Dendritic integration of sensory and reward information facilitates learning

Gwendolin Schoenfeld[1,4], Sepp Kollmorgen[2,4], Christopher Lewis[1], Philipp Bethge[1], Anna Maria Reuss[3,4], Adriano Aguzzi[3,4], Valerio Mante[2,4], and Fritjof Helmchen[1,4†]

[1]Laboratory of Neural Circuit Dynamics, Brain Research Institute, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[2]Institute of Neuroinformatics, University of Zurich and ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[3]Institute of Neuropathology, University Hospital Zürich, Schmelzbergstrasse 12, 8091 Zurich, Switzerland
[4]Neuroscience Center Zürich, University of Zurich and ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
Correspondence: helmchen@hifo.uzh.ch

ABSTRACT – Learning goal-directed behaviours requires integration of separate information streams representing context, relevant stimuli and reward. Dendrites of pyramidal neurons are suitable sites for such integration, but it remains elusive how their responses adapt when an animal learns a new task. Here, we identify two distinct classes of dendritic responses that represent either contextual/sensory information or reward information and that differ in their task- and learning-related dynamics. Using longitudinal calcium imaging of apical dendritic tufts of L5 pyramidal neurons in mouse barrel cortex, we tracked dendritic activity across learning and analyzed both local dendritic branch signals and global apical tuft activity. During texture discrimination learning, sensory representations (including contextual and touch information) strengthened and converged on the reward-predicting tactile stimulus when mice became experts. In contrast, reward-associated responses were particularly strong in the naïve condition and became less pronounced upon learning. When we blocked the representation of unexpected reward in naïve animals with optogenetic inhibition, animals failed to learn until we released the block and learning proceeded normally. Our results suggest that reward signals in dendrites are essential for adjusting neuronal integration of converging inputs to facilitate adaptive behaviour.

Learning enables animals to adapt to environmental challenges and flexibly acquire new behaviours. For learning to occur, distinct information streams in the brain—coding for context, specific sensory stimuli, and outcome value (e.g. reward)—need to be integrated and processed in a novel, meaningful way to link reward predictions to suitable actions. Consistent with this notion, neuronal circuits dynamically reorganize their activity upon task learning[1]. For example, neuronal populations in the barrel cortex of primary somatosensory cortex (S1), especially those projecting to secondary somatosensory cortex (S2), undergo functional changes that reflect behavioural adaptations during texture discrimination learning[2]. Functional adaptations of neuronal circuits ultimately are implemented at the level of single neurons through plasticity mechanisms shaping synaptic integration in neuronal dendrites[3]. Because of their complexity, including nonlinear properties, dendrites possess a large computational power[4,5] and thus are prime candidate sites for adjustments that may be required for learning. Apical dendritic tufts of L5 pyramidal neurons in neocortex, for instance, receive a wide range of functionally distinct inputs, representing sensory stimuli[6,7], motor actions[8], information from higher-order areas[9], and reward[10]. Yet, it remains unknown how dendritic signals in individual branches or entire apical tufts—driven by these diverse inputs—reorganize during learning and to what degree such adaptations are related to behavioural changes upon task learning.

To address these questions, we sparsely expressed GCaMP6f in L5 pyramidal neurons in the S1 barrel cortex of 6–10-week-old adult Rbp4-Cre mice. We specifically targeted L5

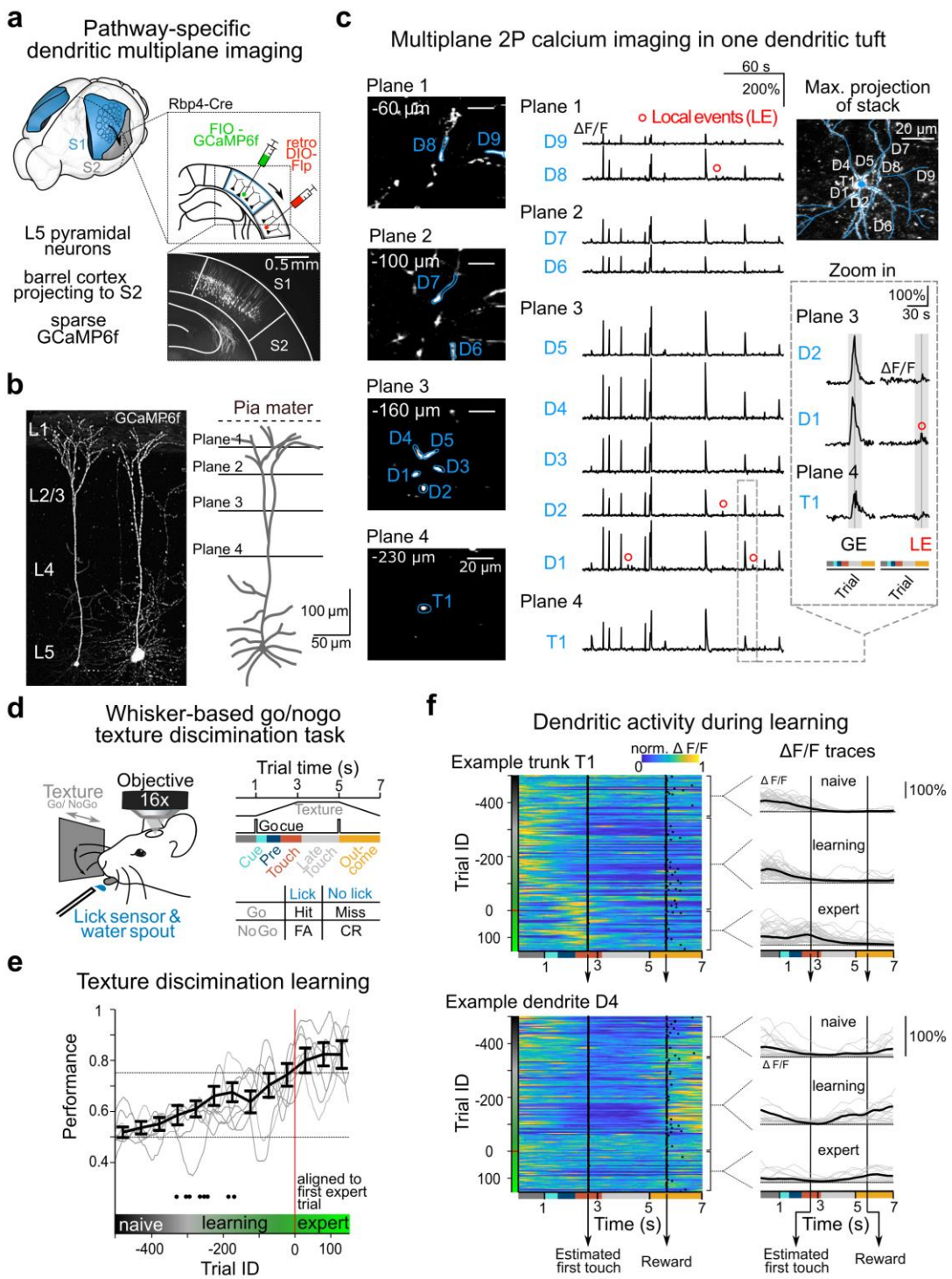


Fig. 1 | Longitudinal multi-plane calcium imaging of L5 dendrites across learning. a, S1→S2 projecting L5 pyramidal neurons labelled with GCaMP6f using a dual-virus approach in Rbp4-cre mice. *Lower right:* Coronal maximum-intensity projection (MIP) of labelled neuronal population in a cleared brain (CLARITY, 120-µm range). **b**, *Left:* Coronal MIP of sparsely GCaMP6f-labelled L5 neurons in barrel cortex in a cleared hemisphere (iDISCO, 150-µm range). *Right:* Schematic of four two-photon imaging planes. **c**, *Left:* Example field-of-view with selected ROIs and respective ΔF/F traces across all imaging planes. Red circles: local events (LE). *Top right*: MIP of the same tuft with arbourization schema. *Lower right:* Example ΔF/F traces of two trials containing global event (GE) and LE, respectively. **d**, Schematic of go/no-go texture discrimination task. *Right:* Trial structure and table of trial outcomes. **e**, Learning curves across 650 trials aligned to first expert trial (n = 8 mice). Black line: mean ± s.d. in 50-trial bins; grey lines: individual mice; dots indicate transitions from naïve to learning period for individual mice. **f**, *Left:* Heatmaps of normalized ΔF/F traces for two example dendrites across learning. *Right:* Example ΔF/F traces for naïve, learning, and expert Hit trials (grey: individual trials; black: mean traces). Arrows indicate approximate times of first touch and reward valve opening.

neurons projecting to S2 by using an intersectional Cre/Flp viral labelling approach (**Fig. 1a,b**; Methods). Because Rbp4-Cre mice are not specific for L5 subtypes, labelled neurons presumably included L5A and L5B neurons sparsely distributed across barrel cortex (**Supplementary Fig. 1**). Through a chronic cranial window and using multi-plane two-photon calcium imaging with an electrically tunable lens[11], we recorded calcium transients in multiple dendritic tuft branches and their parent apical trunks (**Fig. 1b,c**; **-**30 µm to -370 µm depth; 10 Hz effective frame rate). We manually selected trunk cross-sections and dendritic branch segments (continuous stretches of 5 - 20 µm) as regions of interests (ROIs). Multi-plane calcium imaging of multiple dendritic tuft branches and their corresponding parent trunk enabled us to discriminate between 'global events' (GE), with large-amplitude signals in the trunk and most tuft branches, and 'local events' (LE), with typically smaller but detectable calcium signals in only one or few tuft branches but no detectable calcium transient in the trunk (**Fig. 1c**).

To investigate dendritic calcium signals during learning, we trained mice (n = 8) in a go/no-go texture discrimination task[2,12]. In each trial, we presented either a coarse (grit size P100) or a smooth (P1200) sandpaper to the whiskers (go texture: P100, n = 3 mice; P1200, n = 5). Correct licking in go trials (Hits) triggered a water reward whereas correct rejections (CR) in no-go trials, as well as Misses on go trials, were neither rewarded nor punished. False alarms (FA) in no-go trials were mildly punished with white acoustic noise (**Fig. 1d**; Methods). For analysis, we defined five time windows linked to the trial structure (cue, pre-touch, touch, late-touch, and outcome). All mice improved their performance from naïve (50%, i.e. chance level) to expert level (>75%), on average within 1178 ± 368 trials (8-14 days; 1 session per day; 67-209 trials per session; mean ± s.d., n = 8 mice). To compensate for different learning rates, we aligned all individual learning curves to the first expert trial and performed analysis in the time window from 500 trials before to 150 trials after this time point (trial identifier [ID] from -500 to 150; **Fig. 1e**). With learning, mice developed anticipatory whisking preceding the first whisker-texture touch and anticipatory whisking and licking before the outcome window[2] (**Supplementary Fig. 2**). Throughout the entire training period, we measured task-related calcium signals longitudinally in the same dendrites of S1→S2 L5 neurons. Example calcium traces from a dendritic trunk (T1) and a tuft branch (D4) of an individual L5 neuron reveal preferred activity in variable task windows and also illustrate differences in the evolution of dendritic activity profiles across learning (**Fig. 1f**).

For a comprehensive analysis of task- and learning-related dendritic dynamics in our entire data set, we identified and selected calcium traces with significant calcium transients (Methods). Using a UMAP embedding based on similarity of the trial-related calcium traces, we found that both the short time scale (within single trials) and the long time scale (across trials) are reflected in the two-dimensional UMAP plot (**Fig. 2a**). Colour coding according to the onset times of clearly detectable dendritic calcium transients (Methods) showed that dendritic activity was distributed over the entire trial time course, including cue, pre-touch, touch, late-touch, and outcome windows. A small fraction of calcium traces displayed double-transients in both the pre-touch and outcome windows, represented in the middle of the UMAP plot. Colour coding based on trial ID revealed an interesting pattern of overrepresentation of calcium transients with onsets in the pre-touch and touch window in expert trials, whereas in early naïve trials calcium transients in the cue and outcome windows appeared to be more abundant (**Fig. 2a;** for colour coding of other features see **Supplementary Fig. 3**). This overview indicates that S1→S2 L5 dendritic activity reflects various task-related events and that the trial-related temporal pattern of dendritic activation reorganizes during learning.

To evaluate dendritic functional reorganization quantitatively and to examine whether consistent learning-related patterns exist, we developed 'transformation clustering', a statistical approach to identify dendritic tuft

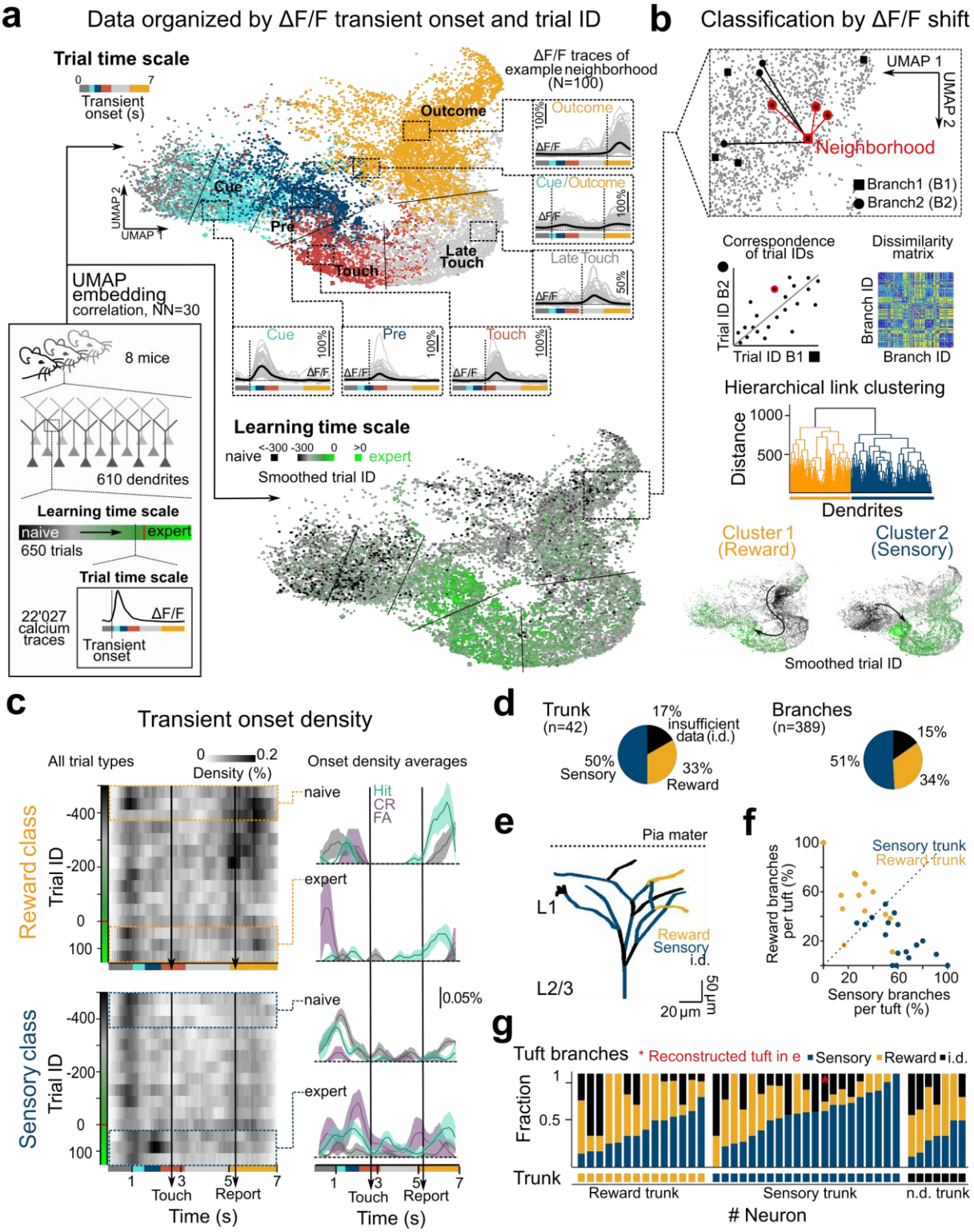


Fig. 2 | Two dendritic subclasses defined by distinct changes in sensory and reward representations. a, *Top:* Correlation-metric-based UMAP embedding of all recorded ΔF/F traces with detectable transient onsets, coloured according to transient onset time. Inserts: ΔF/F traces and average ΔF/F traces associated with local 2D neighbourhoods. *Bottom:* UMAP embedding coloured according to smoothed trial ID (8 mice, 22'027 traces). **b**, Schematic depiction of transformation clustering, identifying two functionally distinct clusters of dendritic responses by their coherent shift over learning. **c**, *Left:* Transient onset density during learning aligned to trial start in the reward and sensory class, Right: Average onset densities in naïve and expert conditions. Only trials with detectable onsets are included. **d**, Percentages of trunk and apical branches per class. **e**, Reconstruction of example tuft from a two-photon image stack. Branches are coloured according to the respective class. **f**, Class assignment of apical trunks in relationship to the percentage of branches classified as sensory and reward. **g**, Fraction of branches per neuron assigned to each cluster (42 tufts; sorted according to trunk class, sub-sorted according to percentage of sensory branches).

branches or apical trunks that exhibited similar learning-associated activity changes. Specifically, we compared the high-dimensional single-trial responses between all possible dendrite pairs across learning using a nearest neighbour method[13], and performed hierarchical link clustering (**Fig. 2b**; we greedily expanded clusters by also assigning dendritic ROIs with sparse activity; Methods and **Supplementary Fig. 4**). We identified two major, functionally distinct classes of change in dendritic responses across learning. One class shows abundant and large calcium transients in the outcome window for the naïve condition, especially upon reward in Hit trials and to a lesser degree upon punishment in FA trials (**Fig. 2c**). These responses decrease in frequency in the expert condition. We refer to this response pattern as the 'reward class', which in our interpretation mainly represents unexpected outcome and is less prominent when the animal has learned the task. In contrast, dendrites belonging to the second functional class display strong activation in naïve animals early in the trial (cue-related), which then shifts towards the pre-touch and touch windows in the expert condition (**Fig. 2c**). This 'sensory class' of responses apparently represents sensory inputs related to both contextual stimuli and the reward-predicting sensory stimulus, i.e. the texture touch. The distinct temporal dynamics across learning for both classes is also evident in the UMAP embedding (**Fig. 2b** and **Supplementary Fig. 3**). Differences across classes are also apparent in the average ΔF/F transients, albeit less pronounced than for transient onset densities (**Supplementary Fig. 5**).

To analyze the tuft composition in terms of functional class, we evaluated the 42 neurons with identified trunk ROI. For both trunks and tuft branches, nearly half of the dendrites belonged to the sensory class and about a third to the reward class (**Fig. 2d**; the remaining dendrites were not classified due to insufficient number of data points). Based on reconstructions from z-stacks we found that all tufts contained dendritic branches of both functional classes (**Fig. 2e-g**) with no apparent dependence on depth below the pia (**Supplementary Fig. 6**). In the majority of neurons, the functional class of the trunk corresponded to the most abundant class in its tuft branches (**Fig. 2f,g**). No difference in morphology was obvious in tufts dominated by either of the functional classes (**Supplementary Fig. 6**). Taken together, we identified two major functional classes of dendritic branches with distinct changes in their activity profiles across learning. One class highlights the representation of unexpected outcomes in the naïve condition, whereas the other develops enhanced activity in the pre-touch and touch windows upon learning, indicating integration of contextual, sensory, and reward inputs such that the reward-predicting sensory stimulus becomes expected and well-represented in the expert condition.

Functional changes in dendritic integration may also be reflected by changes in the abundance of local and global events (LE and GE, respectively). Whereas LE in tuft branches may reflect inputs not sufficiently strong to trigger calcium spikes near the main bifurcation (or lacking sufficient coincident basal dendritic activation[14,15]), GE likely are associated with somatic action potential output, presumably in form of action potential bursts[6,8]. To discriminate LE and GE, we considered for each dendritic tuft the simultaneously recorded ΔF/F signals in trunk and daughter branches. We defined events with significant calcium transients in the apical trunk as GE and events with activation of at least one tuft branch, but not the trunk, as LE (**Fig. 3a-c** and **Supplementary Fig. 7**). These definitions are justified because nearly all tuft branches were co-active with the trunk in GE whereas only a small fraction of branches were engaged in LE (**Fig. 3d**). In 27'300 trials considered in total, we labeled 14.4% as GE, 7.9% as LE, and 2.1% as mixed events (at least one GE and one LE with peak times separated by >1 s). In 73.1% of trials no event was detected, indicating sparse tuft activity (**Supplementary Fig. 7** and Methods; 2.5% of trials were not considered due to data exclusion criteria). If multiple GE or LE occurred in the same trial, the event with the largest amplitude in any ROI was used for further analysis. Overall, the correlation of branch and trunk

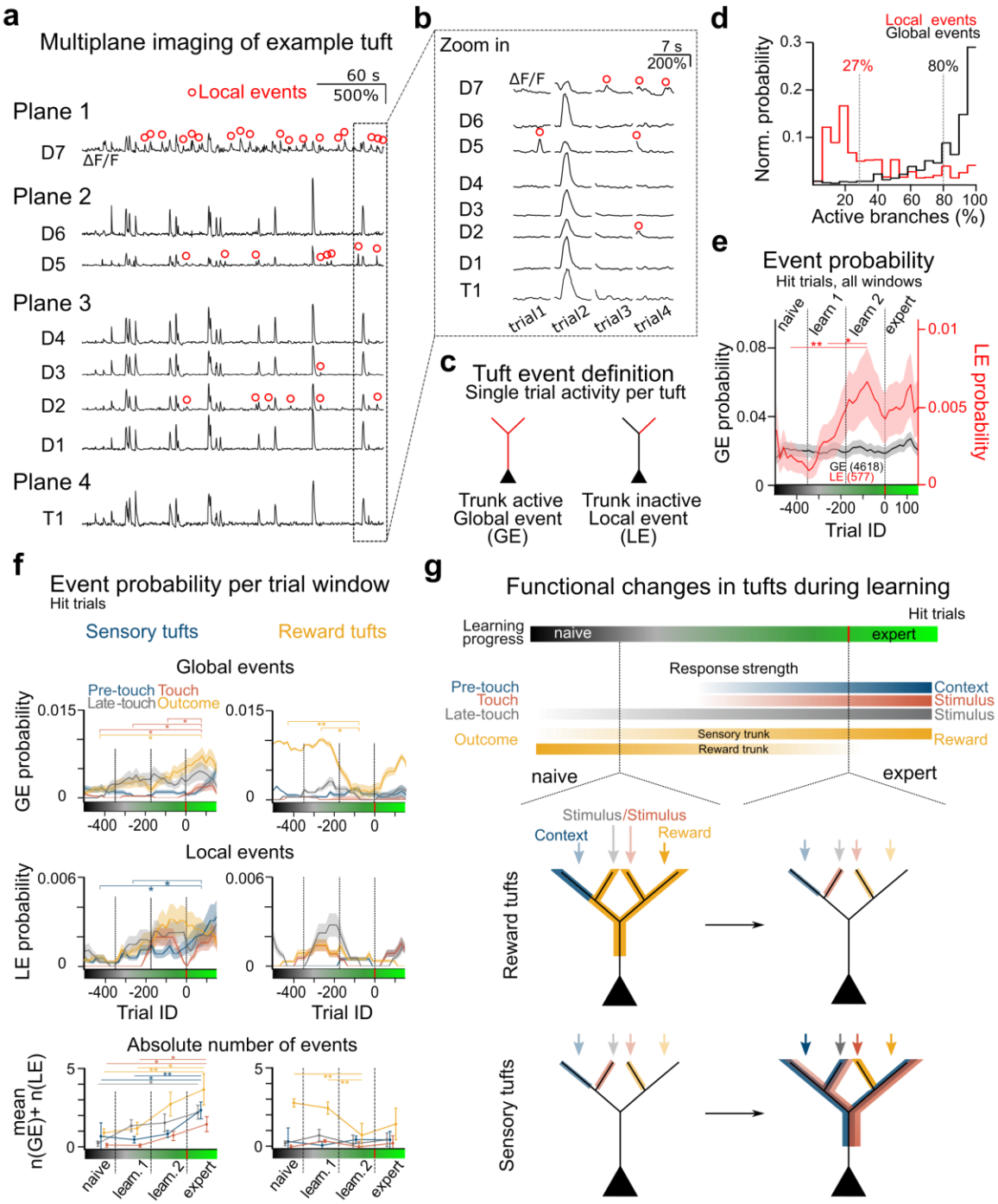


Fig. 3 | Gradient like change of dendritic representations during learning depends on behavioral window. a, Example ΔF/F traces of multiple tuft branches and trunk across all imaging planes (Red: LE). **b**, ΔF/F traces LE and GE. **c**, Schematic depiction of LE and GE. **d**, Percentages of branches with detectable calcium transients during LE and GE. (Dashed lines: mean). **e**, Probability of GE and LE over learning across all tufts in Hit trials. (Shaded area: s.e.m.; 0.003 and 0.04 for naïve vs. learning 2 and learning 1 vs. learning 2, one-way ANOVA) **f**, Probability of GE and LE per pre-touch, touch, late-touch and outcome window in Hit trials. (Shaded area: s.e.m.; Sensory trunks: GE in touch window: 0.03, 0.03 and 0.007 for naïve vs. expert, learning 1 vs. expert and learning 2 vs. expert; GE in outcome window: 0.011 for naïve vs. expert. LE in pre-touch window: 0.04 and 0.04 for naïve vs. expert and learning 1 vs. expert. Reward cluster: GE in outcome window: 0.009 and 0.01 for naïve vs. learning 2 and learning 1 vs. learning 2; one-way ANOVA). *Bottom*: Average number of trials with GE or LE (mean ± s.e.m.; Sensory tufts: 0.02 and 0.005 for naïve vs. expert and learning 1 vs. expert for pre-touch; 0.02, 0.007 for naïve vs. expert and learning 1 vs. expert for touch;0.015 for naïve vs. expert for late-touch; 0.003 and 0.04 for naïve vs. expert and learning 1 vs. expert for outcome; Reward tufts: 0.001 and 0.005 for naïve vs. learning 2 and learning 1 vs. learning 2 for outcome; one-way ANOVA) **g**, Summary of changes in strength of functional representations during learning and cellular working model.

signals in the same tuft was high but decreased with distance from the main bifurcation (**Supplementary Fig. 7**).

Next, we analyzed how the probability of GE and LE in Hit trials changes across learning. Averaged across all tufts, the probability of GE remained rather constant while the probability of LE increased with learning and plateaued once mice reached expert performance (**Fig. 3e**). We further differentiated in which of the salient trial time windows GE and LE occurred, with tufts separated according to their trunk's functional class (sensory or reward). In sensory tufts, LE and GE probabilities increased across learning for all tested windows, starting in the outcome and late-touch window and followed by the pre-touch and touch periods (**Fig. 3f**). Interestingly, LE and GE in sensory tufts also became more abundant in FA trials in expert animals, perhaps reflecting a reward prediction error (**Supplementary Fig. 8**). In reward tufts, GE probability was prominently high in naïve and early learning phase and then decreased with learning, whereas LE probability remained low and relatively unchanged. We also analyzed the discrimination power of GE, as an estimate of neuronal output. Although trunks showed some discriminability for trial types, discriminability of go vs. no-go texture overall was low (**Supplementary Fig. 9**).

We interpret these findings regarding changing event probabilities as learning-related changes of the effective strengths of several input streams converging on the L5 dendritic tufts, carrying information about context (pre-touch), relevant sensory stimulus (touch and late-touch), and reward (outcome window). In **Fig. 3g** we propose a working model, in which the functional class of a dendritic tuft receiving mixed inputs is determined by its dominant input type. Sensory tufts show a strengthening of the representation (more frequent LE and GE) of the task-relevant contextual and tactile inputs with learning whereas reward tufts in naïve animals display strong activity representing unexpected reward. From this model, we hypothesize that this salient outcome representation in reward tufts possibly drives circuit adaptations that are essential for learning, including the observed changes in dendritic integration.

To assess whether unexpected reward representation in the S1→S2 pathway is behaviourally relevant for learning, we densely expressed eArchT3.0 in S1→S2 L5 neurons and applied optogenetic inhibition specifically during the outcome window (**Fig. 4a-c;** n = 5). We validated the inhibitory effect of the optogenetic manipulation on eArchT3.0-expressing S1→S2 L5 neurons using extracellular recordings in anesthetized mice. Laser illumination reduced spontaneous multi-unit activity and generated a pronounced sink in the current source density (CSD) signal in superficial layers including L1, indicating inhibition of dendritic tufts (**Fig. 4d,e**). In addition, laser stimulation reduced the firing rate of single units detected in L5, consistent with the targeted subpopulation (**Fig. 4f**). Having verified the population effect of the perturbation in the anesthetized condition, we applied optogenetic inhibition in awake mice, starting in the naïve condition and continuing during training in the texture discrimination task. We applied optogenetic perturbation for 1'800 trials, well above the average number of trials that the set of mice trained in the 2-photon experiments required to reach expert performance. At the end of this long perturbation period, none of the mice with eArchT3.0-expression had reached expert level (mean 62% ± 6.2%). After 1'800 trials, the optogenetic block was lifted. Only then, eArchT3.0-expressing mice were able to improve their task performance to expert levels, with a time course comparable to mice expressing eYFP or GCaMP6f (**Fig. 4g,h; Supplementary Fig. 10**). Licking and whisking behaviour remained unaffected by the optogenetic perturbation (**Supplementary Fig. 11**). We conclude that processing of the reward signal in the S1→S2 neuronal population, presumably involving dendritic tuft activity, is required for behavioural adaptation.

In summary, dendritic tuft activity in S1→S2 L5 pyramidal neurons displays a spectrum of task-related responses, with two major functional classes distinguished by distinct profiles of

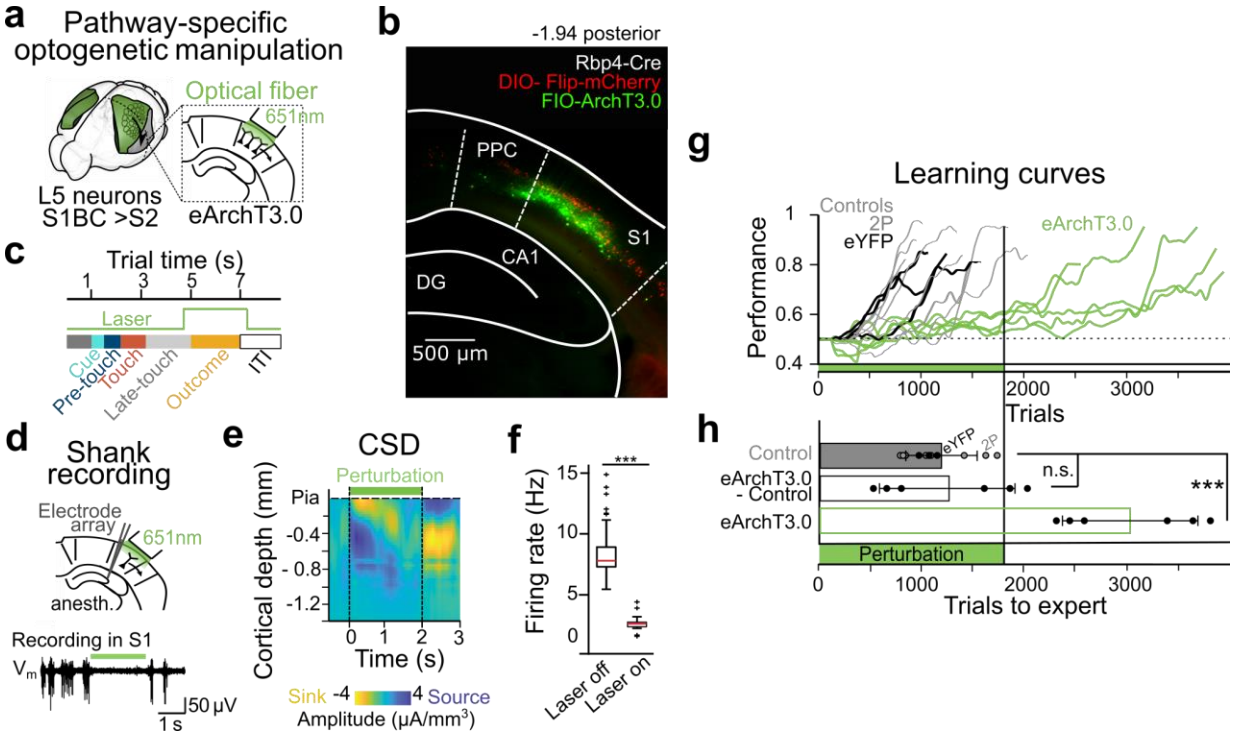


Fig. 4 | Optogenetic suppression of reward representations prevents learning. a, Pathway-specific optogenetic manipulation of S1→S2 L5 neurons using eArchT3.0. **b,** Dense labelling of eArchT3.0 in L5 neurons (confocal image, coronal view, 100 µm). **c**, Trail structure and timing of optogenetic perturbation. **d**, *Top:* Schematic of optogenetic perturbation with simultaneous electrode shank recording in isoflurane-anesthetized mice. *Bottom:* Voltage recording of example unit in S1 BC. **e**, Average current source density across S1 around optogenetic perturbation (isoflurane aesthesia, 3 sessions, 2 mice, 300 trials) ***f,*** Firing rate of L5 units with and without optogenetic manipulation (red line:median, box: 25th and 75th percentile, whiskers: 5th and 95th percentile, outliers: crosses; 5 mW, 43 units, 600 trials, 6 sessions, n = 4 mice, t-test p<0.0001). **g**, Optogenetic perturbation during a texture-discrimination task was carried out for 1'800 trials. Learning curves of eArchT3.0-expressing mice (n = 5) and control mice (from GCaMP6f experiments, n = 8; and YFP controls, n = 3)). **h,** Average number of trials required to reach expert performance in control mice (YFP and two-photon), optogenetically perturbed eArchT mice, and for perturbed mice after subtraction of the 1'800 perturbed trials (1'150 ± 314, 3'019 ± 659, and 1'219 ± 659 trials; mean ± s.d.; p<0.001 for perturbed ArchT-expressing mice vs. control mice; p = 0.89 for corrected perturbed ArchT-expressing mice vs. control mice; Wilcoxon ranksum test).

change during learning. We interpret these classes as arising from salient input streams converging on the apical tufts, which adapt their relative strengths as animals gain task proficiency. Dendrites in the sensory class may receive touch-specific inputs conveyed via thalamus[16,17] but also contextual or anticipatory inputs from posterior association areas[9,18] and anterior motor and premotor areas[7,19,20]. Strengthening of these responses around the time of whisker touch is in line with the recently reported enhancement of behaviourally relevant sensory representations during learning[21]. Interestingly, texture discrimination remained low even in experts. L5 neurons in S1 thus may predominantly represent the presence and saliency of relevant tactile inputs and discriminate choice rather than stimulus identity[22]. Dendrites in the reward class may receive feedback input about behavioural outcome, in particular prediction errors. Dendritic representations of reward in L5 tufts of barrel cortex have been described previously[10], for both unexpected random rewards and rewards delivered in a behavioural trial structure. In our study, reward representations became less pronounced during learning. The learning-related changes in both classes might be due to local plasticity in the tuft itself or in the projections from upstream regions, such as S2[23], postrhinal [9], premotor[20], or orbitofrontal[24] areas. Furthermore, alterations in the

engagement of local interneurons in superficial layers of cortex[25] and neuromodulation of dendritic processing[26,27] could contribute to the weakening of reward representation and the strengthening of the representation of the behaviourally-relevant sensory stimuli. In how far functional classes of tufts correspond to anatomical subdivisions of L5 neurons[28] remains unclear and requires further study.

The existence and relevance of local dendritic branch activations in L5 tufts *in vivo* has recently been debated[14,15,29,30]. Here, we reliably detected local dendritic events in about 8% of all trials. We assume that LE are caused by spatially-restricted synaptic inputs leading to postsynaptic potentials that can cause local regenerative events but fail to invade the whole tuft and initiate calcium spikes at the dendritic nexus. The occurrence of localized calcium events in tuft branches could thus reflect strengthening of such inputs. In addition, LE may directly contribute to local plasticity, as suggested by experiments during motor learning[31]. Finally, our demonstration that inhibition of apical dendritic reward signals suppresses the animal's ability to learn, adds to findings that dendrites are crucial for sensorimotor processing[6,8] and lends support to their hypothesized role as subcellular sites of credit assignment[32]. The flexible integration of diverse information streams in the dendritic tufts of cortical pyramidal neurons provides a powerful means to facilitate dynamic computations in service of adaptive behaviour.

Author contribution

G.S. and F.H. conceived the project and designed the study; G.S. carried out all awake *in vivo* experiments; C.L. performed extracellular recordings in anesthetized mice; G.S. and S.K. analyzed data; S.K. developed transformation clustering; A.M.R. performed iDISCO clearing of brain hemispheres; P.B. and A.M.R. performed light-sheet microscopy on cleared brains; A.A., V.M., and F.H. supervised experiments and analysis; G.S. and F.H. wrote the manuscript with comments from all authors.

Acknowledgement

This work is supported by grants from the Swiss National Science Foundation (project grant 310030B_170269 and Sinergia grant CRSII5_180316 to F.H., 179040 to A.A, Award PP00P3-157529 to V.M..), the European Research Council (ERC Advanced Grant BRAINCOMPATH, project 670757, to F.H.), the University Research Priority Program (URPP) 'Adaptive Brain Circuits in Development and Learning' (AdaBD) of the University of Zurich (to F.H. and V.M.), a Forschungskredit from the University of Zurich (project K-41220-04, C.L.), NOMIS Distinguished Scientist Award to A.A., and Simons Foundation (SCGB 328189 and 543013 to V.M.). The authors thank Henry Lütcke for his support with the CaImAn pipeline, Ladan Egolf for managing transgenic mouse lines, and Dubravka Dujmovic-Göckeritz for genotyping and CLARITY clearing of whole brains.

Supplementary Information is available for this paper.

References

1 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).

2 Chen, J. L. *et al.* Pathway-specific reorganization of projection neurons in somatosensory cortex during learning. *Nat Neurosci* **18**, 1101-1108, doi:10.1038/nn.4046 (2015).

3 Magee, J. C. & Grienberger, C. Synaptic Plasticity Forms and Functions. *Annu Rev Neurosci* **43**, 95-117, doi:10.1146/annurev-neuro-090919-022842 (2020).

4 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).

5 Poirazi, P. & Papoutsi, A. Illuminating dendritic function with computational models. *Nat Rev Neurosci* **21**, 303-321, doi:10.1038/s41583-020-0301-7 (2020).

6 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).

7 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).

8 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).

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

10 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).

11 Grewe, B. F., Voigt, F. F., van 't Hoff, M. & Helmchen, F. Fast two-layer two-photon imaging of neuronal cell populations using an electrically tunable lens. *Biomed Opt Express* **2**, 2035-2046, doi:10.1364/BOE.2.002035 (2011).

12 Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L. & Helmchen, F. Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature* **499**, 336-340, doi:10.1038/nature12236 (2013).

13 Kollmorgen, S., Hahnloser, R. H. R. & Mante, V. Nearest neighbours reveal fast and slow components of motor learning. *Nature* **577**, 526-530, doi:10.1038/s41586-019-1892-x (2020).

14 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).

15 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).

16 Gambino, F. *et al.* Sensory-evoked LTP driven by dendritic plateau potentials in vivo. *Nature* **515**, 116-119, doi:10.1038/nature13664 (2014).

17 Audette, N. J., Bernhard, S. M., Ray, A., Stewart, L. T. & Barth, A. L. Rapid Plasticity of Higher-Order Thalamocortical Inputs during Sensory Learning. *Neuron* **103**, 277-291 e274, doi:10.1016/j.neuron.2019.04.037 (2019).

18 Gilad, A. & Helmchen, F. Spatiotemporal refinement of signal flow through association cortex during learning. *Nat Commun* **11**, 1744, doi:10.1038/s41467-020-15534-z (2020).

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

20 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).

21 Benezra, S. E., Patel, K. B., Campos, C. P., Hillman, E. M. & Bruno, R. M. Learning enhances behaviorally relevant representations in apical dendrites. *bioRxiv* (2021).

22 McGuire, L. M. *et al.* Short Time-Scale Sensory Coding in S1 during Discrimination of Whisker Vibrotactile Sequences. *PLoS Biol* **14**, e1002549, doi:10.1371/journal.pbio.1002549 (2016).

23 Chen, J. L., Voigt, F. F., Javadzadeh, M., Krueppel, R. & Helmchen, F. Long-range population dynamics of anatomically defined neocortical networks. *Elife* **5**, doi:10.7554/eLife.14679 (2016).

24 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).

25 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).

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 Williams, S. R. & Fletcher, L. N. A Dendritic Substrate for the Cholinergic Control of Neocortical Output Neurons. *Neuron* **101**, 486-499 e484, doi:10.1016/j.neuron.2018.11.035 (2019).

28 Takahashi, N. *et al.* Active dendritic currents gate descending cortical outputs in perception. *Nat Neurosci* **23**, 1277-1285, doi:10.1038/s41593-020-0677-8 (2020).

29 Francioni, V. & Harnett, M. T. Rethinking Single Neuron Electrical Compartmentalization: Dendritic Contributions to Network Computation In Vivo. *Neuroscience*, doi:10.1016/j.neuroscience.2021.05.038 (2021).

30 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).

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

32 Richards, B. A. *et al.* A deep learning framework for neuroscience. *Nat Neurosci* **22**, 1761-1770, doi:10.1038/s41593-019-0520-2 (2019).

Methods

All experimental procedures were carried out in accordance with the guidelines of the Federal Veterinary Office of Switzerland and were approved by the Cantonal Veterinary Office in Zurich under license number 234/2018.

Animals and preparations for chronic imaging. We used male and female adult 6-10 week old Rbp4-Cre transgenic mice (n = 8, Tg(Rbp4-cre)KL100Gsat/Mmucd, MGI:4367068, ref. 33,34). For surgical preparation, mice were anesthetized using isoflurane (1.5-2% in O_2) and the body temperature was maintained at 37°C using a heating pad with rectal probe. After exposing the skull, a 4-mm diameter craniotomy was made above the left S1 barrel cortex and S2. Stereotactic injections of AAVretro-hSyn1-chI-FLEX-mCherry_2A_NLS_FLPo virus solution (6.3 x 10^{12} vg/ml, dilution 1:50) was injected into S2 (three injections à 210 nl; AP|ML|DV coordinates from bregma (in mm): -0.7|3.5|-1, -1.2|4.1|-1.5, -1.3|4.5|-1.5). AAV-2.1-hSyn1-fio-GCaMP6f virus solution (1.8 x 10^{12} vg/ml) was injected into L5 of barrel cortex (three injections à 210 nl: -0.7|-3|-0.6, -1.1|-3|-0.5, -1.1|-2.4|-0.6). The craniotomy was sealed with a 4-mm glass cover slip and dental cement (Tetric EvoFlow). A light-weight head-post was fixed on the skull using dental cement. For the 3 days following the surgery, animals were monitored and analgesics (Metacam, 5 mg/kg, s.c.) and antibiotics (Baytril, 10 mg/kg, s.c.) were administered. Animal handling began 5 days after surgery and the first imaging session took place >21 days after virus injection.

Behavioural task and mouse training. The setup for the go/no-go texture discrimination task has been described previously[12,35]. Each trial started with the opening of the laser shutter (Thorlabs, SH05/M) followed after 1 s by an auditory tone (two 2-kHz beeps of 100-ms duration with 50-ms interval). Then, either the rough or smooth texture (P100/P1200 sandpapers) was moved for 2 s towards the whiskers on the right side of the animal's snout. We presented the two texture types randomly but with no more than 3 repetitions. In expert mice, which typically show anticipatory whisking, the first texture-whisker touch typically occurs around 0.5 s before the texture stops[35]. After a 2-s stimulus presentation period the texture was retracted and an auditory tone (4 beeps of 4 kHz; 50-ms duration with 25-ms intervals) signalled the start of the 2-s response period. A water reward was given when the mouse licked in the outcome window after the presentation of the go texture ('Hit'). The first lick during the outcome window triggered the feedback. Licks during the late-touch window were not punished but ignored. A white noise punishment was given for licking in the outcome window for the no-go texture ('False alarm', FA). When the mouse withheld licking after the presentation of the go texture ('Miss') or the no-go texture ('Correct rejection') neither reward nor punishment was given. In the first training session, the identities of go and no-go textures were randomly assigned to the animal and maintained for the whole experiment (go texture: P100 in 3 mice and P1200 in 5 mice).

Animals were kept on a reversed light/dark cycle. After accustoming the mice to the experimenter, habituation to head-immobilization began. We increased head-restraining time with every training session, carrying out two session per day. Mice were water scheduled for behavioural training once they sat quietly for >2 min and were introduced to the experimental setup. Weight, health and water intake were monitored daily. During the first two sessions in the setup, mice only received water reward (~5 µl per repetition). In session 3 and 4 the go-texture presentation was introduced and an automatic water reward was given, to form an association between texture and reward. Once the mice were able to trigger the water reward autonomously, the no-go texture was introduced starting from presentation in 1% of the cases and gradually increasing to 50%. The first imaging session was scheduled when mice licked consistently for both the textures. Imaging sessions were carried out once per day per animal and lasted as long as a mouse actively engaged in the task (63-209 trials per session). For the first 3-5 imaging sessions go and no-go textures were presented each in 50% of the trials. Thereafter, to facilitate learning, presentation of the no-go texture was repeated in trials following an error trial (false alarm or miss). This 'repeat-incorrect' strategy was accounted for in the calculation of behavioural performance by considering the occurrence of the go-texture in a sliding window of 5 trials. Mice learned to differentiate the textures and showed stable expert performance (>75% correct trails) after 12-18 sessions. Performance of each animal was quantified by a state-space smoothing algorithm that provides a learning curve with confidence intervals[36]. The first expert trial and the last naïve trial were identified by an expectation maximization algorithm using a Gaussian state equation. Learning onset (i.e., the last naïve trial) was defined as the trial when the lower 95% confidence interval exceeded 50% correct responses. The first expert trial was defined as the trial, from which on onwards the performance of the animal exceeded chance level with 95% confidence. For analysis of the learning process of all mice, we aligned learning curves to the first expert trial and used a time window of 500 trials before and 150 trials after the first expert trial (trial ID -500 to 150).

Recording of licking and whisking behaviour. Using a 950-nm infrared LED, whisker motion was imaged during the trial at 40 Hz using a high-speed CMOS camera (A504k, Basler). The average whisking angle across all whiskers was analyzed from the videos using a whisker tracking software[37]. The whisker envelope was extracted as the difference between the maximum and minimum whisker angle using the Matlab function *envelope*. The estimated time point of the first touch between whisker and texture was obtained by calculating the time of the average whisker envelope maximum within the pre-touch and touch window across one session. Licking was estimated based on the event rate from the capacitive lick sensor sampled at 100 Hz. The lick rate was calculated based on the number of lick events in a 200-ms sliding window, assuming that an average lick event lasts 4 ms.

Two-photon calcium imaging. In vivo awake calcium imaging was performed using a custom-built two-photon microscope equipped with a Ti:sapphire laser system (Chameleon Ultra, Coherent), a water-immersion objective (CFI LWD 16X/, 0.8 NA; Olympus), a custom-built scanner unit with a 4-kHz resonance scan mirror (CRS 4KHz, Cambridge Technology) and a galvometric mirror (6220H, Cambridge Technology), a Pockel's Cell (Model 350-80-LA-02, Conoptics, Danbury, CT) and a hybrid-photodetector (HPDs, R11322U-40 MOD, Hamamatsu). The microscope was controlled by the custom-written software Scope[23] (http://sourceforge.net). An electrically tunable lens (ETL; Optotune EL-10-30-TC, Optotune AG, Zurich, CH; with an plano-concave offset lens, f = -100 mm, Qioptiq) was imaged on the scan mirrors using a 1:1 telescope of f= 100 mm lenses (AC254-100-B-ML, Thorlabs). For initial identification of GCaMP6f-positive neurons, a volume stack was acquired using 800-nm excitation and a green emission filter (510 ± 42 nm bandpass). For calcium imaging, GCaMP6f was excited at 920 nm. Four imaging planes were identified per animal, spanning from close to the pia mater to below the nexus of L5 tufts (approx. -30 µm to -370 µm). Images were acquired at 10 Hz with 508x168 pixel resolution resulting in a 230 µm x 230 µm field of view. Laser power was adjusted per plane ranging from 10 to 65 mW under the objective. Single trials of >7 s duration were recorded with 4-s inter-trial intervals.

Optogenetic silencing. To transiently suppress dendritic tuft activity of L5 pyramidal neurons in barrel cortex during reward delivery (in the outcome window), we used the same surgical procedure of virus injection and window implantation as described for the calcium imaging experiments. In five mice, we made three injections of undiluted AAVretro-hSyn1-chI-FLEX-mCherry_2A_NLS_FLPo virus solution (6.3 x 10^{12} vg/ml) into S2 and three injections of AAV-1/2-hSyn-chl-dFRT-eArchT3.0_EYFP-dFRT virus solution (5.3 x 10^{12} vg/ml) in S1 barrel cortex (coordinates and volumes as described above). In three additional mice we expressed eYFP (AAV5-EF1a-fDIO-EYFP_WPRE, 4.9 x 10^{12} vg/ml) instead of eArchT3.0 for control. After the implantation of the glass window, a ferrule holding an optical fiber (910 µm) was positioned and secured in place with dental cement above the window centered over barrel cortex. Animal handling and training was carried out as described above. Once mice reliably licked for water in the experimental setup, 561-nm green laser light (5 mW, CW laser Coherent OBIS-561-50 LS) was delivered through the optical fiber in 100% of the trials. The perturbation only occurred during the outcome window of the trial, lasting 2.4 s (4.8 - 7.2 s in trial time). After 1800 trials of laser perturbation, optogenetic silencing was stopped. The optical fiber transmitting the laser light to the behavioural setup was detached from the ferrule placed above the craniotomy. This change preserved similar light conditions and allowed the mouse to behave and learn without the optogenetic manipulation. Experiments were stopped after 8 weeks of experimentation in accordance with our animal licence. For mice that did not reach expert performance levels, but performed above chance levels at this time, the last recorded trial was considered as their first expert trial. Licking behaviour was constantly recorded during the whole

experiment. To determine the effect of optogenetic perturbation on whisking behaviour, we connected the optical fiber to the ferrule in expert mice and applied laser illumination in 50% of trials. We analyzed whisking behaviour as described above and the conditions with and compared without manipulation.

***In vivo* electrophysiological recordings.** To validate the effect of optogenetic perturbation we performed acute *in vivo* recordings in lightly anesthetized mice (n = 3) expressing eArchT3.0 selectively in S1→S2 L5 neurons. At the start of validation experiments, animals were anesthetised with isoflurane (2% for induction and <1.5% during recording), and their body temperature was maintained at 37°C using a heating pad. A small craniotomy (<1 mm diameter) was performed over the area of virus injection in barrel cortex and the brain was covered with silicon oil. A silver wire was placed in contact with the cerebrospinal fluid through a small (0.5 mm) trepanation over the cerebellum to serve as reference electrode. A silicon probe (Atlas Neurotechnologies, 32-contact linear array with 50 µm inter-contact spacing) was inserted into the left cortical hemisphere. The top-most electrode was left in contact with the surface of the brain under visual guidance, to ensure that the probe covered the entire cortical column including superficial L1. A fiber optic cannula was positioned to deliver laser light (561 nm, 5 mW) to the surface of the brain just adjacent to the silicon probe, but not inserted into the brain. After positioning of the silicon probe and cannula, the preparation was left for 30 min to allow the brain and electrode to stabilise. After stabilisation, the broadband voltage was amplified and digitally sampled at a rate of 30 kHz using a commercial extracellular recording system (RHD2000, Intan Technologies). Spontaneous activity was recorded over 1-1.5h long recording sessions divided into trials (7-s duration, laser on for 2 s) separated by 1-s inter-trial intervals, mimicking the awake optogenetic experiments. The raw voltage traces were processed offline using fourth-order Butterworth filters to separate the local field potential (< 400 Hz lowpass filter) and the multi-unit activity (MUA; bandpass filter 0.46-6 kHz). Subsequently, the local field potential was used to compute the current source density in order to localize currents arising from the optogenetic stimulation. The high-pass data were thresholded at 5.5 times the standard deviation across the recording session and the numbers of spikes in windows of interest were counted. In order to combine data across mice, the activity at sites with clear MUA was expressed in percent of the baseline value, i.e. the average spike rate during the period without laser illumination.

Cleared tissue light-sheet microscopy. Two mice were injected with retrograde AAVretro-hSyn1-chI-FLEX-mCherry_2A_NLS_FLPo virus in S2 and AAV5-EF1a-fDIO-EYFP_WPRE virus in S1 (for details see above) and their brains were cleared using the CLARITY protocol[38,39]. In brief, after 4 weeks of expression, mice were perfused and the brains post-fixed for 48 hours in a hydrogel solution (1% paraformaldehyde, 4% acrylamide, 0.05% bis-acrylamide, 0.25% VA044) before the

hydrogel polymerization was induced at 37°C. Then the brains were placed in 40 ml of 8% SDS at room temperature (RT) for approx. 25 days. The brains were put into a refractive index matching solution (RIMS) and equilibrated for 1 day before imaging.

We visualized sparsely GCaMP6f-labelled S1→S2 L5 neurons after clearing brain hemispheres (n = 2) with a custom iDISCO protocol[40]. After 4 weeks of expression, mice were perfused and the brains post-fixed in 4% PFA in PBS for 4.5 hours at 4°C, shaking at 40 rpm. Brain hemispheres were washed in PBS for 3 days at RT and 40 rpm, with daily solution exchange. Samples were dehydrated in serial incubations of 20%, 40%, 60%, 80% methanol (MeOH) in ddH_2O, followed by 2 times 100% MeOH, each for 1 hour at RT and 40 rpm. Pre-clearing was performed in 33% MeOH in dichloromethane (DCM) overnight (o.n.) at RT and 40 rpm. After 2 times washing in 100% MeOH each for 1 hour at RT and then 4°C at 40 rpm, bleaching was performed in 5% hydrogen peroxide in MeOH for 20 hours at 4°C and 40 rpm. Samples were rehydrated in serial incubations of 80%, 60%, 40%, and 20% MeOH in in ddH_2O, followed by PBS, each for 1 hour at RT and 40 rpm. Permeabilization was performed by incubating the mouse hemispheres 2 times in 0.2% TritonX-100 in PBS, each for 1 hour at RT and 40 rpm, followed by incubation in 0.2% TritonX-100 + 10% dimethyl sulfoxide (DMSO) + 2.3% glycine + 0.1% sodium azide (NaN_3) in PBS for 3 days at 37°C and 65 rpm. Blocking was performed in 0.2% Tween-20 + 0.1% heparine (10 mg/ml) + 5% DMSO + 6% donkey serum in PBS for 2 days at 37°C and 65 rpm. Samples were stained gradually with primary polyclonal chicken-anti-GFP antibody (Aves Labs, GFP-1020) and secondary donkey-anti-chicken-AlexaFluor488 antibody (Jackson ImmunoResearch, 703-545-155) 1:400 in 0.2% Tween-20 + 0.1% heparine + 5% DMSO + 0.1% NaN_3 in PBS (staining buffer) in a total volume of 1.5 ml per sample every week for 4 weeks at 37°C and 65 rpm. Washing steps were performed in staining buffer 5 times each for 1 hour, and then for 2 days at RT and 40 rpm. Clearing was started by dehydrating the samples in serial MeOH incubations as described above. Delipidation was performed in 33% MeOH in DCM o.n. at RT and 40 rpm, followed by 2 times 100% DCM each for 30 minutes at RT and 40 rpm. Refractive index (RI) matching was achieved in dibenzyl ether (DBE, RI = 1.56) for 4 hours at RT.

3D stacks of cleared brains and hemispheres were acquired using a mesoSPIM light-sheet microscope[41] (www.mesospim.org). Imaging data were post-processed using custom-written routines in MATLAB. To visualize neurons, local contrast enhancement was performed per slice by subtracting a Gaussian-smoothed version of the slice (4σ). Barrels were visible in the green autofluorescence channel. An anatomical barrel map was fitted to the barrel autofluorescence using the MATLAB functions *cpselect* and *fitgotrans*. 3D volume projection was performed using Imaris (9.8.0, Oxford Instruments).

Confocal histology. After the last awake imaging session mice were administered a lethal dose of pentobarbital (Ekonarcon, Streuli) and transcardially perfused with sterile NaCl (0.9%) followed by

4% paraformaldehyde (PFA, 0.1 M phosphate buffer, pH 7.4). From 100-µm thick coronal brain slices we acquired histological images with a confocal laser-scanning microscope (Olympus FV1000). Coronal sections were registered to Paxinos and Franklin's mouse brain atlas using manually set landmarks using *cpselect* (MATLAB) and aligning the atlas via *fitgeotrans* (MATLAB).

Morphological reconstructions. Anatomical two-photon image stacks of all field of views were aquired before behavioural training using the two-photon microscope with 800-nm laser excitation. 3D reconstructions of imaged dendritic tufts were obtained using the semi-manual interpolation option of the VolumeSegmenter app in MATLAB. Tuft membership was determined based on the morphological stacks as well as on high correlation of calcium signals. In 42 neurons, the trunk could be clearly identified together with its corresponding daughter tuft branches. The comparison of functional class within dendritic tufts as well as the LE/GE analysis were performed on this subset of neurons.

Preprocessing and visualization of calcium imaging data. Motion correction of the acquired movies of GCaMP6f fluorescence was carried out by a custom-written Python pipeline using the NoRMCorre algorithm for non-rigid artefact correction provided by CaImAn[42]. Single non-overlapping dendritic branches were identified and regions of interest (ROIs) were defined manually for each session. A consistent nomenclature was used to identify the same dendritic branches over consecutive sessions. If multiple ROIs along the same branch were identified, only the ROI closest to the trunk was used for further analysis. Calcium indicator fluorescence signals were extracted using custom software routines written in MATLAB (Mathworks). Background fluorescence was estimated in a background ROI as the bottom 1st percentile fluorescence signal across the entire session and subtracted before calculating the relative percentage change of fluorescence from baseline $\Delta F/F = (F-F_0)/F_0$. Baseline fluorescence F_0 was computed as 51st percentile of the fluorescence signal in a 4-s sliding window. ΔF/F traces were smoothed with a 5-point 1st-order Savitsky-Golay filter. Upon visual inspection, we manually excluded calcium traces with obvious artefacts such as motion-induced artefacts, light reflections from the texture, or non-physiological calcium traces. All remaining data were visualized using an Euclidian-distance based UMAP embedding (UMAP embedding 1; **Supplementary Fig. 4**) with a neighbourhood size of 100 data points in the high-dimensional space. For further analyses, detectable transients were defined as fluorescence signals that deviated from baseline by >5.5 standard deviations. The set of trials with detectable ΔF/F transients was visualized using a correlation metric-based UMAP embedding (UMAP embedding 2) and a neighbourhood size of 30 data points in the high-dimensional space. Analysis and data exploration were carried out using custom-written MATLAB code.

To determine the onsets of calcium transients, we found for every threshold-crossing event the peak position in a given trial (highest peak found by MATLAB function *findpeaks* with minimal distance

between peaks of 1 s and >25% ΔF/F peak prominence). Local maxima were not included in further analysis. Based on the determined primary transient peak, the calcium transient onset time point was defined as the minimum of the first derivative of the ΔF/F trace until 1 s prior to the detected peak. Transients with their peak position within the start window (0-1 s of trial time) were not considered as their onset likely occurred during the inter-trial interval. Density maps of calcium transient onsets were derived by calculating the distribution of onset time points in 50-trials bins of from naïve to expert condition.

Functional co-evolution of dendritic signals and transformation clustering. To assess learning-related functional changes of trial-related calcium traces in individual dendritic branches or trunk ROIs we employed the custom-developed approach of "transformation clustering" that is inspired by earlier work using nearest neighbour graphs to understand high dimensional data[13]. Functional changes across the learning time course in one dendritic branch were compared to the changes of any other dendritic branch in our dataset as follows:

Let d and d' denote a pair of dendritic ROIs with trial responses x_i^d and $x_i^{d'} \in \mathbb{R}^{70}$ (7-s recording at 10 Hz), where i denotes the index of the trial (the trial ID) in which the response was recorded. Note that many trials contain no detectable transients (as defined above). Let $i_1, \dots, i_k$ denote the indices of the trials in which dendritic ROI d shows a detectable transient. Let $NN^k(d,\ i,\ d')$ denote the set of k nearest neighbours of the trial response x_i^d among all responses $x_i^{d'}$ for dendritic ROI d' that show detectable transients. Let $NN_{tID}^k(d,\ i,\ d')$ denote the set of trial indices of the nearest neighbours of the trial response x_i^d among the activations for ROI d':

$$NN_{tID}^k(d,\ i,\ d') = \{j \mid \mathrm{x}_\mathrm{j}^{\mathrm{d}'} \in NN^k(d,\ i,\ d')\} \tag{1}$$

Finally, let $<NN_{tID}^k(d,\ i,\ d')>$ denote the average of those trial indices. We define the similarity of d and d' as the correlation coefficient between the two vectors $[i_1, \dots, i_k]$ and $[<NN_{tID}^k(d,\ i_1,\ d')>, \dots, <NN_{tID}^k(d,\ i_k,\ d')>]$. We refer to this correlation coefficient as $Tcc_{dd'}$ (transformation correlation coefficient). To assess significance of the correlation, we compared the actual $Tcc_{dd'}$ to a null hypothesis derived by shuffling trial IDs for dendritic ROI *d'*. Shuffling removes the long term temporal relationship between ROIs d and d'. We define the corrected transformation correlation coefficient, $CT_{dd'}$, as the inverse percentile of $Tcc_{dd'}$ with respect to the null hypothesis distribution generated by many shuffle iterations (e.g. 1'000).

Let $\mathbf{CT}$ denote the square matrix of corrected transformation correlation coefficients for all pairs of dendritic ROIs *d* and *d'*. We compute a symmetric dissimilarity matrix, $\mathbf{D}$, through $\mathbf{D} = -(\mathbf{CT} + \mathbf{CT}^{\mathrm{T}})$, where $\mathbf{CT}^{\mathrm{T}}$ denotes the matrix transpose of $\mathbf{CT}$. We applied hierarchical link clustering to extract transformation clusters. We achieved comparable results using k-means clustering after applying multidimensional scaling to the dissimilarity matrix $\mathbf{D}$.

Note that only dendritic ROIs with more than 40 ∆F/F traces with detectable calcium transients spanning a range of more than 500 trial IDs were included in the initial transformation clustering analysis. The results from this initial clustering were greedily expanded in order to assign cluster IDs also to dendritic branches and apical trunks that had only 6-39 trials with detectable calcium transients (independent of the range of recorded trial IDs). For these dendritic ROIs we calculated corrected transformation correlation coefficients with all other previously classified dendritic ROIs. The 5 previously classified ROIs with highest coefficients values (corresponding to the smallest dissimilarities) were selected and the predominant cluster ID of this set was as assigned to the unclassified ROI.

Detection of local and global events. In the analysis of local events (LE) and global events (GE) we included the 42 neurons with identified trunk ROI. ∆F/F traces in all ROIs of a tuft were compared by binarizing the aligned ΔF/F traces such that all time bins with a calcium transient peak were assigned a '1' and all other time bins zeros, forming a binary matrix with dimensions ROIs × frames. Trials without any detectable transient were labelled 'no event' trials. Trials with any detectable event in the trunk, irrespective of the activity in the tuft branches, were classified as global event (GE) trials. Transients occurring in the tuft branches in a 2-s time window around the trunk event (-1 to +1 s) were considered to be related to the GE. If no detectable transient occurred in the trunk but one or several tuft branches showed transients (with peak times not spread over more than 2 s) this trial was labelled as a LE trial. If several ROIs showed calcium transients during a LE, the ROI with the largest transient amplitude was selected for further analysis. A 'mixed trial' contained a GE and a LE, with the GE peak (in the trunk ROI) and the LE peak (in the tuft ROI with the largest transient amplitude) separated by more than 1 s. All detected LE and GE were double-checked and manually approved. Event probabilities were calculated depending on various parameters, such as tuft cluster identity, trial type and trial window. The numbers of GE or LE events in 16-trial bins were determined and the probability was calculated as the number of events per tuft and per trial. For the statistical analysis, values were averaged for 4 training periods ('naïve': trial ID -500 to -351, 'learning 1': -350 to -176; 'learning 2': -175 to -1; and 'expert': 0 to 150) followed by one-way ANOVAs. The average number of trials with events across tufts was calculated by summing up the number of GE and LE per trial window per period and significance tested with one-way ANOVAs.

Discrimination analysis for apical trunks. Analysis of the trial type discriminability of GE was carried out for 42 trunks. We tested the discrimination power of the trunk population in two trial time windows and in the 4 training periods ('naïve', 'learning 1', 'learning 2', and 'expert'). The first trial time window included pre-touch, touch and a fraction of the late-touch window (2.3-3.5 s in trial time) and the second window was defined around the outcome period (4.9-6.9 s in trial time). We calculated discriminability of Hit vs. CR, Hit vs. FA, and FA vs. CR trials for the tuft pool by determining the number of detectable transients per window and per training period. The receiver-

operating characteristics (ROC) curve was calculated and the area under the curve (AUC) determined using the MATLAB function *perfcurve*. The 95% confidence interval was obtained by shuffling trial type labels 5'400 times (100 times per trunk) and calculating the 95th and 5th percentile of the shuffled data set per bin.

Methods References

33 Gong, S. *et al.* Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J Neurosci* **27**, 9817-9823, doi:10.1523/JNEUROSCI.2707-07.2007 (2007).

34 Gong, S. *et al.* A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**, 917-925, doi:10.1038/nature02033 (2003).

35 Gilad, A., Gallero-Salas, Y., Groos, D. & Helmchen, F. Behavioral Strategy Determines Frontal or Posterior Location of Short-Term Memory in Neocortex. *Neuron* **99**, 814-828 e817, doi:10.1016/j.neuron.2018.07.029 (2018).

36 Smith, A. C. *et al.* Dynamic analysis of learning in behavioral experiments. *J Neurosci* **24**, 447-461, doi:10.1523/JNEUROSCI.2908-03.2004 (2004).

37 Knutsen, P. M., Derdikman, D. & Ahissar, E. Tracking whisker and head movements in unrestrained behaving rodents. *J Neurophysiol* **93**, 2294-2301, doi:10.1152/jn.00718.2004 (2005).

38 Chung, K. & Deisseroth, K. CLARITY for mapping the nervous system. *Nat Methods* **10**, 508-513, doi:10.1038/nmeth.2481 (2013).

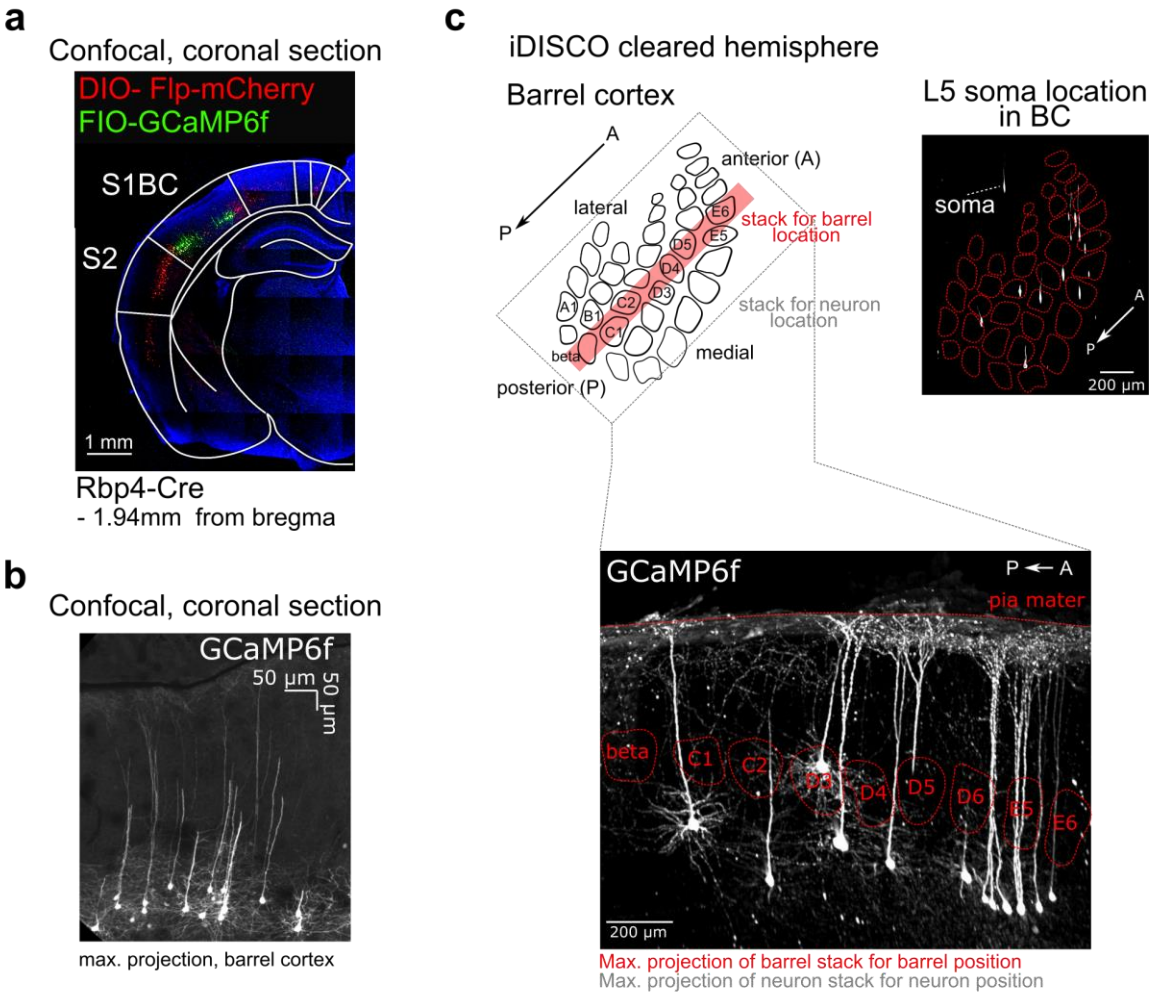
39 Yang, B. *et al.* Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158**, 945-958, doi:10.1016/j.cell.2014.07.017 (2014).

40 Renier, N. *et al.* iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* **159**, 896-910, doi:10.1016/j.cell.2014.10.010 (2014).

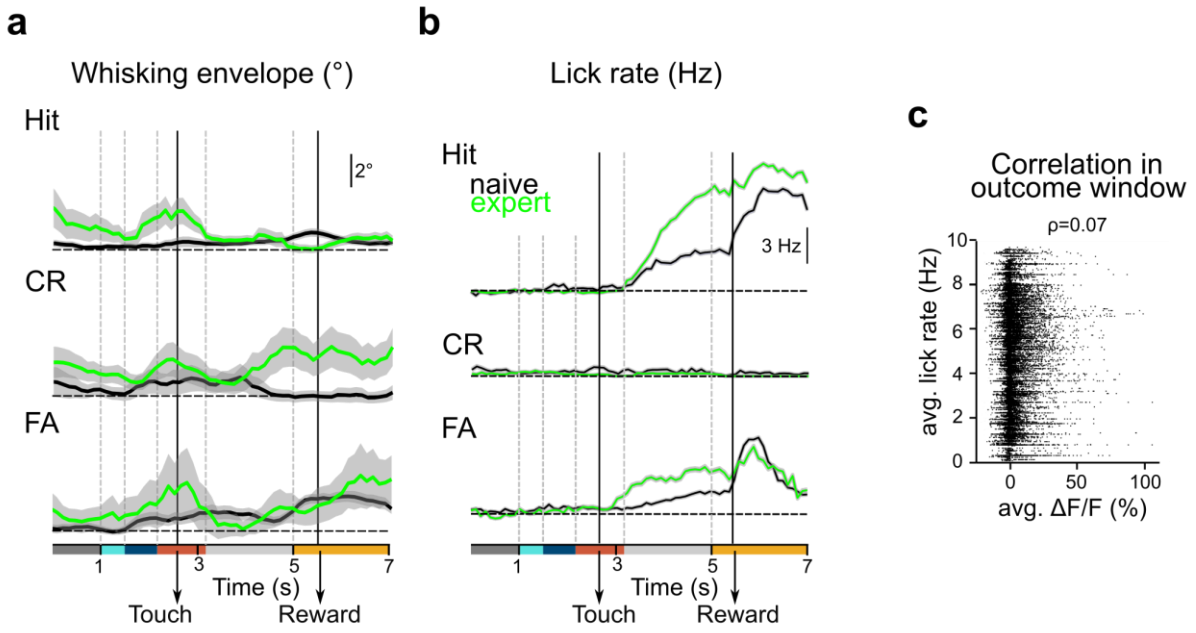
41 Voigt, F. F. *et al.* The mesoSPIM initiative: open-source light-sheet microscopes for imaging cleared tissue. *Nat Methods* **16**, 1105-1108, doi:10.1038/s41592-019-0554-0 (2019).

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

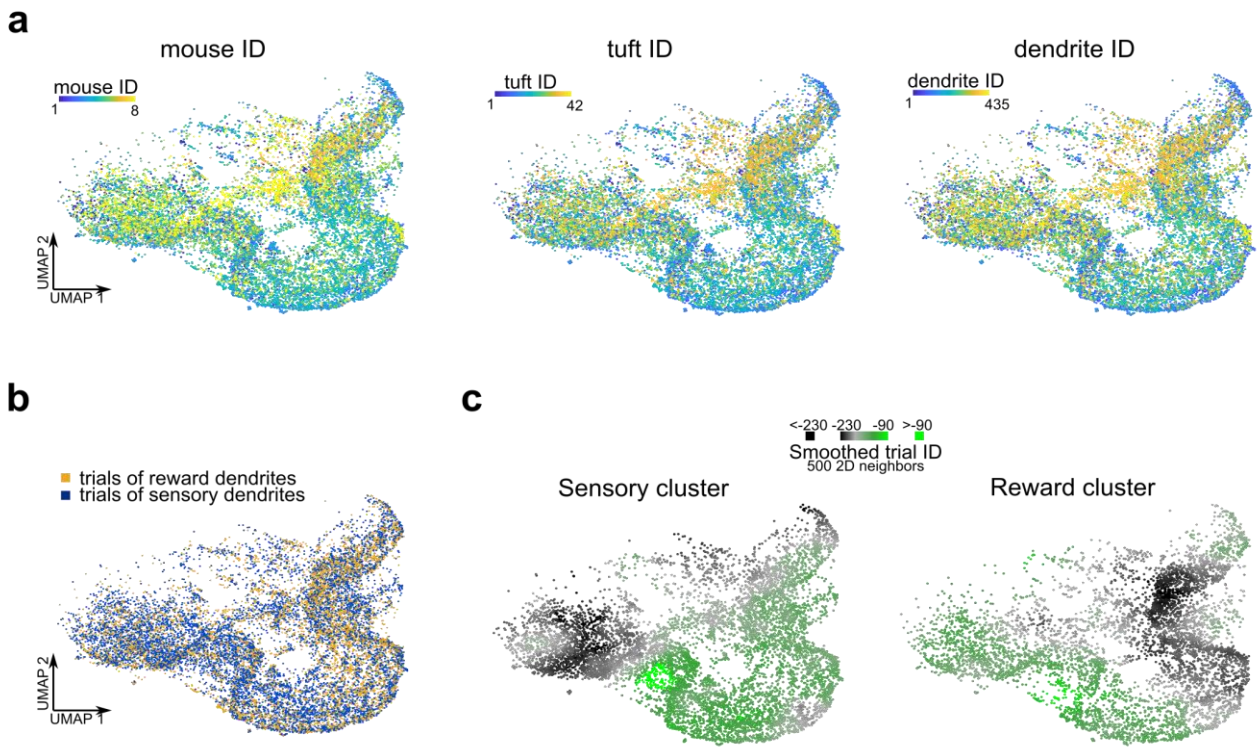
Supplementary Figures



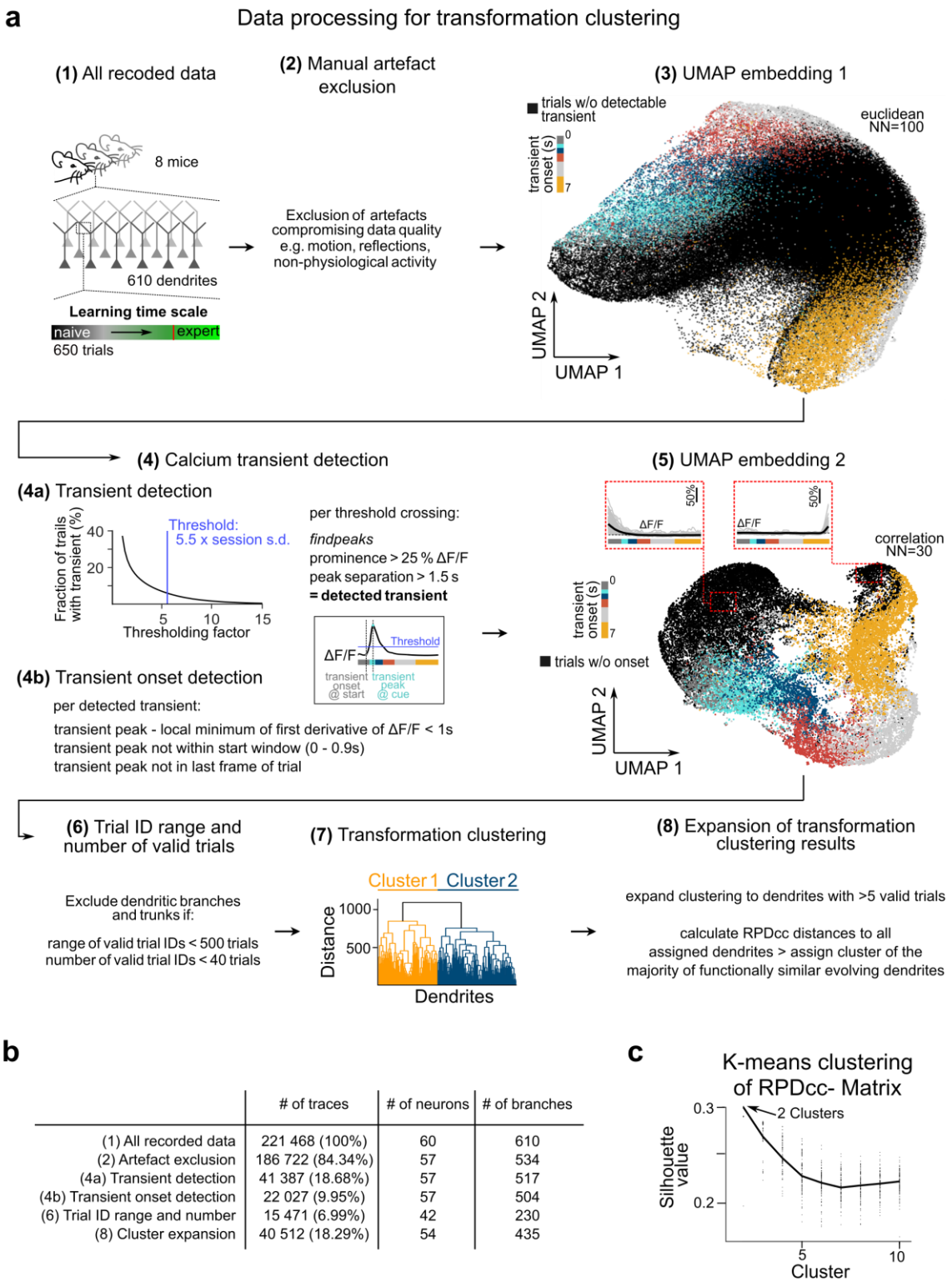
Supplementary Figure 1 | Histological characterization of imaged S1→S2 L5 pyramidal neurons. a, Coronal confocal overview image of GCAMP6f-labelled L5 pyramidal neurons in barrel cortex (green) and retrogradely-labelled mCherry-tagged neurons (red) with projections to S2. **b,** Maximum-intensity projection of a confocal image stack of sparsely labelled L5 neurons expressing GCaMP6f in S1 barrel cortex (150 µm slice). Note that most apical dendrites were cut in this slice preparation. **c,** Visualization of the location of sparsely-labelled GCaMP6f-expressing L5 neurons in barrel cortex from an iDISCO-cleared hemisphere. *Top*: Dorsal view of L5 soma location in barrel cortex, with barrel map overlaid according to L4 autofluorescence. *Bottom*: Sagittal view of L5 neurons in barrel cortex along the A/P axis. Note, the maximum-intensity projection of GCaMP6f-expressing neurons was calculated across the whole of barrel cortex (1.3 mm slice) whereas the barrel map outline was derived from a reduced stack (210 µm slice) as visualized in the schematic above (red strip).



Supplementary Figure 2 | Change of behavioural variables during learning. a, Average whisking envelope in naïve and expert mice split by trial type (6311 naïve Hit, 4952 expert Hit, 28919 naïve CR, 3811 expert CR, 10938 naïve FA and 2769 expert FA trials, 8 mice, mean ± s.e.m.). **b,** Average lick rate in naïve and expert mice split by trial type (same number for trials as in (a); 8 mice, mean ± s.e.m.). **c,** Relationship of mean lick rate and mean ΔF/F levels during the outcome trial window (41387 trials; 8 mice, ρ = 0.07).

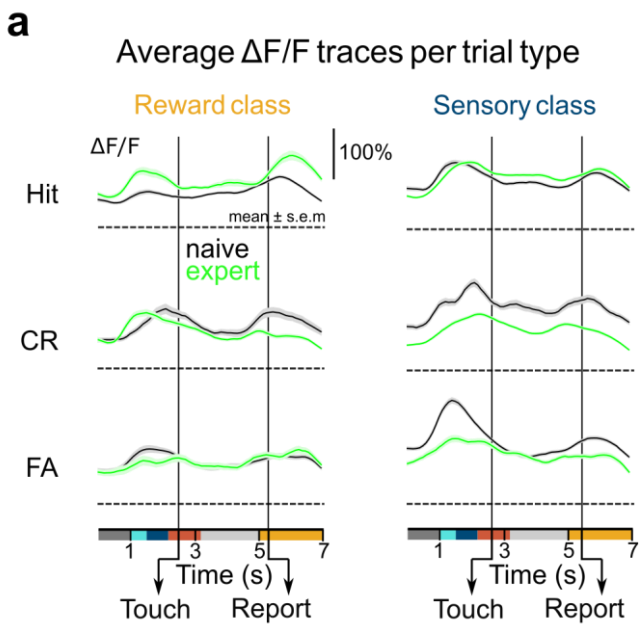


Supplementary Figure 3 | Mapping of various features of trial-related dendritic calcium signals in the UMAP space. a, UMAP embedding of all ΔF/F traces with detectable calcium transient onsets within trials coloured by mouse ID (left), tuft ID (middle), and dendrite ID (right; 22'027 data points, NN=30). **b**, UMAP embedding of ΔF/F traces of sensory and reward class coloured according to their cluster ID. Note that trials of both classes are inter-mixed within the UMAP embedding. The classes can only be separated by their temporal profile of activity changes during learning as shown in **c**, UMAP embedding of ΔF/F traces in dendrites of sensory and reward class, coloured according to smoothed trial ID.

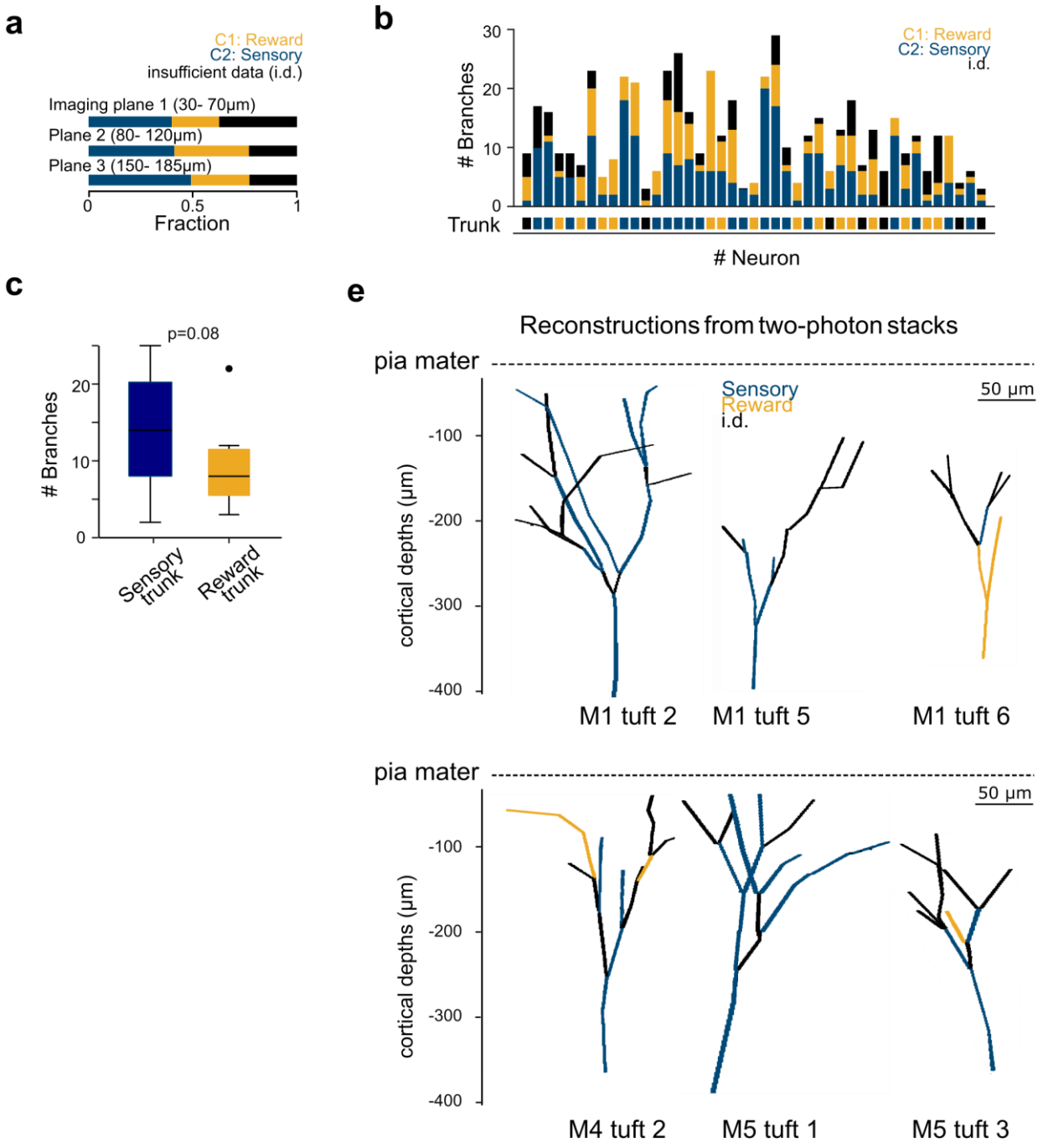


	# of traces	# of neurons	# of branches
(1) All recorded data	221 468 (100%)	60	610
(2) Artefact exclusion	186 722 (84.34%)	57	534
(4a) Transient detection	41 387 (18.68%)	57	517
(4b) Transient onset detection	22 027 (9.95%)	57	504
(6) Trial ID range and number	15 471 (6.99%)	42	230
(8) Cluster expansion	40 512 (18.29%)	54	435

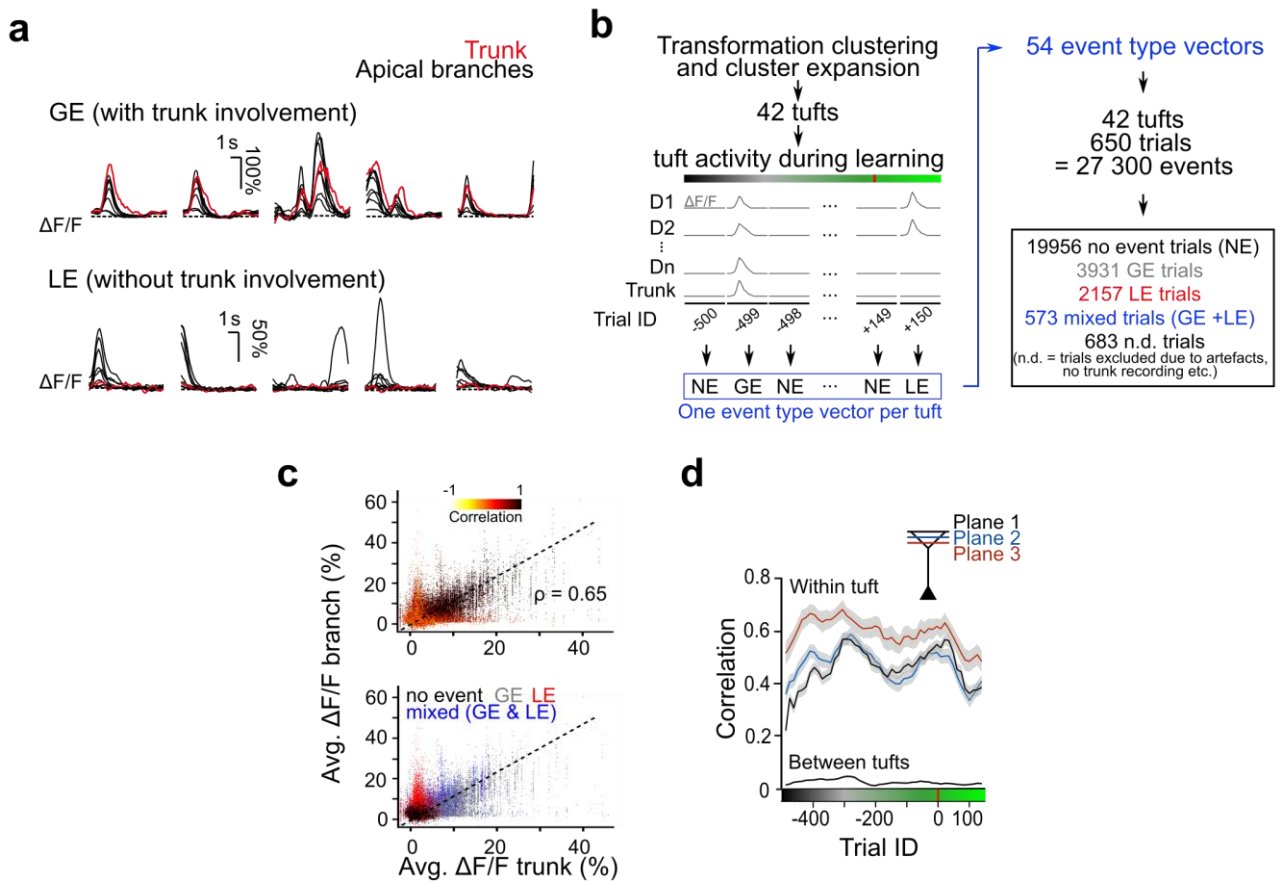
Supplementary Figure 4 | Processing pipeline for dendritic ΔF/F data. a, Flow chart for ΔF/F data processing. **b,** Number of traces, neurons and dendritic branches per processing step. Note not for every neuron a trunk was recorded. Colour in (3) and (5) are based on transient onset times (4b). **c,** Relationship between silhouette values and cluster numbers obtained as an alternative clustering approach for transformation clustering step 7 using k-means clustering. Black line indicates mean silhouette value (k = 2-10 clusters, 100 repetitions each).



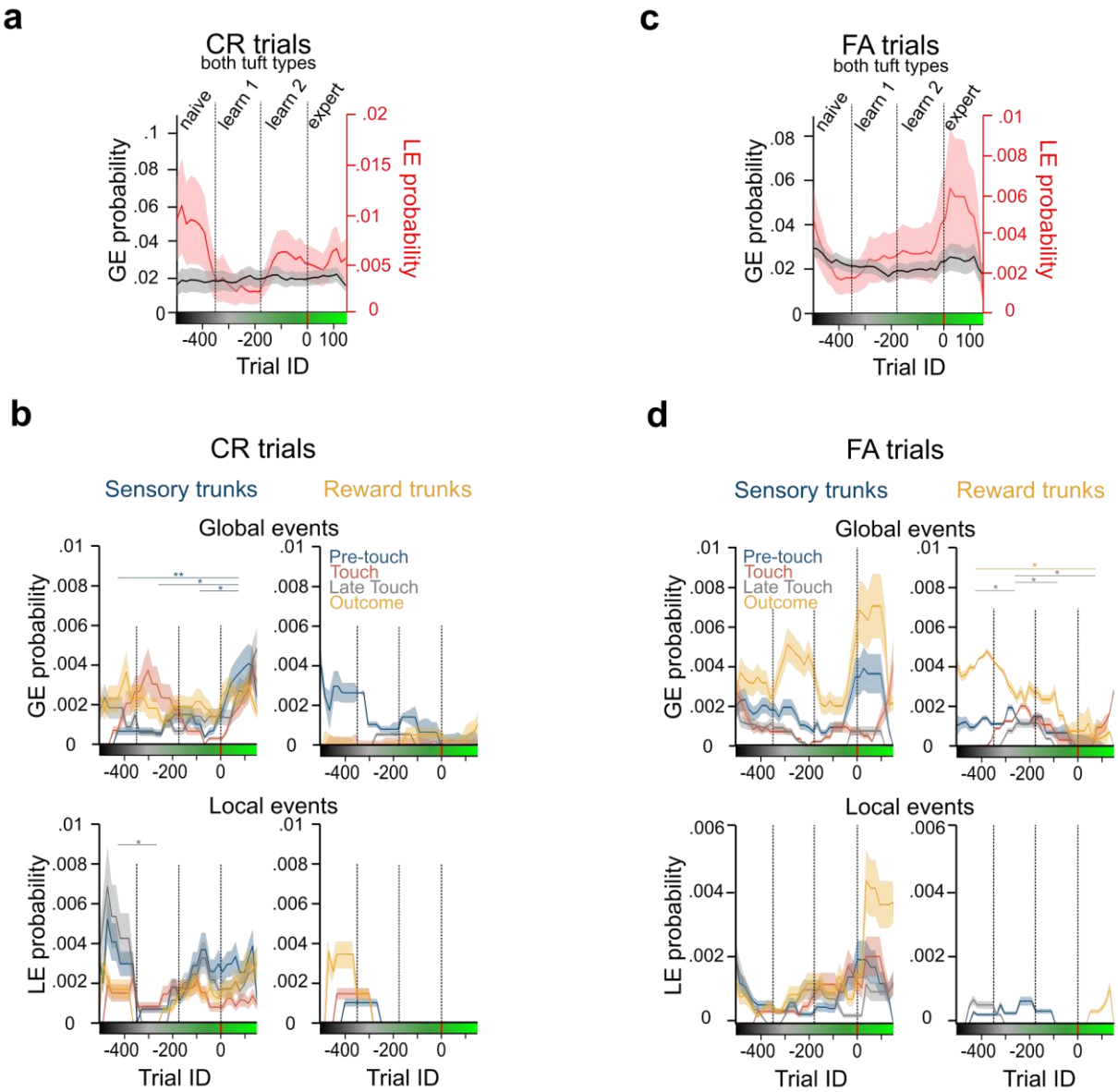
Supplementary Figure 5 | Average ΔF/F traces per trial type and functional cluster. a, Average ΔF/F traces across trial types for dendritic branches belonging to reward (left) and sensory (right) class. Averages include only trials of tuft branches and apical trunks with a detectable calcium transient onset and are shown separately for the naïve (black) and expert (green) condition. Note that average ΔF/F traces reflect both transient abundance and transient amplitudes, while transient onset densities shown in **Fig. 2c** do not consider amplitudes. (Reward dendrites: 1452 naïve Hit, 964 expert Hit, 304 naïve CR, 597 expert CR, 1472 naïve FA, 402 expert FA trails; Sensory dendrites: 1383 naïve Hit, 1671 expert Hit, 566 naïve CR, 1283 expert CR,2377 naïve FA, 714 expert FA trails; n = 8 mice; mean ± s.e.m).



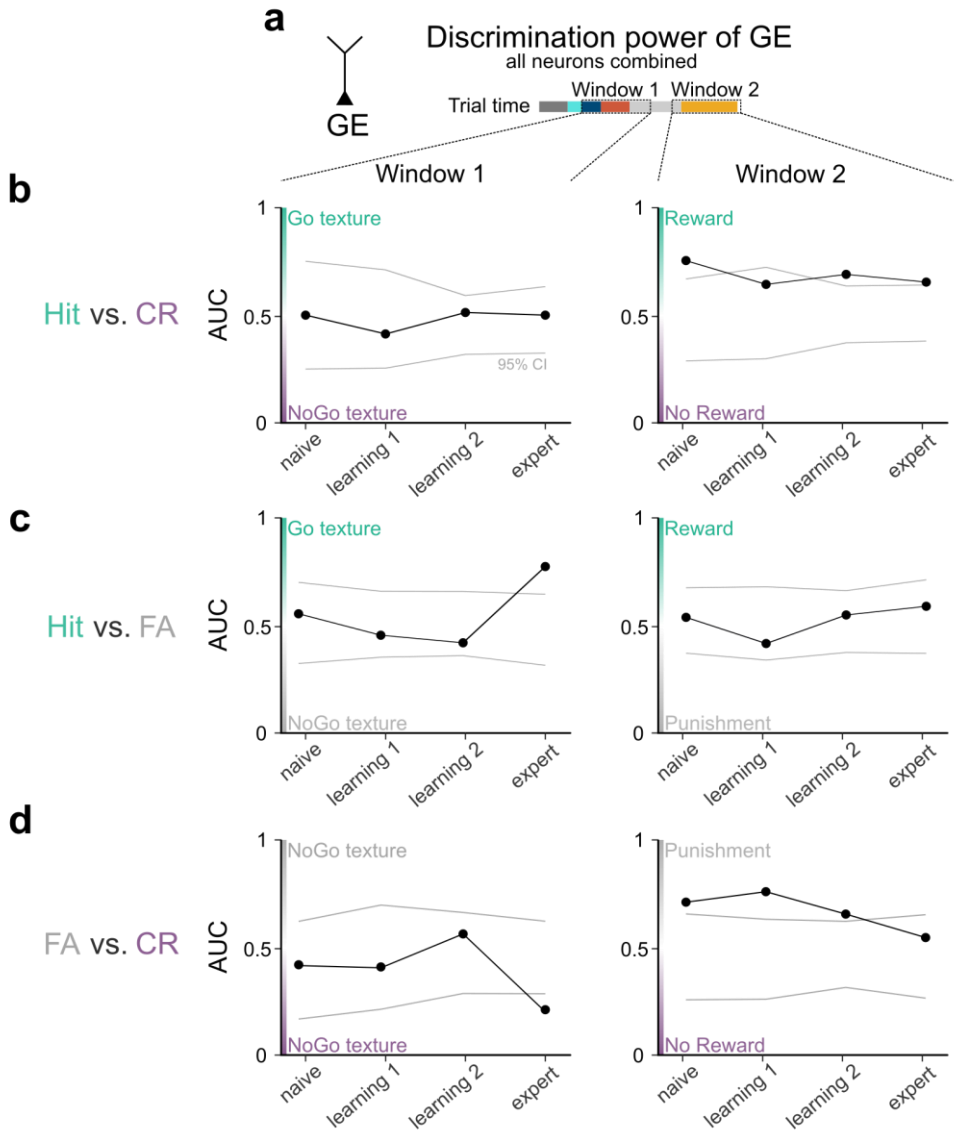
Supplementary Figure 6 | Morphological reconstructions of imaged dendritic tufts and their functional composition. a, Fraction of apical branches per class across imaging planes. **b**, Number of branches per functional class per tuft (n = 42). Dendrites that were not assigned to a class due to insufficient data point are shown in black. **c,** Absolute number of apical in sensory and reward tufts (13 ± 7 and 9 ± 5 branches, respectively; mean ± s.d.; p = 0.08, Student t-test; Black line: median , box: 25th and 75th percentile, whiskers: 5th and 95th percentile, outliers: points. **d,** Examples of L5 apical dendritic tufts reconstructed from two-photon image stacks. Branches are colour-coded according to their functional class.



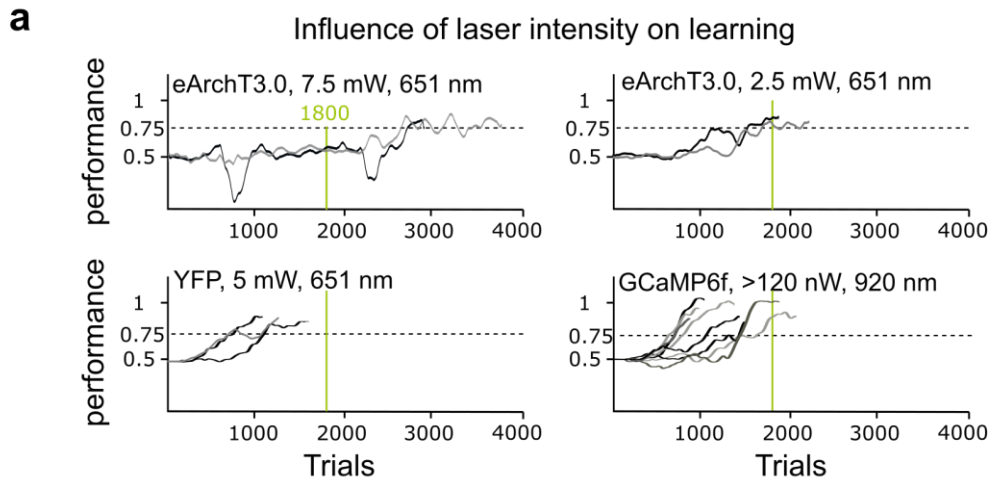
Supplementary Figure 7 | Characteristics of local and global events. a, Example ΔF/F traces of GE (top row) and LE (bottom row) for all ROIs in an example tuft (trunk: red; tuft branches: black). **b**, Data processing flow for event classification. **c**, Mean ΔF/F values for all tuft branches versus the mean ΔF/F value of the corresponding trunk in the same trial (ρ = 0.65 for all data points). *Top*: Colour code according to the trunk-tuft correlation value across each trial. *Bottom*: Colour code indicates the event type. **d**, Correlation between ΔF/F traces in apical branches and trunks across imaging planes within and across tufts during learning (42 tufts, mean ± s.e.m).



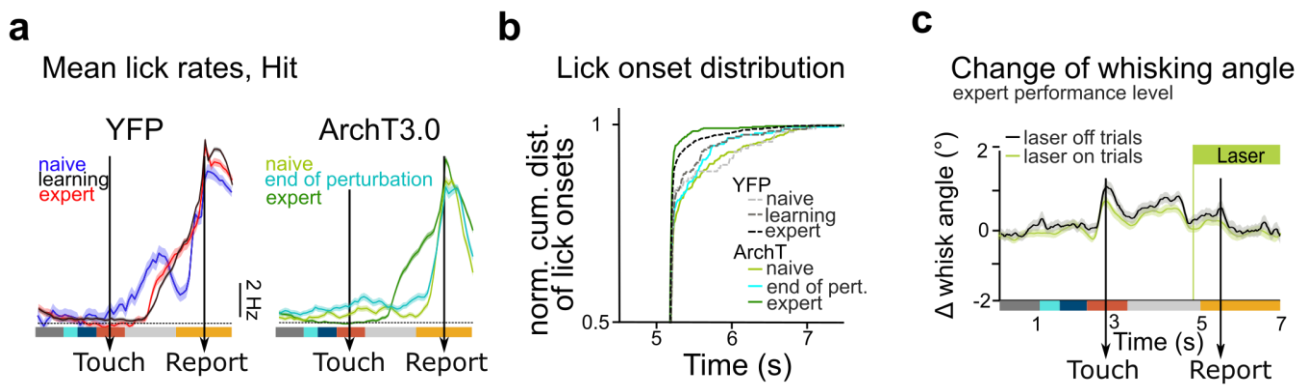
Supplementary Figure 8 | Local and global event probabilities in CR and FA trials. **a**, Overall probability for GE and LE per trial across learning in CR trials(Shaded areas: s.e.m.). **b**, GE and LE probability in CR split by trial window (pre-touch, touch, late-touch and outcome) and trunk class (p= 0.007, 0.025, 0.016 for GE in sensory tufts in pre-touch window between naïve and expert, learning 1 and expert and learning 2, and expert, respectively; p=0.034 for LE in sensory tufts in late-touch window between naïve and learning 1; one-way ANOVA). **c**, Overall GE and LE probability in FA trials. **d**, GE and LE probability in FA trials, split by trial window and trunk class (p = 0.047,0.042 and 0.047 for GE in reward tufts in late touch window between naïve and learning 1, learning 1 and learning 2 and learning 1 and expert; p = 0.031 for GE in reward tufts in outcome window between naïve and expert; one-way ANOVA).



Supplementary Figure 9 | AUC-based discrimination ability of L5 trunk population. a, Discrimination power of global dendritic tuft output (GE) was assessed in two windows (window 1: pre-touch, touch and partially late-touch; window 2: outcome window). **b**, Pooled AUC-based discrimination power for Hit versus CR trials for all recorded trunks(Confidence intervals: 95th and 5th percentile of randomly shuffled data). **c**, Pooled discrimination of Hit versus FA trials. **d**, Pooled discrimination of FA versus CR trials.



Supplementary Figure 10 | Influence of optogenetic perturbation on learning. a, Learning curves of various control experiments with eArchT3.0-, YFP- or GCaMP6f-expressing mice excited by 651-nm or 920-nm laser light with 2.5, 5.0, 7.5 or >120 mW light intensity(Green line : end of the laser perturbation after 1’800 trials).



Supplementary Figure 11 | Behavioural variables during optogenetic manipulation. a, **Average** lick rates during Hit trials in eYFP-expressing control mice (n = 3; 5 mW laser illumination at 561 nm; 321, 401 and 589 trials in the naïve, learning and expert condition) and eArchT3.0-expressing mice (n = 5; 5 mW laser illumination, 1000 trials per condition) during and after perturbation (end of perturbation= last 200 perturbation trials, mean ± s.e.m). **b,** Comparison of normalized cumulative distribution of reward-triggering licking onsets in Hit trials in eArchT3.0- and eYFP-expressing mice. **c,** Average whisking angle in Hit trials in expert mice in laser on and laser off trials (468 and 428 trials; n = 3 mice).