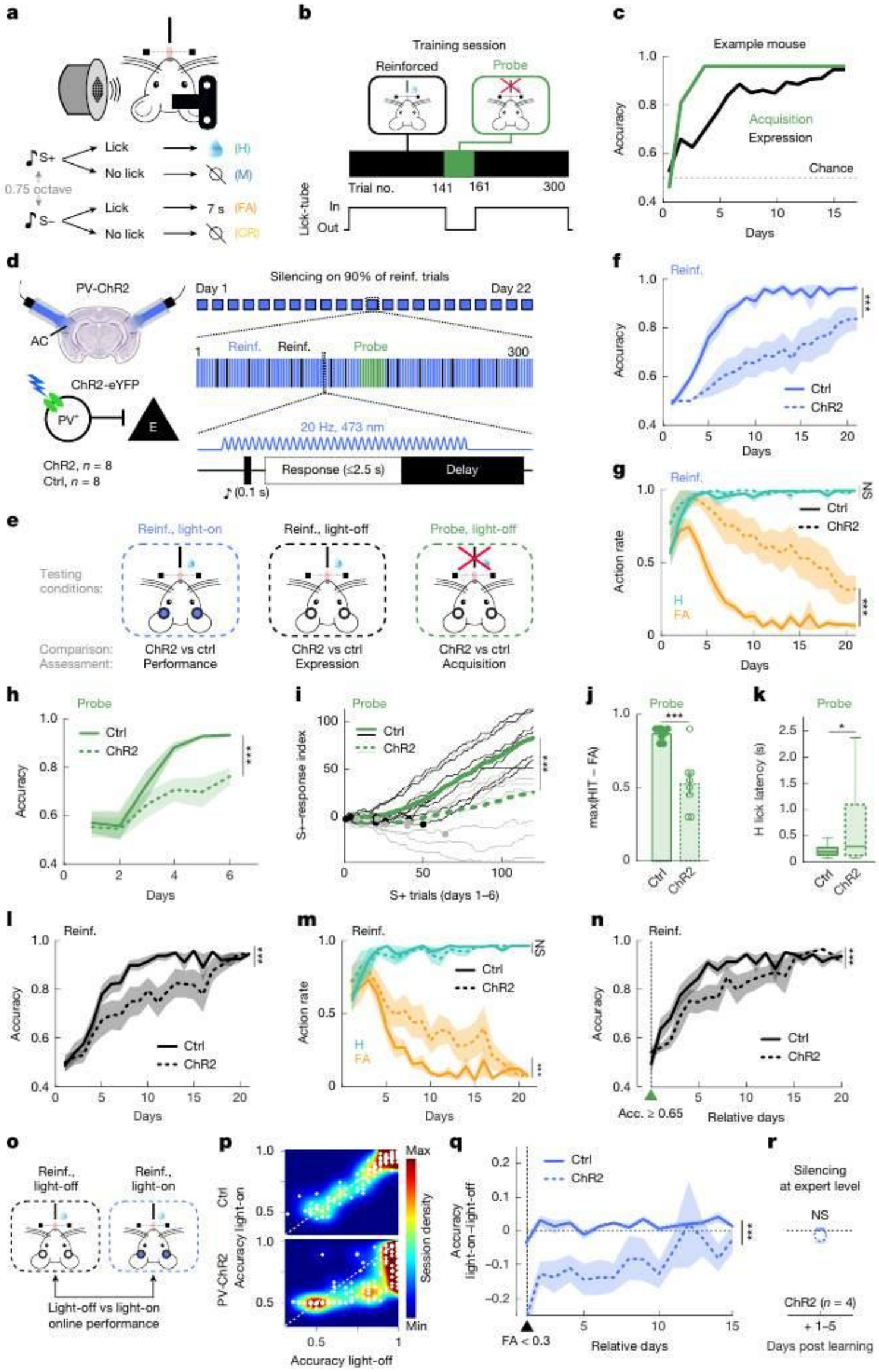


Fig. 1 | AC silencing impairs sound-guided learning and performance during learning. a, Auditory go/no-go task. H, hit; M, miss; FA, false alarm; CR, correct reject. **b**, Task knowledge is probed daily by omitting reinforcement. **c**, Two learning trajectories are revealed: a fast acquisition of task contingencies (measured in probe trials (green)) and a slower performance improvement (measured in reinforced trials (black)). **d**, Probabilistic optogenetic AC silencing over learning. Ctrl, control; PV, parvalbumin; E, excitatory; reinf., reinforcement. **e**, Testing conditions ($n = 8$ control, $n = 8$ PV-ChR2 mice). **f**, Accuracy in reinforced light-on trials (two-way ANOVA, $P = 2.21 \times 10^{-33}$; all ANOVA statistics are reported in Supplementary Table 1). **g**, Action rate in reinforced light-on trials (two-way ANOVA; H, $P = 0.07$; FA, $P = 1.22 \times 10^{-33}$). See also Extended Data Fig. 4. **h**, Accuracy in probe light-off trials (two-way ANOVA, $P = 1.72 \times 10^{-5}$). **i**, S+ response index (two-way ANOVA, $P < 10^{-101}$; Methods). Black and grey lines represent individual mice and dots indicate change points (Methods). **j**, Maximal difference between hit and FA rates in probe light-off trials over the first 6 days ($n = 8$ mice per group, mean ± s.e.m., t-test, $P = 5.34 \times 10^{-4}$). **k**, Hit lick latency in probe light-off trials over the first 6 days ($n = 48$ days per group, t-test on log-transformed data, $P = 0.02$). **l**, Accuracy in reinforced light-off trials (two-way ANOVA, $P = 1.14 \times 10^{-9}$). **m**, Action rate in reinforced light-off trials (two-way ANOVA; H: $P = 0.57$, FA: $P = 1.91 \times 10^{-9}$). **n**, Accuracy in reinforced light-off trials with inter-subject alignment (probe accuracy (acc.) ≥ 0.65 (indicated by green triangle); two-way ANOVA, $P < 10^{-5}$; see Extended Data Fig. 3a–c). **o**, Comparison of light-off versus light-on trials to measure AC silencing effect on on-line performance. **p**, Session density plot of accuracy in reinforced light-on against light-off. See also Extended Data Fig. 3d–g. Max, maximum; min, minimum. **q**, Within-subject accuracy difference in reinforced light-on and light-off trials, aligned to when FA rate <0.3 in reinforced light-off (two-way ANOVA, $P = 3.37 \times 10^{-16}$). **r**, Within-subject accuracy difference in reinforced light-on and light-off trials when silencing at expert level ($n = 4$; t-test, $P = 0.58$). See also Extended Data Fig. 3n–p. See Methods for further description of statistics for all figures. In all figures, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; NS, not significant.

(Extended Data Fig. 1), suggesting that the AC was indeed not needed for perceptual sharpening in the task. Performance was evaluated in each session in reinforced and non-reinforced ('probe') trials (Fig. 1b). Performance in probe trials revealed a rapid acquisition of task contingency knowledge which was expressed only much later in reinforced trials[7] (Fig. 1c). Reinforcement feedback, although critical for learning, paradoxically masked the underlying task knowledge. By combining this behavioural procedure with optogenetics and longitudinal two-photon imaging, we aimed to determine how quickly mice learn stimulus–action contingencies and to define the fundamental role of AC in sound-guided learning.

AC is the default pathway for sound-guided learning

Lesion studies have suggested that the AC may not be essential for learning or execution of cue-guided tasks with simple sensory stimuli[11–14]. However, permanent lesions cannot determine whether the AC is normally used for, or causally produces[15], learning in an intact brain. To address this, we exploited a transient silencing approach to prevent the recruitment of alternative pathways[14,16] while also using a probabilistic design to enable assessment of learning as distinct from performance by measuring behaviour on non-silenced trials, thereby avoiding direct effects of silencing on performance.

We examined the effect of bilateral cortical silencing of the AC throughout learning (Fig. 1d, $n = 8$ control and $n = 8$ parvalbumin channel rhodopsin (PV-ChR2) mice; see Methods for silencing approach and Extended Data Fig. 2 for physiological confirmation of silencing). We probabilistically silenced the AC on 90% of reinforced trials throughout learning ('light-on reinforced'; Fig. 1d), leaving 10% of reinforced ('light-off reinforced') and 100% of probe trials ('light-off probe') with intact AC activity. This design enabled us to assay the effect of cortical silencing on performance (control versus PV-ChR2 performance on light-on reinforced trials) versus acquisition learning (control versus PV-ChR2 performance on light-off probe trials) and expression learning (control versus PV-ChR2 performance on light-off reinforced trials) (Fig. 1e).

We first compared performance in light-on reinforced trials between PV-ChR2 and control mice (Fig. 1e) and observed a large performance impairment in PV-ChR2 mice (Fig. 1f,g). To address whether this performance reduction was accompanied by an impairment in rapid learning, we examined performance during light-off probe trials (Fig. 1e,h–k) when the AC was not silenced and knowledge acquisition can be accurately measured[7]. Accuracy was lower during probe trials in PV-ChR2 mice (Fig. 1h), with delayed S+-response learning (Fig. 1i), lower discrimination (Fig. 1j), and longer lick latency on hit trials (Fig. 1k). Rapid acquisition of task knowledge was therefore impaired in PV-ChR2 mice.

Accuracy was also lower in reinforced light-off trials in PV-ChR2 mice (Fig. 1l,m), even after controlling for their slower task acquisition (Fig. 1n and Extended Data Fig. 3a–c). These impairments were also apparent in response latency and response vigour (Extended Data Fig. 4). Together, these results suggest that the AC is the default pathway for sound-guided reward learning, even when it is not needed for perceptual sharpening.

AC is used during learning but not at expert

We next sought to understand the contribution of AC activity as mice transitioned to expert performance. Transient inactivation of AC in expert mice has led to conflicting results, with some reports showing degradation of sound-guided behaviour[13,14,17,18] and others showing no such degradation[13,19,20]. We exploited our probabilistic silencing strategy and compared performance in light-on (AC silenced) versus light-off (AC functional) reinforced trials within subjects (Fig. 1o). Performance on these two conditions was similar at early periods of training, as performance was poor overall (Fig. 1p). As training progressed, performance remained poor on light-on trials but improved on light-off trials (Fig. 1p), demonstrating that the AC is used for task performance at early and intermediate time point during learning. Surprisingly, this deficit in performance on light-on trials waned (Fig. 1p,q), suggesting that whereas the AC was used during learning, it became dispensable once the mice had mastered the task.

These results have three possible explanations: (1) AC silencing alters tone perception, increasing task difficulty at the perceptual level and thereby delaying learning; (2) repeated exposure to light reduces optical clarity over time owing to factors such as dura regrowth, thereby reducing the silencing effect; or (3) AC silencing serves as a distraction versus interfering with a task-relevant process. To address (1) and (2), we inactivated the AC in a separate group of mice only after mice reached expert performance (Methods), which did not result in a performance deficit[13] (Fig. 1r and Extended Data Fig. 3n–p). To address (3), we bilaterally silenced the visual cortex (VC) throughout learning in PV-ChR2 mice and observed no effect on performance (Extended Data Fig. 3h–m).

Together, these results show that the AC is engaged during learning but is dispensable at expert levels, potentially tutoring subcortical structures that take over once the associations are learned.

Unsupervised discovery of learning-related dynamics

We next sought to understand the nature and dynamics of AC activity underlying learning and performance. We performed longitudinal, two-photon calcium imaging of thousands of AC layer 2/3 excitatory neurons (Fig. 2a,b and Extended Data Fig. 5a–e,i,j) in mice learning the auditory go/no-go task ($n = 5$). A separate group of water-restricted mice was passively exposed to two pure tones over the same duration but with no association with reinforcement ($n = 3$; Extended Data Fig. 5d–f and Methods). This design enabled us to use the passive network as a base case model to isolate learning-related neural dynamics. We tracked the activity of the same neurons across weeks, including pre- and post-learning tuning curve sessions ($n = 4{,}643$ neurons in 8 mice; Fig. 2c and Extended Data Fig. 5f–h and Methods).

From this high-dimensional dataset, we sought to identify single neurons and neuronal ensembles carrying learning-related information, resolve stimulus and non-stimulus-related activity within a given trial, identify changes in representation across trials, and determine outcome-specific dynamics. To do so, we organized our data into a four-dimensional array containing neurons × time in trial × trials across learning × trial outcome (Fig. 2d). To identify shared and distinct variability in neuronal populations recorded in passive ($n = 2{,}339$ neurons 'passive network') and learning ($n = 2{,}304$ neurons, 'learning network') mice, we created a 'megamouse' by combining data from all mice and aligning neural activity to learning phase ($n = 4{,}643$ neurons; Fig. 2e, Extended Data Fig. 6 and Methods). We then used low-rank tensor decomposition to allow unsupervised identification of demixed, low-dimensional neural dynamics across multiple dimensions[21,22] (Extended Data Fig. 7a–c and Methods). The tensor decomposition revealed six neuronal dynamics, each characterized by the four factors of the original tensor (Fig. 2f, Extended Data Figs. 7d,e, 8d and Methods). These six dynamics represented separable computations performed by the auditory cortical networks.

To summarize and obtain an intuitive visualization of the tensor decomposition output, we projected the product of the decomposition into principal component subspace. We observed that learning and passive networks exhibit distinct dynamics (Fig. 2g and Extended Data Fig. 7g,h) and that the neural dynamics of different trial types evolved further apart in the learning versus passive networks (Extended Data Fig. 7i,j). Notably, the identified dynamics were not driven by isolated mice (Extended Data Fig. 7f).

Fig. 2 | Low-rank tensor decomposition reveals learning-related network dynamics. a, Multi-plane, longitudinal two-photon calcium imaging of layer 2/3 excitatory network in the AC (scale bar, 100 µm) during learning ($n = 5$ mice) or passive exposure ($n = 3$ mice; Methods). **b**, Tonotopic organization of the field of view of one example mouse before learning (scale bars, $x = 73$ µm, $y = 76$ µm). See also Extended Data Fig. 5. **c**, Six example cells tracked daily over weeks during behaviour (black triangles, tuning curve sessions). **d**, Neural data are arranged by neurons × time in trial (−1 to +4 s relative to tone onset, vertical line) × trials over time × trial outcomes. **e**, Left, activity from all learning and passive cells is concatenated to create a fourth-order tensor (megamouse). In the third, 'across-trials' dimension, data are aligned across mice according to learning phases (Extended Data Fig. 6 and Methods). **f**, Megamouse tensor decomposition identifies six neuronal dynamics (numbered; Methods) characterized by four factors: neuron, within trial, across trials (x ticks, 20-trial blocks increments) and outcome (see also Extended Data Fig. 7). **g**, Projection of the tensor decomposition output onto principal component subspace. $W_{\mathcal{N}r}$, $W_{\mathcal{W}r}$ and $W_{\mathcal{A}r}$ indicate neuronal, within-trial and across-trial weights for a component r, respectively. **h**, Left, t-distributed stochastic neighbour embedding (t-SNE) projections of neuronal weights. Each dot represents a cell, coloured according to the neuronal dynamic of highest contribution. Right, bars display the proportion of learning and passive cells among the highest contributors for each dynamic. Dynamics 1 and 2 are driven by the passive network (burgundy), whereas dynamics 3–6 are driven by the learning network (blue). **i**, In the passive network, the highest contributing cells in dynamic 1 define cell ensemble 1, and the highest contributing cells in dynamic 2 define cell ensemble 2. Similarly, in the learning network, cell ensembles 3–6 are constituted by the highest contributing cells to dynamics 3–6, respectively. **j**, Absolute weights of cell ensembles across the six identified dynamics. Neurons can participate in more than one dynamic.

For further analyses, we attributed each dynamic to individual neurons on the basis of the maximum weight of the neuron ('unique participation'; Fig. 2h, Extended Data Fig. 8 and Methods). We mapped the six dynamics onto six distinct cell ensembles—that is, groups of neurons that maximally encode a particular network-specific dynamic (Fig. 2i and Extended Data Fig. 8d). It is important to note that individual neurons and corresponding ensembles could exhibit mixed selectivity for the six dynamics, which allows an individual neuron to contribute to multiple, independent computations (Fig. 2j). Overall, non-selective, tone-evoked activity dominated the passive network (Fig. 2f, dynamics 1 and 2), whereas both tone-selective (Fig. 2f, dynamics 3 and 4) and contingency-specific delayed responses (Fig. 2f, dynamics 5 and 6) were present in the learning network. We therefore carefully inspected these separable dynamics.

Learning counteracts tone-evoked habituation

A prevailing view in sensory systems holds that sensory cortices contribute to associative learning through plasticity of the cue representation[5,23–26]. This model posits that individual neurons, via changes in sensory tuning, and neural populations, via cortical map expansion, enhance the representation of behaviourally relevant cues for use by downstream regions[27,28]. These studies, however, measure neural tuning and map expansion outside the task context before and after learning and infer that plasticity of cue representations reflects the mechanistic

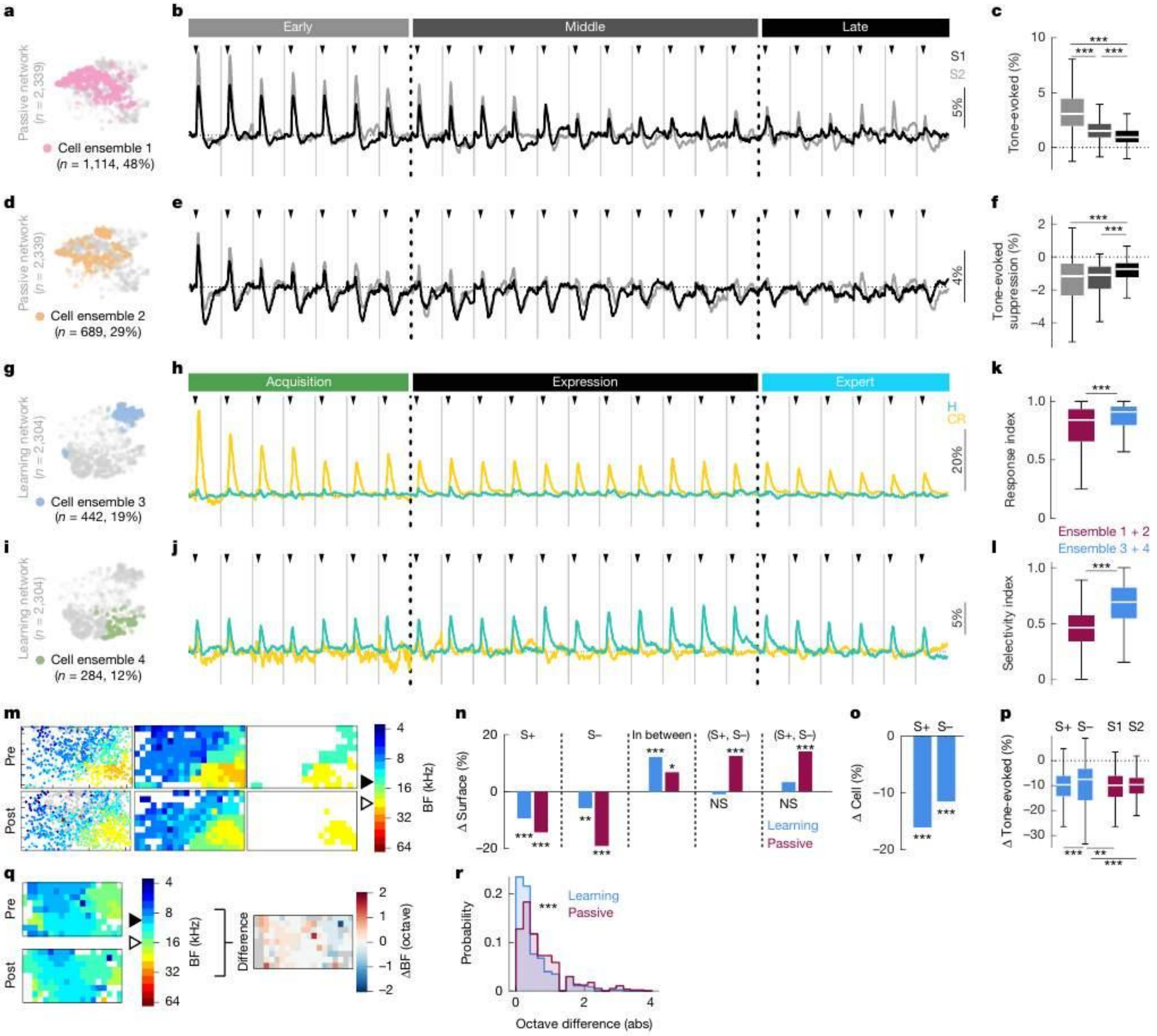


Fig. 3 | Learning counteracts tone-evoked habituation and maintains stimulus selectivity in distinct populations. a, Cell ensemble 1. **b**, Average activity of cell ensemble 1 for S1 and S2 trials across time (80-trial blocks). Black triangles indicate tone onset, grey lines delimit trial blocks. Black dashed lines separate early, middle and late phases (Methods). **c**, Tone-evoked responses of cell ensemble 1 across time phases for S1 and S2 trials combined ($n = 1{,}114$ cells; Friedman test, $P = 1.26 \times 10^{-291}$). **d**, Cell ensemble 2. **e**, Average activity of cell ensemble 2 for S1 and S2 trials. **f**, Tone-evoked responses of cell ensemble 2 across time phases for S1 and S2 trials combined ($n = 689$ cells; Friedman test, $P = 7.32 \times 10^{-121}$). **g**, Cell ensemble 3. **h**, Average activity of cell ensemble 3 for hit (H) and CR trials across learning (80-trial blocks). Black dashed lines separate learning phases (Methods). **i**, Cell ensemble 4. **j**, Average activity of cell ensemble 4 for hit and CR trials. **k**, Response index (response probability over learning; Methods) of cell ensembles 1 and 2 ($n = 1{,}803$ cells) versus cell ensembles 3 and 4 ($n = 726$ cells) (Wilcoxon test, $P = 1.23 \times 10^{-30}$). **l**, Selectivity index (Methods) of cell ensembles 1 and 2 ($n = 1{,}803$ cells) versus cell ensembles 3 and 4 ($n = 726$ cells) (Wilcoxon test, $P = 1.37 \times 10^{-94}$). **m**, Pre (top) and post (bottom) learning tonotopic maps (left), after spatial binning (middle) and restricted to surface with S+ (filled triangle) and S− (open triangle) best frequency (BF; right) (example mouse). **n**, Change in surface representation pre- versus post-task learning (learning) or passive exposure (passive) (binomial proportion tests). **o**, Pre- versus post-learning change in neural responsiveness to S+ and S− (binomial proportion tests). **p**, Pre- versus post-learning change in tone-evoked responses of pre-task S+ and S− responsive neurons ($n = 604$ and 438 cells in learning, $n = 314$ and 614 cells in passive; Kruskal–Wallis test, $P = 2.77 \times 10^{-5}$). **q**, Pre- versus post-learning comparison of local best-frequency differences in tonotopy. **r**, Distribution of map local differences in learning ($n = 389$ pixels) versus passive mice ($n = 166$ pixels; Kolmogorov–Smirnov test, $P = 1.69 \times 10^{-9}$).

role of the sensory cortex. To assess this model, we initially focused on the cell ensembles that exhibited classical stimulus-evoked activity (Fig. 2f), cell ensembles 1–4.

We observed a prominent signature of stimulus-evoked habituation over hundreds to thousands of trials. This habituation dominated activity in passive networks, as seen in cell ensembles 1 and 2 which represented around 77% (1,803 out of 2,339) of all passive cells (Fig. 3a,d). These neurons exhibited stimulus-evoked activation (cell ensemble 1) or suppression (cell ensemble 2), both of which decreased in amplitude over time (Fig. 3b,c,e,f). These cell ensembles were not stimulus-selective and displayed the same dynamic in both stimulus 1 (S1) and stimulus 2 (S2) trials (Fig. 3b,e). These ensembles thus reflected the broad-based

suppression of non-selective neurons after long-term repeated presentation of the same sounds, consistent with previous reports[29].

We observed stimulus-evoked responses that were selective for the S− (cell ensemble 3) or S+ (cell ensemble 4) cues (Fig. 3g–j). Cell ensemble 3 consisted of 19% of the learning network (Fig. 3g) and displayed a modest habituation but mainly a strong preference for the S− throughout learning (Fig. 3h), whereas cell ensemble 4 (12% of total learning cells; Fig. 3j) exhibited S+ selectivity throughout learning (Fig. 3j). Cell ensembles 3 and 4 were more responsive and selective than cell ensembles 1 and 2 (Fig. 3k,l). Stimulus-evoked activity analyses across days of all recorded neurons (n = 7,137) support these results (Extended Data Fig. 9). Therefore, learning counteracted tone-evoked habituation and maintained distinct ensembles that encoded either the S+ or S− selectively.

Learning is not associated with AC map expansion

The sensory cortex is thought to contribute to associative forms of learning by increasing the salience of the conditioned stimuli. This can occur via an increase or change in the representation or selectivity for a particular feature—that is, a detection- or discrimination-based mechanism, respectively. To directly test representational expansion and tuning shifts, we assessed stimulus-evoked responses before (pre-task) and after (post-task) learning, akin to classical measures of tuning and tonotopy. We computed the change in surface area occupied by S+- and S−-preferring cells in tuning curve sessions, outside the task (Fig. 3m). Surprisingly, we observed no increase in the map-level representation of the S+ or S− after learning, and instead observed a modest decrease (Fig. 3m,n). In addition to the best-frequency representation, the fraction of neurons responding to the S+ and S− decreased (Fig. 3o) and the response amplitude of neurons that were initially tuned to the S+ and S− was lower after learning (Fig. 3p). Of note, although we observed no increase in representation to the S+ and S−, learning networks favoured the representation of frequencies between S+ and S−, but not higher or lower as seen in passive networks (Fig. 3n). Finally, using our passive networks as a base case comparison, we calculated the local changes in the tonotopic map structure (Fig. 3q). Learning networks were surprisingly stable and exhibited less local changes than passive networks (Fig. 3r). These pre- versus post-learning changes in responsiveness and tonotopy thus mirrored the responsiveness observed on-line during learning (in dynamics 1 and 2) in a stable, tracked network (n = 4,643 neurons, Fig. 3a–l), and when including all neurons from each session (n = 7,137 neurons) (Extended Data Fig. 9). Together, our results suggest that cortical map expansion and changes in single-neuron tuning are unlikely to be the substrate for associative learning[30].

Partial effects of tone-restricted AC silencing

We next sought to understand whether the maintenance of stimulus selectivity by learning networks contributed to learning and performing the task. We performed daily bilateral silencing of AC during stimulus presentation throughout learning (Extended Data Fig. 10a). Tone-restricted AC silencing impaired task performance throughout learning (Extended Data Fig. 10b–e), task acquisition (Extended Data Fig. 10f–i) and on-line performance during learning, with gradual fading of the effect at expert performance (Extended Data Fig. 10n–q). Accuracy and action rate were not affected in reinforced light-off trials (Extended Data Fig. 10j,k), but PV-ChR2 mice licked more and faster to the S− (Extended Data Fig. 10l,m), suggesting that tone-restricted AC silencing impaired expression, but to a lesser extent than full-trial silencing. These results showed that information carried by the AC network in the tone-evoked window is used during learning. Notably, tone-restricted silencing affected learning less than full-trial silencing across most measures (Fig. 1and Extended Data Fig. 10), suggesting that activity after the tone-evoked window was critical for rapid contingency acquisition and performance during learning.

Rapid emergence of reward prediction activity in AC

The sensory cortex is widely considered to be specialized for perception by interpreting complex sensory objects[31] or adjusting representations of behaviourally relevant stimuli[2,24,28,29,32]. Recent evidence, however, suggests that sensory cortical neurons directly encode non-sensory variables such as movement[33,34], reward[35], reward timing[36], expectation[37] and context[18,29,38–40]. Conjoint representations of sensory and non-sensory variables in the same network could further hone perception or, alternatively, drive more integrative associative processes.

Inspection of within-trial dynamics in learning-driven cell ensembles 5 and 6 suggested that these neurons showed non-canonical activity, with a signal that occurred late in the trial, delayed from the tone-evoked response (Fig. 2f). This late-in-trial signal increased over learning and was trial type selective (Fig. 2f). Cell ensemble 5 (n = 155 cells from the learning network), exhibited late-in-trial activity on hit trials (licking to the S+) that increased with learning (Fig. 4a). This delayed activity was not present on correct S− trials (correct reject (CR)), as neurons exhibited classical stimulus-evoked responses that habituated over learning (Fig. 4b).

We then exploited our multiple trial types to disambiguate the contribution of sensory, motor and reward signals to this dynamic. To assess whether the late-in-trial signal was a delayed sensory activity, we compared activity in hit versus miss trials (Figs. 1a, 4c–e and Extended Data Fig. 11a; see Methods for our approach to generating a balanced set of trials). Cell ensemble 5 did not exhibit late-in-trial activity on miss trials (Fig. 4c–e), discarding the possibility that it reflected a delayed sensory response. We then tested whether this activity reflected reward consumption. We compared cell ensemble activity during hit trials in the reinforced context to the activity during hit trials in the probe context (Fig. 4f), where the mice expected reward and correctly licked to the S+ but the reward was omitted (Fig. 1b). We matched the number of trials between reinforced and probe contexts and controlled for within-session and across-session changes by comparing probe hit trials to reinforced hit trials immediately before and after the probe block. Strikingly, late-in-trial activity was preserved in probe trials (Fig. 4f), indicating that it did not reflect reward consumption. Finally, although movement has been reported to decrease auditory cortical activity[33,41], we examined whether this late-in-trial signal could be driven by licking itself. We first exploited probe hit trials where the lick rate was reduced compared to reinforced hit trials (Fig. 4f). We observed no difference in the late-in-trial neural signal and could thus conclude that the signal was not due to ongoing licking (Fig. 4f). Second, we tested the possibility that this late-in-trial signal was driven by the initiation of a lick bout as compared to the ongoing licking activity. We isolated spontaneous lick bouts between training blocks and observed that the cell ensemble was not lick-responsive (Fig. 4g and Extended Data Fig. 11b). In addition, if lick initiation drove this activity, we would expect to see it on FA trials (incorrect licking to the S−). For this analysis, we focused on FAs that occurred after task acquisition, as these errors are unlikely to be errors due to imperfect task knowledge. We observed no systematic late-in-trial activity on these trials (Extended Data Fig. 11c) even though the licking pattern in FA and probe hit trials were similar (Extended Data Fig. 11d). Taken together, the late-in-trial activity did not reflect stimulus, reward consumption, licking or lick initiation. Instead, these results showed that cell ensemble 5 encoded the higher-order process of reward prediction (RP).

We next sought to identify the precise moment when this reward prediction signal emerged. Initially, some, but not all, of these neurons exhibited classical tone-evoked responses (Extended Data Fig. 11i–k) but then abruptly, and within only 40 hit trials, developed a robust reward-prediction activity (Fig. 4h and Extended Data Fig. 11e). This reward-prediction signal developed over acquisition, strengthened during expression, and then surprisingly receded at expert level when learning is nominally complete (Fig. 4a and

Fig. 4 | Rapid emergence of reward-prediction activity drives learning. **a**, Left, cell ensemble 5 activity ($n = 155$ cells) across learning phases (white dashed lines and coloured rectangles) in hit trials. White trace shows average trial activity. Right, traces at black triangles. **b**, Activity across learning phases in CR trials. **c**, Fraction of cell ensemble 5 ($n = 20$ cells) from one example mouse across consecutive S+ trials. Black dots represent licks. **d**, Activity in hit versus miss trials (Extended Data Fig. 11a). **e**, Area under the curve (AUC) of shaded area in **d** ($n = 155$ cells; Wilcoxon signed-rank test, $P = 6.78 \times 10^{-21}$). a.u., arbitrary units. **f**, Bottom, late-in-trial activity (left; $n = 155$ cells; Friedman test, $P = 0.3071$) and licks in a 1 s post-tone window (right; $n = 406$ trials per conditions; Kruskal–Wallis test, $P = 3.18 \times 10^{-56}$) in hit trials before or after probe hit trials (top traces; scale bars: $x = 0.5$ s, $y = 1\%$ $\Delta F/F$ or 50 licks). **g**, Bottom, activity aligned to lick bout onset outside task events in day 1 of training. Top, lick peristimulus time histogram. **h**, Average activity over the first trials in learning (40-hit-trial blocks). **i**, Activity in probe FA trials. Late-in-trial AUC (grey rectangle) compared to zero ($n = 155$ cells; Wilcoxon signed-rank test, $P = 1.46 \times 10^{-8}$). **j**, Fraction of cell ensemble 5 ($n = 51$ cells) from one example mouse across consecutive reinforced FA trials. Filled circles, reward-prediction (RP+) trials; open circles, RP− trials. **k**, Fraction of cell ensemble 5 ($n = 51$ cells) from the same example mouse in **j** in reinforced FA trials ($n = 423$ trials) after RP classification. **l**, Activity in FA trials with (RP+, $n = 101$) or without (RP−, $n = 322$) RP signal, and in probe FA trials ($n = 19$ trials) reflecting knowledge errors (Extended Data Fig. 12). **m**, Distribution of RP+ and RP− FA trials over learning (binomial proportion tests; acquisition: $P = 1.65 \times 10^{-7}$, expression: $P = 3.32 \times 10^{-10}$, expert, $P = 0.22$). **n**, Trial-specific closed-loop optogenetic AC inactivation over learning. **o**, Performance index (left; $n = 5$ control and $n = 8$ PV-ChR2 mice; two-way ANOVA, $P = 2.11 \times 10^{-21}$; Methods) and hit lick latency (right; $n = 447$ and $n = 316$ trials; Wilcoxon test, $P = 7.7 \times 10^{-4}$) in probe context in post-hit silencing experiments. **p**, Performance index (left, $n = 7$ control and $n = 7$ PV-ChR2 mice; two-way ANOVA, $P = 6.36 \times 10^{-5}$) and hit lick latency (right; $n = 250$ and $n = 171$ trials; Wilcoxon test, $P = 0.004$) in probe context in post-FA silencing experiments.

Extended Data Fig. 11f,g). This longitudinal temporal dynamic mirrored our optogenetic results which demonstrates that the AC is the default pathway for learning but then becomes dispensable at expert levels. Together, these results show that a reward prediction signal rapidly emerges at the timescale of acquisition in AC networks.

Revealing the underlying cognitive drivers of errors

Identifying the cognitive drivers of errors is particularly challenging during learning[4]. Errors during learning are typically considered 'mistakes' while discriminative contingencies, or task knowledge, are still forming. However, errors arise not only from knowledge-related

mistakes, for which animals incorrectly predict reward, but also from factors such as impulsivity, disengagement and exploration. Although detailed behavioural inspection has been a promising route to uncover the nature of errors[10], an alternative approach is to use neural activity itself.

Mice lick in response to the S− (false alarms) during probe trials only at early stages of learning and rapidly withhold thereafter, despite continuing to lick to the S− on reinforced trials. On these probe false alarms, lick rate is strongly reduced and mice do not display signs of impulsivity or exploration[7,10]. Thus, we inferred that probe false alarms are 'knowledge-related' errors. Notably, the neural activity in these trials was similar to hit trials encoding reward prediction (Fig. 4i and Extended Data Fig. 11h), suggesting that the mice were indeed expecting reward.

We reasoned that such knowledge errors should be present not only on probe trials, but also on reinforced trials. We classified individual FA trials in the reinforced context on the basis of the presence of a reward prediction signal (Extended Data Fig. 12a and Methods). We identified a significant proportion of trials that exhibited robust reward prediction activity, but surprisingly many that did not (Fig. 4j,k and Extended Data Fig. 12b). The reward prediction signal was identical to that observed in probe trials (Fig. 4l and Extended Data Fig. 12d), providing confidence that these were indeed knowledge errors. Notably, we found that knowledge errors (RP+) were interspersed with errors that did not elicit reward-prediction activity (RP−; Fig. 4j), suggesting that these RP− false alarms were driven by other factors such as exploration and/or impulsivity[7,10]. Finally, we hypothesized that knowledge errors should occur predominantly during the acquisition phase, when animals are still learning the discriminative contingencies. We computed the fraction of RP+ (knowledge-related errors) and RP− (non-knowledge errors) over time and found that RP+ errors peaked during the acquisition phase of learning, and rarely occurred during expression or expert phases of behaviour (Fig. 4m and Extended Data Fig. 11c). These results demonstrate that the internal cognitive drivers of errors may be accessible from neural data, which is particularly valuable when behaviour alone is insufficient (Extended Data Fig. 12c).

Reward prediction provides the core teaching signal

Learning theory proposes that animals learn from correct actions that are rewarded but also from incorrect actions that are not rewarded[42]. This allows animals to select the appropriate action after reward-predictive (S+) versus non-predictive (S−) cues. Given the presence of the reward prediction activity on correct S+ trials throughout learning and incorrect S− trials early in learning, we reasoned that silencing auditory cortical activity during the post-response period could affect learning and/or performance. To test this, we performed closed-loop probabilistic optogenetic silencing of the AC whereby light was delivered upon lick detection in 90% of either S+ reinforced trials (n = 5 control, n = 8 PV-ChR2 mice) or, in a separate cohort, S− reinforced trials (n = 7 control, n = 8 PV-ChR2 mice; Fig. 4n, Extended Data Figs. 13a, 14a and Methods). No light was delivered in 10% of S+ reinforced trials and 100% of probe trials. Given that the light was delivered after the instrumental lick response, the manipulation could not affect the instrumental behaviour on the current trial, only on subsequent ones. We confirmed this by comparing performance between light-on and light-off trials and observed no difference (Extended Data Figs. 13b–d and 14b–d). In the S+ cohort, post-hit silencing weakened the stimulus–action association (Fig. 4o), delayed cue–response discrimination (Fig. 4o), but did not affect probe accuracy over the first six days (Extended Data Fig. 13e–g). Crucially, the same silencing protocol above the VC (n = 6 PV-ChR2 mice) had no effect on behaviour, confirming that these effects were specific to AC (Extended Data Fig. 13k,l). In the S− cohort, post-FA silencing weakened the stimulus–action association as measured on hit trials (Fig. 4p), robustly delayed cue–response discrimination (Fig. 4p and Extended Data Fig. 13g), and impaired probe accuracy over the first six days (Extended Data Fig. 13e,f). Accuracy of PV-ChR2 mice was lower than control in the reinforced context in both experiments (Extended Data Figs. 13h and 14h), with lower hit rate and higher false alarm rate (Extended Data Figs. 13i and 14i), and longer response latencies on hit trials (Extended Data Figs. 13j and 14j), suggesting an impairment of expression. These closed-loop manipulations showed that AC activity at the time of the reward prediction signal in both hit and FA trials was used by the mouse for the task acquisition and expression. These data also demonstrate that learning is sensitive to cortical silencing on mistakes, suggesting that in a go/no-go paradigm, reward feedback on error trials is crucial to the learning process. Together, these results suggest that reward-prediction activity in auditory cortical networks is used as a teaching signal during learning.

Action suppression encoding enables task performance

A critical requirement in a go/no-go task is the ability to suppress responding to the non-rewarded, S− cue. Here we demonstrate that mice can withhold licking to the S− cue early in learning, as shown in probe trials during the acquisition phase, but continue to lick for hundreds to thousands of trials when being reinforced throughout expression. We explored the extent to which the AC mediates this form of action suppression. Neurons in cell ensemble 6 (n = 704, 31% of learning networks; Fig. 5a), but not non-member cells, exhibited late-in-trial activity when mice correctly withheld from licking on S− trials (CR; Fig. 5b and Extended Data Fig. 15a,b). This signal was stable throughout training despite the increase of CR rate over learning (Fig. 5c and Extended Data Fig. 15c,d). This all-or-none attribute suggested that this late-in-trial activation was tied to performance rather than used for learning. We therefore tested the hypothesis that late-in-trial activation in CR trials reflected action suppression. First, we reasoned that activity in FA and CR trials should be similar until the moment of suppression failure—that is, first lick. We compared the activity of cell ensemble 6 in CR versus FA trials—that is, when mice fail to withhold licking (Methods), exploiting the different response latencies in FA trials (Fig. 5d). We observed that activity dropped abruptly in FA trials at the time of the first lick compared with CR trials (Fig. 5d,e and Extended Data Fig. 15e). Second, if lick suppression is an active contingency-specific process, the late-in-trial activation should be specific for correct rejections for the S− tone, and not observed when the mouse did not lick in response to the S+ tone (miss trials). We controlled for the fact that miss trials were rare and sporadic (Methods) and did not observe late-in-trial activation on miss trials, despite similar peak activity after tone onset in miss and CR trials (Fig. 5f). Third, we reasoned that if this activity reflects the active process of action suppression, the signal should decrease when the mouse is disengaged. We therefore compared late-in-trial activity in CR trials immediately before, during and after short blocks of disengagement (Methods) and observed that the activity dropped significantly when mice transiently disengaged from the task (Fig. 5g). These data suggest that AC integrates a higher-order action-suppression signal.

Finally, we tested whether the action-suppression activity was causal to performance during learning. We reasoned that silencing the AC network throughout S− trials should increase not only the FA rate but also the lick probability (since the action-suppression neurons are silent). By contrast, silencing the AC network only during the stimulus period should increase the FA rate but not the lick probability (Extended Data Fig. 15f). We observed a marked increase in FA rate and lick probability during full-trial silencing in PV-ChR2 mice (Fig. 5h,i). Notably, this effect was not the result of the perception of optogenetic manipulation per se as suppression of the VC in PV-ChR2 mice did not have this effect (Fig. 5h,i). By contrast, restricting silencing to the stimulus period increased FA rate while not affecting lick probability (Fig. 5h,i), suggesting that the late-in-trial activity in CR trials was critical for the

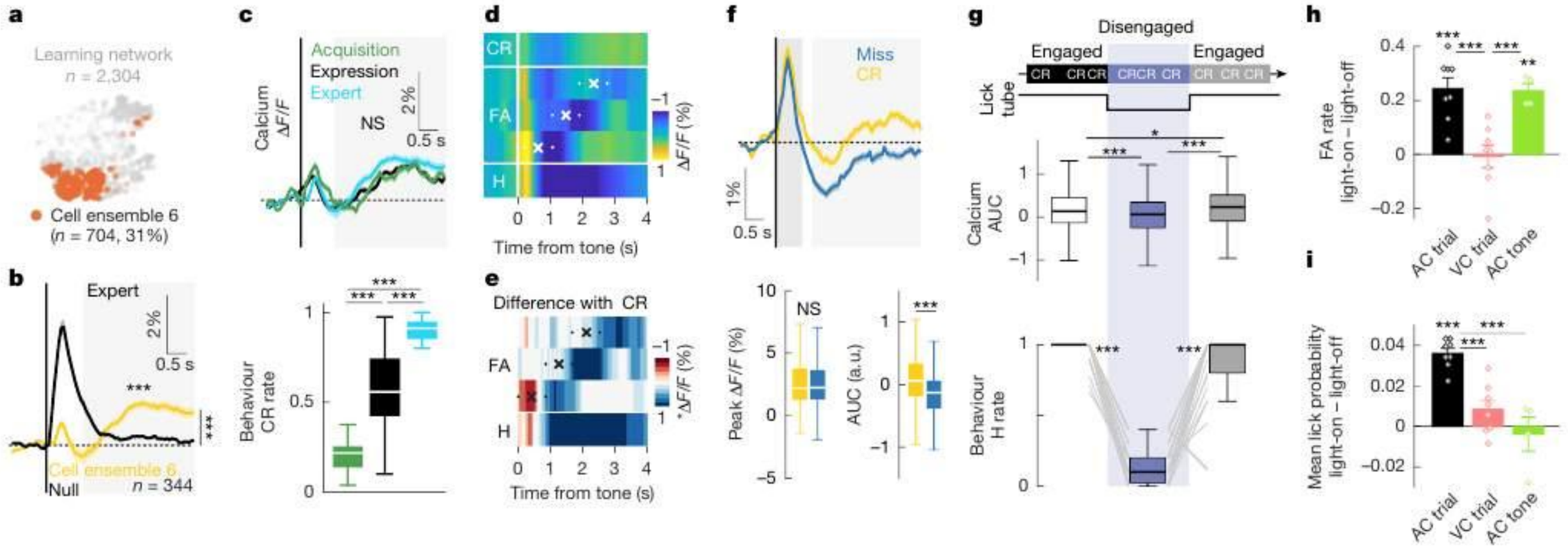


Fig. 5 | Action suppression signals in the AC induce suppression of licking. **a**, Cell ensemble 6. **b**, Average activity of cell ensemble 6 versus cells that do not contribute to this dynamic (null) in CR trials in expert phase (Wilcoxon test, $P = 7.44 \times 10^{-17}$). **c**, Average activity of cell ensemble 6 in CR trials (top; $n = 704$ cells; Kruskal–Wallis test, $P = 0.09$) and CR rate during acquisition (green, $n = 30$ trial blocks), expression (black, $n = 44$ trial blocks) and expert (blue, $n = 24$ trial blocks) phases (bottom; Kruskal–Wallis test, $P = 4.23 \times 10^{-16}$). **d**, Heat map of cell ensemble 6 activity in hit, FA and CR trials. FA trials are binned according to lick latencies (white dots show latency range extremes; white cross represents mean latency). **e**, Heat map of cell ensemble 6 activity in hit and FA trials is significantly different from that of CR trials (Wilcoxon tests). **f**, Top, average cell ensemble 6 activity ($n = 704$ cells) in miss and CR trials. Quantification of tone-evoked activity (bottom left; Wilcoxon signed-rank test, $P = 0.84$) and late-in-trial AUC (bottom right; Wilcoxon signed-rank test, $P = 5.24 \times 10^{-26}$). **g**, Procedure of reinforced and probe CR trial matching (top) and corresponding neural activity (middle; $n = 704$ cells; Friedman test, $P = 1.36 \times 10^{-11}$) and local hit rate (bottom; $n = 30$ matching; Friedman test; $P = 3.45 \times 10^{-11}$). **h**, FA rate difference between light-on and light-off trials in PV-ChR2 mice (mean ± s.e.m.; two-way ANOVA, $P = 7.20 \times 10^{-16}$; *t*-tests compared to 0, AC trial: $P = 4.96 \times 10^{-4}$, VC trial: $P = 0.96$, AC tone: $P = 0.002$). AC or VC were inhibited during the full trial (AC trial, $n = 8$; VC trial, $n = 8$ mice) or AC was silenced during tone presentation only (AC tone, $n = 4$ mice). **i**, Average lick probability in FA light-on versus FA light-off trials (mean ± s.e.m.; $n = 8$ AC full trial, $n = 8$ VC full trial, $n = 4$ AC tone PV-ChR2 mice; two-way ANOVA, $P = 1.18 \times 10^{-5}$; *t*-tests compared to 0, AC trial: $P = 1.94 \times 10^{-6}$, VC trial: $P = 0.10$, AC tone: $P = 0.68$).

maintenance of action suppression. Together, these results showed that action suppression is encoded in AC and is instrumental for performance during learning.

Organization of higher-order contingency ensembles

We next explored the extent to which reward-prediction and action-suppression ensembles mapped onto the underlying stimulus properties of the AC. We exploited the spatial resolution of two-photon imaging to characterize the spatial distribution of reward-prediction and action-suppression neurons. We observed that the two cell ensembles were spatially clustered (Fig. 6a–c), but not related to pre-task stimulus selectivity (Fig. 6d,e). We then examined whether the spatial location of reward-prediction and action-suppression neurons aligned with the underlying tonotopic map. In other words, whether action-suppression neurons have the S− tone for best frequency, and whether reward-prediction neurons were preferentially responsive to S+ tone (Fig. 6f and Methods). We found that this was not the case (Fig. 6g,h), with similar proportions of S+ and S− preferring neurons in reward-prediction and action-suppression cell ensembles (Fig. 6i). Therefore, contingency-related ensembles clustered into spatial domains that were uncoupled from underlying stimulus selectivity and tonotopy, indicating a higher-order functional segregation within AC.

Discussion

Learning-related neural dynamics are traditionally defined as task-specific changes that occur at the timescale of an animal's performance improvements in the learning—that is, a reinforced, context. Using this conceptual and experimental framework, perceptual and instrumental (reward-based) learning and their underlying neural dynamics have been described as slow and gradual—for example, in some cases requiring thousands of trials to learn low information content tasks[2–4]. We took advantage of a recent behavioural paradigm[7,10] that uses non-reinforced probe trials to show that task knowledge emerges more rapidly and earlier than behavioural performance improvement in the learning, reinforced context. Using this behavioural manipulation to quantitatively assess when the mice acquired the task contingencies, we aligned our neuronal recordings to learning stages between animals while preserving trial-based resolution, and took advantage of an unsupervised, dimensionality-reduction method across multiple timescales[21] to identify learning-specific neural dynamics. We observed that reward-prediction activity emerged remarkably quickly—within tens of trials and on the first day of training—in the AC, hundreds to thousands of trials before noticeable performance improvements. The AC thus exhibits latent knowledge of the task, encoded in the network but not behaviourally apparent, with animals experiencing periods when knowledge of environmental contingencies becomes rapidly encoded in the brain, perhaps reflecting an insight-like moment. The latent task knowledge was manifested not as changes in sensory representations but as the emergence of discrete ensembles that encode reward prediction, needed for identifying that a particular cue signals reward availability, and action suppression, needed for suppressing licking on S− trials. These computations were spatially clustered and developed in a manner that was uncoupled from the underlying stimulus-related processing, suggesting a higher-order functional organization. Overall, we find that AC contains separable and causal neural dynamics for both learning and performance.

Although the identified learning-specific dynamics appear fundamental to the role of the AC, the activity of a substantial fraction of neurons (around 30%) is better described by the tone-evoked habituation observed in the passive network. This heterogeneity may reflect differences in connectivity that lead to the selective recruitment of neurons into task-specific ensembles. Future work will be needed to understand how this recruitment occurs by exploring the evolution of the local cortical micro-circuit[29] as well as the local

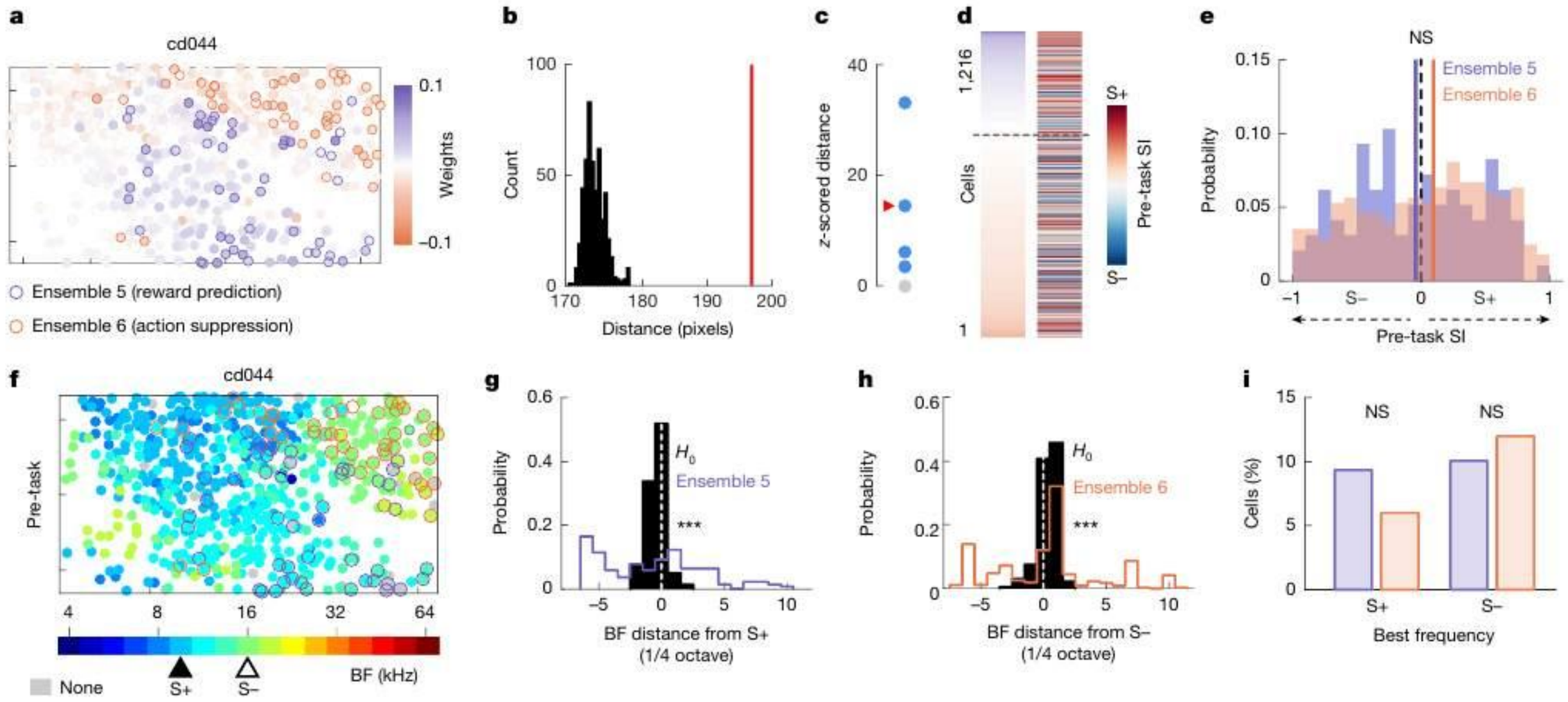


Fig. 6 | Reward-prediction and action-suppression cell ensembles are clustered and uncoupled from stimulus selectivity. a, Spatial distribution of reward-prediction (purple circles) and action-suppression (orange circles) cell ensembles in an example mouse. Colour scale indicates neuronal weights in dynamics 5 (purple) and 6 (orange). **b**, Median of cell distance between cell ensembles (red line) compared to shuffle distribution for the example mouse in **a** (Methods). Null hypothesis: the ensembles are not spatially organized. **c**, *z*-Scored distances between clusters per mouse (blue: significant; grey: non-significant). Red arrow points to the example mouse in **a**. **d**, Neuronal weights in dynamics 5 and 6 of cells from learning mice (left) and their pre-task stimulus selectivity index (SI; right). **e**, Distributions of pre-task stimulus selectivity of cell ensembles 5 ($n = 97$ cells) and 6 ($n = 284$ cells; Kolmogorov–Smirnov test, $P = 0.25$; Wilcoxon test, $P = 0.18$). **f**, Pre-task tonotopic map of the example mice in **a**. Cells are coloured according to their best frequency. Frequencies used as S+ and S− for training are indicated by full and empty triangles, respectively. **g**, Distribution of best-frequency distance from S+ for reward-prediction cell ensemble (purple). Null hypothesis (H_0): reward-prediction cells have a BF as close to S+ as possible (black; Kolmogorov–Smirnov test, $P = 3.81 \times 10^{-9}$; Methods). **h**, Distribution of BF distance from S− for action-suppression cell ensemble (orange). Null hypothesis: action-suppression cells had a BF as close to S− as possible (Kolmogorov–Smirnov test, $P = 9.21 \times 10^{-16}$; Methods). **i**, Proportions of S+ and S−-preferring cells in reward-prediction and action-suppression cell ensembles (binomial proportion tests; S+: $P = 0.17$, S−: $P = 0.53$).

distribution of top-down and ascending neuromodulatory inputs. Keeping a pool of non-task-selective neurons may represent an adaptive mechanism to preserve cortical processing power for future needs.

The detailed input–output circuit that enables reward-prediction and action-suppression computations remains an important area for future exploration. One possibility is that ascending neuromodulatory inputs[18,23,43,44] and top-down projections from motor and frontal regions[32,45–47] serve as critical non-sensory inputs to the sensory cortex. The sensory cortex may then integrate and generate higher-order computations that are incorporated by broader decision-related circuits (for example, frontal cortex, striatum and amygdala) to enable rapid learning and ongoing performance. Contrary to the reward prediction, the action-suppression signal does not evolve over learning, although the fraction of correct rejects does, suggesting that the causal driver of learning to perform occurs outside of AC. Learning to perform is likely to involve two computational features that are potentially subserved by frontal regions: cognitive control to prevent impulsive responses[48], and strategic choice sampling[49] that evolves as the animal becomes increasingly confident that the environment is stable[10]. We propose that AC serves as a conduit, receiving an action-suppression signal during learning and relaying it to downstream regions, such as the striatum, for action selection[50]. At expert levels, the AC is not used to drive task performance[51], as it has potentially tutored subcortical structures to take over, reminiscent of the role of motor cortex in the learning of motor sequences[52]. Putative subcortical candidates include the inferior colliculus[53] and the medial geniculate body (MGB) of the thalamus[54], both of which receive direct feedback projections from AC[55,56]. Moreover, the striatum receives inputs from AC and MGB[57,58]. One possibility is that over learning, MGB-to-striatum synapses become strengthened owing to coincident input from the AC. This in turn removes the necessity of the AC for performance, potentially freeing up cortical resources for future learning.

Our results call for a revision of the classical view of the sensory cortex, according to which its primary role is to process and interpret sensory stimuli. We propose instead that the sensory cortex is better described as a sensory-enriched associative cortex that drives rapid forms of associative learning and where sensory and associative functions are intrinsically intermingled (that is, co-exist within the same network) but computationally separable (Fig. 6). It is important to note that our results do not contradict studies that demonstrate single-neuron tuning curve shifts and tonotopic map plasticity when animals learn perceptually challenging tasks. Our revised model of the sensory cortex would suggest that perceptual sharpening and complex object processing can be subserved by stimulus-related plasticity while the higher-order computations related to associative learning and performance occur in parallel. We expect this view will apply beyond rodents, as rich encoding of non-sensory and task-relevant variables has also been described in sensory cortical areas in humans and non-human primates[59,60].

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-08730-8.

1. Blake, D. T., Strata, F., Churchland, A. K. & Merzenich, M. M. Neural correlates of instrumental learning in primary auditory cortex. *Proc. Natl Acad. Sci. USA* **99**, 10114–10119 (2002).
2. Poort, J. et al. Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* **86**, 1478–1490 (2015).
3. Meister, M. Learning, fast and slow. *Curr. Op. Neurobiol.* **75**, 102555 (2022).
4. Carandini, M. & Churchland, A. K. Probing perceptual decisions in rodents. *Nat. Neurosci.* **16**, 824–831 (2013).
5. Polley, D. B., Heiser, M. A., Blake, D. T., Schreiner, C. E. & Merzenich, M. M. Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. *Proc. Natl Acad. Sci. USA* **101**, 16351–16356 (2004).
6. Gallistel, C. R., Fairhurst, S. & Balsam, P. The learning curve: implications of a quantitative analysis. *Proc. Natl Acad. Sci. USA* **101**, 13124–13131 (2004).
7. Kuchibhotla, K. V. et al. Dissociating task acquisition from expression during learning reveals latent knowledge. *Nat. Commun.* **10**, 2151 (2019).
8. Schreiner, C. E. & Polley, D. B. Auditory map plasticity: diversity in causes and consequences. *Curr. Opin. Neurobiol.* **24**, 143–156 (2014).
9. Sachidhanandam, S., Sreenivasan, V., Kyriakatos, A., Kremer, Y. & Petersen, C. C. Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat. Neurosci.* **16**, 1671–1677 (2013).
10. Zhu, Z. & Kuchibhotla, K. V. Performance errors during learning reflect a dynamic choice strategy. *Curr. Biol.* **34**, 2107–2117.e5 (2024).
11. Ohl, F. W., Wetzel, W., Wagner, T., Rech, A. & Scheich, H. Bilateral ablation of auditory cortex in mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. *Learning* **6**, 347–362 (1999).
12. Romanski, L. M. & LeDoux, J. E. Bilateral destruction of neocortical and perirhinal projection targets of the acoustic thalamus does not disrupt auditory fear conditioning. *Neurosci. Lett.* **142**, 228–232 (1992).
13. Ceballo, S., Piwkowska, Z., Bourg, J. & Bathellier, B. Targeted cortical manipulation of auditory perception. *Neuron* **104**, 1168–1179.e5 (2019).
14. O'sullivan, C., Weible, A. P. & Wehr, M. Auditory cortex contributes to discrimination of pure tones. *eNeuro* **6**, 0340-19.2019 (2019).
15. Barack, D. L. et al. A call for more clarity around causality in neuroscience. *Trends Neurosci.* **45**, 654–655 (2022).
16. Slonina, Z. A., Poole, K. C. & Bizley, J. K. What can we learn from inactivation studies? Lessons from auditory cortex. *Trends Neurosci.* **45**, 64–77 (2022).
17. Jaramillo, S. & Zador, A. M. The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat. Neurosci.* **14**, 246–253 (2011).
18. Kuchibhotla, K. V. et al. Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.* **20**, 62–71 (2016).
19. Gimenez, T. L., Lorenc, M. & Jaramillo, S. Adaptive categorization of sound frequency does not require the auditory cortex in rats. *J. Neurophysiol.* **114**, 1137–1145 (2015).
20. Stoilova, V. V. et al. Auditory cortex reflects goal-directed movement but is not necessary for behavioral adaptation in sound-cued reward tracking. *J. Neurophysio.* **124**, 1056–1071 (2020).
21. Williams, A. H. et al. Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron* **98**, 1099–1115.e8 (2018).
22. Acar, E., Dunlavy, D. M., Kolda, T. G. & Mørup, M. Scalable tensor factorizations for incomplete data. *Chemom. Intell. Lab. Syst.* **106**, 41–56 (2011).
23. Kilgard, M. P. & Merzenich, M. M. Cortical map reorganization enabled by nucleus basalis activity. *Science* **279**, 1714–1718 (1998).
24. Polley, D. B. Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.* **26**, 4970–4982 (2006).
25. Blake, D. T., Heiser, M. A., Caywood, M. & Merzenich, M. M. Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. *Neuron* **52**, 371–381 (2006).
26. Elias, G. A., Bieszczad, K. M. & Weinberger, N. M. Learning strategy refinement reverses early sensory cortical map expansion but not behavior: support for a theory of directed cortical substrates of learning and memory. *Neurobiol. Learn. Mem.* **126**, 39–55 (2015).
27. Buonomano, D. V. & Merzenich, M. M. Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* **21**, 149–186 (1998).
28. Weinberger, N. M. Specific long-term memory traces in primary auditory cortex. *Nat. Rev. Neurosci.* **5**, 279–290 (2004).
29. Kato, H. K., Gillet, S. N. & Isaacson, J. S. Flexible sensory representations in auditory cortex driven by behavioral relevance. *Neuron* **88**, 1027–1039 (2015).
30. Shepard, K. N., Chong, K. K. & Liu, R. C. Contrast enhancement without transient map expansion for species-specific vocalizations in core auditory cortex during learning. *eNeuro* **3**, 0318-16.2016 (2016).
31. Olshausen, B. A. & Field, D. J. in *Problems in Systems Neuroscience* (eds van Hemmen, J. L. & Sejnowski, T. J.) 182–212 (Oxford Univ. Press, 2009).
32. Banerjee, A. et al. Value-guided remapping of sensory circuits by lateral orbitofrontal cortex in reversal learning. *Nature* **585**, 245–250 (2020).
33. Schneider, D. M., Nelson, A. & Mooney, R. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* **513**, 189–194 (2014).
34. 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).
35. Lacefield, C. O., Pnevmatikakis, E. A., Paninski, L. & Bruno, R. M. Reinforcement learning recruits somata and apical dendrites across layers of primary sensory cortex. *Cell Rep.* **26**, 2000–2008.e2 (2019).
36. Shuler, M. G. & Bear, M. F. Reward timing in the primary visual cortex. *Science* **311**, 1606–1609 (2006).
37. Guo, L., Weems, J. T., Walker, W. I., Levichev, A. & Jaramillo, S. Choice-selective neurons in the auditory cortex and in its striatal target encode reward expectation. *J. Neurosci.* **39**, 3687–3697 (2019).
38. Otazu, G. H., Tai, L. H., Yang, Y. & Zador, A. M. Engaging in an auditory task suppresses responses in auditory cortex. *Nat. Neurosci.* **12**, 646–654 (2009).
39. Fritz, J., Shamma, S., Elhilali, M. & Klein, D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* **6**, 1216–1223 (2003).
40. Insanally, M. N. et al. Spike-timing-dependent ensemble encoding by non-classically responsive cortical neurons. *eLife* **8**, e42409 (2019).
41. Yang, Y., Lee, J. & Kim, G. Integration of locomotion and auditory signals in the mouse inferior colliculus. *eLife* **9**, e52228 (2020).
42. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*, 2nd edn (MIT Press, 2018).
43. Zhu, F., Elnozahy, S., Lawlor, J. & Kuchibhotla, K. V. The cholinergic basal forebrain provides a parallel channel for state-dependent sensory signaling to auditory cortex. *Nat. Neurosci.* **26**, 810–819 (2023).
44. Guo, W., Robert, B. & Polley, D. B. The cholinergic basal forebrain links auditory stimuli with delayed reinforcement to support learning. *Neuron* **103**, 1164–1177 (2019).
45. Liu, D. et al. Orbitofrontal control of visual cortex gain promotes visual associative learning. *Nat. Commun.* **11**, 2784 (2020).
46. Winkowski, D. E. et al. Orbitofrontal cortex neurons respond to sound and activate primary auditory cortex neurons. *Cereb. Cortex* **28**, 868–879 (2018).
47. Miller, K. J., Botvinick, M. M. & Brody, C. D. Value representations in the rodent orbitofrontal cortex drive learning, not choice. *eLife* **11**, e64575 (2022).
48. Ridderinkhof, K. R., Wildenberg, W. P. V. D., Segalowitz, S. J. & Carter, C. S. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* **56**, 129–140 (2004).
49. Kolling, N., Behrens, T. E. J., Mars, R. B. & Rushworth, M. F. S. Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
50. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013).
51. Bajo, V. M. et al. Silencing cortical activity during sound-localization training impairs auditory perceptual learning. *Nat. Commun.* **10**, 3075 (2019).
52. Kawai, R. et al. Motor cortex is required for learning but not for executing a motor skill. *Neuron* **86**, 800–812 (2015).
53. Quass, G. L., Rogalla, M. M., Ford, A. N. & Apostolides, P. F. Mixed representations of sound and action in the auditory midbrain. *J. Neurosci.* **44**, e1831232024 (2024).
54. Hasegawa, M., Huang, Z., Paricio-Montesinos, R. & Gründemann, J. Network state changes in sensory thalamus represent learned outcomes. *Nat. Commun.* **15**, 7830 (2024).
55. Winer, J. A., Diehl, J. J. & Larue, D. T. Projections of auditory cortex to the medial geniculate body of the cat. *J. Comp. Neurol.* **430**, 27–55 (2001).
56. Saldaña, E., Feliciano, M. & Mugnaini, E. Distribution of descending projections from primary auditory neocortex to inferior colliculus mimics the topography of intracollicular projections. *J. Comp. Neurol.* **371**, 15–40 (1996).
57. Ledoux, J. E., Farb, C. R. & Romanski, L. M. Overlapping projections to the amygdala and striatum from auditory processing areas of the thalamus and cortex. *Neurosci. Lett.* **134**, 139–144 (1991).
58. Chen, L., Wang, X., Ge, S. & Xiong, Q. Medial geniculate body and primary auditory cortex differentially contribute to striatal sound representations. *Nat. Commun.* **10**, 418 (2019).
59. FitzGerald, T. H., Friston, K. J. & Dolan, R. J. Characterising reward outcome signals in sensory cortex. *NeuroImage* **83**, 329–334 (2013).
60. Huang, Y., Heil, P. & Brosch, M. Associations between sounds and actions in early auditory cortex of nonhuman primates. *eLife* **8**, e43281 (2019).

Methods

Animals

All procedures were approved by Johns Hopkins University Animal Care and Use Committee (MO20A272). Male and female double (PV-ChR2; test mice) or single (PV-Cre or flox-ChR2; control mice) transgenic mice between 6 and 12 weeks of age at the start of experiments were used for the optogenetic experiments[13,61]. PV-cre (Jackson Laboratory, strain 017320), flox-ChR2 (Ai32, Jackson Laboratory, strain 012569) and PV-ChR2 mice were bred in-house. PV-ChR2 mice were obtained by crossing male PV-cre$^{+/-}$ mice with female flox-ChR2$^{+/+}$ mice or by crossing male flox-ChR2$^{+/+}$ mice with female PV-cre$^{+/-}$ mice. To obtain the PV-cre$^{+/-}$ line, we bred female PV-cre$^{+/+}$ mice with male C57BL/6J mice (Jackson Laboratory, strain 000664). Offspring genotypes were confirmed by PCR (Lucigen EconoTaq Plus GREEN 2X) and using two-photon imaging to observe expression of the reporter protein (GFP, see 'Optogenetic experiments'). Male C57BL/6J (Jackson Laboratory, strain 000664) aged between 6 and 12 weeks at the start of experiments were used for two-photon calcium imaging experiments. Mice were group housed in standard plastic cages with food available ad libitum and maintained on a 12-h reversed light–dark cycle at stable temperature (19.5–22 °C) and humidity (35–38%). Experiments took place during the dark phase. Mice were kept on a mild water restriction diet (>85% of body weight) after surgery and throughout task training.

Surgical procedures

Mice were anaesthetized with isoflurane (5% at induction and maintained at 2% during surgery) and their body temperature was maintained at ~35 °C throughout the surgery.

Calcium imaging experiments. Mice were injected (34 gauge, 25.4 mm, 12° bevel needle; Hamilton Company) with 1 µl of AAV9-CaMKII-GCaMP6f (Addgene #100834-AAV9, dilution 1:15) at 0.75 µl min^{-1} (microinjection pump, Harvard Apparatus) in the left primary AC (centred at 1.75 mm anterior to the intersection of the lambdoid and interparietal–occipital sutures, dorsal–ventral (DV): −200 µm). Above the injection coordinates, a cranial window was implanted, replacing a circular piece of skull with a 3-mm diameter cover glass slip (Warner Instruments) that was secured in place using a mix of dental cement and Krazy Glue. A custom-made, three-point stainless steel headpost was secured to the skull with C&B Metabond dental cement (Parkell). The headpost consisted on a two-point kinematic fixation on the right side of the head, prolonged by a rod encircling the cranial window and descending at ~45° ventrally on the left. Mice were given a two-week recovery period to allow weight recovery and viral expression.

Optogenetic experiments. Three-millimetre-diameter glass coverslips were implanted bilaterally over AC (centred at 1.75 mm anterior to the intersection of the lambdoid and interparietal–occipital sutures, on the ridge line of the temporal bone). Custom-made aluminium funnels were implanted above each cranial window. The role of these funnels was threefold: (1) to precisely centre the end of the patch cord on the cranial windows; (2) to hold the patch cord perpendicular to the cranial window (optimizing in-depth light diffusion); and (3) to fix the distance between the patch cord and the cranial window to allow identical light delivery across days. A custom-made, two-point stainless steel headpost was fixed onto the skull with C&B metabond dental cement (Parkell) and dental cement. Mice were allowed to recover for at least one week following surgery.

Optogenetic silencing verification experiments. For silencing verification experiments (Extended Data Fig. 2, $n = 2$), PV-ChR2 mice were injected with 1 µl of AAV-CaMK2-GCaMP6f (Addgene #100834-AAV9, dilution 1:15) at 0.75 µl min^{-1} in the left primary AC (centred at 1.75 mm anterior to the intersection of the lambdoid and interparietal–occipital sutures, DV: −200 µm) and implanted with a 3-mm cover glass slip and a custom-made, two-point stainless steel headpost. Mice were given a two-week recovery period to allow weight recovery and viral expression.

Auditory go/no-go task

All mice (optogenetic and two-photon imaging) underwent the same habituation and training procedures. Mice were randomly chosen to be part of learning or passive groups for two-photon imaging, whereas their genotype determined their assignment to experimental groups for optogenetic experiments. After recovery from surgery, mice were water-restricted for at least 5 days so that their weight stabilized at 85% of their ad libitum weight. During this period, mice were handled daily. Mice were then head-fixed and placed in the experimental context, where they were trained to lick from a lick tube or water cup to receive a drop of water (3 µl). No tone was presented during lick training. Lick training session ended after 30 min or when 1 ml of water was consumed. After 2 days of lick training, mice were trained on the auditory go/no-go task for at least 15 days.

Mice were trained to lick to a target (S+) tone to receive a water drop (3 µl) and withhold licking to the foil (S−) tone to avoid a timeout. Auditory stimuli were three-quarter octave-spaced pure tones. Target and foil tones were presented pseudo-randomly and counterbalanced every 20 trials. Each trial consisted of a no lick period (1 s), tone presentation (100 ms), dead period (200 ms), response period (2.5 s) and a delay period: hit: 4 s (to enable full licking of the reward), miss and correct reject: 2 s, false alarm: 7 s (timeout). In this learning context, called the 'reinforced' context, the lick tube delivering water was positioned within reach of the tongue. By contrast, in the 'probe' context, the lick tube was moved out of tongue and whisker reach by an automated actuator. The blocks of probe trials were interspersed between reinforced trials and no additional delay was introduced by lick-tube movement. Importantly, we have shown that the performance gap observed between probe and reinforced trials early in learning is not driven by the change in the sensory context induced by the absence of the lick tube in the probe context[7]. For passive exposure controls (passive mice), the same protocol was applied but the tones were always presented when the lick tube was out of reach and the water not delivered.

Optogenetic experiments

Mice were trained in the go/no-go task for 300 trials every day: 280 trials in the reinforced context interspersed with a short block of 20 non-reinforced (probe) trials starting at trial 141. Head-fixation habituation, lick training and go/no-go task training took place in custom-made, sound-attenuated behavioural boxes (ambient noise level ~53 dB SPL) controlled with custom MATLAB programs interfacing with Bpod State Machines (Sanworks). Pure tones (4,757 and 8,000 Hz) were delivered through an electrostatic speaker driver (TDT) to a free field electrostatic speaker (TDT) at an intensity of 70 dB SPL and licks were detected through an infrared beam. Blue light (453 nm, DPSS laser, Opto-Engine) was delivered in a 20-Hz sinewave generated by Arduino. The power recorded at the end of the patch cord (splitter branching fibre-optic patch cords, Doric Lenses) was 6–8 mW. When dispersed over a diameter of 3 mm, that yields a light intensity of 0.85–1.13 mW mm^{-2} at the cortical surface. Sound amplitude, water drop size, and laser power were calibrated at the beginning of each experiment. To dissociate the effect of AC silencing on behaviour from its consequence on the learning process, we used a probabilistic approach whereby no light was delivered during probe trials and a subset of reinforced trials. These light-off trials were critical to assess behaviour when AC was available again.

Full-trial experiment ($n = 8$ PV-ChR2, $n = 8$ control mice, $n = 8$ PV-ChR2 VC): light was turned on on 90% of reinforced trials pseudo-randomly (18 trials, 9 S+ and 9 S−, every 20-trial block). In light-on trials, the light was turned on 100 ms before tone onset and stayed on for ~2.5 s for all

trial types (hit: 2.5 s post operant lick, CR and miss: stop at the end of response window, FA: 2.5 s post first lick).

Expert-only full-trial experiment ($n = 4$ PV-ChR2 mice): mice were trained for 18 days without light. Afterwards and for 5 days, from day 19 to 23, the light was turned on following the 'full-trial experiment' protocol or on 90% of reinforced trials consecutively.

Tone experiment ($n = 4$ PV-ChR2, $n = 3$ control mice): light was turned on on 90% of reinforced trials pseudo-randomly (18 trials, 9 S+ and 9 S−, were light-on every 20-trial block). In light-on trials, the light was turned on 100 ms before tone onset and turned off at tone offset.

Post-hit experiment ($n = 8$ PV-ChR2, $n = 5$ control mice, $n = 6$ PV-ChR2 VC): we used a closed-loop lick-triggered stimulation approach, whereby light was turned on after a rewarded lick on 90% of reinforced trials pseudo-randomly (light could be turned on on 9 over 10 S+ trials every 20-trial block). In light-on trials, the light was turned on 70 ms after the first lick detection (to allow the lick cycle to complete and the tongue to retract) and 100 ms before reward delivery and stayed on for 2.5 s.

Post-false alarm experiment ($n = 7$ PV-ChR2, $n = 7$ control mice): we used a closed-loop lick-triggered stimulation approach, whereby light was turned on after a non-rewarded lick on 90% of reinforced trials pseudo-randomly (light could be turned on on 9 over 10 S− trials every 20-trial block). In light-on trials, the light was turned on 70 ms after the first lick detection (to allow the lick cycle to complete and the tongue to retract) for 2.5 s.

At the end of the experiments, mice were anaesthetized (isoflurane 5% at induction and 2% during surgery; body temperature maintained at ~35 °C) and the left funnel was drilled out. Mice were then put under the two-photon microscope and the field of view was excited at 980 nm. Green fluorescence was detected in test mice (ChR2-eYFP) but not in control mice. This procedure allowed to confirm mice genotypes and to assess cell health (Extended Data Fig. 2g). *z*-stacks were collected (unidirectional, 30.98 Hz; magnification 1.7× or 2.0×; range: 450 µm, step: 10 µm, 50 frames per step; depth from brain surface 420–445 µm) to generate 3D reconstruction (ImageJ).

Longitudinal two-photon calcium imaging during learning

Two-photon fluorescence of GCaMP6f was excited at 980 nm using a mode-locked Ti:Sapphire laser (Spectra-Physics) and detected in the green channel (GFP emission). Imaging was performed with a two-photon resonant-scanning microscope (Neurolabware) equipped with a water immersion objective (16×, 0.8 NA, Nikon) tilted to an angle of 40–50° to image AC. The arm of the microscope was enclosed in a custom-made sound-attenuated box. An electronically tunable lens was used to record near-simultaneously two planes in layer 2/3 (150–250 µm below dura, 50-µm spaced, at 15.96 Hz per plane, with a laser power of ≤40 mW). Images were collected at 1.7× or 2× (583 × 387 µm^2) magnification using ScanBox (Neurolabware) and task events (sounds, rewards, licks and frames) were recorded using a digitizer (Digidata 1550b). Pure tones were delivered through an electrostatic speaker driver (RZ6, TDT) to a free field electrostatic speaker (TDT) located at ~5 cm from the right ear at intensity of 70 dB SPL. Licks were detected through an infrared beam. Scanner noise (8 kHz) was attenuated using a custom-made foam sound enclosure directly surrounding the mouse and the resonant scanner was set to continuous throughout the recording session (to avoid any scanning onset-related activity). A custom MATLAB program interfaced with RPvdsEx to control task events. Mice were placed in a plastic tube and head-fixed via a two-point pneumatic clamp on the right and a one-point, 360°-rotational clamp on the left (at 45–50° in the horizontal plane). The whole behavioural platform was installed on a rotation platform so that the field of views could be precisely retrieved one day to the next. Imaging fields were retrieved every day before task training by visual inspection (see also 'Pre- and post- task tonotopic mappings'). Typically, mice were trained for three blocks of 80–100 trials, with either 2 blocks of 10 probe trials interleaved in two of these 3 blocks, or 1 block of 20 probe trials. The field of view was adjusted in between blocks to compensate for *z*-drift, if necessary. An additional 10,000 frames of spontaneous activity were recorded in a separate block at the end of each behaviour session.

Pre- and post-task tonotopic mapping

One day before lick training, mice were placed under the microscope and were presented with a set of 17 pure tones (duration 100 ms), three-quarter octave-spaced, in a pseudo-random order ranging from 4 to 64 kHz at 70 dB SPL. Target and foil tones were selected for the go/no-go task as pure tones that were similarly represented in the recorded neuronal population. The same mapping procedure took place immediately after or one day after the last behaviour session, and 7 and 14 days later.

Two-photon calcium imaging and one-photon blue light stimulation for silencing verification

To validate our optogenetic silencing protocol and determine light power to use for efficient and reliable silencing of cortical networks, we recorded calcium activity of layer 2/3 pyramidal cells while stimulating ChR2-expressing PV interneurons with blue light (Extended Data Fig. 2a,b). Two-photon imaging was performed as indicated in 'Longitudinal two-photon calcium imaging during learning', except that only one plane was recorded (15.49 Hz, 150–250 µm below dura, 1.7× or 2× magnification, laser power ≤40 mW). A mounted LED (490 nm, M490L4, Thorlabs) and a LED driver (Thorlab, LEDD1B) were used to deliver blue light at 6 different power levels over the AC. Pure tones (4–64 kHz, 70 dB SPL) and complex sounds were played (100 ms duration each, 100-frame intervals) and blue light was delivered in a counterbalanced manner. On a silencing trial, a trigger command is sent 100 ms before sound onset from Clampex to the Tower electronics (Scanbox) that generates control signals for the LED and the photomultiplier tube (PMT) shutter (LED on for 1 ms, PMTs off for 9 ms, repeat for 5 frames; Extended Data Fig. 2c). The first pulse was triggered 68 ms before the onset of the sound, and the stimulation continued for a total of 320 ms (Extended Data Fig. 2c). To estimate the LED powers at the cortical surface (in mW mm^{-2}), we measured the LED power coming out of the objective and estimated the cortical surface illuminated to be 2 mm (16× Nikon objective), leading to LED powers ranging from 0 to 3.15 mW mm^{-2}.

Non-rigid registration and cell segmentation were performed using suite2p[62] (https://github.com/MouseLand/suite2p). Fluorescence of each putative neuron ($n = 454$) was extracted and converted to $\Delta F/F$ by taking the mean activity as the baseline. We aligned neural responses to tone presentation, and quantified the effect of optogenetic silencing by comparing the mean activity of each neuron across all repetitions of sound presentations at different light powers (Extended Data Fig. 2d,e). Only $\Delta F/F$ in frames immediately following light presentation were considered for quantification to avoid light contamination of the signal.

Calcium imaging preprocessing

Upon acquisition, images were cropped (to remove artefact bands on plane 1 due to the electronically tunable lens) and converted to HDF5 files. Non-rigid registration (suite2p[62]; https://github.com/MouseLand/suite2p) was run on the concatenated movie of all files recorded for a given mouse. All motion-corrected movies were visually inspected. Because recordings were made over weeks for a given dataset, our dataset could contain cells only weakly active overall. We, therefore, opted for manual detection of regions of interest (ROIs) rather than a semi-automatic one that uses cell activity to detect ROIs (for example, suite2p cell registration). Manual ROI drawing was done in ImageJ using mean enhanced and maximum projection images. We identified 7,137 ROIs in 8 mice, with an average of 892 ± 109 ROIs per mouse. The stability of each ROI throughout the entire recording was then carefully assessed using a custom GUI in MATLAB (MathWorks), that included ROI redrawing to adjust its shape when necessary. Overall, 2,332 out

of 3,935 cells were tracked every day of the task training in learning mice (mean proportion of 67.3 ± 7.5% of total ROIs per mouse), and 2,321 out of 3,202 cells were tracked every day of passive exposure in passive mice (mean proportion of 87.6 ± 6.2% of total ROIs per mouse). Fluorescence activity from the ROIs was extracted using custom functions (MATLAB). Raw fluorescence of each cell was then normalized as:

$$\Delta F/F = F_{all} - \eta_{all}$$

where

$$F_{all} = F_1 \frown F_2 \frown \ldots F_n$$

where the symbol $\frown$ represents a concatenation, n is the number of files, $F_i = \frac{F - F_0}{F_0}$ where F is the raw fluorescence extracted from recording file i and F_0 is the median of this time series. η_{all} is the median of F_{all} over a sliding window of ~3 min. To compare calcium activity across trials, baseline fluorescence (activity during the inter-trial interval, before tone onset) was subtracted from the trial activity, so the $\Delta F/F$ reflected changes of intensity to the original intensity before trial onset.

Data analysis

Statistics. Analyses were performed in MATLAB (MathWorks), using custom programs, FMAToolbox (M. Zugaro; http://fmatoolbox.sourceforge.net) and Tensor Toolbox for MATLAB (https://www.tensortoolbox.org/). Descriptive statistics are reported as mean ± s.e.m. when the underlying distribution is Gaussian-shaped (Jarque–Bera test) or median ± standard error of the median otherwise. Unless indicated otherwise, bars represent median ± standard error of the median, box plots represent median (centre line), upper and lower quartiles (box limits) and 1.5× interquartile range (whiskers), and all statistical tests were two-sided. Student's t-test was used for two-group comparisons of Gaussian distributions, paired t-test was used for paired Gaussian distributions. For non-Gaussian distributions of independent data, two-group comparisons were made using Wilcoxon rank sum tests. Wilcoxon signed-rank tests were used for two-group comparisons of non-Gaussian paired data or to compare medians of non-Gaussian distributions to single values. For non-Gaussian distributions of independent (non-paired) data, multiple comparisons were made using Kruskal–Wallis test and differences between groups were assessed using Wilcoxon rank sum test with Bonferroni correction for post hoc analysis. For paired data following a non-Gaussian distribution, Friedman test was used, with signed-rank tests with Bonferroni correction for post hoc analyses for assessing differences between groups. Multiple Gaussian distributions were compared using ANOVA (independent distributions) or repeated measure ANOVA (paired distributions) with Bonferroni correction for post hoc analysis. Two-way ANOVAs were performed to evaluate the effects of two independent variables on data and their interaction. All ANOVA statistics are reported in Supplementary Table 1. Proportions were compared using the binomial proportion test. Distributions were compared using the Kolmogorov–Smirnov test. No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those generally utilized in the field. Data collection and analysis were not performed blind to the conditions of the experiments.

Behaviour analysis. Rare non-learner mice were excluded and massive drops in performance after reaching high performance (accuracy >0.7) were not analysed. Accuracy in probe and reinforced context was computed as $(n_H + n_{CR})/(n_{S+} + n_{S-})$, where n_H, n_{CR}, n_{S+} and n_{S-} are the number of hit, correct reject, S+ and S− trials, respectively. To have trial-resolution assessment of behaviour, we also computed response index curves (Fig. 1i), which reflected the latency to respond to the cues compared to local, spontaneous licking rate[6,63]. Response index curves were computed for the two cues (S+ and S− trials) separately as the latency to lick in a 2.5-s window before the cue onset minus the latency to lick in the response window (2.5 s after cue onset). If no lick was detected in either of these windows, the latency was set to the window duration (2.5 s). Therefore, for a given trial, the response index ranges from −2.5 to +2.5, with positive values indicating that the response to the cue was shorter than the local spontaneous licking rate of the mouse, negative values indicating a decrease of licking in response to the cue, and values around 0 indicated that the cue did not affect the response rate. Performance index (Fig. 4o,p) was computed as the difference between S+ and S− cumulative response index curves. From the S+ response index, we identified the change point (CP)[6,63]—that is, the trial after which there is a consistent expression of cued behaviour (Fig. 1i). We used the method described in ref. 63, itself a variation of the method used in ref. 6. In brief, a recursive algorithm successively run over each data point i of the cumulative S+ response index curve and performs the following steps: (1) draws a straight line from trial i to trial 0 or the previous true CP, whatever is the closest to i and identifies the point that deviates maximally from this line as a putative CP; (2) calculates the strength of the evidence that it is a true CP—that is, the log of the odds against the null hypothesis of no change (the logit). If logit >1.3, the putative CP becomes a true CP[6,63]. As multiple CPs can be identified on a single curve, we reported in Fig. 1i only the first CP associated with a positive change of the slope of the cumulative behavioural responses[63].

Best frequency. Single cell responses to the 17 tones presented were evaluated with paired t-test comparing pre- versus post-tone mean activity (over 10 frames, ~626 ms). Bonferroni correction for the number of sounds ($n = 17$) was applied. For each cell, the peak amplitude response to each tone was determined as the maximum value of the averaged traces in the ten-frame post-tone window. A neuron's best frequency was determined as the pure tone for which the peak amplitude response was the highest among significant responses only.

Tone-evoked responses across days. Evolution of tone-evoked responses in the reinforced context was analysed using all cells recorded (Extended Data Fig. 9) but the conclusions held when restricted to cells tracked every day. Response to S+ and S−, or stimulus 1 (S1) and stimulus 2 (S2) for passive mice, were analysed separately with paired t-tests comparing pre- versus post-tone mean activity (in 11-frame windows, ~688 ms). A cell was considered tone-responsive in a given day if it significantly responded to either S+/S1 or S−/S2. Given that response profiles were identical to S1 and S2, responses to the two tones were sometimes represented together (Extended Data Fig. 9c–g).

Tone-evoked responses, responsiveness, response index and stimulus SI. Tone-evoked responses were defined as the mean $\Delta F/F$ in a 11-frame window (~688ms) post-tone onset. Responsiveness was defined as the proportion of cells exhibiting a significant tone response (paired t-tests; Extended Data Fig. 9c–g). To compute response indices (Fig. 3k), the peak of the average $\Delta F/F$ for hit and S− trials (FA trials until mid-expression, CR trials after that) in 80-trial blocks was calculated, followed by the proportion of blocks with significant (peak $\Delta F/F$ >2% of baseline) response throughout learning. The response index of a neuron was computed as the average response probability in hit and S− trials over learning. Stimulus selectivity was computed for each neuron in 80-trial blocks over learning and defined as:

$$SI = \frac{|S^+ - S^-|}{|S^+| + |S^-|},$$

where S^+ is the peak $\Delta F/F$ in the tone-evoked response window on hit trials, S^- is the peak $\Delta F/F$ in the tone-evoked response window on S− trials. SI could therefore range from 0 to 1, with 1 indicating maximal selectivity for either the S+ or the S−. Values of S^+ and S^- lower than 2% were set to zero, and SI in blocks where S^+ and S^- were both equal to

zero was set to zero. The selectivity index of a neuron was its average SI over learning (Fig. 3l).

Stimulus decoding. For each mouse, cue identity was decoded across trial frames from activity of cells tracked across all days using linear discriminant analysis with fivefold cross-validation (Extended Data Fig. 1). Tone decoding accuracy in the tone-evoked window referred to the mean accuracy in the tone-evoked window (11 frames post-tone onset; Extended Data Fig. 1e). Chance accuracy level was estimated by decoding cue identity across trial after randomly shuffling cue identity across trials (n = 20 shuffles per day per mouse).

Data organization and tensor decomposition. To analyse our high-dimensional dataset, we took advantage of tensor decomposition[64–66], a method that enables unbiased and interpretable descriptions of dynamic changes at multiple timescales, also referred to as tensor component analysis[21]. Here we used it not only to reveal within and across-trial dynamics[21], but also to identify shared and distinct variability in cell networks recorded from learning and passive mice. We organized calcium traces into a fourth-order tensor (or four-dimensional array) with four axes corresponding to individual neurons (recorded in learning and passive mice), time within trial, trials over time, and trial types. We then fit a tensor CANDECOMP/PARAFAC (CP) decomposition model[66–68] to identify in an unsupervised way a set of low-dimensional components describing variability along each of these four axes (also referred here as factors; Extended Data Fig. 7a).

Data organization. We first built two arrays for learning and passive data separately and combined them afterwards. Only data from the reinforced context was taken for learning mice. We filtered out disengagement periods (hit rate ≤0.5 in a 20-trial block), sometimes occurring during the last dozens of trials of the day and associated with significant changes in neuronal dynamics compared to engaged state[18,29,38,39,69,70]. For both learning and passive data, $\Delta F/F$ of each trial was selected from −1 s to +4 s relative to tone onset (2nd tensor dimension). With 4,643 cells tracked all days, 75 frames per trial, ~300 trials per day over 15 days, our dataset approximated 1,567,000,000 data points. To reduce computation time, trials of identical types (hit, miss, FA or CR) within 20-trial blocks were averaged together. In other words, from a given 20-trial block, up to 4 trial traces could be obtained (4th tensor dimension). Because of the exclusion of disengaged periods and the tendency of the mice to lick, miss trials were too rare in the learning group to be considered without adding significant noise and were excluded. As a result, the 4th tensor dimension dissociated S+ (hit trials for learning data, miss trials for passive data), FA and CR trials. Finally, a crucial goal of this analysis was to be able to identify neural dynamics associated with task learning, and more precisely to isolate any dynamics associated with task contingency acquisition (measured in the probe context) or performance improvement (measured in the reinforced context). To this end, we aligned the trial traces to learning phases (3rd tensor dimension). First, we identified acquisition, expression and expert phases in our five learning mice (see Extended Data Fig. 6). The acquisition phase started at the first trial of training and continued until maximum accuracy was reached in probe or when accuracy was ≥0.65 in probe and ≤0.70 in reinforced trials. This marked the beginning of expression phase, which continued until expert phase started at the second day of high and stable performance. Data in between acquisition and expert phases was part of the expression phase. Evolution of individual mouse performance per identified phases is quantified in Extended Data Fig. 6f. Resultant megamouse performance (that is, pooled performance in 20-trial block across mice) is shown in Extended Data Fig. 6d,e. Second, because these phases varied in duration across animals, we identified the mouse with the minimum number of trial traces in a given phases and downsampled the number of trial traces of the other mice to match this number. Downsampling was performed by preserving the duration/performance range in each mouse (that is, keeping first and last trial traces) and removing trial traces at consistent intervals in between, such as the overall learning evolution of the phase was preserved. Third, each passive mouse was assigned with the learning phases of a learning mouse, and the same downsampling procedure was used. Finally, the two four-dimensional arrays containing learning and passive data, respectively, were concatenated in the first (neurons) dimension (referred as the 'megamouse' tensor) and $\Delta F/F$ traces were z-scored. Because passive mice essentially did not lick, any data for FA trials for passive cells were zeroed out. Any missing entries of the megamouse tensor were also zeroed out.

Tensor decomposition. To deal with incomplete data (absence of FA trials in passive mice and possible missing CR early in learning or missing FA at expert level for learning mice), we fitted an R-component weighted CP model[22] to our megamouse tensor. In brief, CP decomposition decomposes a tensor into a sum of rank-one tensors. For a third-order tensor $X \in \mathbb{R}^{I \times J \times K}$, we wish to write it as:

$$X \approx \sum_{r=1}^{R} a_r \otimes b_r \otimes c_r,$$

where $\otimes$ represents the vector outer product, $a_r \in \mathbb{R}^I$, $b_r \in \mathbb{R}^J$ and $c_r \in \mathbb{R}^K$ for $R = 1, \ldots R$, and $a_r \otimes b_r \otimes c_r$ is a rank-one tensor. With perfect data we would obtain equality; however, in practice the presence of noise prevents it. We can use the Kruskal operator to simplify the previous expression[71,72]:

$$\sum_{r=1}^{R} a_r \otimes b_r \otimes c_r \equiv [\![A, B, C]\!],$$

where factor matrices $A \in \mathbb{R}^{I \times R}$, $B \in \mathbb{R}^{J \times R}$ and $C \in \mathbb{R}^{K \times R}$, with

$$A^{(n)} = [a_1^{(n)} \ldots a_R^{(n)}] \text{ for } n = 1, \ldots, 3$$

To fit the CP decomposition model to data, we used the CP-WOPT (CP weighted optimization) algorithm[22] that uses a first-order optimization approach to solve the weighted least-squares problem—that is, minimize the error function

$$f(A, B, C) = \frac{1}{2} \| W(X - [\![A, B, C]\!]) \|^2,$$

where W is a nonnegative weight tensor with same size as X defined as

$$w_{ijk} = \begin{cases} 1 \text{ if } x_{ijk} \text{ is known,} \\ 0 \text{ if } x_{ijk} \text{ is missing,} \end{cases} \quad \text{for all } i = 1, \ldots, I, j = 1, \ldots, J, k = 1, \ldots, K$$

The weighted least-squares objective function is solved over all the factor matrices simultaneously.

In practice, the rank R of a tensor is generally not known and is not easily determined[73]. To fit the CP models and choose the number of components, we closely followed the pipeline detailed in ref. 21. In brief, we ran models 20 times with different random initializations for different numbers of low-dimensional components $R = 1, \ldots, 6$. We used two metrics to compare and assess models: (1) the (normalized) weighted squared reconstruction error, computed for each fitted model, defined as:

$$\frac{1}{2} \left(\frac{\| W(X - [\![A, B, C]\!]) \|^2}{\| WX \|^2} \right)$$

and (2) a similarity score[21,74], quantifying the match between two fitted models—that is, how similar are the components resulting from two different runs. Let's consider the Kruskal form of the tensor X (or ktensor)

$$X = \sum_{r=1}^{R} \lambda_r a_r \otimes b_r \otimes c_r,$$

where λ_r is the scaling factor after rescaling a_r, b_r and c_r to be unit length. Considering two tensors $[\![A, B, C]\!]$ and $[\![D, E, F]\!]$,

$$\max_{\omega \in \Omega} \frac{1}{R} \sum_{r=1}^{R} \text{penalty} \times (a_r' d_{\omega(r)} \otimes b_r' e_{\omega(r)} \otimes c_r' f_{\omega(r)}),$$

with

$$\text{penalty} = 1 - \frac{|\lambda_r - \lambda_{\omega(r)}|}{\max(\lambda_r, \lambda_{\omega(r)})}$$

where Ω is the set of all permutations of the R components, and ω is a particular permutation. With increasing number of components R, considering all possible matches is exponentially expensive and can be computationally prohibitive and factors were matched in a greedy fashion to identify good alignment (although not necessarily optimal). Similarity for each model fit was computed with respect to the best-fit model with the same number of components. Adding more components caused models to be less reliably identified (lower similarity score). For a given number of components R, the model fits were also visually inspected and compared. With our dataset, models with similarity scores above 0.8 were qualitatively similar while consistency dropped for values closed to 0.5. Therefore, a decomposition into four components was chosen for our dataset. The output of our decomposition was therefore a set of four components, each composed of four factors (that is, weight vectors): (1) neuron factor (W_N), reflecting cell ensembles; (2) within-trial factor (W_W), indicating when the activity occur in the trial; (3) across-trial factor (W_A), reflecting the evolution profile over learning/time at trial resolution; and (4) outcomes factor (W_O), reflecting contribution of sensory, motor and cognitive variables. When R is small, increasing number of components demixed the activity until providing redundant information (when $R > 4$ for this tensor). Importantly, other types of decomposition were run, and other tensors (individual mouse, passive and learning data separately) were decomposed, and they all converged into the same description of the data.

Identification of learning-related dynamics. *Quantification*. To determine whether the low-dimensional dynamics described by the tensor decomposition were selectively attributed to the cells from learning or passive mice, we analysed the neuronal factor—that is, the neuronal weights (W_N) of the four components. We first compared the contribution of learning and passive networks to the highest (absolute) neuronal weights across components (Extended Data Fig. 7d and Extended Data Fig. 8c). Given that no constraint was applied on the sign of the weights, a given component could describe up to two distinct dynamics. We therefore also analysed positive and negative neuronal weights separately (Extended Data Figs. 7e,f and 8d) and obtained the same results: components 1 and 2 described dynamics largely driven by the passive network while components 3 and 4 described neural dynamics driven by the learning network. Importantly, we verified that this effect was not driven only by one mouse: for each component, we compared the neuronal weights of cell populations recorded in each mouse of a group (for example, passive) and compared it to the other group (for example, learning) (Extended Data Fig. 7f). Because the components described different neuronal dynamics, this result therefore implied that learning and passive networks contained different low-dimensional dynamics.

Visualization. To visualize how the revealed neural dynamics maps onto our two experimental groups (learning and passive), we used two different dimensionality-reduction approaches to project the data into a two- or three-dimensional space. First, we used t-SNE on the neuronal weight matrix W_N of size $N \times R$, where N is the number of cells in tensor and R the number of components (Fig. 2h,i and Extended Data Fig. 7g). Second, we used principal component analysis on different combinations of factors: $W_N \otimes W_W$ (Extended Data Fig. 7h), $W_N \otimes W_W \otimes W_A$ (Fig. 2g), $W_N \otimes W_W \otimes W_O$ (Extended Data Fig. 7i), and $W_N \otimes W_W \otimes W_A \otimes W_O$ (Extended Data Fig. 7j), and projected learning and passive data separately into the same principal component subspace.

Unique participation and defining cell ensembles. For visualization and quantification purposes, we attributed each neural dynamic to unique cell ensembles based on neurons' weights (Extended Data Fig. 8a). As indicated earlier, factor weights could be positive or negative and therefore up to two distinct dynamics could be represented per component. With this in mind, each neuron i was associated with a two digit code [componentID, sign]—that is, a unique dynamic, where componentID is the component where the $|W_N|$ of the neuron i was maximal. This approach therefore filtered out non-participating (that is, low weighted) neurons in describing neuronal dynamics, as illustrated in Extended Data Fig. 8b. Finally, in order to assess the nature of encoding of these cell ensembles, cell ensembles 1 and 2 were restricted to cells recorded in the passive mice, while cell ensembles 3 to 6, describing dynamics of components 3 and 4, were restricted to cells recorded in learning mice (Fig. 2i).

Comparison of calcium responses between trial outcomes with a time-changing signal. For each $\Delta F/F$ comparison between different trial types, both the number of trials taken ('how many') and the trial numbers ('when') were matched between groups to control for time/learning effect and power/noise difference (Figs. 4d–f, 5f,g and Extended Data Fig. 11h).

Analysis of licks outside task events. Lick bouts outside task events were defined as lick bouts that preceded the first tone presentation at the beginning of each behavioural block. The analysis was restricted to the first day of training, to remove learning confound as much as possible (Fig. 4g and Extended Data Fig. 11b). A lick bout was defined as a succession of at least 3 licks with less than 1 s interval in between each lick. In addition, it had to be preceded by a 1 s no lick period, used to z-score the traces.

Classification of false alarm trials based on reward-prediction activity. For each learning mouse, we trained a two-class support vector machine (SVM) algorithm to decode trial identity (matched hit and CR trials) from late-in-trial activity (single trial AUCs) of neurons part of cell ensemble 5. This decoding gave us access to a misclassification rate (for each class and global), representing the noise level in the data (Extended Data Fig. 12a,b,e). We then used this trained SVM to classify FA trials, reasoning that if a reward-prediction signal is present during an FA trial, it will be decoded as a hit trial. In each mouse, the proportion of FA trials with an RP signal was higher than the misclassification rate of the decoder (Extended Data Fig. 12e).

Isolating brief disengagement periods during behaviour. Once mice acquire task contingencies and start increasing their correct rejection in the reinforced context, they generally stop behaving in the probe context[7] (hit rate close to zero; for example, Extended Data Fig. 6). We therefore found these periods by looking for probe blocks with hit rate <0.4 (Fig. 5g).

Pre- versus post-behaviour changes in tonopy. To assess how learning and passive exposure affected the cortical tonotopic map, we compared best-frequency surfaces from tuning curve recording sessions before and after learning (see 'Pre- and post- task tonotopic mapping'). We first split the field of views in 30 × 30 pixels (~41 × 41 μm) and computed the best-frequency mode of the local neuronal population in each of those

pixel blocks (Fig. 3o). We estimated the change in surface before and after behaviour as:

$$\Delta\text{surface} = \left[\left(\frac{n_T}{n_{\text{any}}}\right)^{\text{post}} - \left(\frac{n_T}{n_{\text{any}}}\right)^{\text{pre}}\right] \times 100,$$

where n_T is the number of pixel block with T best-frequency mode and n_{any} is the number of pixel block with any best frequency. In our analysis, T could be the S+, S−, tones in between S+ and S−, and tones with lower or higher frequency than S+ or S− (Fig. 3n). We also evaluated best-frequency mode differences before and after behaviour in pixel blocks (Fig. 3q).

Spatial clustering of contingency-related cell ensembles. To assess the spatial distribution of reward-prediction and action-suppression cell ensembles (referred to here as clusters), we compared the distance between the two ensembles to a random spatial organization (Fig. 6a,b). To do so, we computed the median of between-cluster cell distances and compared it to a median distribution obtained with cell ensemble identity shuffles ($n = 500$). This allowed us to assess the clustered nature of these two cell ensembles while preserving the spatial cell distribution in the fields of view. We considered the cell ensembles significantly clustered if the median distance of the cell ensembles was >97.5% of the shuffle distribution. Because of the different statistics of cell distribution inside a field of view for each mouse, comparing raw cell ensembles distances between mice was prohibited. Instead, we computed a z-scored distance for each mouse by subtracting the mean and dividing by the standard deviation of the shuffle distribution to the data median distance (Fig. 6c).

Pre-task stimulus SI. For cells with positive tone-evoked responses to both S+ and S− in pre-task tuning curve session, pre-task SI (Fig. 6d,e and Extended Data Fig. 5i–k) was computed as:

$$\text{SI} = \frac{S^+ - S^-}{S^+ + S^-},$$

where S^+ is the peak $\Delta F/F$ in the tone-evoked response window to the S+ tone and S^- is the peak $\Delta F/F$ in the tone-evoked response window to the S− tone. SI could therefore ranged from −1 to 1, with 1 indicating total selectivity for the S+, −1 indicating total selectivity for the S−, and zero an absence of selectivity (similar response to both tone).

Assessing the relationship between tonotopic map and contingency organization. To assess whether reward-prediction cells were S+-preferring cells and action-suppression cells were S−-preferring cells before training started, we generated two separate statistical tests (Fig. 6g,h). First, we tested the hypothesis that the reward-prediction cell ensemble emerged from S+-preferring cells. We constructed a distribution of best-frequency distance to S+ if the null hypothesis was true—that is, if reward-prediction cells were to have a best frequency the closest to S+ given the field of view statistics (Fig. 6g). Separately, we tested the hypothesis that the action-suppression cell ensemble emerged from S−-preferring cells. We constructed a distribution of best-frequency distance to S− if the null hypothesis was true—that is, if action-suppression cells were to have a best frequency the closest to S− given the field of view statistics (Fig. 6h). Finally, we compared the proportion of S+- and S−-preferring cells among reward-prediction and action-suppression cell ensembles and observed no differences (Fig. 6i).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

61. Li, N. et al. Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *eLife* **8**, e48622 (2019).
62. Pachitariu, M. et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at *bioRxiv* https://doi.org/10.1101/061507 (2017).
63. Vega-Vilar, M., Horvitz, J. C. & Nicola, S. M. NMDA receptor-dependent plasticity in the nucleus accumbens connects reward-predictive cues to approach responses. *Nat. Commun.* **10**, 4429 (2019).
64. Hitchcock, F. L. The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.* **6**, 164–189 (1927).
65. Hitchcock, F. L. Multilple invariants and generalized rank of a p-way matrix or tensor. *J. Math. Phys.* **7**, 39–79 (1927).
66. Kolda, T. G. & Bader, B. W. Tensor decompositions and review. *SIAM Rev.* **51**, 455–500 (2009).
67. Harshman, R. Foundations of the parafac procedure: models and conditions for an "explanatory" multimodal factor analysis. *UCLA Work. Pap. Phon.* **16**, 1– 84 (1970).
68. Carroll, J. D. & Chang, J. J. Analysis of individual differences in multidimensional scaling via an *n*-way generalization of "Eckart-Young" decomposition. *Psychometrika* **35**, 283–319 (1970).
69. Atiani, S. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* **82**, 486–499 (2014).
70. Franceschi, G. D. & Barkat, T. R. Task-induced modulations of neuronal activity along the auditory pathway. *Cell Rep.* **37**, 110115 (2021).
71. Kruskal, J. B. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Its Appl.* **18**, 95–138 (1977).
72. Kolda, T. G. *Multilinear Operators for Higher-Order Decompositions* https://doi.org/10.2172/923081 (Sandia National Laboratories, 2006).
73. Acar, E., Dunlavy, D. M. & Kolda, T. G. A scalable optimization approach for fitting canonical tensor decompositions. *J. Chemometrics* **25**, 67–86 (2011).
74. Tomasi, G. & Bro, R. A comparison of algorithms for fitting the parafac model. *Comput. Stat. Data Anal.* **50**, 1700–1734 (2006).

Acknowledgements The authors thank P. Janak, C. Honey, S. Mysore, D. Lee and the members of the Kuchibhotla laboratory for comments on the manuscript; and scidraw.io for illustrations. This work was supported by the Johns Hopkins Kavli NDI Distinguished Postdoctoral Fellows Program (C.D.), National Institute of Health K99 DA059024 (C.D.), Johns Hopkins Kavli NDI Distinguished Graduate Student Fellowship (Z.Z.), National Institute of Health R00 DC015014 and R01 DC018650 (K.K.), National Science Foundation CAREER 2145247 (K.K), NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP) (K.K.) and Brain and Behavior Research Foundation NARSAD (K.K).

Author contributions C.D. and K.K. designed the experiments and data analyses. C.D. and Z.Z. performed the two-photon experiments. C.D., Z.Z. and K.F. preprocessed calcium imaging data. C.D., Z.W., K.F., A.W. and S.E. performed the optogenetic experiments. C.D. analysed the data. C.D. and K.K. wrote the manuscript.

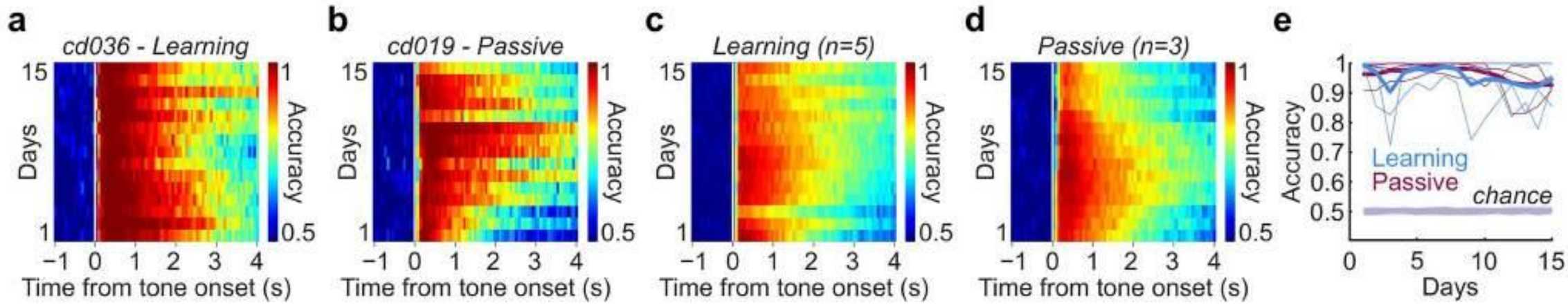
Competing interests The authors declare no competing interests.

Additional information

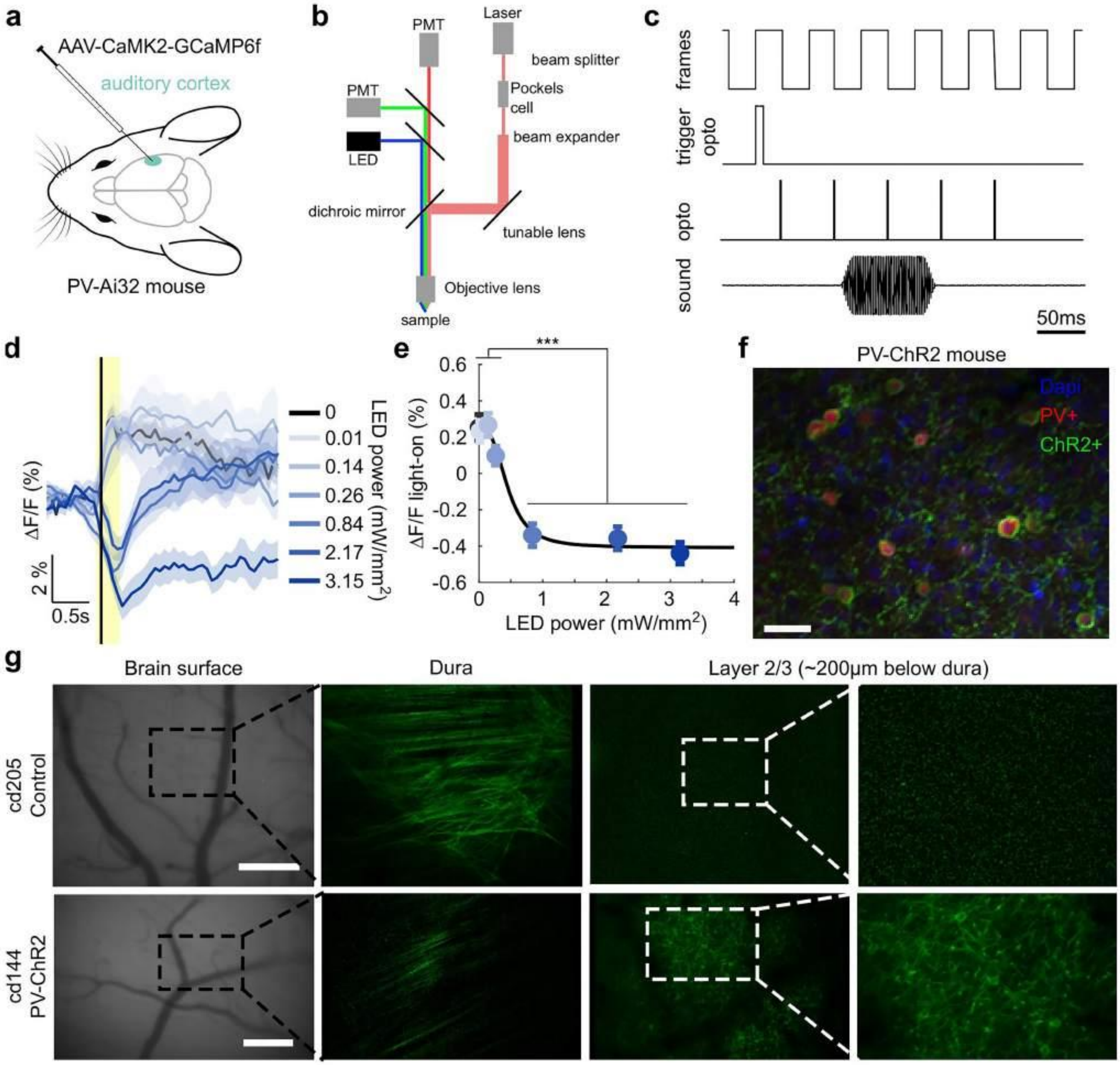
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-08730-8.

Correspondence and requests for materials should be addressed to Céline Drieu or Kishore Kuchibhotla.

Peer review information *Nature* thanks Shihab Shamma and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

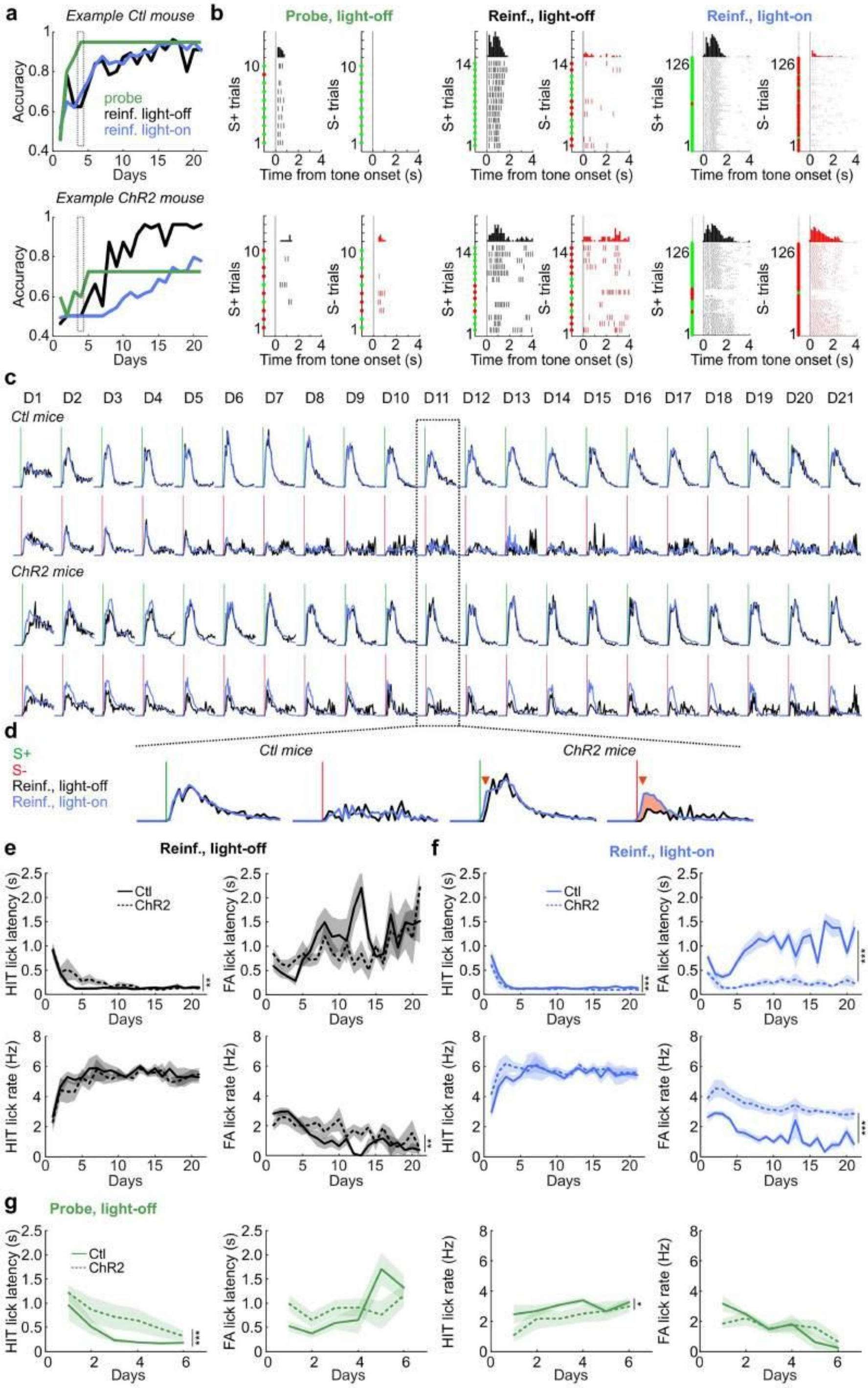


Extended Data Fig. 1 | Stimulus decoding in the auditory cortex is at ceiling from Day 1 of learning. a, Stimulus decoding is at ceiling on Day 1 and remains high throughout learning (example mouse). Only the cells tracked across all days were used to decode tone identity. **b**, Stimulus decoding is at ceiling on day 1 and remains high throughout passive exposure over 15 days (example mouse). **c**, Average decoding accuracy for all Learning mice ($n = 5$). **d**, Average decoding accuracy for all Passive mice ($n = 3$). **e**, Evolution of tone decoding accuracy in the tone-evoked window across days for Learning and Passive mice compared to chance level (trial shuffle, see Methods).



Extended Data Fig. 2 | Activating PV+ neurons in the auditory cortex robustly suppresses stimulus-evoked activity of excitatory neurons. a, PV-ChR2 mice ($n = 2$) were injected with AAV-CaMKII-GCaMP6f to allow simultaneous one-photon excitation of PV cells and two-photon recordings of pyramidal cell population. **b**, Schematic of simultaneous widefield optogenetics and two-photon imaging. **c**, Optogenetic activation was locked to frame acquisition. **d**, Trial-averaged ΔF/F (mean ± s.e.m.) aligned to tone onset (black vertical line) of an example neuron at different intensity of LED power (blue scale). Yellow rectangle indicates period of light delivery. **e**, Effect of optogenetic silencing as a function of LED power (median ± s.e.median; $n = 454$ neurons; Friedman test, $P \sim 0$). ΔF/F at powers 0-0.26 mW/mm² are all significantly different from ΔF/F at powers 0.84-3.15 mW/mm² (post-hoc comparisons with Tukey-Kramer test, $^{***}P < 0.001$). Black line is the logistic fit. **f**, Immunostaining of PV-ChR2 mice auditory cortex showing ChR2 expression in PV cells (PV+ and ChR2+ colocalization; scale bar, 25 µm). **g**, Post-task imaging of a representative control (top) and a representative test (PV-ChR2, bottom) mouse used in AC silencing experiments (scale bars, 100 µm). Note that no fluorescence below the dura is detected in control mice.

Extended Data Fig. 3 | AC, but not VC, full trial silencing impairs expression and on-line performance during learning but not at expert level. a, Assessment of the impact of AC full trial silencing over learning on expression by controlling for the delay in acquisition. **b**, Cumulative distribution function (CDF) of mice as function of the day to reach an accuracy ≥0.65 in probe trials. **c**, Cumulative distribution function (CDF) of mice as function of the relative number of days to reach accuracy (acc.) criteria of >0.7 (left), >0.8 (middle), and >0.9 (right) in reinforced light-off trials after reaching an accuracy ≥0.65 in probe trials. Black and dark gray vertical lines correspond to when CDF was reach for acc.>0.7 and >0.8, respectively. **d**, Comparing action rate and accuracy between reinforced light-off versus reinforced light-on trials to assess the impact of AC silencing on on-line performance. **e**, Hit (solid line) and FA (dashed line) of an example control mouse (top) and an example PV-ChR2 mouse (bottom) in reinforced light-off (black) and reinforced light-on (blue) trials across learning. **f**, Averaged action rate in reinforced light-off (black) and reinforced light-on (blue) trials per day for control (top) and PV-ChR2 (bottom) groups. **g**, Accuracy in light-on reinforced trials from the day when FA<0.3 in light-off reinforced trials. Note how PV-ChR2 mice (gray lines) increase accuracy (positive slopes) with light-on, showing that performance impairment fades away. **h**, Silencing of the visual cortex in 90% of the reinforced trials throughout learning (n = 8 PV-ChR2 mice). **i**, Comparison of reinforced light-off versus light-on trials shows no deficit when silencing the VC demonstrating the specificity of the effects of AC silencing. **j**, Accuracy in reinforced light-off and light-on trials across days (two-way repeated measures ANOVA, groups: P = 0.50). **k**, Accuracy in reinforced light-off and light-on trials (n = 168 sessions; Wilcoxon signed rank, P = 0.41). **l**, Difference in accuracy in reinforced light-on versus light-off trials per session. **m**, Difference in accuracy in reinforced light-on versus light-off trials across days in visual cortex PV-ChR2 mice (dashed line) versus auditory cortex control mice (solid line) (two-way ANOVA, groups: P = 0.13). **n**, Probabilistic optogenetic silencing of the auditory cortex at expert level. Silencing starts once stable performance is reached. **o**, Accuracy in probe light-off (green), reinforced light-off (black) and reinforced light-on (blue) trials. Silencing is performed from day 19 to 23. **p**, Accuracy in reinforced light-off and light-on trials (n = 20 days, paired t-test, P = 0.602).

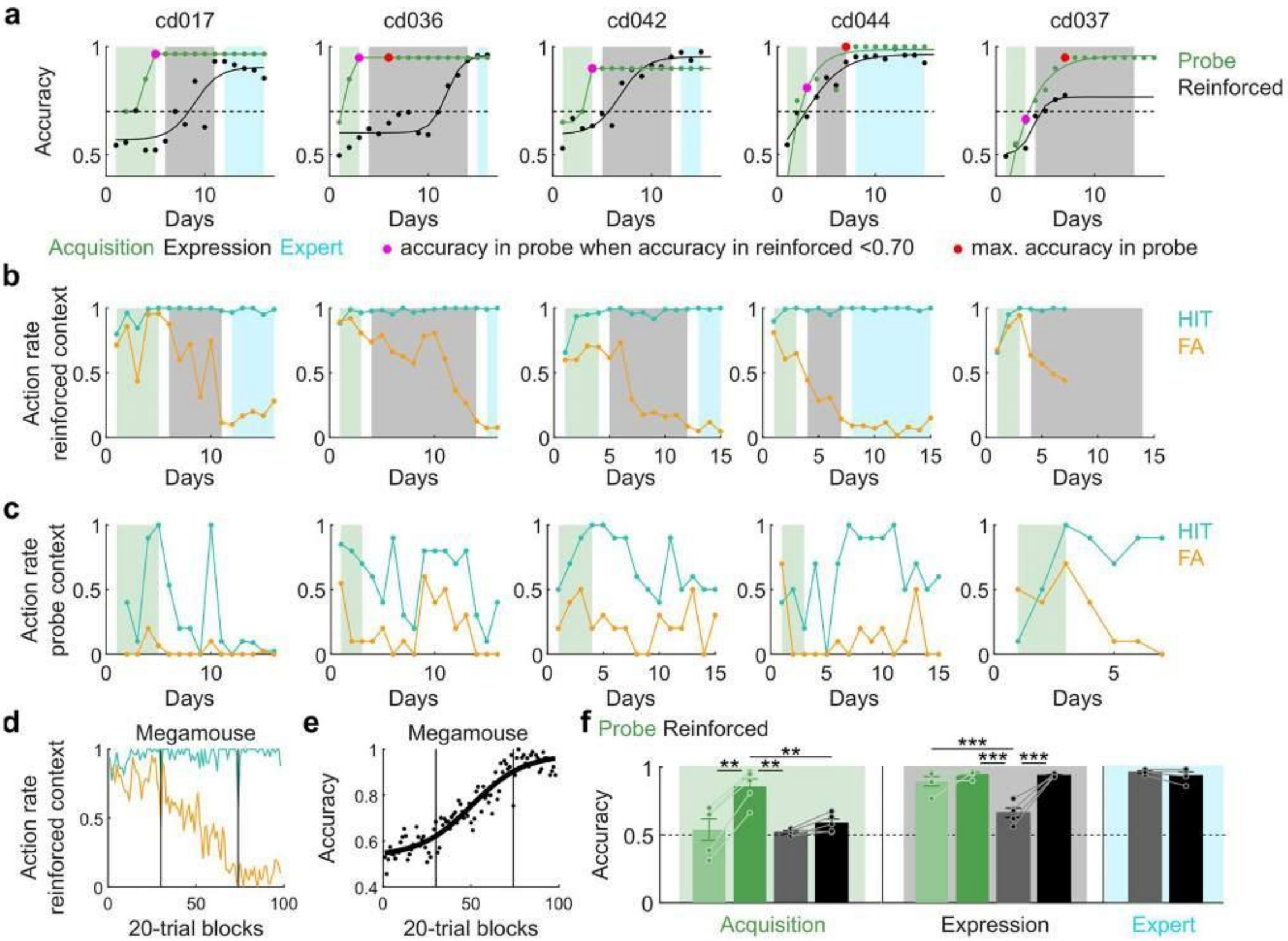


Extended Data Fig. 4 | See next page for caption.

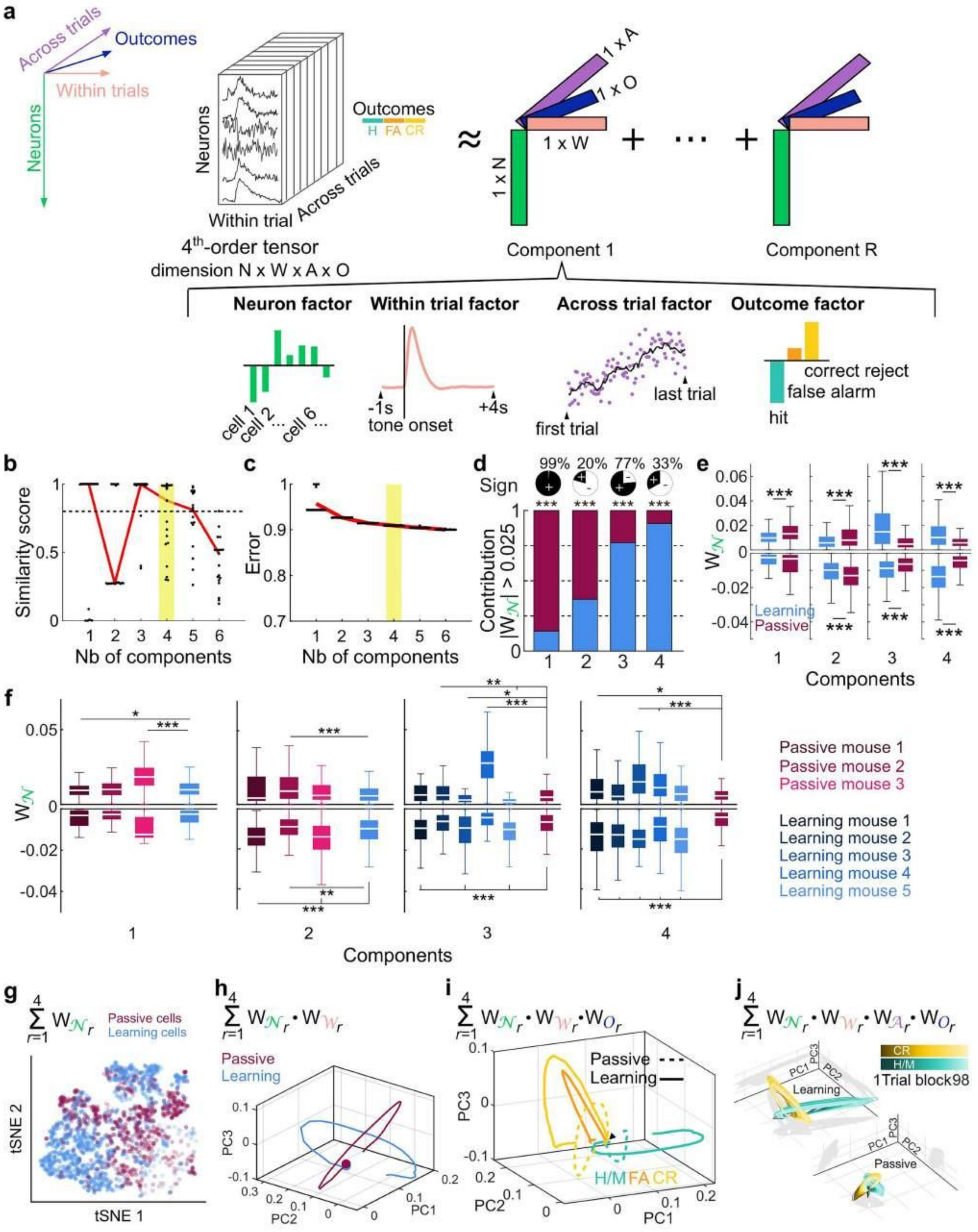
Extended Data Fig. 4 | Effect of AC full trial silencing on lick patterns. **a**, Example control (top) and PV-ChR2 (bottom) mice accuracy in probe light-off, reinforced light-off and reinforced light-on trials across day. Dashed rectangle indicates day where licks in **b** are extracted from. **b**, Lick raster plots from day 4 from the example mouse from **a** in probe light-off (left), reinforced light-off (middle) and reinforced light-on (right) trials, split into S+ (black, left) and S− (red, right) trials. Green and red dots indicates correct and incorrect trials, respectively. Note the difference in discrimination in all contexts between control and PV-ChR2 mice. **c**, Average lick probability across training days for control ($n = 8$) and PV-ChR2 ($n = 8$) mice in response to S+ (vertical green line) and S− (vertical red line) tones, in reinforced light-off (black) and light-on (blue) trials. **d**, Insets showing faster lick latencies (red arrow heads) in response to both tones and higher lick probability in response to the S− (incorrect licking) in reinforced light-on compared to light-off in PV-ChR2 mice (right). Light has no effect on lick structure in control mice (left). **e**, Lick latencies (top) and lick rate (bottom) in response to S+ (HIT trials; left) and S− (false alarm (FA) trials; right) tones in reinforced light-off trials (two-way ANOVAs, HIT lick latencies, $P = 0.0046$; FA Lick latencies, $P = 0.18$; HIT lick rate, $P = 0.087$; FA lick rate, $P = 0.0071$). **f**, Lick latencies (top) and lick rate (bottom) in response to S+ (HIT trials; left) and S− (FA trials; right) tones in reinforced light-on trials (two-way ANOVAs, HIT lick latencies, $P < 10^{-3}$; FA Lick latencies, $P < 10^{-57}$; HIT lick rate, $P = 0.0531$; FA lick rate, $P < 10^{-44}$). **g**, Lick latencies (left) and lick rate (right) in response to S+ (HIT) and S− (FA) tones in probe light-off trials (two-way ANOVAs, HIT lick latencies, $P < 10^{-3}$; FA Lick latencies, $P = 0.338$; HIT lick rate, $P = 0.017$; FA lick rate, $P = 0.89$).

Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Auditory cortex imaging and cell tracking. a, After surgery, animals underwent a 10-day recovery period after which water restriction started. Tonotopic mapping (tuning curve session) of the auditory cortex took place 5 days later under the two-photon microscope, followed by two days of lick training under the two-photon microscope. These two sessions also allowed for habituation to head fixation and context. Behaviour sessions started the following day for 15 or 16 days, after which tonotopic mapping sessions took place at day +1, +7 and +15 post learning. **b**, One behavioural session consisted of three blocks of 80 or 100 trials, and a baseline session (no tone presented). Two groups of mice were imaged under the two-photon microscope: the Passive group (top; $n = 3$) was presented with two pure tones but was never rewarded (lick tube out), and the Learning group ($n = 5$) was rewarded (3μl water drop) if licking in the response window after the S+ tone. Two probe blocks of 10 trials each were introduced in two of the three reinforced blocks. **c**, Trial structure. After a no-lick period of 1s, a 100-ms tone was played, followed by a 200-ms dead period and a ≤2.5s response period. The length of the delay period was of 2s after a miss (M, no lick after S+) or a correct reject (CR, no lick after S−), 4s after a hit (H, lick after S+) and 7s after a false alarm (FA, lick after S−). **d**, Tonotopic organization of the field of view of one example mouse before learning (same as 2b; left; scale bar, 73 µm). Cells are colored according to their best frequency and tone-evoked responses of example cells circled in black to 17 pure tones ranging from 4 to 64 kHz are displayed on the right. **e**, Tone-evoked activity (top) and proportion of responsive cells (bottom) to pure tones. S+ and S− (filled and unfilled triangles, respectively) are chosen for training in the task based on their equal representation in the field of view in **d**. **f**, Two planes recorded in one example mouse. Cells are colored according to the number of days tracked among the 19 recording sessions in this mouse (scale bar, 73 µm). **g**, Distribution of number of tracked days per cells in **f**. **h**, Cumulative distribution of tracked cells according to the percentage of recording sessions. Data for mouse in **f** is the light blue line. **i**, Pre-task stimulus selectivity index of neurons with S+ or S− as best frequency (BF). Shuffle distribution ($n = 500$) generated from randomly selected neurons. **j**, Underlying distribution of boxplots in **i**. **k**, Relative best frequency as a function of pre-task stimulus selectivity index. The two measures are significantly correlated (Pearson's correlation coefficient −0.49, $P = 3.93.10^{-73}$). The y-axis represents the BF distance to the S+ (set at 0) and S− (set at +0.75) because mice were trained with different S+ and S−, and the S+ could be higher or lower than the S−.

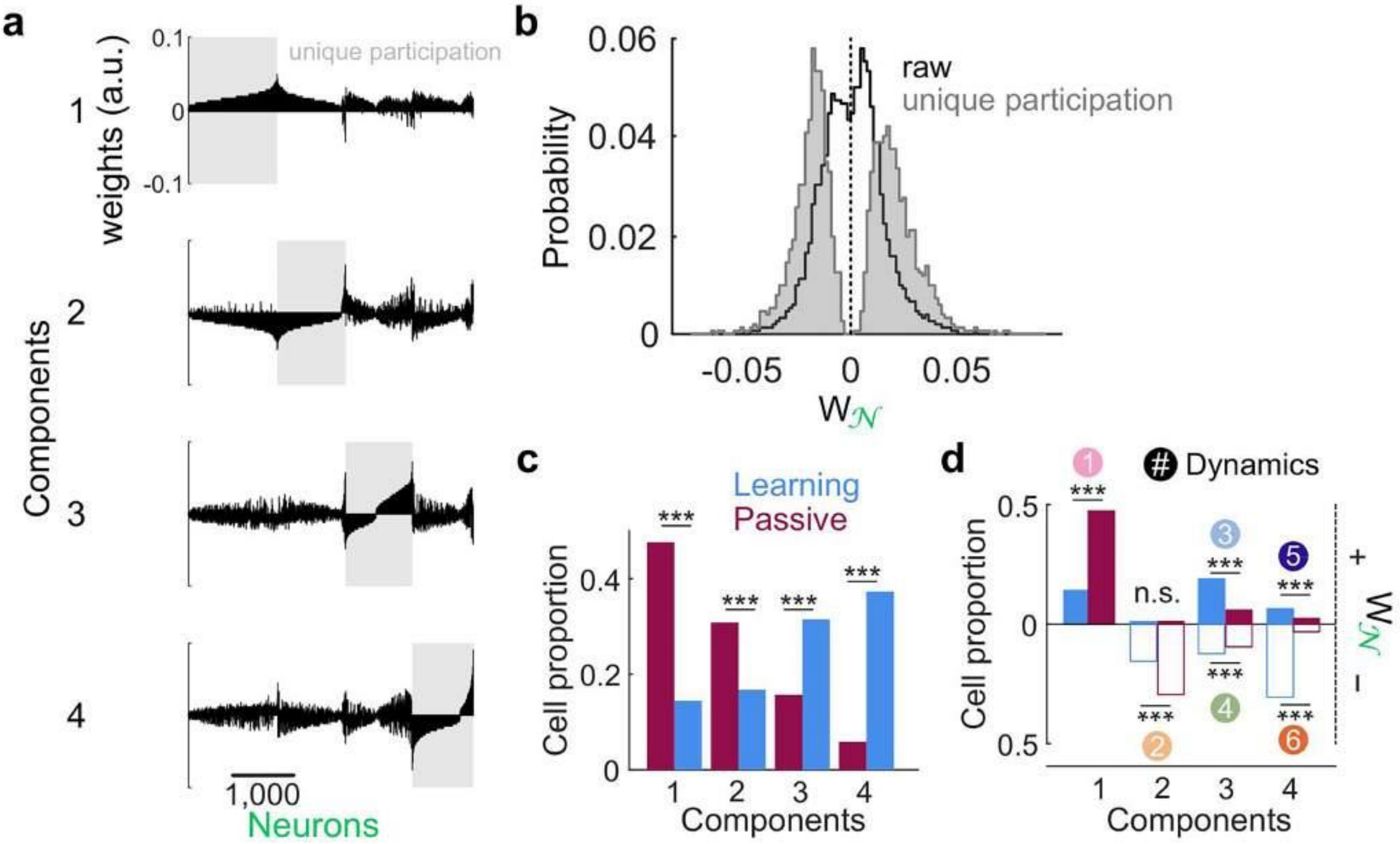


Extended Data Fig. 6 | Inter-subject performance alignment for megamouse tensor. a, Accuracy in probe and reinforced contexts across days of all learning mice. **b**, Action rate in reinforced context across days of all learning mice. **c**, Action rate in probe context across days of all learning mice. Please note that we fixed the probe performance at the maximum discrimination that was followed by a decrease in hit rate due to extinction. **d**, After the alignment procedure, action rate from the megamouse (all learning mice pooled) in reinforced context across learning phases. **e**, Megamouse accuracy in reinforced context across learning phases. **f**, Accuracy difference between the start (light color) and the end (dark color) of the three learning phases in probe (green) and reinforced (black) contexts ($n = 5$ mice). Acquisition is characterized by an increase of accuracy in probe trials (paired t-test, $P = 5.47.10^{-4}$) but not in reinforced trials (paired t-test, $P = 0.07$), expression corresponds to an increase of accuracy in reinforced trials (paired t-test, $P = 0.008$) and expert is when accuracy in reinforced trials is high and stable (paired t-test, $P = 0.27$).



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Low-rank tensor decomposition. a, Data are organized into a fourth-order tensor with dimensions N × W × A × O. Tensor decomposition approximates the data as a sum of outer products of four vectors. Each outer product contains a neuron factor (green rectangles), within trial factor (pink rectangles), across trial factor (blue rectangles) and outcome factor (purple rectangles). Each set of low-dimensional factors (i.e. component) describes the activity of group of neurons within and across trials according to trial outcomes. **b**, Similarity score as a function of model components. Each dot shows the similarity of a single optimization run compared to the best-fit model within each category. Yellow rectangle indicates the number of components selected. **c**, Model reconstruction error as a function of the number of components, where each dot corresponds to a different optimization run. **d**, Neuronal contribution (learning (blue) vs passive (burgundy) cells) per components (binomial proportion tests, all $P < 0.001$). **e**, Positive and negative neuronal weights across components in cell population recorded in learning mice (Learning) and in passive mice (Passive) (Wilcoxon tests; component (comp.) 1(+), $P = 7.80.10^{-62}$, $n = 2211, 2299$ cells; comp. 1(−), $P = 0.53$, $n = 93, 40$ cells; comp. 2(+), $P = 0.0001$, $n = 474, 355$ cells; comp. 2(−), $P = 1.89.10^{-30}$, $n = 1830, 1984$ cells; comp. 3(+), $P = 1.45.10^{-63}$, $n = 988, 946$ cells; comp. 3(−), $P = 7.10.10^{-21}$, $n = 1316, 1393$ cells; comp. 4(+), $P = 1.31.10^{-34}$, $n = 621, 1639$ cells; comp. 4(−), $p = 6.10.10^{-106}$, $n = 1683, 700$ cells). **f**, Positive and negative neuronal weights across components and individual mice (KW tests; comp. 1(+), $P = 1.50.10^{-216}$, $n = 562, 547, 1190, 2211$ cells; comp. 1(−), $P = 0.31465$, $n = 12, 21, 7, 93$ cells; comp. 2(+), $P = 3.32.10^{-5}$, $n = 21, 238, 96, 474$ cells; comp. 2(−), $P = 4.18.10^{-56}$, $n = 553, 330, 1101, 1830$ cells; comp. 3(+), $P = 3.40.10^{-155}$, $n = 55, 317, 75, 521, 20, 946$ cells; comp. 3(−), $P = 1.53.10^{-38}$, $n = 142, 312, 173, 25, 664, 1393$ cells; comp. 4(+), $P = 2.95.10^{-40}$, $n = 62, 84, 90, 290, 95, 1639$ cells; comp. 4(−), $P = 6.93.10^{-113}$, $n = 135, 545, 158, 256, 589, 700$ cells). **g**, t-SNE of neuronal weights. Note how learning and passive cell populations are largely non-overlapping. **h**, Projection of neuronal × within trial weights of learning and passive network activity into principal component space. **i**, Projection of neuronal × within trial × trial outcome weights of learning and passive network activity into principal component space. **j**, Projection of neuronal × within trial × across trials × trial outcome (H/M and CR only) weights of learning and passive network activity into principal component space.

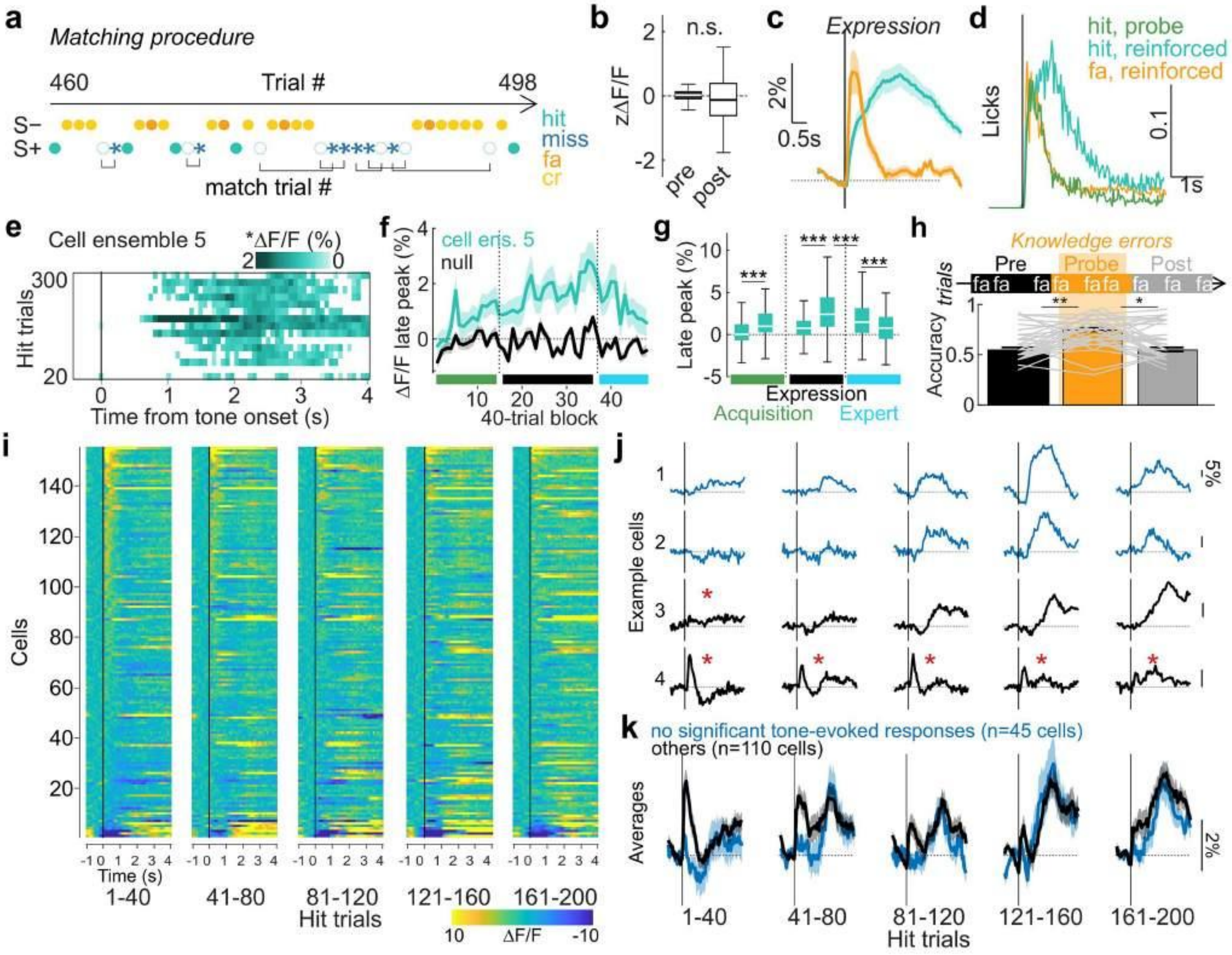


Extended Data Fig. 8 | Defining unique cell ensembles based on neuronal weights. a, Neuronal weights in the four components. Each neuron is attributed to a given dynamic according to its highest absolute weights, i.e. highest contribution. As a result, each dynamic is attributed to a unique cell ensemble (gray rectangles). **b**, Neuronal weights distribution before (raw, black) and after unique contribution attribution (gray). **c**, Learning and passive cell proportion among components after unique attribution (binomial proportion tests). **d**, Learning and passive cell proportion among components and given neuronal weight sign after unique attribution. In other words, proportion of cells from learning and passive networks describing the tensor-revealed neuronal dynamics (binomial proportion tests). ***$P < 0.001$, n.s.: not significant.

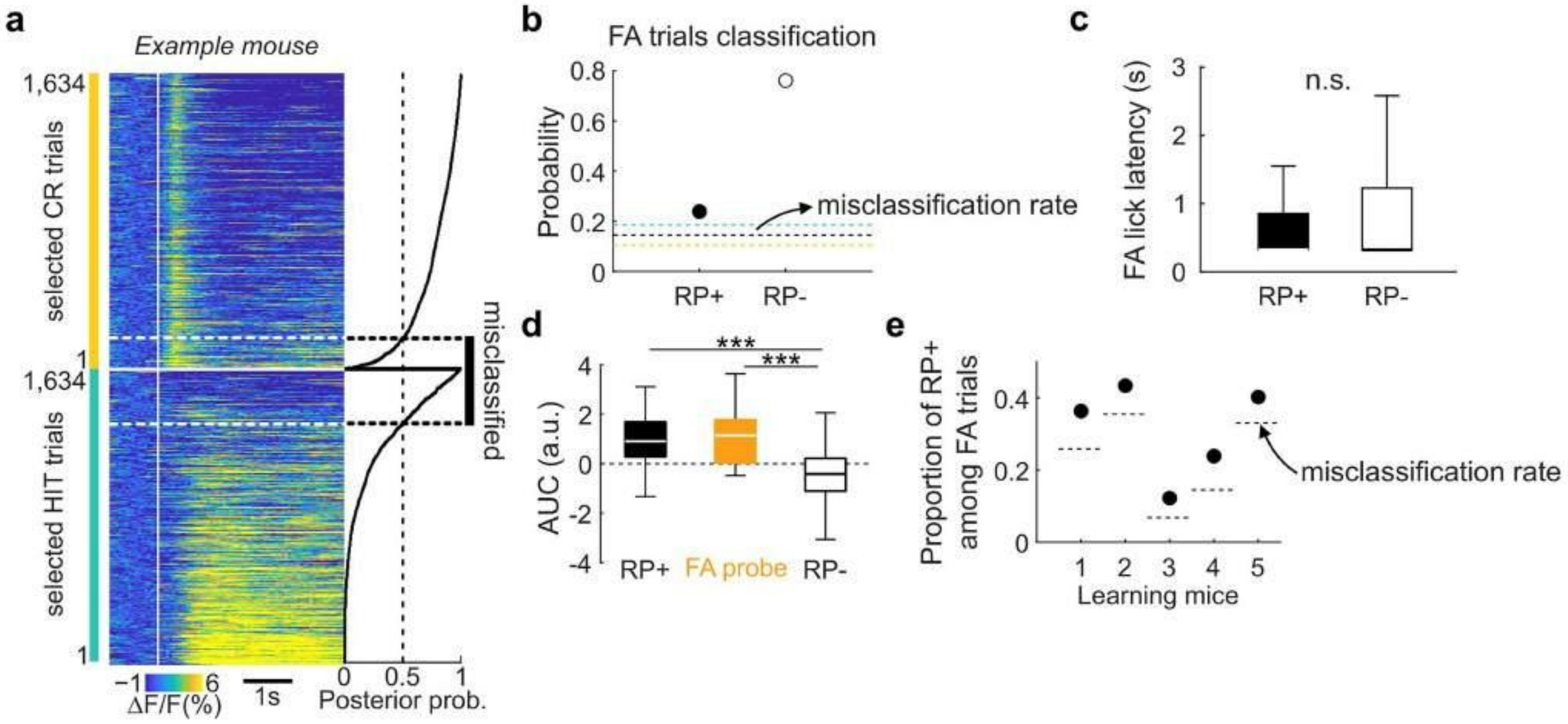
Extended Data Fig. 9 | Learning counteracts tone-evoked habituation. **a**, Tone-evoked responses to S+ and S− in learning mice across days for all cells recorded. **b**, Tone-evoked responses to S1 and S2 in passive mice across days for all cells recorded. **c**, Proportion of tone-responsive cells across days among passive and learning cells (binomial proportion tests). **d**, Averaged proportion of tone-responsive cells in passive and learning networks (mean ± s.e.m.; $n = 15$ days, t-test, $P = 3.89.10^{-5}$). **e**, Proportion of tone-responsive cells in days 1-5 versus days 11-15 in learning and passive networks (mean ± s.e.m.; $n = 5$ days per groups, two-way ANOVA, Time × Group, $P = 1.73.10^{-7}$). **f**, Proportion of cells responsive to S+ and S− in learning network and S1, S2 or *S1 or S2* (S) in passive network. **g**, Averaged proportion of cells responsive to S+, S− or S (mean ± s.e.m.; $n = 15$ days, one-way ANOVA, $P = 1.93.10^{-6}$).

Extended Data Fig. 10 | See next page for caption.

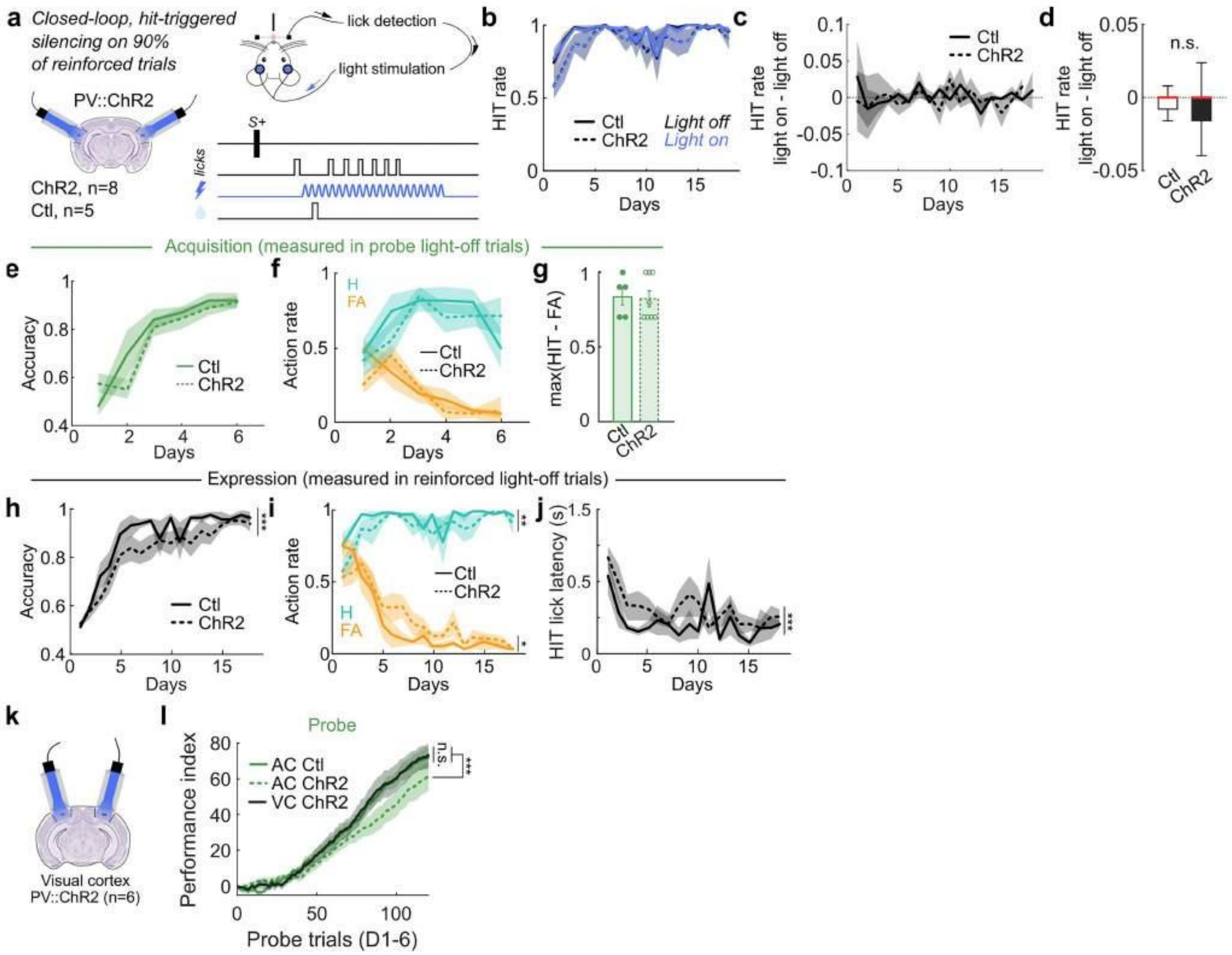
Extended Data Fig. 10 | AC silencing restricted to sound presentation impairs audiomotor learning and on-line performance during learning. **a**, Probabilistic optogenetic silencing of the auditory cortex during learning. Light-on periods were restricted to sound presentation only (see Methods). **b**, Accuracy in reinforced light-on trials (two-way ANOVA, $P<10^{-9}$). **c**, Action rate in reinforced light-on trials (two-way ANOVAs, HIT, $P=0.89$; FA, $P<10^{-8}$). **d**, False alarm lick rate in reinforced light-on trials (two-way ANOVA, $P<10^{-14}$). **e**, False alarm lick latency in reinforced light-on trials (two-way ANOVA, $P<10^{-26}$). **f**, Accuracy in probe light-off trials (two-way ANOVA, $P=0.038$). **g**, Action rate in probe light-off trials (two-way ANOVAs, HIT, $P=0.017$; FA, $P=1$). **h**, HIT lick latency in probe light-off trials (two-way ANOVA, $P=0.0032$). **i**, Maximal difference between hit and false alarm rates in probe light-off trials over the first 6 days ($n=3$ control mice, $n=4$ PV-ChR2 mice, t-test, $P=0.40$). **j**, Accuracy in reinforced light-off trials (two-way ANOVA, $P=0.20$). **k**, Action rate in reinforced light-off trials (two-way ANOVAs, HIT, $P=0.15$; FA, $P=0.48$). **l**, FA lick latency in reinforced light-off trials ($n=54$ control days, $n=72$ PV-ChR2 days, two-way ANOVA, $P=0.0013$). **m**, FA lick rate in reinforced light-off trials ($n=54$ control days, $n=72$ PV-ChR2 days, two-way ANOVA, $P=4.44.10^{-4}$). **n**, Comparison of light-off versus light-on trials to measure auditory cortex silencing effect on on-line performance. **o**, Session density plot of accuracy in reinforced light-on against light-off. Top, control; bottom, PV-ChR2. **p**, Accuracy in light-on reinforced trials from day where FA < 0.3 in light-off reinforced trials. Note the general trend for PV-ChR2 mice (gray lines) to increase accuracy (positive slopes), i.e. performance impairment fades away. **q**, Within subject difference between accuracy in reinforced light-on and light-off aligned to the day where false alarm rate <0.3 in reinforced light-off.



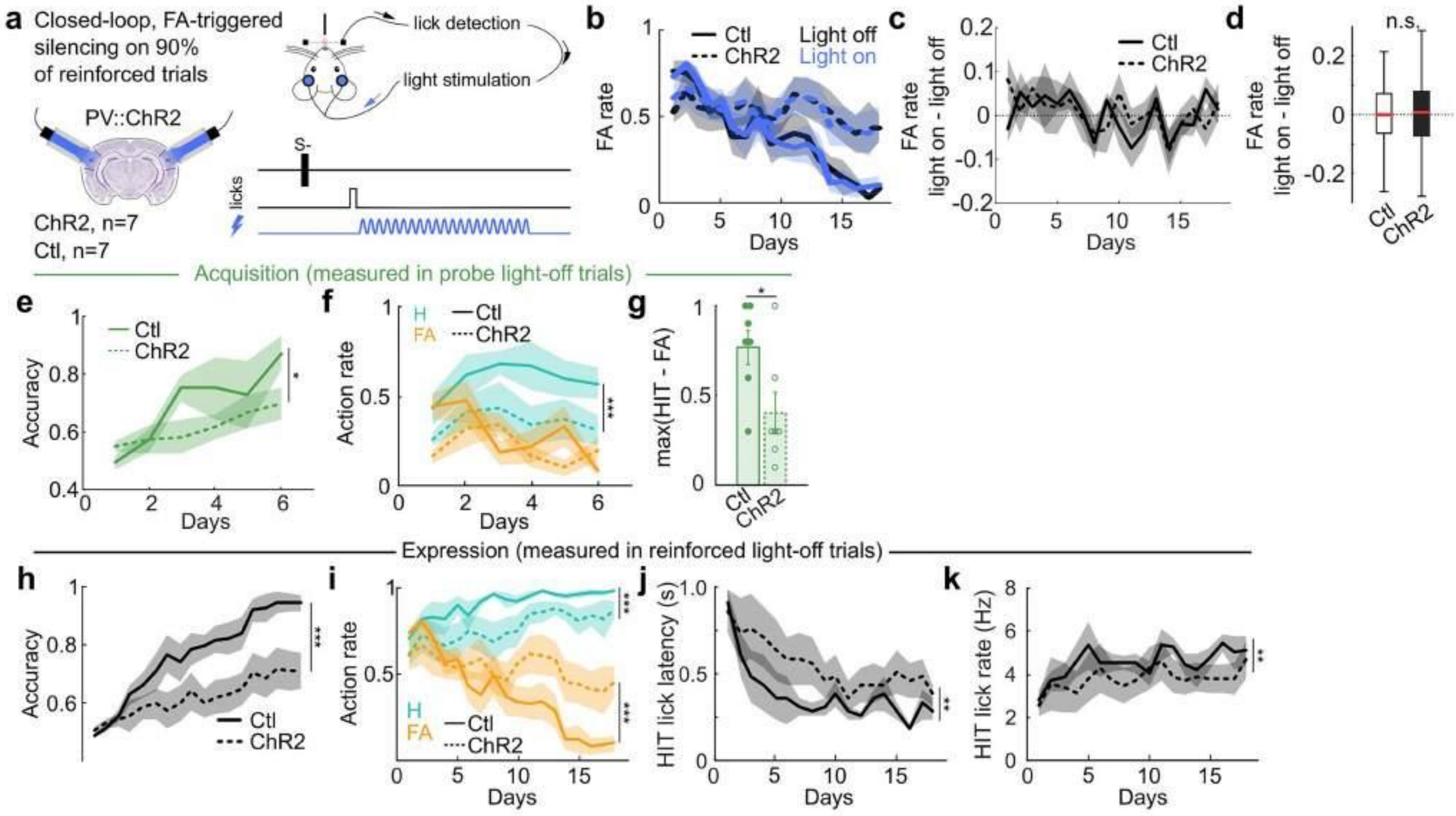
Extended Data Fig. 11 | Emergence of reward prediction signal. a, Procedure of hit and miss trial matching. **b**, Quantification of z-scored calcium activity 1s pre- vs 1s post non-task lick bout onset in Fig. 4g ($n = 105$ cells, Wilcoxon test, $P = 0.20$). **c**, Average cell ensemble 5 activity in reinforced hit (green) and FA (orange) trials over expression phase. **d**, Lick PSTHs aligned to tone onset of FA trials in expression and hit trials in probe context. **e**, Cell ensemble 5 activity over the first 300 hit trials (20-trial blocks). Only significant activity (and higher than null population, see Methods) is represented. Note the emergence of a stable late-in-trial signal after 40 hit trials onwards. **f**, Late-in-trial activity in hit trials across learning phases of cell ensemble 5 ($n = 155$ cells, green) and low weighted cells (null, black). **g**, Quantification of **f**, taking first and last two 40-hit trial blocks in each learning phase (KW test, $P = 1.05.10^{-23}$). **h**, Procedure of reinforced and probe FA trial matching and corresponding local accuracy quantification related to Fig. 4i (see Methods; $n = 34$ matches, repeated measures ANOVA, $P = 3.16.10^{-4}$). **i**, Activity of all neurons in cell ensemble 5 over the first five blocks of 40-reinforced hit trials in learning. Traces are ordered according to tone-evoked response in the first block. **j**, Four example cells showing the heterogeneity in tone-evoked responses in cell ensemble 5: cells 1 and 2 do not exhibit significant tone-evoked responses over the first five blocks, while cell 3 is tone responsive (red star) only in the first block and cell 4 is tone responsive over all 5 blocks. **k**, Average activity (mean ± s.e.m.) of cells without tone-evoked responses in the five blocks (blue) vs with tone-evoked responses in at least one block (black). Note that the amplitude of the late-in-trial signal is similar between groups. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, n.s.: not significant.



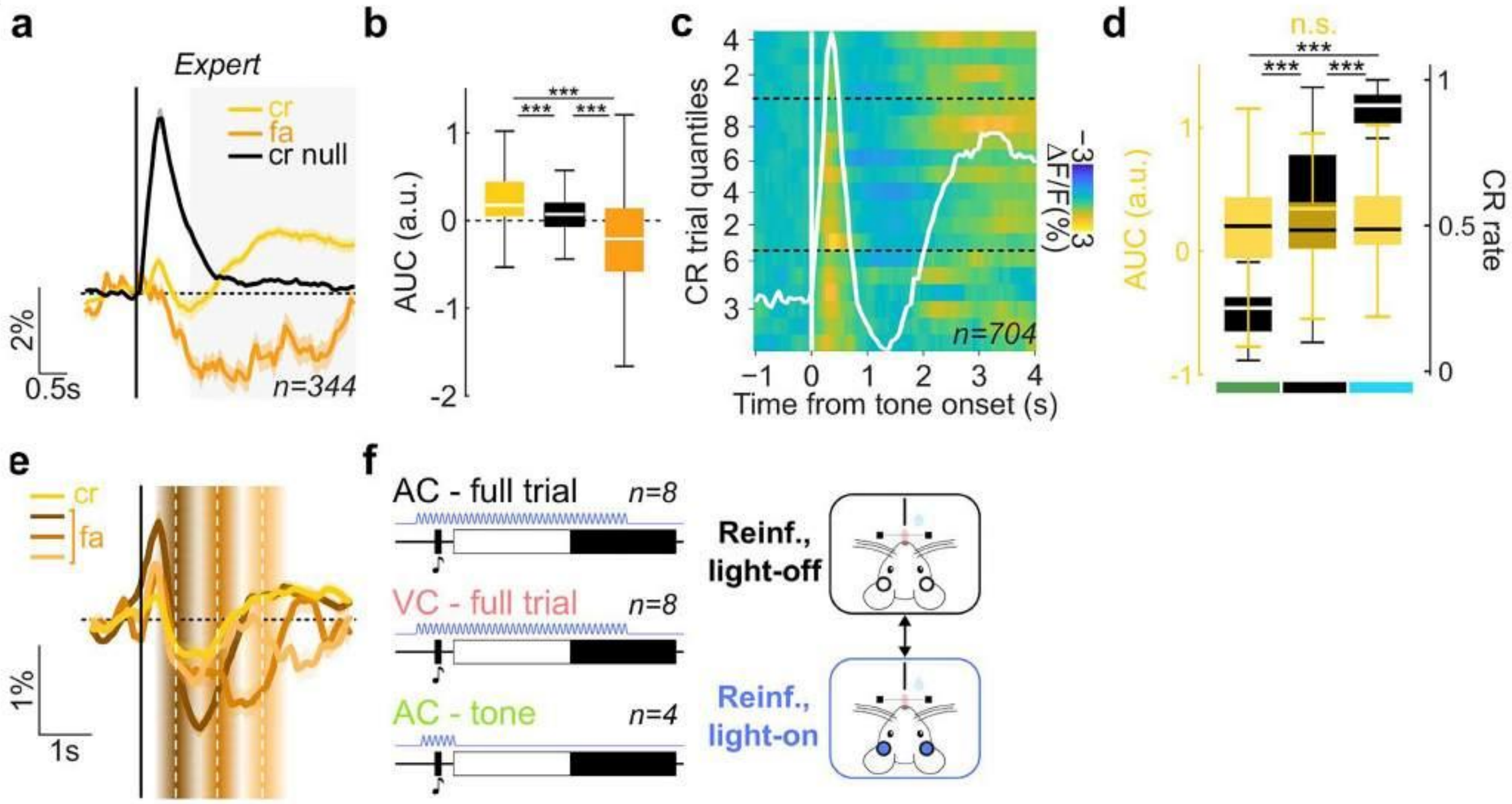
Extended Data Fig. 12 | Reward prediction signal on error trials. **a**, Classification of hit versus CR trials in the reinforced context from the AUC post-tone of a fraction of cell ensemble 5 ($n = 51$ cells) recorded in the example mouse showed in Fig. 4j–l. Right: posterior probability of being part of CR class. **b**, Proportion of RP+ and RP− FA trials from the example mouse showed in Fig. 4j–l. **c**, No difference in lick latency was observed between RP+ ($n = 101$) and RP− ($n = 322$) FA trials (Wilcoxon test, $P = 0.83$). **d**, AUC quantification of RP+ ($n = 101$), RP− ($n = 322$) and probe ($n = 19$) FA trials (KW test, $P = 9.76.10^{-28}$). **e**, Proportion of RP+ among all FA trials and misclassification rate in each learning mice. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, n.s.: not significant.



Extended Data Fig. 13 | Post-hit silencing over learning. a, Experimental design of optogenetic silencing of AC activity throughout learning post hit only. **b**, Hit rate across days in control (Ctl) and test (PV-ChR2) mice in reinforced light-on or light-off trials across days. **c**, Difference in hit rate in reinforced light-on versus light-off trials across days. **d**, Difference in hit rate in reinforced light-on versus light-off trials ($n = 90$ control sessions, $n = 144$ PV-ChR2 sessions, Wilcoxon test, $P = 0.13$). **e**, Accuracy in probe light-off trials (two-way ANOVA, $P = 0.3$). **f**, Action rate in probe light-off trials (two-way ANOVAs, HIT, $P = 0.58$; FA, $P = 0.47$). **g**, Maximum difference between hit and false alarm trials over the first 6 days in probe light-off trials ($n = 5$ control mice, $n = 8$ PV-ChR2 mice, t-test, $P = 0.86$). **h**, Accuracy in reinforced light-off trials (two-way ANOVA, $P = 7.33.10^{-5}$). **i**, Action rate in reinforced light-off trials (two-way ANOVAs, HIT, $P = 0.0068$; FA, $P = 0.023$). **j**, Hit lick latency in reinforced light-off trials (two-way ANOVA, $P = 0.0019$). **k**, Silencing of visual cortex (VC) activity throughout learning post hit only. **l**, Performance index in probe trials for AC control ($n = 5$), AC PV-ChR2 ($n = 8$) and VC PV-ChR2 ($n = 6$) (two-way ANOVA, $P = 1.90.10^{-32}$).



Extended Data Fig. 14 | Post-FA silencing over learning. a, Experimental design of optogenetic silencing of AC activity throughout learning post false alarm (FA) only. **b**, False alarm rate across days in control (Ctl) and test (PV-ChR2) mice in reinforced light-on or light-off trials across days. **c**, Difference in false alarm rate in reinforced light-on versus light-off trials across days. **d**, Difference in false alarm rate in reinforced light-on versus light-off trials ($n = 126$ control and $n = 126$ ChR2 sessions, t-test, $P = 0.76$). **e**, Accuracy in probe light-off trials (two-way ANOVA, $P = 0.03$). **f**, Action rate in probe light-off trials (two-way ANOVAs, HIT, $P = 1.07.10^{-4}$; FA, $P = 0.1$). **g**, Maximum difference between hit and false alarm trials over the first 6 days in probe light-off trials ($n = 7$ control and $n = 7$ ChR2 mice, t-test, $P = 0.028$). **h**, Accuracy in reinforced light-off trials (two-way ANOVA, $P = 2.17.10^{-14}$). **i**, Action rate in reinforced light-off trials (two-way ANOVAs, HIT, $P = 4.28.10^{-8}$; FA, $P = 1.31.10^{-5}$). **j**, Hit lick latency in reinforced light-off trials (two-way ANOVA, $P = 7.99.10^{-4}$). **k**, Hit lick rate in reinforced light-off trials (two-way ANOVA, $P = 0.005$).



Extended Data Fig. 15 | A signal for action suppression in Learning network. **a**, Average activity of cell ensemble 6 ($n = 344$ cells) or low weighted cells ($n = 532$ cells, null, black) in CR and FA trials in expert phase. **b**, Quantification of late-in-trial activity ($n = 344, 532, 344$ cells, KW test, $P = 4.76.10^{-44}$). **c**, Average cell ensemble 6 activity across learning phases. CR trials were split into 6, 9 and 4 quantiles over acquisition, expression and expert phases, respectively. **d**, Quantification of late-in-trial activity (left axis; $n = 704$ cells, KW test, $P = 0.09$) and CR rate (right axis; $n = 30, 44, 24$ trial blocks, KW test, $P = 4.23.10^{-16}$) over learning phases. **e**, Averaged ensemble 6 activity in FA and CR trials. FA trials are split according to lick latencies (white dashed line, mean latency; graded rectangles, latency range extrema). **f**, Silencing protocols compared in Fig. 5h,i.