Learning enhances behaviorally relevant representations in apical dendrites

Sam E. Benezra[1,3,4], Kripa B. Patel[2,3,4], Citlali Pérez Campos[2,3,4], Elizabeth M.C. Hillman[1,2,3,4], and Randy M. Bruno[1,3,4,5,6*]

[1] Department of Neuroscience, Columbia University, New York, NY 10027
[2] Departments of Biomedical Engineering and Radiology, Columbia University, New York, NY 10027
[3] Kavli Institute for Brain Science, Columbia University, New York, NY 10027
[4] Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10027
[5] Department of Physiology, Anatomy & Genetics, University of Oxford, Oxford, OX1 3PT

[6] Lead contact

* Correspondence: randy.bruno@dpag.ox.ac.uk (R.M.B.)

SUMMARY

Learning alters cortical representations and improves perception. Apical tuft dendrites in Layer 1, which are unique in their connectivity and biophysical properties, may be a key site of learning-induced plasticity. We used both two-photon and SCAPE microscopy to longitudinally track tuft-wide calcium spikes in apical dendrites of Layer 5 pyramidal neurons in barrel cortex as mice learned a tactile behavior. Mice were trained to discriminate two orthogonal directions of whisker stimulation. Reinforcement learning, but not repeated stimulus exposure, enhanced tuft selectivity for both directions equally, even though only one was associated with reward. Selective tufts emerged from initially unresponsive or low-selectivity populations. Animal movement and choice did not account for changes in stimulus selectivity. Enhanced selectivity persisted even after rewards were removed and animals ceased performing the task. We conclude that learning produces long-lasting realignment of apical dendrite tuft responses to behaviorally relevant dimensions of a task.

INTRODUCTION

Learning and memory depend on the ability of biological networks to alter their activity based on past experience. For example, as animals learn the behavioral relevance of stimuli in a sensory discrimination task, neural representations of those stimuli are enhanced[1-7], potentially improving the salience of information relayed to downstream areas. Studies in primary somatosensory (S1)[8] and visual cortex[2] have revealed that top-down signals from distant cortical regions can modify sensory representations during learning, although the cellular and circuit mechanisms underlying this plasticity remain unclear.

Cortical layer 1, comprised mainly of apical tuft dendrites of layer 5 (L5) and layer 2/3 pyramidal neurons, may be a key site driving the enhancement of sensory representations during learning. Apical tufts are anatomically well positioned for learning, receiving top-down signals from numerous cortical and thalamic areas[9-11]. While L5 distal tufts are electrically remote and far from the soma, they are in close proximity to the highly electrogenic calcium spike initiation zone at the main bifurcation of the apical dendrite, and form a separate biophysical and processing compartment from the proximal dendrites[12-16]. Top-down signals arriving at the tuft can trigger tuft-wide dendritic calcium spikes in L5 neurons[17], which can modulate synaptic plasticity across the entire dendritic tree[18] and potently drive somatic burst firing[15,19-23]. Consistent with this observation, L5 apical dendrite activity is highly correlated with somatic activity[24,25]. Therefore, by strongly influencing somatic activity, L5 apical dendritic calcium spikes can play an important role in modulating cortical output. Several neuromodulators can augment the excitability of the apical tuft and increase the likelihood of eliciting calcium spikes[26,27], which could be a substrate for control of plasticity by behavioral state. Consistent with these ideas, we recently demonstrated that during behavioral training with positive reinforcements, apical tufts in sensory cortex acquire associations that extend beyond their normal sensory modality[28]. In mouse models of dementia and Alzheimer's disease[29,30], tuft dendrites exhibit degeneration which may contribute to the cognitive and memory deficits.

L5 pyramidal neurons are the major source of output from cortex, targeting numerous subcortical structures that affect behavior. The activity of apical dendrites is known to correlate with stimulus intensity, and manipulating L5 apical dendrites and their inputs impacts performance of sensory tasks[17,31-33]. Apical dendritic calcium spikes of pyramidal cells could be a crucial cellular mechanism in learning-related plasticity and behavioral modification[18,34,35]. However, sensory representations of apical tufts, as well as possible changes across learning, have received little attention.

To address this question, we used two-photon microscopy and a new high-speed volumetric imaging technique called Swept Confocally-Aligned Planar Excitation (SCAPE)[36,37] to longitudinally track the activity of GCaMP6f-expressing L5 apical tufts in barrel cortex during a sensory discrimination task. We found that apical tufts underwent extensive dynamic changes in selectivity for task-relevant stimuli as performance improved, even though only one of the stimuli was unrewarded. These changes in responses persisted even after animals disengaged from the task, demonstrating that learning induced long-lasting changes in tuft sensory representations. Animals that were exposed to the same stimulation protocol without any reinforcement did not develop enhanced representations. Our results show for the first time that reinforcement learning expands apical tuft sensory representations along behaviorally relevant dimensions.

RESULTS

Direction discrimination behavior

We devised an awake head-fixed mouse conditioning paradigm that enables controlled investigation of reinforcement effects across learning (Fig.1A,B). In addition to discriminating tactile objects, rodents are known to sense wind direction using their whiskers[38,39] and can be trained to discriminate different directions of whisker deflections[40,41]. With this in mind, we directed brief (100-ms) air puffs at the whiskers in either of two directions: rostrocaudal (backward) or ventrodorsal (upward). One of the directions was paired with a water reward delivered 500 ms after the air puff and thus constituted a conditioned stimulus (CS+). No reward was given for the other direction (CS-).

Licking and whisking were monitored throughout the session (Fig.1C,D). Stimuli elicited a brief passive whisker deflection followed by active whisking over the subsequent ~1.5 seconds

(analyzed below, Fig.6). Any anticipatory licks prior to reward delivery were counted as a response. Typically, on the first session, mice exhibited few anticipatory licks to either stimulus (Fig.1C, top, grey shading). By session 2 or 3, mice had learned an association between whisker deflection and reward, but could not discriminate the CS+ and CS- (middle). Within a week (by sessions 7-9), every mouse we tested learned to reliably lick to the CS+ while withholding licks to the CS-, performing substantially above chance after a single week of training (Fig.1C, bottom; Fig.1E,F). Thus, mice rapidly learned to discriminate the direction of whisker stimuli in our behavioral task.

Overall stimulus-evoked activity is unbiased and stable across conditioning

To investigate the effects of reinforcement learning on apical tuft activity, we imaged apical tufts (433 x 433 μm field of view) across conditioning days as well as on an unrewarded pre-conditioning day to measure naïve stimulus responses and an unrewarded post-conditioning day to detect any long-lasting changes in responses (Fig.1B). Mice remained water-restricted on the post-conditioning day and continued licking for reward toward the beginning of the session (see below). We virally delivered the gene for Cre-dependent GCaMP6f[42] in the barrel cortex of Rbp4-Cre mice, which labels a heterogeneous population of pyramidal neurons comprising approximately 50% of layer 5[28,43,44]. By targeting our injections to layer 5B, we predominantly labeled thick-tufted pyramidal neurons (see Methods). Using intrinsic signal imaging, we mapped the location of the C2, D2, and gamma whisker barrel columns and identified an overlapping region in layer 1 with sufficient GCaMP6f expression (Fig.2A). The air puff nozzles were aimed toward the whiskers corresponding to this region. Dendritic activity was longitudinally recorded from the same field-of-view (horizontal location and depth) in layer 1 across all sessions (Supplementary Movie 1).

To extract calcium signals from individual cells, we segmented tufts using CaImAn, a sparse non-negative matrix factorization method that clusters pixels according to their temporal correlation[45] (see Methods), and analyzed regions of interest exhibiting apical tuft structure (Fig.2B; 65 ± 15 tufts per mouse; mean ± SD). Individual segmented tufts were substantial in their spatial extent (>100 μm), reflecting tuft-wide voltage-gated calcium spikes rather than branch-specific N-methyl-D-aspartate (NMDA) receptor-mediated spikes. All calcium analyses hereafter refer to tuft-wide calcium spikes. Average responses to an event include failures. In many tufts, the CS+ and CS- reliably evoked an influx of calcium that robustly activated the tuft (examples in Fig.2C). Successful calcium events across tufts averaged 28% ΔF/F, consistent with previous studies of layer 5 apical dendrites[17,31]. Interestingly, during intermediate but not early learning, the average population response to the CS+ exhibited a two-peak structure (Supp Fig.1, session 4) similar to tuft reward-related signals we observed previously in barrel cortex[28]. By the last-rewarded and post sessions, the second CS+ peak was no longer visible, which could be an endpoint of mice learning that the conditioned stimulus predicts the upcoming reward.

Reward can alter somatic receptive fields in the auditory, visual, and somatosensory cortex of both rodents and non-human primates such that rewarded stimulus representations become more robust after learning[4,5,28,46], although cortical sensory responses can remain unchanged during learning[47]. We investigated whether calcium responses to the CS+ increased in the tuft population as animals learned its association with reward (Fig.2). Average responses of tufts to the CS+ and CS- were similar during the pre-conditioning session (Fig.2D; $p = 0.20$, signed rank test, $n = 440$ pre tufts and 418 post tufts), indicating that there was no inherent bias in the population toward a particular stimulus in naïve animals. Surprisingly, even after learning, responses to the CS+ and CS- were similar on the last- and post-conditioning sessions ($p = 0.62$, 0.64, respectively, signed rank test, Fig.2D,E), revealing that no bias develops for the CS+ among dendritic tufts. Only a minority of tufts exhibited statistically significant (see Methods) average responses to air puff

stimuli (CS+ responsive: 26 ± 8%; CS- responsive: 25 ± 8%; mean ± SD across all sessions). When we excluded responses that were not statistically significant, we again found no difference between the average response amplitudes to the CS+ and CS- on the pre, last-rewarded, and post sessions ($p = 0.65$, 0.31, and 0.69, respectively, rank sum test; data not shown). Similarly, the probability of transients in response to CS+ versus CS- (see Methods) did not differ during pre-conditioning or post-conditioning sessions ($p = 0.66$ and $p = 0.44$, respectively, data not shown). Therefore, reinforcement learning in our paradigm does not bias tuft representations toward the rewarded stimulus.

While a bias for the CS+ did not develop after learning, we wondered whether overall tuft responses to both conditioned stimuli increased as animals learned the task. Linear regression analysis revealed that conditioning session number was a poor predictor of both CS+ and CS- amplitudes (All tufts R^2, CS+: 0.0064, CS-: 0.0035, Fig.2E; Significantly responding tufts R^2, CS+: 0.014, CS-: 0.014, data not shown). We did find a small but significant decrease in amplitude from pre to last for CS+ ($p < 0.01$) and CS- ($p < 10^{-7}$), but this was not permanent: amplitudes did not significantly differ between the pre and post sessions (Fig.2D; $p = 0.53$, 0.33, CS+ and CS- respectively, Wilcoxon rank sum test). Taken together, these findings demonstrate that reinforcement learning does not robustly bias the magnitudes of tuft calcium responses to either stimulus at the population level.

Development of tuft selectivity with task learning

While learning produced no bias in overall tuft activity, learning might enhance selectivity for conditioned stimuli. Barrel cortex neurons are tuned to the angle of whisker deflection[48-50], indicating that the sets of synaptic connections activated by the CS+ and CS- may be overlapping but should not be identical. Therefore, the possibility exists that responses to the CS+ and CS- can change independently of each other. To examine this, we compared the amplitude of the

155 average response to CS+ and CS- trials for all segmented tufts on the pre, last-rewarded, and post
156 sessions (Fig.3A; n = 7 mice; 465 pre, 442 last-rewarded, and 430 post tufts). In agreement with
157 our previous analysis, we found no significant bias in response amplitude toward CS+ or CS-
158 during any of the three sessions (Fig.3A; Pre: $p = 0.20$; last-rewarded: $p = 0.43$; Post: $p = 0.64$,
159 sign-rank test). Under naïve conditions during the pre session, most tufts that responded to air
160 puff stimuli did not strongly prefer the CS+ or CS- (Fig.3A, left). Surprisingly, on the last-
161 rewarded session and the unrewarded post-conditioning session, we observed a prominent shift in
162 the response distribution, where many tufts exhibited more selective responses to one stimulus or
163 the other (Fig.3A, middle and right).
164
165 Plasticity can occur after repeated exposure to stimuli even in the absence of reinforcements[51-55].
166 To test whether enhanced selectivity depended on reinforcement, we imaged a separate group of
167 similarly water-restricted mice that were repeatedly exposed to the same stimuli for the same
168 number of days but without any reward. These mice only received water in their home cage
169 following each imaging session, but never during stimulus presentation. Repeated exposure mice
170 exhibited a stable distribution of response selectivity over time (Fig.3B; a separate cohort of 7
171 mice; 317, 313, and 321 tufts on Day 1, Day 8, and Day 9, respectively). These results suggest
172 that reinforcement learning, and not simply repeated stimulus exposure, drives apical tufts to
173 become more selective for either the CS+ or CS-.
174
175 To directly quantify the response selectivity of tufts, we computed a selectivity index (SI; see
176 Methods) ranging from -1 (exclusively CS- responsive) to 1 (exclusively CS+ responsive) for
177 each tuft. Initially in both the conditioned and repeated exposure mice, the SI distribution was
178 centered around zero, indicating that most tufts in naïve animals did not strongly prefer either
179 stimulus (Fig.3C,D, left panels). Consistent with our other analyses (Fig.2D), the mean SI
180 remained close to zero for each of the three sessions (Fig.3C and Supp.Fig.2D; -0.049, -0.001,

and 0.003 for pre-conditioning, last rewarded, and post-conditioning days, respectively; one-way ANOVA $p = 0.37$), confirming that learning produced no overall bias toward one particular stimulus among the population. During learning, the SI distribution of conditioned but not repeated exposure mice shifted markedly, whereby a much greater proportion of neurons were highly selective for either the CS+ or CS- (Fig.3C,D, middle and right panels, |SI| pre versus last-rewarded: $p < 10^{-6}$, |SI| pre versus post: $p < 10^{-5}$; Wilcoxon rank sum test). These effects can even be observed within individual mice (Supp.Fig.2). Notably, different tufts within the same animal exhibited opposite changes in selectivity (Supp.Fig.2A,B). Learning significantly increased tuft selectivity in individual conditioned mice, but not repeated exposure mice (Supp.Fig.2C). The degree of enhancement in tuft selectivity was closely correlated with conditioned animals' ability to discriminate stimuli across sessions (Fig.3E; Pearson's $R = 0.60$, $p < 10^{-5}$).

Whereas selectivity magnitude (|SI|) only considers the amplitude of tuft responses to CS+ and CS-, their discriminability also depends on their variability. For example, a large difference in CS+ and CS- responses would not be discriminable if the variability of those responses were very high; a small difference might be discriminable if the variability were low. We therefore additionally calculated a d-prime metric of neural discriminability that normalizes differences in response magnitudes to each stimulus by their variability (see Methods). Similar to selectivity magnitude, we found that neural discriminability was correlated with behavioral performance (Fig.3F). In conditioned animals, neural discriminability of CS+ and CS- responses of tufts increased significantly across learning (Fig.3G, blue; first-rewarded versus last-rewarded: $p < 10^{-3}$, pre versus post: $p < 10^{-4}$; Wilcoxon rank sum test). By contrast, neural discriminability of tuft responses in the repeated exposure mice decreased slightly with progressive exposure to the stimuli (Fig.3G, gray; Day 1 versus Final: $p < 0.01$). Finally, we asked whether the ability to decode stimulus identity on a trial-by-trial basis increased after learning. To test this, we trained a support vector machine (SVM) to decode stimulus identity from tuft population activity (see

Methods). We found that decoder performance increased significantly when comparing Pre and First sessions to Post and Last sessions (Supp.Fig3A; sign-rank test, $p = 0.002$), whereas decoder performance did not improve over time in the repeated exposure mice (Supp.Fig.3B; sign-rank test, $p = 0.22$). Taken together, these results show that enhanced stimulus representations can emerge in apical tufts, but require reinforcement.

The above analyses rely on the accurate measurement of calcium spikes from individual tufts. While two-photon microscopy acquires images with high resolution and speed, the imaging field is restricted to a single focal plane. This method can only measure calcium signals from a thin cross-section of the three-dimensionally complex apical structures. Indeed, many of the spatial components extracted from our two-photon data were comprised of dendritic branches that cross the imaging plane at different locations (Supp.Fig.4A), which makes it difficult to determine whether the segmentation software accurately extracted signals from one tuft or erroneously merged multiple tufts. For the same reasons, a single apical tuft could be falsely classified as two different tufts. Such errors could mislead our interpretation of selectivity in the population, especially given that a single apical tuft can exhibit non-homogenous branch-specific events[15,56,57].

To confirm that our interpretation was not due to segmentation errors, we repeated the conditioning experiment using a new, high-speed volumetric imaging approach called SCAPE[36,37], which allowed us to monitor calcium across entire apical tufts (Supplementary Movie 2). These three-dimensional datasets ($300 \times 1050 \times 234$ μm field of view) encompassed large portions of the apical tree which included branches converging on their bifurcation points in layer 2, enabling us to identify whole apical trees unambiguously (Fig.4A,B; Supp.Fig.4B). CaImAn effectively demixed overlapping trees in these three-dimensional volumes. Using SCAPE microscopy, we imaged tuft activity of two additional mice conditioned with the same

behavioral paradigm (Fig.4C). Comparison of tuft responses to the CS+ and CS- on the pre, last-rewarded, and post sessions (Fig.4D; 241 pre, 215 last-rewarded, 150 post tufts in 2 mice) revealed again that task learning induced significant increases in tuft selectivity (Fig.4E; pre versus last-rewarded: $p < 10^{-5}$, pre versus post: $p < 10^{-4}$, Wilcoxon rank sum test of |SI|). On average, the SI magnitudes were similar between tufts imaged using 2-photon microscopy and SCAPE (mean ± s.e.m. |SI| for 2-photon versus SCAPE; pre: 0.41±0.01 versus 0.40±0.02; last-rewarded: 0.54±0.02 versus 0.54±0.02; post: 0.51±0.02 versus 0.53±0.03). These data demonstrate that the effects in our two-photon dataset are not caused by errors in segmentation, but rather reflect changes at the level of individual dendritic tufts. Our results, based on two different imaging approaches, clearly demonstrate that reinforcement increases stimulus selectivity at the level of the entire apical tuft.

Selective tufts emerge from both initially unresponsive and responsive populations

The striking effect of reinforcement learning on tuft response selectivity could develop in several ways. For example, initially unresponsive tufts could develop a robust response to either stimulus after learning (e.g., Fig.5A, top). Conceivably, tufts that were initially unselective in naïve animals could also maintain their response to one stimulus while losing their response to the other (e.g., Fig.5A, middle). Either or both scenarios could lead to the increase in neurons that are selective for stimulus direction. To investigate which changes in individual tufts underlie population-wide improvements in stimulus selectivity, we longitudinally tracked the same set of tufts across all sessions and compared their selectivity in pre- and post-conditioning sessions for both conditioned and repeated exposure mice.

First, we categorized tufts that were unresponsive to either stimulus on the first imaging session, which accounted for the large majority of tufts (Fig.5E; conditioned: 458/603; repeated exposure:

334/457), and compared their response to the CS+ and CS- on the last session to determine if they became selective (Fig.5B, see Methods). Stimulus-unresponsive tufts, while on average less active than responsive ones (median calcium events per minute: 2.65 versus 3.66 for stimulus-unresponsive and responsive tufts, respectively; $p < 10^{-40}$, Wilcoxon rank sum test; Supplementary Fig.4), were not silent, with many undergoing tuft-wide calcium influx several times per minute. Silent tufts that are never active during the session may not have been detected in our imaging, but we were able to detect tufts that discharged as few as 3 voltage-gated calcium spikes over a 30-minute behavioral session. Interestingly, in both the conditioned and repeated exposure mice, approximately 40% of initially unresponsive tufts developed a response to at least one stimulus by the last session, becoming either selective or unselective (Fig.5B). However, in conditioned animals, the proportion of initially unresponsive tufts that became selective was significantly larger than in repeated exposure mice (Fig.5B; $p = 0.04$, 2-sample t-test comparing mice). Furthermore, while the proportion of selective and unselective tufts in this category was similar for conditioned animals, unselective tufts were more common in repeated exposure mice (Fig.5B; $p = 0.03$, paired t-test).

Next, we analyzed tufts that were initially responsive and either selective (Fig.5C; conditioned: 56/603, RE: 43/457) or unselective (Fig.5D; conditioned: 89/603, repeated exposure: 80/457). In these smaller categories, we found no significant differences in the outcome of selectivity between the two groups of animals. Together, these results indicate that, while both stimulus exposure and reinforcement can alter tuft tuning, the presence of reward increases the likelihood that initially unresponsive tufts develop selectivity for either the CS+ or CS- (summarized in Fig.5E).

While a greater proportion of tufts from the conditioned animals were selective during the final session (20.2% versus 10.3% of tufts from conditioned and repeated exposure mice, respectively),

we wondered whether conditioning also impacted the degree of selectivity. Note that some tufts had very small yet statistically different CS+ and CS- response amplitudes and were thus classified as selective despite a small SI. First, we compared the SI of initially unresponsive tufts on the final imaging session (Fig. 5F). Supporting our results in Fig. 5B, the SI distribution was shifted toward the tails in conditioned, but not repeated exposure mice, indicating that reward enhances selectivity for either the CS+ or CS- in this subset (|SI| conditioned versus repeated exposure: $p < 10^{-5}$, Wilcoxon rank sum test, n = 199 and 110 tufts, respectively).

Next, we compared the |SI| of all tufts that were categorized as selective during the last imaging session in conditioned and repeated exposure mice (Fig. 5G). Interestingly, we found that even among selective tufts, the |SI| distribution in conditioned mice was significantly greater than in repeated exposure mice (p = 0.006, Wilcoxon rank sum test, n = 122 and 47 tufts, respectively), indicating that while selective tufts are present after both conditioning and repeated stimulus exposure, the magnitude of selectivity is stronger after conditioning.

We then quantified the change in |SI| of all tufts that were responsive in both the first and last sessions by computing the difference between the two sessions (Fig. 5H). Tufts in conditioned mice exhibited a greater increase in |SI| across sessions compared to repeated exposure mice (p = 0.01, Wilcoxon rank sum test, n = 48 and 42 tufts, respectively), demonstrating that the magnitude of selectivity in initially responsive tufts increases after reinforcement learning.

Finally, we found that the degree of selectivity of tufts that eventually became unresponsive on the last session was overall similar between the two groups (Fig.5I, |SI| conditioned versus repeated exposure: p = 0.06, Wilcoxon rank sum test, n = 97 and 81 tufts, respectively). However, tufts that became unresponsive were more likely to be initially highly selective in the conditioned group than in the repeated exposure group (19 tufts with initial |SI| > 0.75 / 97 tufts ending as

unresponsive in the conditioned group versus 3/81 in the repeated exposure group; $p = 0.0013$, Z approximation to binomial). Therefore, learning can involve a loss of responsivity in a small subset of well-tuned tufts.

In summary, our longitudinal analyses revealed that reinforcement learning biases initially unresponsive tufts toward becoming selective and enhances the selectivity of tufts that are initially responsive.

Neither movement nor behavioral choice account for enhanced selectivity

Several plausible factors could underlie the changes in selectivity we observed across learning. For instance, movements like whisking are correlated with layer 5 somatic action potentials[58-60] and might have impacted calcium activity in the apical tuft. To investigate whether whisking could account for the changes in tuft selectivity, we imaged the whiskers with a high-speed camera and computed whisking amplitude (see Methods) while mice underwent conditioning and two-photon imaging (Fig.6A). First, we considered whether animals changed their whisker movements in response to conditioned stimuli over the course of learning. We computed the peak of the mean stimulus-aligned whisking amplitude for the CS+ and CS- (Fig.1C, left; Fig.6B) for each session in five mice. Although conditioning alters licking behavior (Fig.1C,E), the magnitudes of whisker movements following both stimuli were stable across sessions (Fig.6B; CS+: $p = 0.44$; CS-: $p = 0.45$; linear regression). We also computed the standard deviation (SD) of stimulus-evoked whisker amplitude across trials for all sessions (Fig. 6C). While the whisking amplitude became slightly more reliable (decreased SD) across sessions ($p < 10^{-4}$), the change in reliability across sessions was similar for CS+ and CS- ($p = 0.53$). Therefore, whisking is similar on both trial types throughout learning.

We next examined whether whisking was correlated with tuft calcium activity by comparing stimulus-triggered averages and intertrial interval (ITI) whisk-triggered averages of all tufts during post-conditioning. Whisking amplitude was similar between spontaneous ITI whisking bouts and evoked whisking responses to stimuli (n = 115 and 617 events, respectively; p = 0.53, Wilcoxon rank sum test). In contrast to air puff stimuli, ITI whisking bouts were not associated with a robust calcium response (Fig.6D).

To quantify the relationship of whisking and sensory stimuli to tuft calcium spikes, we performed a linear regression analysis (see Methods) on 322 tufts using calcium influx as the response variable and either stimulus or whisking amplitude as a single predictor variable (Fig.6E). Air puff stimuli more reliably predicted calcium influx than whisking amplitude for each of virtually all tufts ($p < 10^{-12}$, sign rank test). These results are consistent with other studies that found either only weak or no correlation between whisking and L5 tuft calcium spikes in S1[28,31,32]. Furthermore, we found no relationship between the whisking response and the median SI magnitude on a given session (Fig.6F, whisking to CS+ p = 0.22, CS- p = 0.78). Therefore, changes in whisker movement cannot account for the changes in selectivity during learning that we observed.

Finally, the possibility remains that other task-related signals relaying information about reward expectation and behavioral choice could impact apical tuft activity and drive increases in selectivity. To test this, we compared tuft responses to the CS- in false alarm trials (FA; mouse incorrectly licked for reward) and correct rejection trials (CR; mouse correctly withheld licks) to determine if their activity was modulated by behavioral choice. Notice that these two trial types have the same sensory input but involve different choices. (The corresponding analysis for CS+ trials is not technically possible for lack of sufficient Miss trials after the first conditioning day, an issue also observed in[1]. A future experiment in which the stimulus strengths are substantially

reduced would drastically increase the error rates, enabling a comparison between Hit and Miss trials.) Tufts were classified as behaviorally modulated if the FA response was significantly different from the CR response, and were not behaviorally modulated if CR and FA responses were statistically indistinguishable (e.g. Fig.7A). Behaviorally modulated tufts accounted for only ~10% of the total tuft population in both early and late learning (50/395 in early; 35/406 in late learning).

To test whether these behaviorally modulated tufts contributed to increased selectivity during learning, we excluded them and compared selectivity of the remaining behaviorally-insensitive tufts. We found that selectivity increased significantly from early to late learning (Fig.7B,C; median |SI| of 345 tufts early versus 371 tufts late learning: 0.38 versus 0.47, $p = 0.02$, Wilcoxon rank sum test), similar to our previous analysis of the entire population. Licking, like whisking, was a relatively poor predictor of tuft calcium influx (Supp.Fig.6A,B). Because some behaviorally modulated tufts may not have been statistically detectable, we used multivariate linear regression to disentangle stimulus responses from licking and whisking, which may have been confounded with choice. Median coefficients for licking and whisking were on average 3.3 times smaller than median stimulus coefficients for the first rewarded, last rewarded, and post sessions (all $p < 10^{-6}$, Wilcoxon rank sum test). Even after we factored out possible effects of movements, CS+ and CS- coefficients were enhanced by learning but not repeated exposure (Supp.Fig.6C,D), consistent with our other analyses. Together, these results demonstrate that enhanced selectivity during learning cannot be explained by non-sensory signals related to the animals' behavior.

Enhanced selectivity in barrel cortex is long-lasting when mice exclusively use whiskers

Mice could conceivably exploit other sensory cues to learn and perform the task, such as auditory cues from the air nozzles or non-whisker tactile cues from air current eddies contacting the fur or

skin. To determine which mice exclusively used their whiskers to distinguish the CS+ and CS-, we trimmed all whiskers after the post-conditioning session and assessed performance in five mice (Figure 8). Performance in each of the five mice decreased after whisker trimming, indicating that each used some whisker information. Three mice performed the task exclusively with their whiskers, falling to chance levels after the whisker trim ("whiskers only"). Two other mice still performed the task above chance after the whisker trim, indicating that they were not exclusively using their whiskers and exploited information from multiple sensory streams ("whiskers + other senses").

We examined whether these two different behavioral strategies impacted tuft selectivity. Both the "whiskers only" and "whiskers + other senses" groups exhibited enhanced tuft selectivity in the last-rewarded session relative to pre-conditioning. This effect was more pronounced in the "whiskers only" mice (Fig.8A,B, left and middle; whiskers only: median |SI| of 180 pre tufts versus 169 last-rewarded tufts: 0.36 versus 0.59, $p < 10^{-3}$; "whiskers + other senses": median |SI| of 144 pre tufts versus 155 last-rewarded tufts: 0.39 versus 0.50, $p = 0.01$). Surprisingly, enhanced selectivity persisted during the post-conditioning session for the "whiskers only" group but not the "whiskers + other senses" group (Fig.8A,B right panels; whiskers only: median |SI| of pre versus 167 tufts post: 0.36 versus 0.58; $p < 10^{-3}$; whiskers + other senses: median |SI| of 155 pre versus post tufts: 0.39 versus 0.42; $p = 0.45$). Therefore, tuft selectivity in barrel cortex is enhanced regardless of behavioral strategy, but outlasts conditioning only when mice rely solely on their whiskers to perform the task.

We further examined this persistence of enhanced tuft selectivity as experienced mice stopped performing the task. While the entire post-conditioning session was unrewarded, mice initially expected rewards and licked for many CS+ trials in the first half of the session. By the second half of the session, the probability of a lick occurring during the CS+ extinguished, approaching

zero (Fig.8C). We compared the selectivity of tufts during the first and second halves of the post-conditioning sessions of mice that exclusively used their whiskers and found no difference in the two distributions (Fig.8D, $p = 0.94$, Wilcoxon rank sum test of $|SI|$), demonstrating that selectivity of the population remained stable throughout the session. Taken together, these results demonstrate that enhanced stimulus selectivity of apical tuft dendrites after reinforcement learning is long lasting, persisting even after mice cease performing the task and expecting reward.

DISCUSSION

Our study is the first to investigate how learning a discrimination task alters apical tuft activity. Using both novel volumetric whole-tuft imaging and conventional planar microscopy, we discovered that L5 apical tufts acquire enhanced representations of multiple stimuli during learning. Rather than simply retuning tufts toward the rewarded stimulus, learning enhanced selectivity for both stimuli, suggesting that tufts are aligning themselves to the behaviorally relevant stimulus dimensions. These enhanced sensory representations persist even after mice cease performing the task. In contrast, representations are slightly degraded by mere repeated exposure to stimuli outside of a task. Consistent with previous studies[28,31], we found that movement in and of itself has little direct impact on tuft spikes, indicating that increased selectivity of apicals reflects alterations in sensory coding as animals learn. This sensitization of tufts to behaviorally relevant sensory dimensions may be a general feature of all sensory cortical areas.

Tuft spikes enhance plasticity of synaptic inputs that occur over behavioral (seconds-long) timescales[18,34]. These new behaviorally relevant tuft representations may therefore prime subsequent plasticity of synapses across the entire pyramidal neuron. Additionally, tuft events

potently modulate somatic burst firing and enhance how somata respond to their basal inputs[15,61]. As learning and plasticity increase apical selectivity for a behaviorally relevant axis, tuft events will unavoidably amplify somatic burst output along the same axis. This could enable action potential output of L5 cells in primary sensory cortex to directly drive behavioral responses via projections to movement related areas, such as the corticostriatal, corticopontine, and corticotrigeminal pathways. Thus, tuft spikes have the potential to modify somatic output, both in the present and in the future.

An open question is whether enhanced stimulus representations in apical tufts are required for learning this task. One way to address this question would be to silence tuft activity during and after learning by optogenetically activating NDNF-positive interneurons in layer 1[62]. This approach is not ideal as NDNF interneurons also inhibit other cells such as Layer 2/3 pyramidal cells, PV interneurons,[63] and possibly the axons of Layer 5 pyramidal cells, which are known to densely innervate layer 1. Because this manipulation is not specific to layer 5 apicals, the results would be difficult to interpret. Focal illumination of inhibitory opsins in tufts has also been used to assess tuft function[64], but balancing tuft against soma silencing remains challenging and complicates interpretation. Better tools for selective targeting of apicals would be extremely useful for addressing such issues.

Enhanced Representation of Behaviorally Relevant Stimuli

Enhancing the representation of relevant stimulus dimensions rather than a singularly important stimulus, such as a rewarded event, has multiple benefits for behavior. In our paradigm, both the CS+ and CS- are predictive of whether or not a reward will occur in the future. Explicitly encoding both stimuli could allow sensory cortical areas to directly elicit actions. In the context of this task, CS+ preferring tufts in barrel cortex may trigger anticipatory licking while CS-

preferring tufts could suppress licking. L5 cells in sensory cortex via their output to striatum, pons, brain stem, and spinal cord would thereby be able to directly and rapidly drive action without further cortical processing, such as by frontal areas including motor cortex[32,65]. Such rapid sensory-motor transformations by primary sensory areas may be critical for natural time-constrained behavior.

Furthermore, learning produced a representation in which the degree of selectivity for the two stimuli was continuous and uniformly distributed. Exclusively CS+ or CS- selective apicals never dominated the population. Continuous degrees of selectivity across the population, rather than discrete representations, may allow the system to be more robust to the variability caused by active movements that alter sensory input. A continuous distribution may also facilitate future adjustments of neural representations as subjects continue to learn a task or encounter new tasks. The uniformity we observed may reflect that neurons are high-dimensional, being sensitive to mixtures of variables[60,66-68], only one of which might be altered here by learning. The uniform distribution of selectivity corresponds to a full range of pessimism to optimism concerning stimulus predictions of upcoming rewards. Recent work shows that behavioral performance benefits from reinforcement learning that incorporates the distribution of reward probabilities rather than just the average expected reward value[69]. L5 corticostriatal synapses could theoretically afford a plastic substrate for acquiring the necessary distribution of reward probabilities.

Surprisingly, past studies in which mice were trained to associate one or more stimuli with a reward typically show that cortical representations are stronger for the rewarded stimulus [1,3,5]. In contrast to these studies of layer 2/3 somatic activity, our experiments revealed that the overall tuft calcium response to the CS+ and CS- at the population level did not change significantly after animals learned the task (Fig.2). Instead, representations for both stimuli were enhanced by

individual tufts developing selectivity for either the CS+ or the CS- (Fig.3). This divergence in phenomena may result from several important differences between our work and the aforementioned studies.

First, enhanced selectivity for both rewarded and unrewarded stimuli could be a phenomenon that is unique to the apical dendritic tufts. In addition to local inputs, the apical tufts of pyramidal cells in S1 receive long-range top-down input from several sources, including motor cortex[31,70], secondary somatosensory cortex[11], and secondary thalamus[9,10,71]. Frontal areas, such as prefrontal cortex, indeed have enhanced representations of the CS+ and CS- after learning[47]. In contrast, input to the somata is dominated by the local cortical area and primary thalamus[72,73]. While somato-dendritic coupling can be strong in L5 neurons[25], it is asymmetric; at least 40% of somatic transients attenuate in a distance-dependent manner along the apical trunk and distal tufts[24]. The non-overlapping anatomical inputs and asymmetric coupling together could produce different learning-related effects on apical tuft and somatic stimulus representations.

Second, learning-related changes may manifest differently in layer 2/3, the usual focus of previous studies[1,3], and layer 5 pyramidal cells, the tufts of which we studied. With the exception of a small population of corticostriatal cells, most excitatory cells in layer 2/3 project to other cortical areas to affect further cortical processing[74,75]. In contrast, many L5 cells project to subcortical structures including the thalamus, superior colliculus, and brainstem, which may directly trigger behavioral responses[76-78]. In discrimination paradigms, both stimuli are relevant to behavior. In our task, the CS+ prompted licking to obtain a reward, and the CS- suppressed licking that would have no benefit. Thus, an enhanced representation of both stimuli in layer 5 would be advantageous for animals to perform the task efficiently. Recently, it was shown that apical dendrite activation of subcortical-targeting pyramidal tract L5 cells, but not intratelencephalic L5 cells that are more like L2/3 cells in their connectivity, determines the

detection of tactile stimuli[32]. The Rbp4-Cre mice we used in this study labels a heterogenous population of layer 5 pyramidal cells, comprising both pyramidal tract and intratelencephalic neurons. In the future, it would be interesting to examine whether learning has different effects on the sensory representations of these two populations. Moreover, direct comparisons of the layers would be particularly informative.

Finally, it is possible that learning-related changes in sensory representations manifest differently between a somatosensory modality and a visual modality, the latter being the focus of previous studies. To our knowledge, we are the first to show changes of sensory representations in somatosensory cortex within a discrimination paradigm. Mice are known to rely more heavily on their tactile senses than vision[79]. Their heavy reliance on whisker-mediated touch may make it advantageous to develop sensory representations of a larger variety of relevant tactile stimuli, in this case, both the CS+ and CS-.

Candidate Plasticity Mechanisms

Enhanced selectivity could be due to changes in local synaptic connectivity, long-range inputs, or both. Learning may strengthen and weaken synapses onto barrel cortex neurons from ascending thalamocortical input or from neighboring cells. Such local plasticity could enhance CS+ or CS- responsiveness. Alternatively or additionally, other cortical regions encoding task context could via long-range inputs reconfigure barrel cortex to respond more strongly to these stimuli. The present results do not completely distinguish between these two scenarios because long-range inputs may still encode the context while the mouse is in the behavioral apparatus. However, we found that enhanced representations persist after mice are no longer engaged in the task and receiving rewards. This result suggests that enhanced representations may be a product of local plasticity in sensory cortex that alters receptive fields.

Even in the absence of reward, repeated exposure to stimuli can drive plasticity in sensory cortex and alter response tuning. For instance, repeated exposure to oriented gratings can alter the orientation tuning of cells in primary visual cortex[51-53], and overstimulation of whiskers induces plasticity at dendritic spines and alters whisker representations in somatosensory cortex[54,80]. Our results demonstrate that at the population level enhanced representations developed only when stimuli were behaviorally relevant. Our longitudinal analysis revealed that while the response dynamics of some tufts changed after repeated stimuli presentations, overall selectivity of the population did not increase when rewards were omitted (Figs.3&5). This raises the question: What are the mechanisms that drive enhanced selectivity under rewarded conditions? In one possible scenario, reward delivery causes the release of neuromodulators that augment the activity of apical tufts. Cortical layer 1 is innervated by cholinergic afferents from the nucleus basalis[81] and adrenergic afferents from the locus coeruleus[82], the main source of acetylcholine and norepinephrine, respectively. Salient events such as reward and arousal lead to the release of these neuromodulators in cortex[83,84], which could increase the excitability of apical dendrites by recruiting disinhibitory circuits or directly influencing dendritic currents[26,27,83,85]. In this model, the release of reward-driven neuromodulators promotes plasticity and an enhanced representation of temporally aligned sensory inputs. This phenomenon was demonstrated in auditory cortex, where tones paired with stimulation of the nucleus basalis shifted the tuning of neurons toward the frequency of the paired stimulus[86].

Why are representations of the CS- equally enhanced when there is no associated reward? One explanation is that, as mice learn that the CS- indicates absence of reward, the CS- effectively signals punishment and acquires negative value. Acetylcholine is released in response to aversive stimuli, and can activate disinhibitory microcircuits that reduce inhibition onto pyramidal cells and may be essential for learning[87,88]. Thus, it is possible that both the CS+ and CS- representations are enhanced by neuromodulatory mechanisms tied to reward and punishment,

569 respectively. An open question is whether the outcome is due to reinforcement learning or the
570 behavioral state brought on by the reinforcers rather than their valence. Sensory cortical plasticity
571 may not be tied to reinforcer valence. Our paradigm creates an environment where mice benefit
572 from being attentive and engaged in order to maximize reward while minimizing effort. Previous
573 work has shown that active engagement in a visual discrimination task was associated with
574 significantly higher selectivity in layer 2/3 cells in visual cortex[1]. Task engagement may lead to a
575 sustained increase in neuromodulator release throughout the conditioning session, priming the
576 apical dendrites for plasticity and the development of selective responses for task-relevant stimuli
577 as they learn.
578
579 What determines whether a particular tuft eventually becomes selective for the CS+ or CS-? Our
580 longitudinal analysis revealed that many tufts that were initially unresponsive to either stimulus
581 developed a highly selective response to either the CS+ or the CS- (Fig.5). In these tufts, stimulus
582 preference after learning might be seeded by initially weak, directionally selective inputs on to
583 the neuron that already exist prior to conditioning and that are potentiated by the learning process.
584 We also found tufts that initially exhibited robust responses to both stimuli and either lost or
585 significantly reduced their response to one stimulus after learning. The reduction of an apical
586 response to a particular stimulus could be driven by local disynaptic inhibition between L5
587 pyramidal cells mediated by the apical-targeting Martinotti cells[89-91]. Through this mechanism, L5
588 neurons that are selective for a particular stimulus could inhibit responses to that stimulus in
589 neighboring L5 apical tufts. Experiments that assess the tuning of excitatory and inhibitory inputs
590 onto apical dendrites as a function of learning could test such mechanisms.
591
592 In addition to demonstrating increased tuft selectivity with learning, we replicated a surprising
593 phenomenon in a previous instrumental behavior in which a population of apical tufts exhibit
594 activity around the time of reward[28]. This reward-related activity was observed in four out of the

595 seven conditioned animals only during CS+ trials and was most prominent during intermediate
596 conditioning sessions, when most animals were still performing at chance levels, and disappeared
597 completely by the final conditioning session (Supp.Fig.1). Other than this transient effect,
598 unconditioned stimuli did not appear to elicit calcium responses, consistent with our previous
599 findings[28]. The disappearance of this reward-related peak might be attributable to the reward
600 becoming predictable in later stages of learning. In previous classical conditioning experiments,
601 dopaminergic cells exhibit responses to rewards early in learning due to the novelty of an
602 unexpected stimulus. These responses are lost after extended training, as animals learn the
603 association between the CS and reward[92,93]. While dopaminergic terminals are sparse in primary
604 sensory areas, they are not entirely absent, nor are dopaminergic receptors. Furthermore, the
605 excitability of the apical tuft is sensitive to noradrenaline[26]. Interestingly, noradrenergic neurons
606 in the locus coeruleus exhibit a similar phenomenon to dopaminergic neurons, where responses
607 shift from temporal alignment with the reward to a predictive conditioned stimulus after
608 learning[94]. Such mechanisms could explain why reward-related activity is restricted to early-to-
609 intermediate learning in our paradigm.
610
611 **Global versus local dendritic spikes**
612 Apical dendrites exhibit not only global spikes that elicit calcium influx across the entire tuft,
613 which we exclusively analyzed here, but also local events known as NMDA spikes, which
614 typically engage short (<30-μm) segments of individual dendritic branches[15,31,57]. These local,
615 NMDA receptor-dependent events can promote prolonged plasticity within individual dendritic
616 branches in the absence of backpropagating actions potentials, a feature that is unique to the
617 apical dendrites[16]. In motor cortex, branch-specific NMDA spikes are crucial for establishing the
618 long-lasting plasticity necessary for learning[56], and depolarization provided by multiple local
619 NMDA spikes is thought to be essential for the generation of a global calcium spike triggered by
620 distal synaptic inputs[15]. We focused this study on global tuft-wide calcium events, rather than

local events. Local events are more difficult to unambiguously identify in planar imaging[95], and their existence *in vivo* is still an open question for L5 apicals in barrel cortex[31,57]. Nonetheless, they may play important roles in plasticity processes that eventually lead to the emergence of global tuft spike selectivity for stimuli. Volumetric microscopy studies, the feasibility of which we showed here, are needed to further investigate the existence of local events in such behaviors as well as examine possible relationships between local and global tuft events during reinforcement learning. However, it would be essential to verify that seemingly spatially overlapping local and global events derive from the same dendritic tree, which requires greater resolution than was practical for the present study.

To analyze activity of individual tufts, we segmented these structures based on spatiotemporal covariance[45]. This method does not discount the possibility of errors where one tuft is split erroneously into two trees, or where two highly correlated tufts are merged. With this in mind, we used volumetric imaging SCAPE microscopy, which allowed us to visualize the apicals in three dimensions and unambiguously screen for such artifacts. The results from SCAPE are quantitatively similar to those from two-photon microscopy, and confirm that our observation of enhanced selectivity with learning is not an artifact of planar imaging.

Stability of learned tuft representations

In contrast to previous studies of discrimination learning[1-3], we included an unrewarded post-conditioning session to examine whether learning-related effects persisted through extinction. Our results show that post-conditioning selectivity of the apical population remains significantly higher than pre-conditioning, even after animals stop licking in response to the CS+ (Fig.8). Interestingly, the effects of learning are much more pronounced in animals that relied exclusively on their whiskers to perform the task. In animals that apparently used other sensory modalities, we observed a modest increase from the pre to last-rewarded session, which seemed to be largely

absent by the post-conditioning session. Considering that these animals were additionally exploiting other sensory areas to perform, selectivity may have been more widely distributed and thus diluted in barrel cortex, diminishing the effect and its stability. How long selectivity persists in the neuronal population after conditioning and which factors influence stability are interesting open questions for future study.

Conclusion

In summary, we have shown for the first time that reinforcement learning enhances representations along behaviorally relevant dimensions in apical tufts. Our results suggest that dendritic calcium spikes are an important cellular mechanism underlying the changes in sensory encoding that occur with learning, and provide an avenue for further investigation of cellular and circuit mechanisms underlying plasticity induced by perceptual experience and reinforcement. This cellular compartment may be key to understanding pathology in some cognitive, memory, and learning disorders.

ACKNOWLEDGEMENTS

We thank Venkatakaushik Voleti for help with the design, construction, and alignment of the SCAPE microscope; Dan Kato, Georgia Pierce, and Jung Park for help with pilot experiments; Eftychios Pnevmatikakis and Johannes Friedrich for advice on dendrite segmentation; and Larry Abbott, Stefano Fusi, Ashok Litwin-Kumar, Chris Rodgers, Georgia Pierce, Gordon Petty, and Dan Kato for comments on the manuscript. Funding was provided by a Wellcome Trust Discovery Award, an Academy of Medical Sciences Professorship, NIH/NINDS R01 NS069679, and NIH/NINDS R01 NS094659 (RMB); a Kavli Institute for Brain Science Postdoctoral Fellowship (SEB); NIH/NINDS/NIMH/BRAIN U01 NS094296, UF1 NS108213, U19 NS104649, and RF1 MH114276 (EMCH).

AUTHOR CONTRIBUTIONS

SEB and RMB conceived of the behavioral and two-photon imaging experiments. EMCH and RMB conceived of the SCAPE imaging experiments. SEB built the behavioral apparatus, EMCH, KBP, and CC designed, built, and maintained the SCAPE microscope, and RMB built and maintained the two-photon and intrinsic signal microscopes. SEB performed the experiments and analyzed the data with input from RMB and EMCH. SEB and RMB wrote the manuscript.

DATA AVAILABILITY

Due to the large volume of data (~80TB), data are maintained by the authors and available upon request.

683 **MAIN FIGURES**

684

685 **Fig. 1 | Mice rapidly learn to discriminate stimulus direction in head-fixed paradigm. a,** A
686 water droplet is paired with air puffs in one direction (CS+) but not the other (CS-). Licking in
687 anticipation of water is assessed in the response window just after CS+ or CS- and prior to water
688 delivery for the CS+ (grey bar). **b,** Experimental timeline. 2-3 weeks after virus injection, naive
689 tuft responses to stimuli are recorded (pre). The CS+ is then paired with water for 8-9 days (blue).
690 On the last day, stimuli are presented without reward (post). In a separate group of mice, the same
691 stimuli are presented over 9 days in the absence of reward (unrewarded group). **c,** Lick rasters for
692 three different sessions in one example mouse. On session 9, the CS+ but not the CS- reliably
693 elicits licks. **d,** Mean baseline-subtracted whisking amplitude aligned to the CS+ (red) and CS-
694 (navy) across sessions 1, 2, and 9 of an example mouse. **e,** Learning curve demonstrates rapid
695 learning. Mean probability of at least one lick in the response window across sessions. **f,**
696 Behavioral performance of each mouse in the rewarded group (M1 – M7).

697

698 **Fig. 2 | Overall tuft response to stimuli is unbiased and relatively stable across conditioning.**
699 **a,** Dendritic activity was recorded in layer 1 (i) in the C1/C2 barrel columns (ii). (i) Two-photon
700 image ~60 µm deep relative to pia. Dashed yellow lines denote C1 and C2 boundaries from
701 intrinsic imaging. Single cell reconstruction in left panel from[50]. (ii) Tangential section through
702 layer 4 showing barrels stained with streptavidin-Alexa 647 and GCaMP6f-expressing apical
703 trunks. Red circles indicate location of 2-photon lesions to mark the imaging region for post-hoc
704 analysis. **b,** Overlay of five segmented pseudo-colored tufts from imaging field in A(i). **c,** Time
705 courses of calcium responses of example tufts in **b** to three air puffs (dashes). **d,** Amplitude for
706 CS+ (red) and CS- responses (blue), computed for each segmented tuft in the first 1.5 s post-
707 stimulus (grey points), do not differ within or across sessions. Colored lines indicate median. **e,**
708 Same as in **d**, showing data for all conditioning sessions.

Fig. 3 | Reinforcement learning, but not stimulus exposure, enhances tuft selectivity for CS+ and CS- stimuli. a, Across the indicated sessions, individual tufts (circles) exhibit larger biases to CS+ or CS- (pooled across all conditioned mice). **b,** Repeated exposure to stimuli does not bias individual tufts to CS+ or CS-. **c,** Conditioning reshapes distribution of selectivity indices for tufts from Normal on pre-conditioning session to uniform on post-conditioning session. **d,** Distribution of tuft selectivity indices remains Normal throughout all repeated exposure sessions. **e,** Selectivity (median SI magnitude of tufts for each session) increases with behavioral performance of 6 animals. **f**, Same as e, but with neural discriminability plotted on the y axis. **g,** Neural discriminability (mean ± sem) of tufts, pooled across all animals on each session, increases with conditioning and decreases with repeated exposure.

Fig. 4 | High-speed volumetric imaging of apical tufts confirms the emergence of enhanced selectivity after learning. a, Top and side view of four example tufts segmented from volumetric SCAPE imaging. **b,** Time courses of calcium activity from example tufts in **a** during five presentations of air puff stimuli (dashes). **c,** Performance across all conditioning sessions of two mice that were imaged with SCAPE. **d,** Across the indicated sessions, individual SCAPE-imaged tufts (circles) exhibit larger biases to CS+ or CS-. **e,** Conditioning reshapes selectivity distribution from Normal to uniform.

Fig. 5 | Longitudinal tracking reveals that reward enhances the selectivity of both initially unresponsive and responsive tufts. a, Three example tufts that were longitudinally tracked across learning. Top row: An initially unresponsive tuft develops a robust response to the CS+ but not the CS- after learning. Middle row: A responsive but unselective tuft loses its robust CS+ response and becomes selective for the CS-. Bottom row: A CS- selective neuron becomes unresponsive to both stimuli. **b,** Tufts that were unresponsive during the first session were

longitudinally tracked to the last session. Plotted is the mean proportion of selective and unselective neurons across all animals in the conditioned (black bars) and repeated exposure (grey bars) groups. **c,d,** Same analysis as **b** for initially selective (**c**) and unselective (**d**) tufts. Two-sample t-test was used for comparisons between conditioned and repeated exposure groups. Paired t-test was used for comparisons within a group. * $p < 0.05$. **e,** Total tuft counts from first to last session within the 3 response categories for either conditioned (left) or repeated exposure (right) groups. **f,** SI of responsive tufts on the last session that were initially unresponsive during the first session. Conditioned tufts have enhanced selectivity compared to repeated exposure. **g,** Tufts that were selective on the last session are more selective if conditioned (black) rather than undergoing repeated exposure (grey). **h,** Tufts that responded on both pre and post sessions tend to have higher selectivity if conditioned rather than undergoing repeated exposure. **i,** SI of responsive tufts on the first session that later became unresponsive during the last session.

Fig. 6 | Whisking is only weakly correlated with tuft activity and cannot account for changes in selectivity during learning. a, Whisking amplitude aligned to calcium activity of three example tufts in one session. Green shading indicates periods of whisking. Red and navy ticks indicate CS+ or CS- delivery, respectively. **b,** Mean whisking response of five mice to CS+ (red) and CS- (navy) does not change across sessions during learning (mean ± s.e.m.). **c,** Mean standard deviation of whisking decreases for both CS+ and CS- across learning, but CS+ and CS- do not differ. **d,** Event-triggered averages of 322 tufts on the post-conditioning day (grey traces - individual tufts, black inset - population average) are responsive to stimuli but relatively unmodulated by whisking. **e,** R^2 values for linear models predicting calcium from stimuli (y axis) are consistently greater than those predicting calcium from whisking (x axis). Each circle represents a tuft. (n = 322 tufts) **f,** Magnitude of tuft selectivity does not correlate with mean whisking amplitude during CS+ (left) and CS- trials (right) on that session.

Fig. 7 | Behavioral responses do not account for enhancement of stimulus selectivity during learning. a, Mean stimulus responses of four tufts during hit (red), CR (cyan), and FA (black) trials. Top row: Example tufts whose responses are not behaviorally modulated (CR is similar to FA). Bottom row: Example tufts with behaviorally modulated responses (CR and FA differ). **b,** Selectivity index (SI) distribution changes from early (left) and late learning sessions (right) even when tufts with behaviorally modulated responses (CR≠FA) are excluded. **c,** Median SI magnitude of tufts in each of six animals (from panel **b**) increases from early to late learning sessions.

Fig. 8 | Apical tufts in barrel cortex of mice performing the task exclusively with their whiskers undergo long-lasting changes in selectivity. a, SI histograms of mice performing the task exclusively with their whiskers exhibit increased selectivity across pre-conditioning, last-rewarded, and post-conditioning sessions. **b,** Relative to pre-conditioning, mice using their whiskers and other sensory cues to perform the task have increased selectivity during the last rewarded session, but not the post-conditioning session. **c,** The probability of anticipatory licks in response to the CS+ extinguishes across post-conditioning blocks (of 20 trials each). **d,** Tuft selectively remains uniformly distributed during post-conditioning trial blocks 1-2 (top) while licking is extinguishing, and blocks 3-4 (bottom) in which licking is extinguished.

METHODS

All experiments complied with the NIH Guide for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee of Columbia University. Sixteen C57BL/6 mice ranging in age from 77 to 316 days old (mean of 123 days at the time of imaging) were used in these experiments. Six were male, and 10 female. Our results were observed in both male and female individuals, and no sex difference was detected.

Surgery

Animals were administered dexamethasone (1 mg/kg) via intramuscular injection 1-4 hours prior to surgery to reduce edema. Anesthesia was induced with 3% isoflurane in oxygen and maintained at 1%. Mice were head-fixed in a stereotax, and a subcutaneous injection of bupivacaine (0.5%, 0.1 mL) was administered under the scalp. Buprenorphine (0.05 mg/kg) was injected subcutaneously on the back. The scalp was cut, and the skull was covered with a thin layer of Vetbond. A circular craniotomy (3-mm diameter) centered at 1.5 mm posterior and 3.5 mm lateral to bregma was made using a dental drill. The dura was kept moist using artificial cerebrospinal fluid.

For both two-photon and SCAPE microscopy, Rbp4-Cre_KL100 mice were injected with 100 nL of virus (initial titer ~$2x10^{13}$ cfu/mL, diluted 1:4 in artificial cerebrospinal fluid) encoding GCaMP6f in a Cre recombinase-specific manner (AAV1-CAG-flex-GCaMP6f, UPenn Vector Core). The virus was injected in layer 5B of the barrel cortex (1.0 mm deep to the pia) using a pulled pipette (20-30 μm ID) fastened on a Nanoject III, which was mounted on a manipulator angled at ~30° from vertical. The virus was delivered via four injections of 100 nL each, spaced at least 400 μm apart. The depth was chosen to maximize labeling of thick-tufted pyramidal

804 neurons. In pilot experiments, we found that placing injections 1.0 mm deep resulted primarily in
805 thick-tufted labeling whereas at more superficial depths (e.g., 0.8 mm deep) we obtained mainly
806 thin-tufted tufts, consistent with ref [96]. The dura was then removed, and a thin cover glass was
807 implanted and sealed using superglue. A custom metal head plate was implanted on the skull
808 using dental cement. Twenty-four hours after surgery, carprofen (5 mg/kg) was administered
809 subcutaneously. Imaging and behavioral training commenced 3 weeks after surgery.
810
811 **Behavior**
812 Animals in both rewarded 'conditioning' and unrewarded 'repeated exposure' groups were water
813 restricted for 2 days prior to starting imaging and habituated to head fixation for ~10 minutes on
814 each of these 2 days. They were subsequently given ~1 mL of water per day for 9 days either by
815 pairing water rewards with a specific stimulus (conditioning group), or in their cage following the
816 imaging session (repeated exposure group). Mice were head restrained in a custom-made
817 behavioral apparatus by positioning the body in a 3D-printed chamber and fastening the head
818 plate to metal posts flanking the chamber. Air puff stimuli (10 psi measured before a control
819 solenoid, 100 ms) were delivered from two nozzles (cut P200 pipette tips) positioned toward the
820 distal tips of the whiskers, in either the rostrocaudal or ventrodorsal direction. Nozzles were
821 oriented to prevent air jets from stimulating other parts of the face. One of these directions (CS+)
822 was paired with a water reward (10 μL), delivered through a lick port 0.5 seconds after the
823 stimulus onset. The particular direction (rostrocaudal vs ventrodorsal) used as the CS+ was
824 randomized and counterbalanced across mice. Approximately 180 stimuli were presented over the
825 course of a 30-minute imaging session (8-12-s intertrial interval). The probability of CS+ or CS-
826 delivery was 50%. In preliminary experiments, we found that an auditory mask helped prevent
827 mice from exploiting auditory cues to discriminate the two stimuli: a third air nozzle was
828 positioned close to the mouse and was active throughout the session.
829

During the first session (pre-conditioning), stimuli were delivered in the absence of reward to assess neural and behavioral responses in naïve animals. In the following 7-9 days, the CS+ was paired with reward. Licks for rewards were detected with a capacitance-based touch sensor (Sparkfun). A trial response was registered when one or more licks were elicited within a 0.5-second response window following the stimulus and before reward delivery. To determine whether behavioral performance was above chance, we computed 95% confidence intervals using the 'binofit' function in MATLAB. During the final session (post-conditioning), stimuli were delivered in the absence of reward. Animals in the unrewarded group received the same two stimuli across 9 days without reward pairing. Behavioral experiments were performed with the Arduino-based OpenMaze open-source behavioral system, whose designs are fully described at www.openmaze.org. Whisking was monitored at 125 fps with a camera (Sony PS3eye) and automatically tracked using published software [97].

Intrinsic signal optical imaging and two-photon imaging

Intrinsic signal optical imaging and two-photon imaging were performed on a Sutter movable objective microscope. The locations of whisker barrels in S1 were identified using intrinsic signal optical imaging. Single whiskers in isoflurane-anesthetized mice were stimulated at 5 Hz using a piezoelectric bimorph while recording the reflectance of 700-nm long-pass incandescent light with a Rolera CCD camera (QImaging) through a low-magnification objective (Zeiss 5X/0.16NA). Movies were collected using software custom-written in Labview (National Instruments). Regions of reflectance change were referenced to an image acquired under green illumination.

Two-photon imaging was conducted on the same microscope under the control of the ScanImage software package (V. Iyer, Janelia Farms). All calcium imaging data was collected by two-photon microscopy except for those in figure 4. Scanning during awake conditions was performed at 30

fps using a Chameleon Ultra II laser (Coherent) tuned to 920 nm, precompensated for group velocity dispersion and focused through a 20x/1.0NA water immersion lens (Zeiss). Aquasonic clear ultrasound gel was used for the immersion medium. Emitted light was collected with an HQ535/50 filter (Chroma) and GaAsP photomultiplier tubes (Hamamatsu Photonics). Apical tuft tufts in Layer 1 were imaged at depths of 40-80 μm from the pial surface (1.5x digital zoom in ScanImage which yielded a 433 x 433 μm field of view, 512 x 512 pixels).

SCAPE imaging

High-speed volumetric imaging was performed using a custom SCAPE microscope as previously described, including for dendritic tufts[36,37,98]. Briefly, the cortex was illuminated with an oblique light sheet through a Olympus XLUMPLFLN 20XW 1.0 NA water immersion objective with a 2-mm working distance. Fluorescence excited by this sheet (extending in the *y*-*z*′ direction) was collected by the same objective lens. A galvanometer mirror in the system was positioned to both cause the oblique light sheet to scan from side to side across the sample (in the *x* direction) but also to de-scan returning fluorescence light. This optical path results in an intermediate, de-scanned oblique image plane that is stationary yet always co-aligned with the plane in the sample that is being illuminated by the scanning light sheet. Image rotation optics and a fast sCMOS camera (Andor Zyla 4.2+) were then focused to capture these *y*-*z*′ images (750 x 200 pixels) at >1000 frames per second as the sheet was repeatedly scanned across the cortex in the *x* direction. All other system parts, including the objective and sample stage, were stationary during high-speed 3D image acquisition. Data were reformed into a 3D volume by stacking successive *y*-*z*′ planes according to the scanning mirror's *x* position and de-skewing to correct for the oblique sheet angle. This rotation of the image volume is responsible for its rectangular appearance despite the camera's square frames. The resulting volumes were large enough to encompass many GCaMP6f-labeled tufts in barrel cortex,

In this study, the stationary objective lens in SCAPE was configured on a manual rotation mount and set to 20º-30º away from the standard upright configuration, so the optical axis was perpendicular to the cranial window to achieve optimal performance without tilting the head of the animal. A 488-nm laser (Coherent OBIS) was used for excitation (<10 mW at the sample) with a 500-nm long-pass filter in the emission path. To achieve optimal spatiotemporal resolution and volume rate, the sample was imaged with an *x*-direction scanning step of 3 μm over a 300 × 1050 × 234 μm field of view (*x*-*y*-*z*, 3.0 × 1.40 × 1.17 μm per voxel, 100 x 750 x 200 voxels) at 10 volumes per second (VPS). Our imaging involves no special practical considerations or limitations of field of view or resolution, beyond the usual imaging goal of maximizing FOV while maintaining sufficient resolution to discern structures of interest (dendrites).

Analysis

Two-photon movies were motion corrected using the NormCorre package [99] in MATLAB. Spatial and temporal components for individual tufts imaged by two-photon and SCAPE were segmented using CaImAn v1.8.3, which employs large-scale sparse non-negative matrix factorization [45,100]. CaImAn inherently corrects for background signal. All further analyses used custom-written routines implemented in MATLAB. Spatial components with tuft structural characteristics were identified and analyzed, while neuropil components were discarded.

To quantify a tuft's response to stimuli, the mean stimulus-aligned ΔF/F was computed across all CS+ or CS- trials and corrected by the mean ΔF/F of the second before the trial. Probability of transients was obtained by taking each trial's ΔF/F in the first 1.5 seconds following either the CS+ or CS- and fitting these data with a univariate mixture of two Normal distributions: $(1-p)\mathrm{N}(\mu_1, \sigma_1) + p\mathrm{N}(\mu_2, \sigma_2)$. The smaller Normal reflects the distribution of failures, and the larger

Normal the distribution of transient amplitudes following the stimulus. The parameter *p* captures the probability of transients.

From these data, a selectivity index (SI) was defined as $(F_{CS+} - F_{CS-}) / (F_{CS+} + F_{CS-})$, in which F_{CS+} and F_{CS-} are the mean stimulus-aligned amplitudes (ΔF/F) to the CS+ and CS- within the first 1.5 seconds, respectively. This yielded values that range from -1 (exclusively CS- responsive) to 1 (exclusively CS+ responsive). Neural discriminability was defined as $d' = |F_{CS+} - F_{CS-}| / \sqrt{((\sigma^2_{CS+} + \sigma^2_{CS-})/2)}$ where σ^2_{CS+} is the variance of the response amplitudes in F_{CS+} and σ^2_{CS-} is the variance of the response amplitudes in F_{CS-}.

For longitudinal analysis, tufts were categorized as stimulus responsive if they met two criteria: 1) Across all trials, the mean ΔF/F 1.5 seconds before and 1.5 seconds after the stimulus were significantly different according to the Wilcoxon rank sum test, for either the CS+ or CS-, and 2) the average response amplitude for that stimulus was greater than 0.04 ΔF/F. Tufts with a significant response to only one stimulus were categorized as highly selective and their |SI| was set to 1. To classify tufts as behaviorally modulated, the mean ΔF/F of the first 1.5 seconds after the stimulus was computed for false alarm and correct rejection trials and compared with a rank sum test. Only sessions with at least 12 false alarm trials were used for this analysis. If the two distributions were significantly different, the tuft was classified as behaviorally modulated.

Custom MATLAB software was used to compute the median whisker angle, and whisking amplitude was computed as described previously [101]. The median angle was bandpass filtered from 4 to 30 Hz and passed through a Hilbert transform to calculate phase. We defined the upper and lower envelopes of the unfiltered median whisking angle as the points in the whisk cycle where phase equaled 0 (most protracted) or π (most retracted), respectively. Whisking amplitude

was defined as the difference between these two envelopes. Periods of whisking were defined as times where whisking amplitude exceeded 20% of maximum for at least 250 ms. Periods of time where amplitude exceeded this threshold for less than 250 ms were considered ambiguous and excluded from analysis of whisking versus quiescence. The whisking-triggered average for each tuft was computed by aligning the calcium signal to the start times of whisking periods during inter-trial intervals (2-8 seconds after stimulus delivery).

For the linear regression analysis, we excerpted the calcium timeseries 2 seconds before and 6 seconds after each stimulus onset. The whisking amplitude signal was frame aligned to the calcium signal according to the lag of the calcium-whisking cross-correlation peak for each tuft. Whisking amplitude was then normalized to the max, yielding values that ranged from 0 to 1. The stimulus predictor variable was a binary vector with an 800-msec 'on' period (24 frames) centered at the stimulus time. The timing of the stimulus variable was then aligned to the calcium signal according to the latency of peak of the mean $\Delta F/F$ of the first 1.5 seconds relative to the stimulus. The lick predictor variable was a binary vector with 'on' periods denoting lick bouts. Lick bouts were defined as periods of time where the mouse elicited at least 2 licks, with a maximum gap of 200 ms, and therefore had variable lengths.

For support vector machine (SVM) analysis, the mean $\Delta F/F$ was computed for a pre-stimulus epoch (1 second immediately preceding the stimulus, used as a negative control) and a post-stimulus epoch (0.1 – 1.1 seconds after the stimulus) for each trial. Binary SVMs were trained separately for each epoch using the MATLAB function fitcsvm. For each iteration, 75% of trials were randomly chosen to train the SVM, and decoder performance was tested on the remaining 25% of trials. Decoder performance for each session was averaged across 10 iterations.

All statistical tests were two-sided. T-tests were used for Normally distributed data. Otherwise non-parametric tests were applied.

References

1 Poort, J. *et al.* Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron* **86**, 1478-1490, doi:10.1016/j.neuron.2015.05.037 (2015).

2 Liu, D. *et al.* Orbitofrontal control of visual cortex gain promotes visual associative learning. *Nat Commun* **11**, 2784, doi:10.1038/s41467-020-16609-7 (2020).

3 Henschke, J. U. *et al.* Reward Association Enhances Stimulus-Specific Representations in Primary Visual Cortex. *Curr Biol* **30**, 1866-1880 e1865, doi:10.1016/j.cub.2020.03.018 (2020).

4 Beitel, R. E., Schreiner, C. E., Cheung, S. W., Wang, X. & Merzenich, M. M. Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals. *Proc Natl Acad Sci U S A* **100**, 11070-11075, doi:10.1073/pnas.1334187100 (2003).

5 Goltstein, P. M., Coffey, E. B., Roelfsema, P. R. & Pennartz, C. M. In vivo two-photon Ca2+ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. *J Neurosci* **33**, 11540-11555, doi:10.1523/JNEUROSCI.1341-12.2013 (2013).

6 David, S. V., Fritz, J. B. & Shamma, S. A. Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc Natl Acad Sci U S A* **109**, 2144-2149, doi:10.1073/pnas.1117717109 (2012).

7 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, doi:10.1038/nn1141 (2003).

8 Banerjee, A. *et al.* Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature* **585**, 245-250, doi:10.1038/s41586-020-2704-z (2020).

9 Zhang, W. & Bruno, R. M. High-order thalamic inputs to primary somatosensory cortex are stronger and longer lasting than cortical inputs. *eLife* **8**, doi:10.7554/eLife.44158 (2019).

10 Rubio-Garrido, P., Perez-de-Manzo, F., Porrero, C., Galazo, M. J. & Clasca, F. Thalamic input to distal apical dendrites in neocortical layer 1 is massive and highly convergent. *Cereb Cortex* **19**, 2380-2395, doi:10.1093/cercor/bhn259 (2009).

11 Cauller, L. J., Clancy, B. & Connors, B. W. Backward cortical projections to primary somatosensory cortex in rats extend long horizontal axons in layer I. *J Comp Neurol* **390**, 297-310 (1998).

12 Amitai, Y., Friedman, A., Connors, B. W. & Gutnick, M. J. Regenerative activity in apical dendrites of pyramidal cells in neocortex. *Cereb Cortex* **3**, 26-38 (1993).

13 Yuste, R., Gutnick, M. J., Saar, D., Delaney, K. R. & Tank, D. W. Ca2+ accumulations in dendrites of neocortical pyramidal neurons: an apical band and evidence for two functional compartments. *Neuron* **13**, 23-43 (1994).

14 Schiller, J., Schiller, Y., Stuart, G. & Sakmann, B. Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. *J Physiol* **505 (Pt 3)**, 605-616 (1997).

15 Larkum, M. E., Nevian, T., Sandler, M., Polsky, A. & Schiller, J. Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science* **325**, 756-760, doi:10.1126/science.1171958 (2009).

16 Sandler, M., Shulman, Y. & Schiller, J. A Novel Form of Local Plasticity in Tuft Dendrites of Neocortical Somatosensory Layer 5 Pyramidal Neurons. *Neuron* **90**, 1028-1042, doi:10.1016/j.neuron.2016.04.032 (2016).

17 Manita, S. *et al.* A Top-Down Cortical Circuit for Accurate Sensory Perception. *Neuron* **86**, 1304-1316, doi:10.1016/j.neuron.2015.05.006 (2015).

18 Roelfsema, P. R. & Holtmaat, A. Control of synaptic plasticity in deep cortical networks. *Nat Rev Neurosci* **19**, 166-180, doi:10.1038/nrn.2018.6 (2018).

19 Larkum, M. E. & Zhu, J. J. Signaling of layer 1 and whisker-evoked Ca2+ and Na+ action potentials in distal and terminal dendrites of rat neocortical pyramidal neurons in vitro and in vivo. *J Neurosci* **22**, 6991-7005, doi:20026717 (2002).

20 Larkum, M. E., Senn, W. & Luscher, H. R. Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cereb Cortex* **14**, 1059-1070, doi:10.1093/cercor/bhh065 (2004).

21 Schwindt, P. & Crill, W. Mechanisms underlying burst and regular spiking evoked by dendritic depolarization in layer 5 cortical pyramidal neurons. *J Neurophysiol* **81**, 1341-1354, doi:10.1152/jn.1999.81.3.1341 (1999).

22 Larkum, M. E., Zhu, J. J. & Sakmann, B. Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. *J Physiol* **533**, 447-466 (2001).

23 Manita, S., Miyakawa, H., Kitamura, K. & Murayama, M. Dendritic Spikes in Sensory Perception. *Front Cell Neurosci* **11**, 29, doi:10.3389/fncel.2017.00029 (2017).

24 Francioni, V., Padamsey, Z. & Rochefort, N. L. High and asymmetric somato-dendritic coupling of V1 layer 5 neurons independent of visual stimulation and locomotion. *Elife* **8**, doi:10.7554/eLife.49145 (2019).

25 Beaulieu-Laroche, L., Toloza, E. H. S., Brown, N. J. & Harnett, M. T. Widespread and Highly Correlated Somato-dendritic Activity in Cortical Layer 5 Neurons. *Neuron* **103**, 235-241 e234, doi:10.1016/j.neuron.2019.05.014 (2019).

26 Labarrera, C. *et al.* Adrenergic Modulation Regulates the Dendritic Excitability of Layer 5 Pyramidal Neurons In Vivo. *Cell Rep* **23**, 1034-1044, doi:10.1016/j.celrep.2018.03.103 (2018).

27 Brombas, A., Fletcher, L. N. & Williams, S. R. Activity-dependent modulation of layer 1 inhibitory neocortical circuits by acetylcholine. *J Neurosci* **34**, 1932-1941, doi:10.1523/JNEUROSCI.4470-13.2014 (2014).

28 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 e2002, doi:10.1016/j.celrep.2019.01.093 (2019).

29 Luebke, J. I. *et al.* Dendritic vulnerability in neurodegenerative disease: insights from analyses of cortical pyramidal neurons in transgenic mouse models. *Brain Struct Funct* **214**, 181-199, doi:10.1007/s00429-010-0244-2 (2010).

30 Tsai, J., Grutzendler, J., Duff, K. & Gan, W. B. Fibrillar amyloid deposition leads to local synaptic abnormalities and breakage of neuronal branches. *Nat Neurosci* **7**, 1181-1183, doi:10.1038/nn1335 (2004).

31 Xu, N. L. *et al.* Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature* **492**, 247-251, doi:10.1038/nature11601 (2012).

32 Takahashi, N. *et al.* Active dendritic currents gate descending cortical outputs in perception. *Nat Neurosci*, doi:10.1038/s41593-020-0677-8 (2020).

33 Takahashi, N., Oertner, T. G., Hegemann, P. & Larkum, M. E. Active cortical dendrites modulate perception. *Science* **354**, 1587-1590, doi:10.1126/science.aah6066 (2016).

34 Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S. & Magee, J. C. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science* **357**, 1033-1036, doi:10.1126/science.aan3846 (2017).

35 Doron, G. *et al.* Perirhinal input to neocortical layer 1 controls learning. *Science* **370**, doi:10.1126/science.aaz3136 (2020).

36 Bouchard, M. B. *et al.* Swept confocally-aligned planar excitation (SCAPE) microscopy for high speed volumetric imaging of behaving organisms. *Nat Photonics* **9**, 113-119, doi:10.1038/nphoton.2014.323 (2015).

37 Hillman, E. M. *et al.* High-speed 3D imaging of cellular activity in the brain using axially-extended beams and light sheets. *Curr Opin Neurobiol* **50**, 190-200 (2018).

38 Yu, Y. S., Graff, M. M., Bresee, C. S., Man, Y. B. & Hartmann, M. J. Whiskers aid anemotaxis in rats. *Sci Adv* **2**, e1600716, doi:10.1126/sciadv.1600716 (2016).

39 Yu, Y. S., Graff, M. M. & Hartmann, M. J. Mechanical responses of rat vibrissae to airflow. *J Exp Biol* **219**, 937-948, doi:10.1242/jeb.126896 (2016).

40 Nakamura, S., Narumi, T., Tsutsui, K. & Iijima, T. Difference in the functional significance between the lemniscal and paralemniscal pathways in the perception of direction of single-whisker stimulation examined by muscimol microinjection. *Neurosci Res* **64**, 323-329, doi:10.1016/j.neures.2009.04.005 (2009).

41 Bernhard, S. M. *et al.* An automated homecage system for multiwhisker detection and discrimination learning in mice. *PLoS One* **15**, e0232916, doi:10.1371/journal.pone.0232916 (2020).

42 Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295-300, doi:10.1038/nature12354 (2013).

43 Kozorovitskiy, Y., Saunders, A., Johnson, C. A., Lowell, B. B. & Sabatini, B. L. Recurrent network activity drives striatal synaptogenesis. *Nature* **485**, 646-650, doi:10.1038/nature11052 (2012).

44 Glickfeld, L. L., Andermann, M. L., Bonin, V. & Reid, R. C. Cortico-cortical projections in mouse visual cortex are functionally target specific. *Nat Neurosci* **16**, 219-226, doi:10.1038/nn.3300 (2013).

45 Giovannucci, A. *et al.* CaImAn an open source tool for scalable calcium imaging data analysis. *Elife* **8**, doi:10.7554/eLife.38173 (2019).

46 Pakan, J. M. P., Currie, S. P., Fischer, L. & Rochefort, N. L. The Impact of Visual Cues, Reward, and Motor Feedback on the Representation of Behaviorally Relevant Spatial Locations in Primary Visual Cortex. *Cell Rep* **24**, 2521-2528, doi:10.1016/j.celrep.2018.08.010 (2018).

47 Wang, P. Y. *et al.* Transient and Persistent Representations of Odor Value in Prefrontal Cortex. *Neuron* **108**, 209-224 e206, doi:10.1016/j.neuron.2020.07.033 (2020).

48 Bruno, R. M., Khatri, V., Land, P. W. & Simons, D. J. Thalamocortical angular tuning domains within individual barrels of rat somatosensory cortex. *J Neurosci* **23**, 9565-9574 (2003).

49 Bruno, R. M. & Sakmann, B. Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* **312**, 1622-1627, doi:10.1126/science.1124593 (2006).

50 Ramirez, A. *et al.* Spatiotemporal receptive fields of barrel cortex revealed by reverse correlation of synaptic input. *Nat Neurosci* **17**, 866-875, doi:10.1038/nn.3720 (2014).

51 Yao, H. & Dan, Y. Stimulus timing-dependent plasticity in cortical processing of orientation. *Neuron* **32**, 315-323, doi:10.1016/s0896-6273(01)00460-3 (2001).

52 Dragoi, V., Sharma, J., Miller, E. K. & Sur, M. Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nat Neurosci* **5**, 883-891, doi:10.1038/nn900 (2002).

53 Dragoi, V., Sharma, J. & Sur, M. Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron* **28**, 287-298, doi:10.1016/s0896-6273(00)00103-3 (2000).

54 Zhang, Y., Cudmore, R. H., Lin, D. T., Linden, D. J. & Huganir, R. L. Visualization of NMDA receptor-dependent AMPA receptor synaptic plasticity in vivo. *Nat Neurosci* **18**, 402-407, doi:10.1038/nn.3936 (2015).

55 Chu, M. W., Li, W. L. & Komiyama, T. Balancing the Robustness and Efficiency of Odor Representations during Learning. *Neuron* **92**, 174-186, doi:10.1016/j.neuron.2016.09.004 (2016).

56 Cichon, J. & Gan, W. B. Branch-specific dendritic Ca(2+) spikes cause persistent synaptic plasticity. *Nature* **520**, 180-185, doi:10.1038/nature14251 (2015).

57 Palmer, L. M. *et al.* NMDA spikes enhance action potential generation during sensory input. *Nat Neurosci* **17**, 383-390, doi:10.1038/nn.3646 (2014).

58 Derdikman, D. *et al.* Layer-specific touch-dependent facilitation and depression in the somatosensory cortex during active whisking. *J Neurosci* **26**, 9538-9547, doi:10.1523/JNEUROSCI.0918-06.2006 (2006).

59 de Kock, C. P. & Sakmann, B. Spiking in primary somatosensory cortex during natural whisking in awake head-restrained rats is cell-type specific. *Proc Natl Acad Sci U S A* **106**, 16446-16450, doi:10.1073/pnas.0904143106 (2009).

60 Rodgers, C. C. *et al.* Sensorimotor strategies and neuronal representations for shape discrimination. *Neuron* **109**, 2308-2325 e2310, doi:10.1016/j.neuron.2021.05.019 (2021).

61 Larkum, M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci* **36**, 141-151, doi:10.1016/j.tins.2012.11.006 (2013).

62 Abs, E. *et al.* Learning-Related Plasticity in Dendrite-Targeting Layer 1 Interneurons. *Neuron* **100**, 684-699 e686, doi:10.1016/j.neuron.2018.09.001 (2018).

63 Cohen-Kashi Malina, K. *et al.* NDNF interneurons in layer 1 gain-modulate whole cortical columns according to an animal's behavioral state. *Neuron* **109**, 2150-2164 e2155, doi:10.1016/j.neuron.2021.05.001 (2021).

64 Ranganathan, G. N. *et al.* Active dendritic integration and mixed neocortical network representations during an adaptive sensing behavior. *Nat Neurosci* **21**, 1583-1590, doi:10.1038/s41593-018-0254-6 (2018).

65 Park, J. M. *et al.* Deep and superficial layers of the primary somatosensory cortex are critical for whisker-based texture discrimination in mice. *bioRxiv*, 2020.2008.2012.245381, doi:10.1101/2020.08.12.245381 (2020).

66 Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585-590, doi:10.1038/nature12160 (2013).

67 Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M. & Harris, K. D. High-dimensional geometry of population responses in visual cortex. *Nature* **571**, 361-365, doi:10.1038/s41586-019-1346-5 (2019).

68 Kim, J., Erskine, A., Cheung, J. A. & Hires, S. A. Behavioral and Neural Bases of Tactile Shape Discrimination Learning in Head-Fixed Mice. *Neuron* **108**, 953-967 e958, doi:10.1016/j.neuron.2020.09.012 (2020).

69 Dabney, W. *et al.* A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, 671-675, doi:10.1038/s41586-019-1924-6 (2020).

70 Petreanu, L. *et al.* Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* **489**, 299-303, doi:10.1038/nature11321 (2012).

71 Wimmer, V. C., Bruno, R. M., de Kock, C. P., Kuner, T. & Sakmann, B. Dimensions of a Projection Column and Architecture of VPM and POm Axons in Rat Vibrissal Cortex. *Cereb Cortex* **20**, 2265-2276, doi:bhq068 [pii] 10.1093/cercor/bhq068 (2010).

72 Feldmeyer, D. Excitatory neuronal connectivity in the barrel cortex. *Front Neuroanat* **6**, 24, doi:10.3389/fnana.2012.00024 (2012).

73 Constantinople, C. M. & Bruno, R. M. Deep cortical layers are activated directly by thalamus. *Science* **340**, 1591-1594, doi:10.1126/science.1236425 (2013).

74 Petersen, C. C. & Crochet, S. Synaptic computation and sensory processing in neocortical layer 2/3. *Neuron* **78**, 28-48, doi:10.1016/j.neuron.2013.03.020 (2013).

75 Yamashita, T. *et al.* Diverse Long-Range Axonal Projections of Excitatory Layer 2/3 Neurons in Mouse Barrel Cortex. *Front Neuroanat* **12**, 33, doi:10.3389/fnana.2018.00033 (2018).

76 Llinas, R., Ribary, U., Contreras, D. & Pedroarena, C. The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci* **353**, 1841-1849, doi:10.1098/rstb.1998.0336 (1998).

77 Krauzlis, R. J., Lovejoy, L. P. & Zenon, A. Superior colliculus and visual spatial attention. *Annu Rev Neurosci* **36**, 165-182, doi:10.1146/annurev-neuro-062012-170249 (2013).

78 Parvizi, J. & Damasio, A. Consciousness and the brainstem. *Cognition* **79**, 135-160, doi:10.1016/s0010-0277(00)00127-x (2001).

79 Petty, G. H. & Bruno, R. M. Attentional modulation of secondary somatosensory and visual thalamus of mice. *eLife* **13** (2024).

80 Feldman, D. E. & Brecht, M. Map plasticity in somatosensory cortex. *Science* **310**, 810-815, doi:10.1126/science.1115807 (2005).

81 Mechawar, N., Cozzari, C. & Descarries, L. Cholinergic innervation in adult rat cerebral cortex: a quantitative immunocytochemical description. *J Comp Neurol* **428**, 305-318 (2000).

82 Freedman, R., Foote, S. L. & Bloom, F. E. Histochemical characterization of a neocortical projection of the nucleus locus coeruleus in the squirrel monkey. *J Comp Neurol* **164**, 209-231, doi:10.1002/cne.901640205 (1975).

83 Chubykin, A. A., Roach, E. B., Bear, M. F. & Shuler, M. G. A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* **77**, 723-735, doi:10.1016/j.neuron.2012.12.039 (2013).

84 Thiele, A. & Bellgrove, M. A. Neuromodulation of Attention. *Neuron* **97**, 769-785, doi:10.1016/j.neuron.2018.01.008 (2018).

85 Hangya, B., Ranade, S. P., Lorenc, M. & Kepecs, A. Central Cholinergic Neurons Are Rapidly Recruited by Reinforcement Feedback. *Cell* **162**, 1155-1168, doi:10.1016/j.cell.2015.07.057 (2015).

86 Froemke, R. C., Merzenich, M. M. & Schreiner, C. E. A synaptic memory trace for cortical receptive field plasticity. *Nature* **450**, 425-429, doi:10.1038/nature06289 (2007).

87 Letzkus, J. J. *et al.* A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* **480**, 331-335, doi:10.1038/nature10674 (2011).

88 Gasselin, C., Hohl, B., Vernet, A., Crochet, S. & Petersen, C. C. H. Cell-type-specific nicotinic input disinhibits mouse barrel cortex during active sensing. *Neuron*, doi:10.1016/j.neuron.2020.12.018 (2021).

89 Berger, T. K., Silberberg, G., Perin, R. & Markram, H. Brief bursts self-inhibit and correlate the pyramidal network. *PLoS Biol* **8**, doi:10.1371/journal.pbio.1000473 (2010).

90 Kapfer, C., Glickfeld, L. L., Atallah, B. V. & Scanziani, M. Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nat Neurosci* **10**, 743-753, doi:10.1038/nn1909 (2007).

91 Naka, A. & Adesnik, H. Inhibitory Circuits in Cortical Layer 5. *Front Neural Circuits* **10**, 35, doi:10.3389/fncir.2016.00035 (2016).

92 Ljungberg, T., Apicella, P. & Schultz, W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* **67**, 145-163, doi:10.1152/jn.1992.67.1.145 (1992).

93 Pan, W. X., Schmidt, R., Wickens, J. R. & Hyland, B. I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J Neurosci* **25**, 6235-6242, doi:10.1523/JNEUROSCI.1478-05.2005 (2005).

94 Bouret, S. & Sara, S. J. Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. *Eur J Neurosci* **20**, 791-802, doi:10.1111/j.1460-9568.2004.03526.x (2004).

95 Sheffield, M. E. & Dombeck, D. A. Calcium transient prevalence across the dendritic arbour predicts place field properties. *Nature* **517**, 200-204, doi:10.1038/nature13871 (2015).

96 Oberlaender, M. *et al.* Cell type-specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissal cortex. *Cereb Cortex* **22**, 2375-2391, doi:10.1093/cercor/bhr317 (2012).

97 Clack, N. G. *et al.* Automated tracking of whiskers in videos of head fixed rodents. *PLoS Comput Biol* **8**, e1002591, doi:10.1371/journal.pcbi.1002591 (2012).

98 Voleti, V. *et al.* Real-time volumetric microscopy of in vivo dynamics and large-scale samples with SCAPE 2.0. *Nat Methods* **16**, 1054-1062, doi:10.1038/s41592-019-0579-4 (2019).

99 Pnevmatikakis, E. A. & Giovannucci, A. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *J Neurosci Methods* **291**, 83-94, doi:10.1016/j.jneumeth.2017.07.031 (2017).

100 Pnevmatikakis, E. A. *et al.* Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **89**, 285-299, doi:10.1016/j.neuron.2015.11.037 (2016).

101 Petty, G. H., Kinnischtzke, A. K., Hong, Y. K. & Bruno, R. M. Effects of arousal and movement on secondary somatosensory and visual thalamus. *bioRxiv*, 2020.2003.2004.977348, doi:10.1101/2020.03.04.977348 (2020).

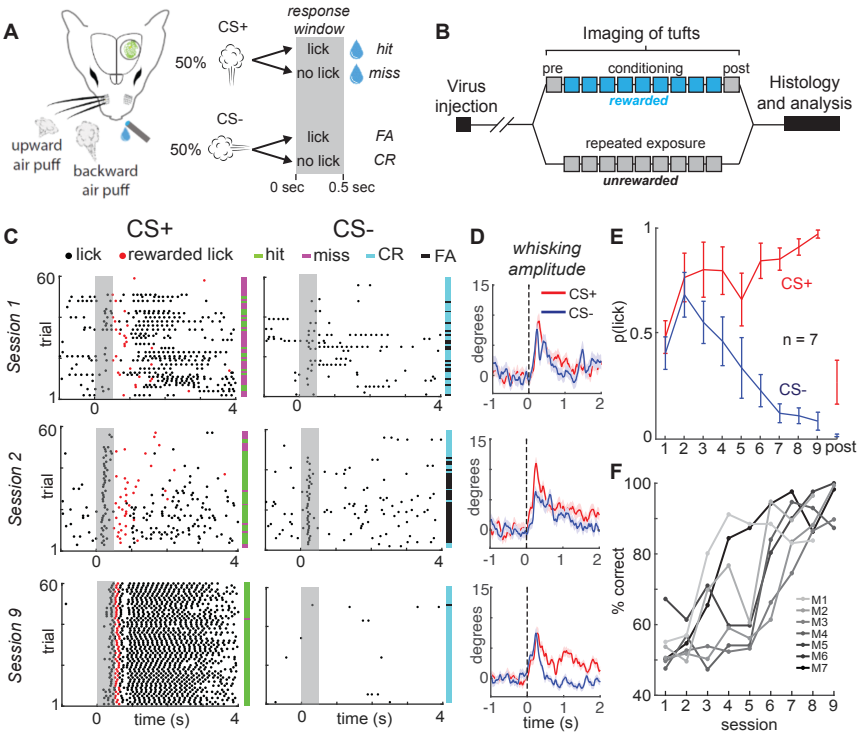

Fig. 1 | Mice rapidly learn to discriminate stimulus direction in head-fixed paradigm. **a,** A water droplet is paired with air puffs in one direction (CS+) but not the other (CS-). Licking in anticipation of water is assessed in the response window just after CS+ or CS- and prior to water delivery for the CS+ (grey bar). **b,** Experimental timeline. 2-3 weeks after virus injection, naive tuft responses to stimuli are recorded (pre). The CS+ is then paired with water for 8-9 days (blue). On the last day, stimuli are presented without reward (post). In a separate group of mice, the same stimuli are presented over 9 days in the absence of reward (unrewarded group). **c,** Lick rasters for three different sessions in one example mouse. On session 9, the CS+ but not the CS- reliably elicits licks. **d,** Mean baseline-subtracted whisking amplitude aligned to the CS+ (red) and CS- (navy) across sessions 1, 2, and 9 of an example mouse. **e,** Learning curve demonstrates rapid learning. Mean probability of at least one lick in the response window across sessions. f, Behavioral performance of each mouse in the rewarded group (M1 – M7).

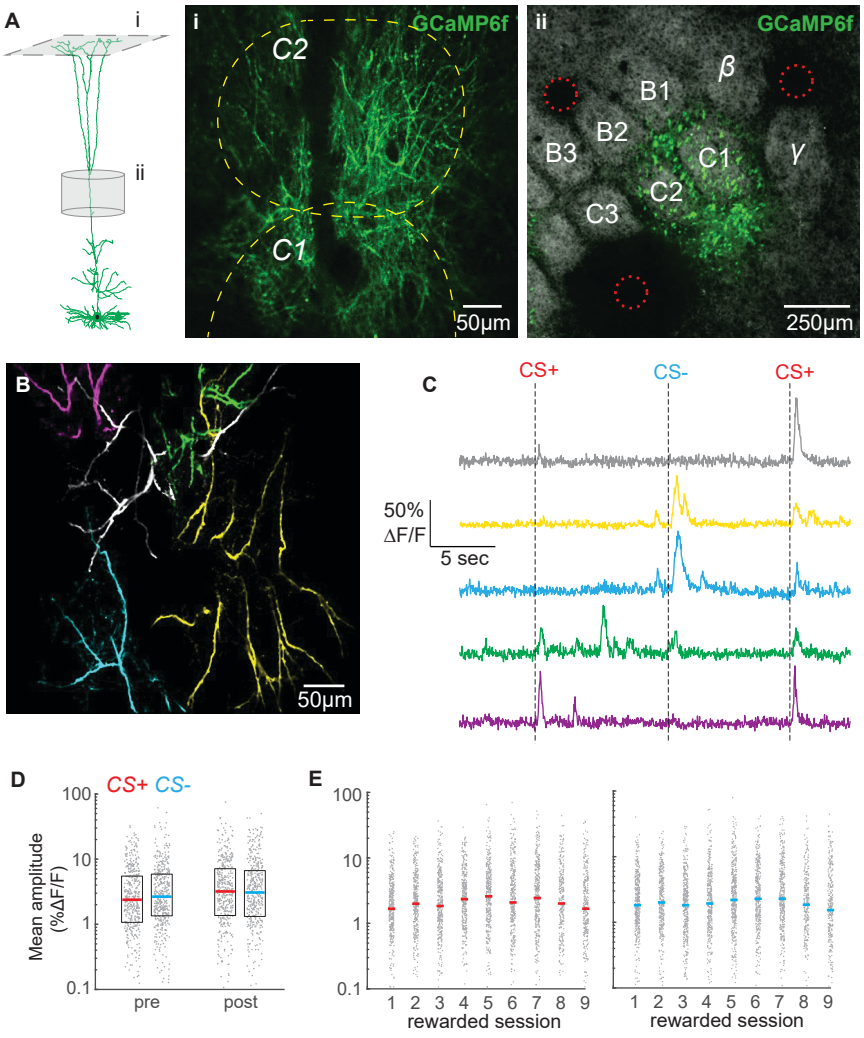

Fig. 2 | Overall tuft response to stimuli is unbiased and relatively stable across conditioning. a, Dendritic activity was recorded in layer 1 (i) in the C1/C2 barrel columns (ii). (i) Two-photon image ~60 µm deep relative to pia. Dashed yellow lines denote C1 and C2 boundaries from intrinsic imaging. (ii) Tangential section through layer 4 showing barrels stained with streptavidin-Alexa 647 and GCaMP6f-expressing apical trunks. Red circles indicate location of 2-photon lesions to mark the imaging region for post-hoc analysis. b, Overlay of five segmented pseudo-colored tufts from imaging field in A(i). c, Time courses of calcium responses of example tufts in b to three air puffs (dashes). d, Amplitude for CS+ (red) and CS- responses (blue), computed for each segmented tuft in the first 1.5 s post-stimulus (grey points), do not differ within or across sessions. Colored lines indicate median. e, Same as in d, showing data for all conditioning sessions.

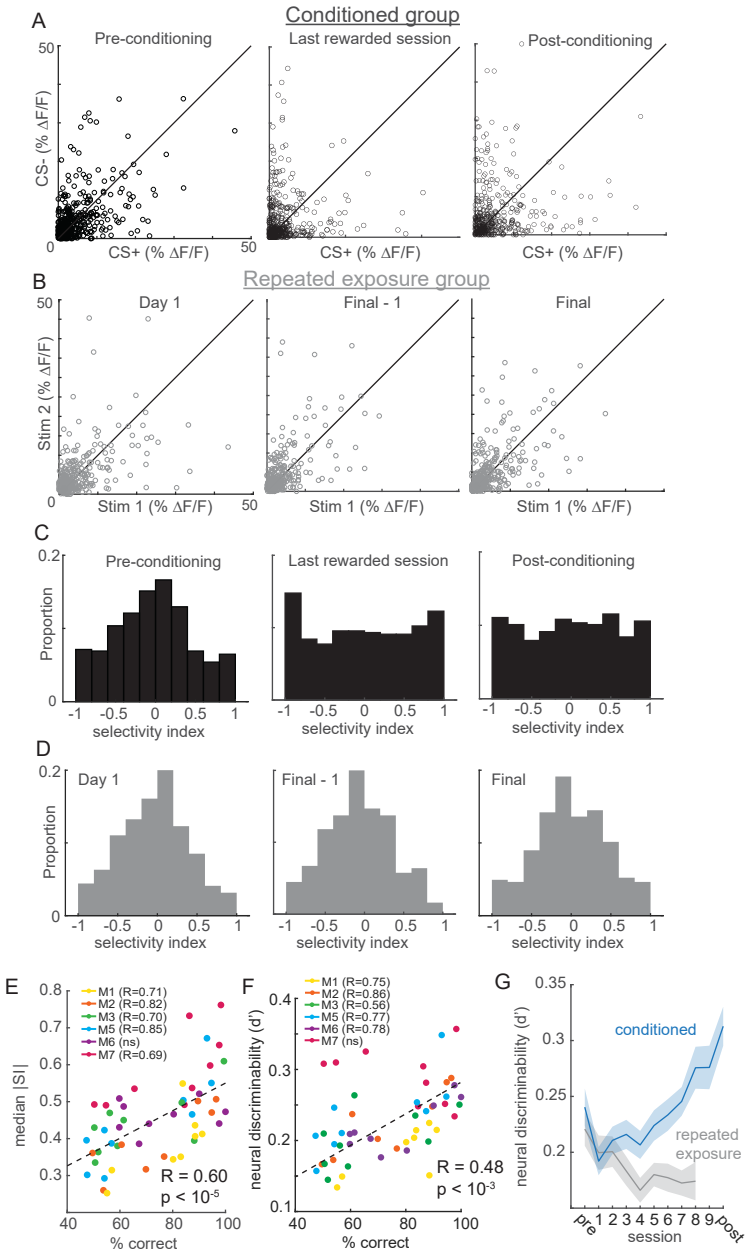

Fig. 3 | Reinforcement learning, but not stimulus exposure, enhances tuft selectivity for CS+ and CS- stimuli. a, Across the indicated sessions, individual tufts (circles) exhibit larger biases to CS+ or CS- (pooled across all conditioned mice). b, Repeated exposure to stimuli does not bias individual tufts to CS+ or CS-. c, Conditioning reshapes distribution of selectivity indices for tufts from Normal on pre-conditioning session to uniform on post-conditioning session. d, Distribution of tuft selectivity indices remains Normal throughout all repeated exposure sessions. e, Selectivity (median SI magnitude of tufts for each session) increases with behavioral performance of 6 animals. f, Same as e, but with neural discriminability plotted on the y axis. g, Neural discriminability (mean ± sem) of tufts, pooled across all animals on each session, increases with conditioning and decreases with repeated exposure.

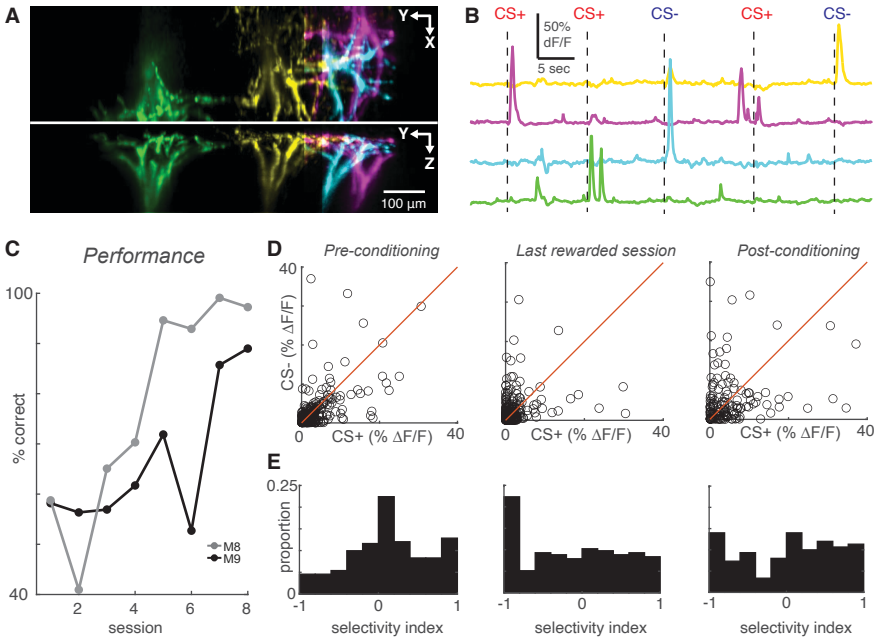

Fig. 4 | High-speed volumetric imaging of apical tufts confirms the emergence of enhanced selectivity after learning. a, Top and side view of four example tufts segmented from volumetric SCAPE imaging. **b,** Time courses of calcium activity from example tufts in **a** during five presentations of air puff stimuli (dashes). **c,** Performance across all conditioning sessions of two mice that were imaged with SCAPE. **d,** Across the indicated sessions, individual SCAPE-imaged tufts (circles) exhibit larger biases to CS+ or CS-. **e,** Conditioning reshapes selectivity distribution from Normal to uniform.

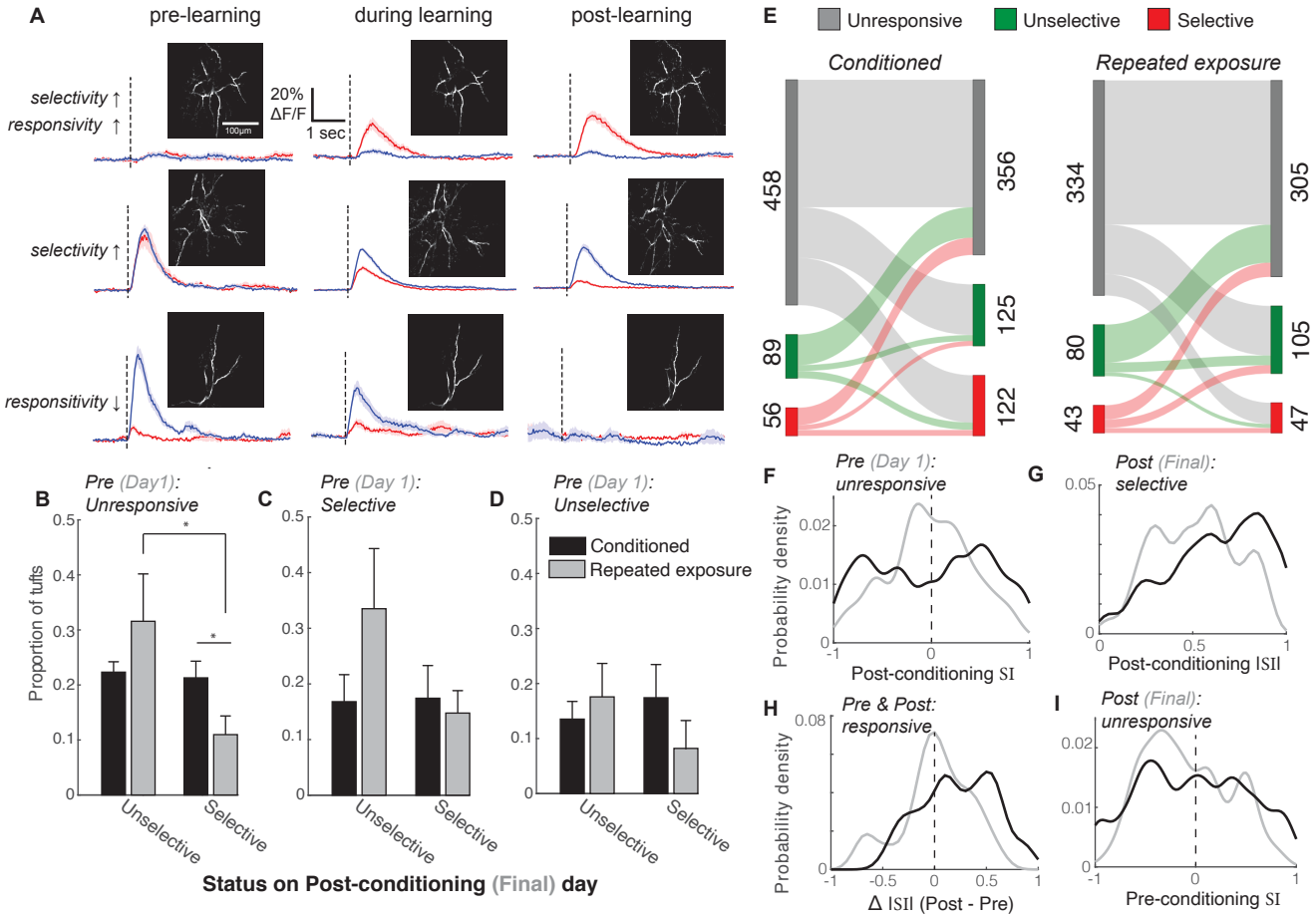

Fig. 5 | Longitudinal tracking reveals that reward enhances the selectivity of both initially unresponsive and responsive tufts. a, Three example tufts that were longitudinally tracked across learning. Top row: An initially unresponsive tuft develops a robust response to the CS+ but not the CS- after learning. Middle row: A responsive but unselective tuft loses its robust CS+ response and becomes selective for the CS-. Bottom row: A CS- selective neuron becomes unresponsive to both stimuli. **b,** Tufts that were unresponsive during the first session were longitudinally tracked to the last session. Plotted is the mean proportion of selective and unselective neurons across all animals in the conditioned (black bars) and repeated exposure (grey bars) groups. **c,d,** Same analysis as **b** for initially selective (**c**) and unselective (**d**) tufts. Two-sample t-test was used for comparisons between conditioned and repeated exposure groups. Paired t-test was used for comparisons within a group. * $p < 0.05$. **e,** Total tuft counts from first to last session within the 3 response categories for either conditioned (left) or repeated exposure (right) groups. **f,** SI of responsive tufts on the last session that were initially unresponsive during the first session. Conditioned tufts have enhanced selectivity compared to repeated exposure. **g,** Tufts that were selective on the last session are more selective if conditioned (black) rather than undergoing repeated exposure (grey). **h,** Tufts that responded on both pre and post sessions tend to have higher selectivity if conditioned rather than undergoing repeated exposure. **i,** SI of responsive tufts on the first session that later became unresponsive during the last session.

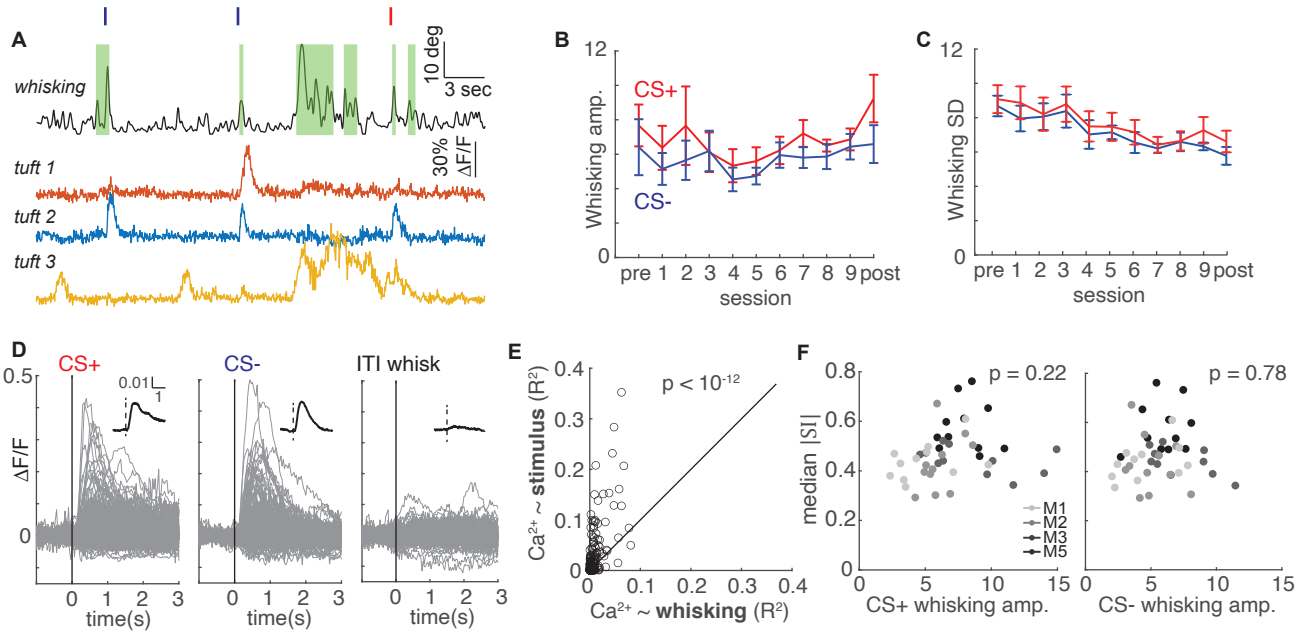

Fig. 6 | Whisking is only weakly correlated with tuft activity and cannot account for changes in selectivity during learning. a, Whisking amplitude aligned to calcium activity of three example tufts in one session. Green shading indicates periods of whisking. Red and navy ticks indicate CS+ or CS- delivery, respectively. **b,** Mean whisking response of four mice to CS+ (red) and CS- (navy) does not change across sessions during learning (mean ± s.e.m.). **c,** Mean standard deviation of whisking decreases for both CS+ and CS- across learning, but CS+ and CS- do not differ. **d,** Event-triggered averages of 322 tufts on the post-conditioning day (grey traces - individual tufts, black inset - population average) are responsive to stimuli but relatively unmodulated by whisking. **e,** R2 values for linear models predicting calcium from stimuli (y axis) are consistently greater than those predicting calcium from whisking (x axis). Each circle represents a tuft. **f,** Magnitude of tuft selectivity does not correlate with mean whisking amplitude during CS+ (left) and CS- trials (right) on that session.

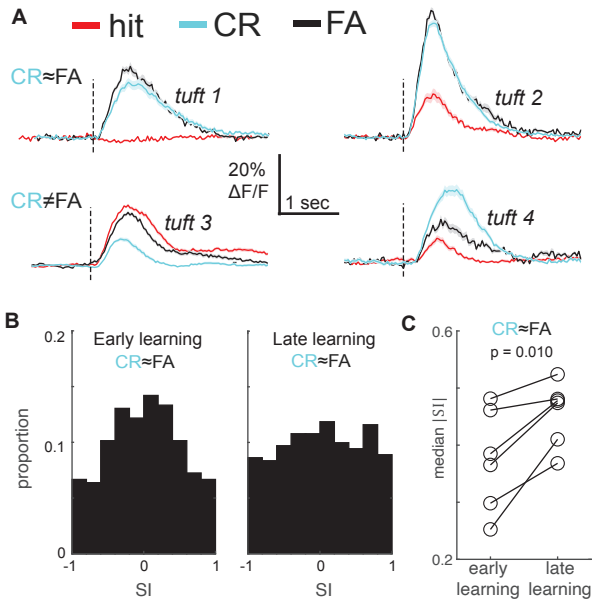

Fig. 7 | Behavioral responses do not account for enhancement of stimulus selectivity during learning. a, Mean stimulus responses of four tufts during hit (red), CR (cyan), and FA (black) trials. Top row: Example tufts whose responses are not behaviorally modulated (CR is similar to FA). Bottom row: Example tufts with behaviorally modulated responses (CR and FA differ). **b,** Selectivity index (SI) distribution changes from early (left) and late learning sessions (right) even when tufts with behaviorally modulated responses (CR≠FA) are excluded. **c,** Median SI magnitude of tufts in each of six animals (from panel **b**) increases from early to late learning sessions.

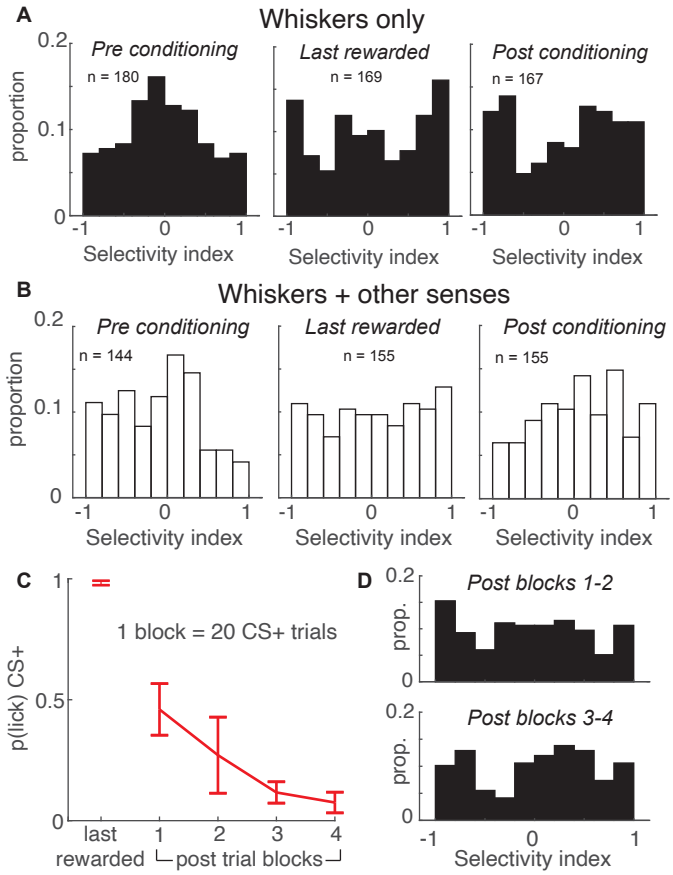

Fig. 8 | Apical tufts in barrel cortex of mice performing the task exclusively with their whiskers undergo long-lasting changes in selectivity. a, SI histograms of mice performing the task exclusively with their whiskers exhibit increased selectivity across pre-conditioning, last-rewarded, and post-conditioning sessions. **b,** Relative to pre-conditioning, mice using their whiskers and other sensory cues to perform the task have increased selectivity during the last rewarded session, but not the post-conditioning session. **c,** The probability of anticipatory licks in response to the CS+ extinguishes across post-conditioning blocks (of 20 trials each). **d,** Tuft selectively remains uniformly distributed during post-conditioning trial blocks 1-2 (top) while licking is extinguishing, and blocks 3-4 (bottom) in which licking is extinguished.