Dynamic encoding of perceptual features in primary somatosensory cortex upon reversal learning

Ronan Chéreau[1], Tanika Bawa[1,2], Leon Fodoulian[1], Alan Carleton[1], Stéphane Pagès[1], Anthony Holtmaat[1,3]

[1]Department of Basic Neurosciences and the Center for Neuroscience, CMU, University of Geneva, Rue Michel Servet 1, 1211, Geneva, Switzerland

[2]Lemanic Neuroscience Doctoral School, University of Geneva, Switzerland

[3]Lead contact

Correspondence: anthony.holtmaat@unige.ch

Keywords: Somatosensory cortex, texture discrimination, reversal learning, plasticity, higher-order functions, neuronal selectivity.

SUMMARY

Neurons in primary sensory cortex encode a variety of stimulus features upon perceptual learning. However, it is unclear whether the acquired stimulus selectivity remains stable when the same input is perceived in a different context. Here, we monitored the activity of individual neurons in the mouse primary somatosensory cortex in a reward-based texture discrimination task. We tracked their stimulus selectivity before and after changing reward contingencies and found that many neurons are dynamically selective. A subpopulation of neurons regains selectivity contingent on stimulus-value. These value-sensitive neurons forecast the onset of learning by displaying a distinct and transient increase in activity, depending on past behavioral experience. Thus, stimulus encoding during perceptual learning is dynamic and largely relies on behavioral contingencies, even in primary sensory cortex.

INTRODUCTION

The mammalian cortex encodes a myriad of sensory signal characteristics which are represented by neuronal assemblies, each with a preference for specific stimulus parameters (Holtmaat and Caroni, 2016; Makino et al., 2016). It is believed that these assemblies are organized in a hierarchical fashion. First-order sensory areas encode lower-order stimulus features, such as textures coarseness (Chen et al., 2013a; Garion et al., 2014; Safaai et al., 2013), object orientation and direction (Hubel and Wiesel, 1959), and sound frequency (Stiebler et al., 1997), whereas more complex features and contextual aspects of a stimulus are encoded by higher-order cortices (Akrami et al., 2018; Felleman and Van Essen, 1991; Hwang et al., 2017; Pho et al., 2018). Nonetheless, the coding in primary sensory cortices can exhibit higher levels of complexity, expressing non-sensory-related signals such as attention (Francis et al., 2018), anticipation (Poort et al., 2015) and behavioral choice (Chen et al., 2015; Francis et al., 2018; Pho et al., 2018; Poort et al., 2015; Yang et al., 2016). Perceptual learning initially shapes the stimulus selectivity and response properties of primary sensory neurons, which may contribute to a reliable detection of particular features, and thereby improve perception (Goltstein et al., 2013; Pecka et al., 2014; Poort et al., 2015; Schoups et al., 2001). However, it is unclear as to whether the stimulus preference of those neurons remains stable when the reward contingencies are changed. To study this, we monitored the evolution of

stimulus selectivity of primary somatosensory cortical (S1) layer 2/3 (L2/3) neurons in mice that learned to discriminate between a rewarded and non-rewarded texture, and reassessed their selectivity upon reversal learning.

RESULTS

Selectivity for rewarded stimuli increases with learning

We trained mice on a head-fixed 'Go/No-go' texture discrimination task, similar to previous designs (Chen et al., 2013a) (Figure 1A). Thirsted animals were incited to lick a spout during a 2-second (s) presentation of a P120 sandpaper (Go stimulus), in order to trigger the supply of a water reward at the end of texture presentation ('hit' trials). The animals needed to withhold from licking upon presentation of a P280 sandpaper (No-go stimulus) to avoid a 200-ms white noise and a 5-s timeout period. A failure to withhold from licking was scored as a 'false alarm' (FA) trial. Mice learned to discriminate between the two stimuli (Figure S1). They typically started at chance level (naïve) and reached above 70% performance, i.e. the expert criterion, within 3-7 days (expert) (Figure 1B). Trimming of the whiskers reduced the performance to chance level (Figure 1C), indicating that the mice had fully relied on somatosensory input and had not used additional sensory information to solve the task. To confirm that S1 is required for this task, we suppressed its activity using a local injection of the γ-aminobutyric acid receptor (GABAR) agonist muscimol. This significantly decreased the performance of mice that had initially learned the task (Figure 1D).

In order to monitor the stimulus selectivity of S1 neurons during texture discrimination learning we co-expressed the genetically encoded calcium sensor GCaMP6s and the cell filler mRuby2, predominantly in excitatory L2/3 neurons using adeno-associated viral vectors (Figure 1E, Figure S2) (Rose et al., 2016). Single-cell responses were recorded using two-photon laser scanning microscopy (2PLSM) (Figure 1, E and F). In trained mice, a fraction of the neurons displayed different response amplitudes upon presentation of either texture, similar to previous studies (Figure 1F and Figure S3) (Chen et al., 2013a; Chen et al., 2015). We used the mean response amplitudes over the 2-s texture presentation period for both stimuli to calculate a discriminability index (DI) based on a receiver-operating characteristic (ROC) curve analysis (Takahashi et al., 2016). In expert mice, we observed that a larger fraction of

the recorded population was selective for the Go P120 as compared to the No-go P280 texture (Figure 1G). On average, the fraction of selective neurons per mouse increased with learning (Figure 1H). This increase was attributable to an increase in the fraction of Go-selective neurons (Figure 1I).

To confirm that the Go-selective neurons primarily responded to texture presentations rather than to licking-related motor behavior, we aligned the normalized calcium traces of all hit trials either to the 1st lick or to the onset of the texture presentation. For the majority of neurons, the rise in the calcium trace occurred immediately after the texture presentation and preceded the 1st lick with large jitter (Figure 1, J to K). This suggests that the activity of the Go-selective neurons was better time-locked to texture presentation than to the licking response. To control for general responses that are associated with the choice to lick (Guo et al., 2014; Li et al., 2016), we monitored task-related calcium transients, as well as whisking and licking activity, prior to the training at which the sound cue, texture presentation, and spout opening were presented independently of one another (Figure S4A). We first categorized neurons based on their calcium response amplitude upon sound or texture presentation, and found that 35.4% of the neurons exhibited touch-related activity while only 0.6% responded to sound (Figure S4B). We then built a random forests model to assess if the activity of some neurons could predict whisking and/or licking rates (Figure S4, C to E). 10.3% of the neurons predicted the animal's whisking rate whereas only 2% predicted licking rates (Figure S4F), much lower than the fraction of Go-selective neurons (Figure 1I). Taken together, the data suggests that the large number of Go-selective neurons that we observed after training could not be attributed to licking-related neuronal activity.

Stimulus selectivity is dynamic upon reversal learning

Next, we investigated if S1 neurons expressed additional task-related selectivity. Previous studies using comparable paradigms have reported that S1 neurons exhibit selectivity not only for the tactile stimulus but also for the behavioral choice (Chen et al., 2013a; Chen et al., 2015; Yang et al., 2016). However, such selectivity could result from the persistent coupling of stimulus and choice rather than the stimulus per se. Therefore, it is unclear whether neurons would remain selective for a stimulus when the reward contingencies are changed. In order

to test this, expert mice were continued to be trained on the same textures but with reversed reward contingencies (Figure S5A). We monitored neuronal activity and calculated their DIs in expert sessions before and after reversal, and observed a variety of changes (Figure 2A). The average ΔDI over the recorded population turned negative immediately upon the reversal, i.e. selectivity for the previously rewarded texture was initially maintained. However, the ΔDI reverted to positive values after relearning (Figure S6A-B), indicating that the population re-acquired selectivity for the rewarded texture. We found that those neurons that exhibited touch-related activity contributed the most to this effect (Figure S6C). Moreover, a substantial fraction of the neurons lost or gained selectivity (19 and 18% respectively) upon reversal learning (Figure 2B and Figure S5B-C, classified as 'gained' and 'lost'), indicating that sensory stimulus encoding in the primary somatosensory cortex is largely dynamic. Such changes could be the result of network plasticity. To assess this we calculated the level of co-fluctuation in spontaneous activity within each group, which may reflect the level of mutual connectivity (Cohen and Kohn, 2011; Cossell et al., 2015; Ko et al., 2011). Upon reversal, the level of co-fluctuation increased for gained neurons and decreased for lost neurons (Figure 2C). This may indicate that reversal learning promotes the rewiring of local synaptic circuits.

About 12% of neurons remained either selective for the original textures or for the original Go or No-go contingencies (Figure 2B and Figure S5D, classified as 'texture' and 'Go/No-go', respectively). We questioned whether the latter group simply displayed a continued preference for the behavioral choice (i.e. indicative of the upcoming action selection). To test this, we focused our analysis on the first reversal session in which the animals typically abandoned their previous behavioral strategy and made inconsistent choices. This naïve phase allowed us to parse out whether a neuron's activity was mediated by the presented texture or by the upcoming choice. For each neuron, we calculated a choice index (CI), based on a ROC curve analysis comparing the response amplitudes between lick and no-lick trials. From this analysis, we could distinguish two classes of neurons (Figure 2D), one for which the CI remained stable, and one for which the CI was altered during the naïve phase. We conclude that while the former class was indicative of behavioral choice irrespective of the presented stimulus or licking behavior (Figure S7), the latter was transiently uncoupled from choice and thus signaled the value that was associated with the texture (i.e. indicative of the expected absence or presence of the upcoming reward). The neurons that signaled stimulus value

consequently re-acquired selectivity for the reward contingencies of the textures (Figure 2E and Figure S5E-F).

Value-selective neurons display error history activity that precedes learning

Next, we hypothesized that the subclasses of S1 neurons that acquired or re-acquired selectivity adjust their stimulus-evoked response amplitudes during learning, reflecting the behavioral error history (Helmchen F. and Banerjee A., personal communication) (Hattori et al., 2019; Lacefield et al., 2019). Previous work suggests that a correct trial that follows an incorrect trial is considered more instructive for the animal than two consecutive correct trials (Akrami et al., 2018; Hattori et al., 2019; Hwang et al., 2017). To test this, we focused our analysis on those consecutive trials in which mice were actively licking upon texture presentation (i.e. hits and FAs), hence ensuring that they were engaged in the task. We compared the mean response amplitudes of hit trials that were preceded by a FA trial ($R_{hit(post\ FA)}$) to those that were preceded by a hit trial ($R_{hit(post\ hit)}$) (Figure 3A). Trial sequences across mice were aligned to the point at which the reversal learning had reached the expert criterion (70% performance; Figure 3B). We observed that the sampling strategy of the texture remained relatively constant across reversal learning (Figure S8). Averaged Hit and FA rates separated from one another ~140 trials before the expert criterion, indicating the moment at which mice started to improve their performance, termed learning onset (Figure 3C, black arrow head). For the choice neurons and the neurons that gained selectivity we did not observe any difference between the $R_{hit(post\ hit)}$ amplitudes and the $R_{hit(post\ FA)}$ amplitudes. In contrast, for the contextual value-selective neurons, the average $R_{hit(post\ FA)}$ response amplitudes became larger than the $R_{hit(post\ hit)}$ amplitudes, at ~260 trials before the expert criterion, and ~120 trials before learning onset (Figure 3, C and E, red arrowhead). The two types of responses became similar again when mice performed above the expert criterion. We used the normalized difference between $R_{hit(post\ FA)}$ and $R_{hit(post\ hit)}$ responses as an error history index (STAR Methods), and observed that a large fraction of the contextual value-selective neurons exhibited a transient increase in the error history as compared to the other neuronal classes that we had identified (Figure 3, E and F, and Figure S9). Such a temporal profile, anticipating the learning onset could not necessarily be deduced

from the DI evolution (Figure S10). Altogether, these results suggest that neurons encoding the contextual value of a stimulus may be indicative of the upcoming improvement in discrimination (i.e. learning).

DISCUSSION

Previous studies have indicated that perceptual learning increases the reliability and selectivity of neuronal responses in primary sensory areas. As a consequence, the population activity stabilizes, which may improve perception (Goltstein et al., 2013; Pecka et al., 2014; Peron et al., 2015; Poort et al., 2015; Schoups et al., 2001). We extend on this work by showing that task-related stimulus feature selectivity of neurons in the primary somatosensory cortex is dynamic upon reversal learning of a simple texture discrimination task. We found that a small fraction of the neurons is stably selective for texture or behavioral choice, while a large fraction loses, gains, or first loses and then regains selectivity when the reward contingencies of the stimuli is reversed. This is remarkable, since upon reversal learning the capacity of the cortex to discriminate lower-order sensory features does not need to be altered in order for the mouse to perform the task. Nonetheless, the finding is congruent with the idea that learning continuously optimizes sensory representations in cortex, and strongly depends on the context (Chen et al., 2015; Makino et al., 2016; Poort et al., 2015). In our study, the altered reward contingency could represent such a contextual change, since the animal has to suddenly shift its attention to the opposite texture to obtain a reward. Our data imply that the simple reversal of reward contingencies disorganizes the existing map, which is then extensively reshaped with relearning. The reshaping of this map could be the result of plasticity mechanisms that also underlie the experience-dependent tuning of neuronal response properties in primary sensory cortex (Rose et al., 2016). In this case, Hebbian plasticity may drive the phenomenon, with the result that similarly tuned neurons become more strongly interconnected (Cohen and Kohn, 2011; Cossell et al., 2015; Ko et al., 2011). This is supported by our finding that both, neurons that gain selectivity and those that lose selectivity show higher co-fluctuation in spontaneous activity during the time they are selective.

Similar to previous studies we found that selectivity is shaped by behavioral choice (Chen et al., 2013a; Chen et al., 2015; Pho et al., 2018; Yang et al., 2016). For example, the population of neurons that became selective for the rewarded texture was larger than for the non-rewarded texture. The reversal learning paradigm allowed us to assess the stability of this coupling, e.g. whether the Go-selective neurons stably respond to the rewarded texture, even during the naïve reversal phase, or whether they lose selectivity shortly upon reversal and then re-built it with relearning. We find that more than half of the Go-selective neurons belonged to the latter class. Thus, their sensory responses were transiently uncoupled from the animal's choice, and primarily depended on whether the presented texture was associated with the upcoming reward (or not). In future experiments it will be interesting to test whether repeated reversal learning would continue to renew the selective population, or whether the population would revert back to the original activity configuration.

Previous studies indicate that the value of a sensory stimulus is encoded by higher-order areas such as the posterior parietal, orbitofrontal, and retrosplenial cortex (Akrami et al., 2018; Hattori et al., 2019; Hwang et al., 2017; Sul et al., 2010). Our data shows that value-encoding is also an attribute of a population of neurons in S1. The instructive cues for this selectivity could be manifold. For example, they could be provided by direct feedback from the aforementioned higher-order cortical areas, or they could be derived from sub-cortical areas that are implicated in attention and behavioral updating during learning (Wimmer et al., 2015). Modulatory reinforcement signals that are associated with behavioral outcome could also play a major role (Pi et al., 2013; Watabe-Uchida et al., 2017). Indeed, reward-related response modulation has been observed in S1 (Lacefield et al., 2019), and was found to promote cortical plasticity processes related to visual response tuning in primary visual cortex (Goltstein et al., 2013). We found that the value-neurons gradually regained their selectivity with relearning, which would be congruent with the idea that reward-related plasticity mechanisms contribute to shaping perceptual representations in cortex.

We also found that value neurons transiently displayed enhanced response amplitudes dependent on the animal's behavioral error history. During the naïve reversal phase these neurons showed higher responses in hit trials if the hit trial was preceded by a false alarm trial. This phenomenon was prominent during the transition from the naïve to expert reversal phase and forecasted the increase in behavioral performance. We speculate that the omission

of reward-associated signals during a false alarm trial directs the animal's attention towards the newly rewarded texture. Elevated attentional signals have been shown to modulate sensory-driven responses in visual cortex (Pooresmaeili et al., 2014). Thus, the attentional signals may be read out by the value neurons, which in turn reshape the texture selectivity of surrounding neurons. Together, this may enhance sensory perception.

SUPPLEMENTAL INFORMATION

Supplemental information includes ten figures.

ACKNOWLEDGMENTS

We thank Ariel Gilad for advice on the behavioral paradigm, Fritjof Helmchen and Abhishek Banerjee for discussions and suggesting the error history analysis, Jose Manuel Nunes for advice on assessment of selection criteria in the prediction model, Sebastien Pellat for technical support and engineering, Laura Bussien and Elodie Husi for assistance with histology. We thank Pieter Roelfsema for comments on the manuscript. This project was supported by the Swiss National Science Foundation (grants 31003A-153448, 31003A_173125, CRSII3_154453, and NCCR Synapsy 51NF40-158776), and a gift from a private foundation with public interest through the International Foundation for Research in Paraplegia.

AUTHORS CONTRIBUTIONS

RC, SP and AH designed the experiments. RC and TK performed the experiments. RC analysed the data and LF performed random forests modelling. RC and AH wrote the manuscript. All of the authors edited the manuscript.

DECLARATION OF INTERESTS

Authors declare no competing interests.

FIGURES

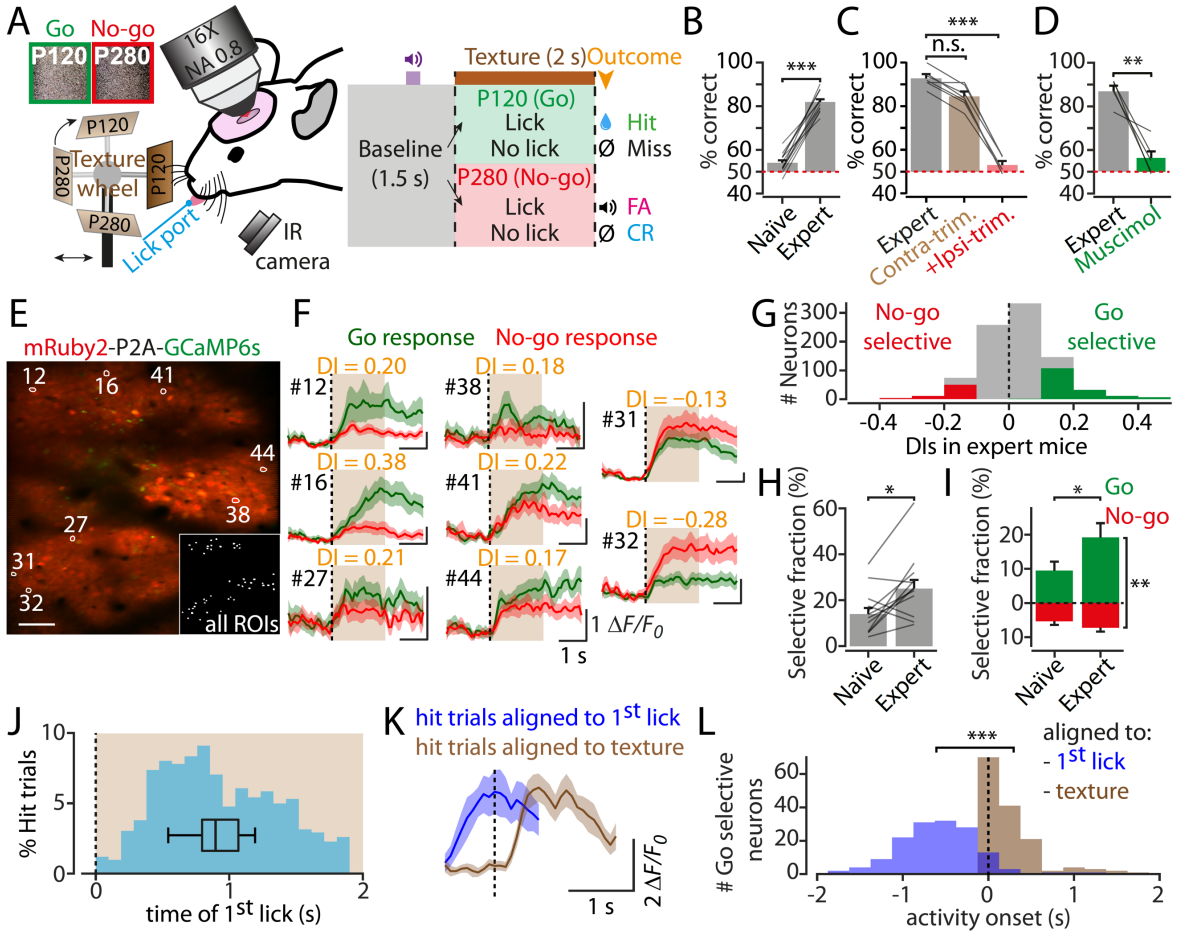

Figure 1. Increase in texture selectivity of S1 neurons during discrimination learning.

(A) Behavioral setup and experimental time-course. Mice were rewarded for licking the water spout upon random presentations of a P120 (Go stimulus) but not a P280 (No-go stimulus) sandpaper.

(B) Discrimination performance across learning (*N*=12 mice, Wilcoxon rank sum test, fraction of correct trials: naïve, 54.0 ± 1.2% (mean ± s.e.m.), expert, 81.8 ± 1.3%, ***$P<10^{-8}$).

(C) Effect of whisker trimming on behavioral performance (*N*=7 mice, Friedman test and post-hoc Dunn's test, fraction of correct trials: expert, 92.7 ± 1.5% correct vs. contralateral trimming, 84.4 ± 1.7%, *P*=0.33; or vs. +ipsilateral trimming, 53.0 ± 1.3% correct, ***P=0.0009).

(D) Acute injection of muscimol in S1 of expert animals reduced performance (*N*=5 mice, Wilcoxon rank sum test, fraction of correct trials: expert, 79.4 ± 1.9% correct, after muscimol, 52.5 ± 6.0%, **P=0.0079).

(E) Field of view containing mRuby2/GCaMP6s-expressing neurons. Single neurons were selected as regions of interest (ROIs) based on mRuby2 expression (inset), and recorded during training. Scale bar: 100μm.

(F) Examples of average calcium responses to the Go-P120 stimulus (green) and No-go P280 stimulus (red) from neurons in (E). Yellow values above each trace represent the discriminability indices (DI, bootstrap, $P<0.05$; STAR Methods).

(G) Distribution of DIs in expert mice (N=12 mice, in green, Go selective neurons: 146/875; in red, No-go selective neurons: 57/875).

(H) The fraction of selective neurons before and after learning (N=12 mice, 875 neurons, naïve, 14.0 ± 2.6% (mean ± s.e.m.); expert, 25.0 ± 3.8% selective neurons; Wilcoxon rank sum test, *P=0.03).

(I) The fraction of Go (green) and No-go (red) selective neurons before and after learning (N=12 mice, 875 neurons, Wilcoxon rank sum tests, Go-selective neurons: naïve, 9.5 ± 2.6%, expert, 19.2 ± 4.2%, *P=0.02; No-go selective neurons: naïve, 5.3 ± 1.1%, expert, 7.2 ± 1.2%, P=0.33; Go vs. No-go selective fractions: naïve, P=0.36, expert, **P=0.006). (B to D, H and I) Lines represent individual mice and bars and error bars represent averages and s.e.m.

(J) The distribution of lick reaction times over all Hit trials across expert mice (N=923 trials). Box plot shows mean, 25th-75th percentiles, and the minimum and maximum average reaction times over mice (N=12 mice).

(K) Example of the average calcium response in a Go-selective neuron of an expert mouse, aligned to texture presentation (brown) and the 1st lick (blue).

(L) Distribution of the response onsets for all Go selective neurons in expert sessions aligned to the texture presentation (brown) or to the 1st lick (blue, N=12 mice, 146 neurons, average onset times: relative to texture, 0.29 ± 0.04 s; relative to 1st lick, -0.61 ± 0.04 s, paired t-test, ***P=2e-46).

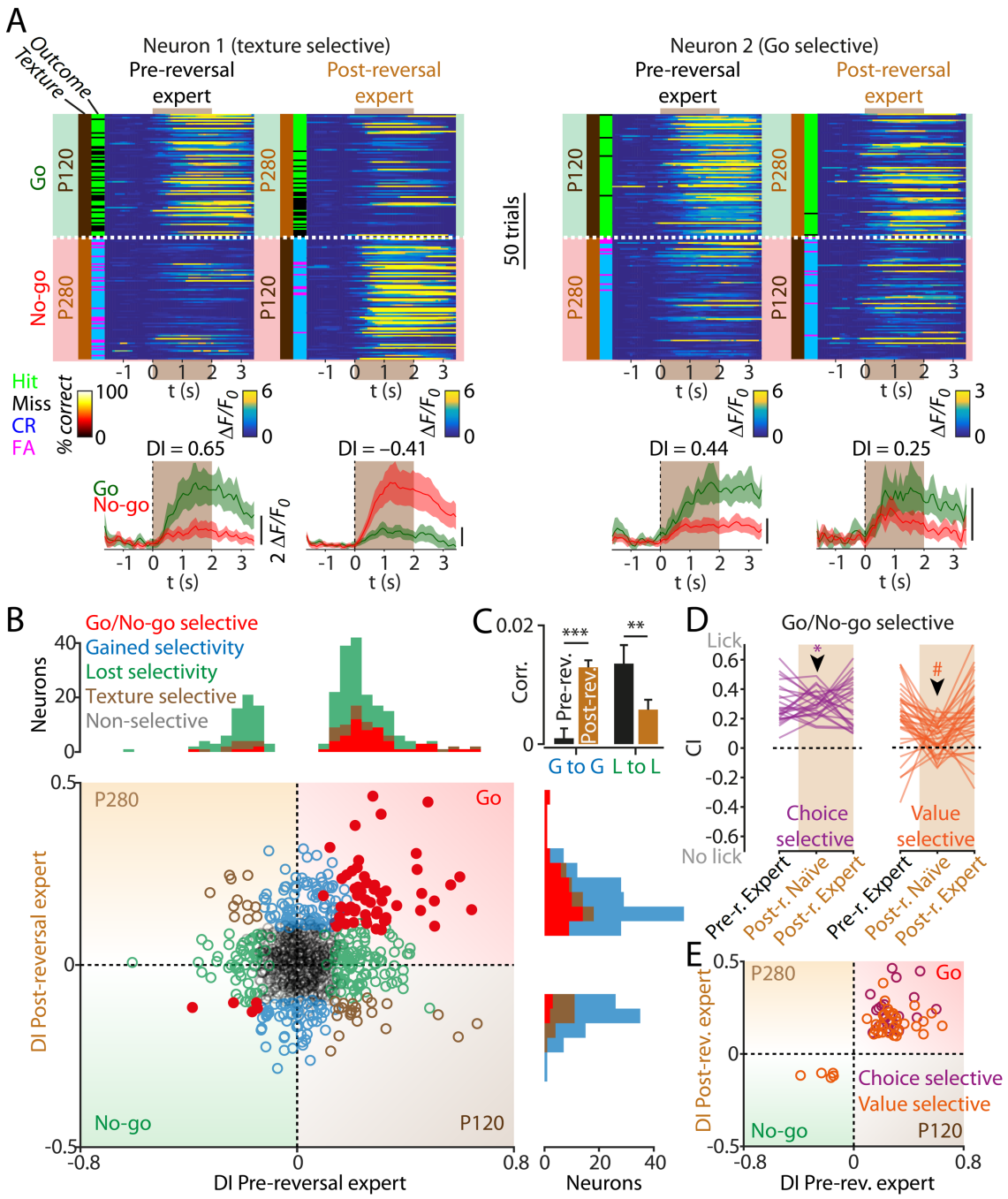

Figure 2. Origins of neuronal selectivity revealed by texture reversal learning.

(A) Single trial responses of two example neurons from expert sessions pre and post texture reversal aligned to texture presentation and sorted by stimulus type (160 trials from these expert sessions were displayed). Corresponding trial outcomes and performance are also shown. The neuron on the left remains selective for the P120 texture whereas the neuron on the right remains selective for the Go stimulus. Below, average calcium response with s.e.m. of these trials for Go and No-go stimuli and their DI value indicated above.

(B) DIs of neurons in expert sessions pre and post reversal. Neurons in the P120 and P280 quadrants are stimulus selective (brown, *N*=33 neurons), preferentially responding to the same texture pre and post reversal. Neurons in the Go and No-go quadrants are stimulus type selective (red, *N*=60 neurons), preferentially responding to the same stimulus pre and post reversal. Other neurons lose (green, *N*=154 neurons) or gain (blue, *N*=144 neurons) selectivity, or remain nonselective (grey, *N*=412 neurons from 12 mice) throughout training. Top and right histograms show the distributions of DIs of selective neurons pre and post reversal.

(C) Average spontaneous activity correlations between pairs of neurons that gained (G to G) or lost selectivity (L to L) comparing pre-reversal expert and post reversal expert (G to G: 224 pairs, ***P=6e-9; L to L: 360 pairs, **P=0.0014 using paired t-tests).

(D) Changes in choice index (CI) of Go/No-go stimulus selective neurons in pre-reversal expert sessions, post-reversal naïve and expert sessions. Individual neurons CI were tested for significance, using random permutations, in the post-reversal naïve sessions. Neurons for which the CI was significant in the post-reversal naïve sessions and remained of the same sign as for pre- and post-reversal expert sessions (*) could be classified as choice selective neurons (purple). Other neurons for which the CI was either not significant or of the opposite sign in the post-reversal naïve sessions (#) as for pre- and post-reversal expert sessions could be classified as value selective neurons (orange).

(E) Scatter plot, same as (B) showing the DI of choice and value selective neurons in pre and post-reversal expert sessions.

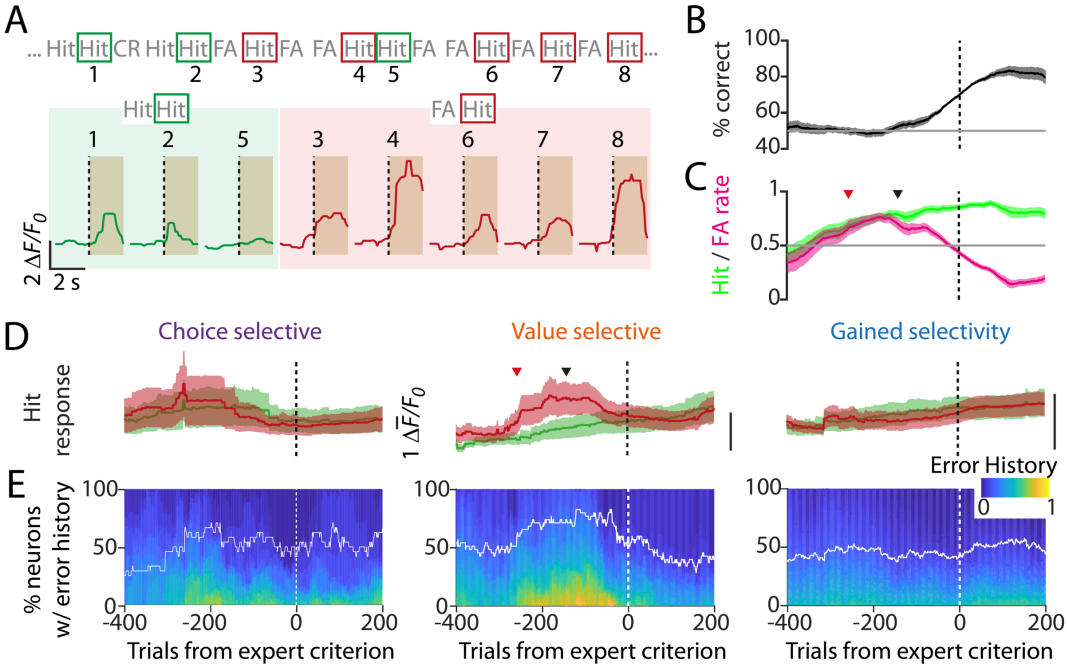

Figure 3. Behavioral error history is signaled by value selective neurons during reversal learning prior to performance increase.

(A) Example of a value selective neuron responses in Hit trials before the post-reversal performance increase. Hit trials preceded by a False Alarm trial (FA, dark red) contained larger responses than Hit trials that were preceded by another Hit trial (dark green). Numbers show the position of the Hit trials in the behavioral outcome sequence above.

(B) Average performance curves across mice realigned to the expert criterion (70% correct, dotted line) post-reversal, calculated with a rolling window of 200 trials (*N*=10 mice that displayed value selective neurons).

(C) The corresponding average Hit and FA rate averages. (B to C) Shaded areas represent the 95% confidence intervals.

(D) The corresponding average calcium response (ΔF/F) in Hit trials that followed upon another Hit trial (green) or upon a FA trial (red) for choice selective (*N*=24 neurons from 6 mice), value selective (*N*=35 neurons from 10 mice) and gained selectivity (*N*=144 neurons from 12 mice) neuronal classes. (C to D) Arrowheads indicate the trials at which the 95% confidence intervals separates for hit and FA rates (black) and the Hit-Hit and FA-Hit calcium responses (red).

(E) Percentage of neurons per class with a significant error history rate (white line) above the cumulative distributions of error history values.

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-GABA	Abcam	Ab86186
Donkey anti-mouse coupled to Alexa Fluor 647	Thermo Fisher Scientific	A32787
Bacterial and Virus Strains		
AAV1.hSyn1.mRuby2.GSG.P2A.GCaMP6s	Addgene	50942-AAV1
Chemicals, Peptides, and Recombinant Proteins		
medetomidine	Orion Pharma	Dormitor
midazolam	Roche	Dormicum
fentanyl	Sinetica	Fentanyl
muscimol	Thermo Fisher Scientific	Bodipy-TMR-X
Experimental Models: Organisms/Strains		
Mouse/C57BL/6J	Janvier Labs	https://www.janvier-labs.com/rodent-research-models-services/
Software and Algorithms		
ImageJ	The NIH	https://imagej.nih.gov/ij/
MATLAB	Mathworks	https://www.mathworks.com/
Scanimage 2016b	Pologruto et al., 2003	http://www.scanimage.org

CONTACT FOR REAGENT AND RESOURCE SHARING

Information and request for reagents may be directed and will be fulfilled by the Lead Contact Anthony Holtmaat (anthony.holtmaat@unige.ch).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

All behavioral experiments were performed during the dark phase. All procedures were conducted in accordance with the guidelines of the Federal Food Safety and Veterinary Office of Switzerland and in agreement with the veterinary office of the Canton of Geneva (licence numbers GE/28/14, GE/61/17, and GE/74/18). C57BL/6J male mice (Janvier Labs) aged 6 weeks were group housed on a 12-hour light cycle (lights on at 8:00 AM) with littermates until surgery. Two weeks after surgery, animals were housed under standard conditions, with an inverted light-dark cycle 7-8 d before the first training session.

METHODS DETAILS

Surgery and intrinsic optical imaging

Stereotaxic injections of adeno-associated viral (AAV) vectors were carried out on 6-week-old male C57BL/6 mice. A mix of O_2 and 4% isoflurane at 0.4 l/min was used to induce anaesthesia followed by an intraperitoneal injection of MMF solution consisting of 0.2 mg/kg medetomidine (Dormitor, Orion Pharma), 5 mg/kg midazolam (Dormicum, Roche) and 0.05 mg/kg fentanyl (Fentanyl, Sinetica) diluted in sterile 0.9% NaCl. AAV1.hSyn1.mRuby2.GSG.P2A.GCaMP6s (Penn Vector Core; 100 nl) (Rose et al., 2016) was delivered to L2/3 of the right C2 barrel-related column in S1 (1.4 mm posterior, 3.5 mm lateral from bregma, 300 µm below the pia). For long-term *in vivo* calcium imaging, a 3 mm diameter cranial window was implanted, as described previously (Holtmaat et al., 2009).

Two weeks after surgery, the C2 barrel column was mapped again using intrinsic optical imaging to confirm the location of mRuby2/GCaMP6s expression. To do this, a mix of O_2 and 4% isoflurane at 0.4 l/min was used to induce anaesthesia followed by an intraperitoneal injection of MM solution consisting of 0.2 mg/kg medetomidine and 5 mg/kg midazolam diluted in sterile 0.9% NaCl. The C2 whisker was inserted into a capillary connected to a piezo actuator. Intrinsic signal from the C2 barrel column was collected during repeated whisker stimulation (1 s at 8 Hz). A 100-W halogen light source connected to a light guide system with a 700-nm interference filter was used to illuminate the cortical surface through the cranial window. Reflectance images 300 µm below the surface were acquired using a 2.7× objective and the Imager 3001F (Optical Imaging, Mountainside, NJ) equipped with a 256 × 256 pixels

array charge-coupled device (CCD) camera. The built-in Imager 3001F analysis program was used to visualise the responses and produce an intrinsic signal image by dividing the stimulus signal by the pre-stimulus baseline signal. An image of the vasculature was then acquired using a 546-nm interference filter, and superimposed to the intrinsic signal image. This reference image was used later to select an appropriate field of view (FOV) using 2PLSM. After this procedure, a metal post was implanted laterally to the window using dental acrylic to restrict head movement during behavior and imaging.

Habituation and water deprivation

Mice were handled and accustomed to be head restrained on the training setup for 10-15 min during 4-5 d. Water deprivation started 3-5 d before the first training session and ceased at the end of the training. Weight was monitored daily during this period and the amount of water given was adjusted to prevent them from losing more than 15% of their original weight. Altogether, mice received a minimum of 1 ml of water per day corresponding to the amount they drank during the training as rewards plus the amount that the experimenter provided outside of the training sessions.

Texture discrimination task

Mice were trained to discriminate between two commercial-grade sandpapers (P120 and P280) in a Go/No-go paradigm as described previously (Chen et al., 2013a). The control of the devices and the recording of behavioural parameters were performed using a data acquisition interface (PCI 6503, National Instruments) and custom-written LabWindows/CVI software (National Instruments). Licks were detected electrically. Mice remained on a metallic plate that maintained an electrical potential difference with the licking spout. The electrical circuit was closed when mice touched the spout with their tongue, producing a 1.2-µA current that was detected by the acquisition interface. Whisking activity was measured with an optical barrier that detected the changes in intensity when whiskers swept through it at a sampling frequency of 10 kHz. To achieve this, an 850 nm LED beam was used as light source (HIR204C, Everlight Electronics) and an 860 nm phototransistor (PT 202C, Everlight Electronics) was used to detect intensity variations through a 1-mm hole placed 60 mm away from the light source.

Whisking activity was quantified as the frequency at which individual whiskers crossed the light beam placed ∼1 mm in front of and centered on the presented texture. The licking and whisking rates were calculated as the average number of events over a sliding window of 100 ms and normalized per second.

Sandpapers were attached onto a four-arm wheel (2 × 2 of the same sandpapers) mounted on a stepper motor (T-NM17A04, Zaber) and a motorized linear stage (T-LSM100A, Zaber) to move textures in and out of reach of whiskers. At the start of each trial, the wheel spun for a random amount of time in the back position of the linear stage (approximately between 0.5 and 1 s) and stopped between 2 textures positions. To present a texture the linear stage first moved to the front position and then the stepper motor rapidly slid the sandpaper into the whisker field at approximately 15 mm from the snout with an angle of 70° relative to the rostro-caudal axis. In the first phase of the training, the coarser P120 sandpaper was the Go-stimulus (i.e. target stimulus for which the mouse was incited to lick the spout in order to receive a water reward) and the P280 sandpaper was the No-go stimulus (i.e. non-target stimulus for which the mouse was incited to refrain from licking the spout). Initially, mice were trained to trigger a 4-6 µl sucrose water reward (100 mg/ml) by licking the spout during the presentation of the Go-stimulus (P120). Then, they were gradually familiarized with the No-go stimulus (from 0 to 30% of the trials) within two sessions (one session per day, 150-300 trials per session). Imaging started when Go and No-go stimuli were pseudo-randomly presented with 50% probability for each trial type with a maximum of 4 consecutive presentations of the same stimuli. A trial consisted of a 1-s pre-stimulus period followed by a 3-kHz auditory cue for 200 ms, a delay period of 500 ms after which the texture reached the whiskers within 150 ms and remained there for 2 s before being retracted. Licking during the Go stimulus presentation triggered a water reward at the end of the 2-s of presentation, and the corresponding trial was scored as a 'hit'. Licking during the No-go stimulus presentation triggered a 500-ms white noise at the end of the 2-s of presentation plus a 5-s time out period and the trial was scored as a 'false alarm' (FA). In the absence of a lick during stimulus presentation, trials were scored as a 'miss' or a 'correct rejection' (CR) for Go and No-go stimuli respectively. To prevent the mice from compulsive licking during training, in addition to the aforementioned rules, mice had to show a 2-fold increase in the licking rate during stimulus presentation as compared to the pre-stimulus baseline period to get rewarded on Go stimulus

presentation. Around 250-400 trials per session were performed (1 session per day) at a rate of ~6 trials/min.

The overall performance of the animal was calculated as the percentage of correct trials (hits + CRs) over an entire session or over a sliding window of 200 trials (Figure 3 and Figure S9). The hit and FA rates were calculated as $N_{hit}/(N_{hit} + N_{miss})$ and $N_{FA}/(N_{FA} + N_{CR})$ respectively where N is a number of trials for an entire session or over a sliding window of 200 trials (Figure 3 and Figure S9). Mice were considered experts when the average performance per session reached a level of 70% correct trials (the expert criterion) over two consecutive sessions. In the second phase of training (i.e. reversal learning), Go and No-go stimuli were inverted (P280 and P120 respectively) and mice were trained until they reached the same expert criterion again in two consecutive sessions.

2PLSM

We used a custom built 2-photon laser scanning microscope mounted onto a modular *in vivo* multiphoton microscopy system (https://www.janelia.org/open-science/mimms-10-2016) equipped with an 8-kHz resonant scanner and a 16× 0.8NA objective (Nikon, CFI75), and controlled with Scanimage 2016b (Pologruto et al., 2003) (http://www.scanimage.org). Fluorophores were excited using a Ti:Sapphire laser (Chameleon Ultra, Coherent) tuned to $\lambda =$ 980 nm that was slightly underfilling the back aperture of the objective to extend the depth of field to 5 µm. Fluorescent signals were collected with GaAsP photomultiplier tubes (10770PB-40, Hamamatsu) separating mRuby2 and GCaMP6s signals with a dichroic mirror (565dcxr, Chroma) and emission filters (ET620/60m and ET525/50m, respectively, Chroma). Fast volumetric imaging was performed at 11.5 Hz using a piezo z-scanner (P-725 PIFOC, Physik Instrumente) for moving the objective over the z-axis. Each acquisition volume consisted of 5 contiguous planes (with 5-µm steps between planes) of 400 × 400 µm (512 × 256 pixels) allowing post-hoc z-motion correction which may be generated by licking-induced brain motion artifacts (Chen et al., 2013b).

QUANTIFICATION AND STATISTICAL ANALYSIS

Image processing

Image acquisitions were processed using custom-written MATLAB scripts and ImageJ (http://rsbweb.nih.gov/ij/). Lateral and axial motion corrections were performed using the mRuby2 signal as a reference. First, rigid lateral movement vectors were calculated based on individual trial movies from the average z-projection of the 20-µm imaged volumes using the NoRMCorre MATLAB toolbox (Pnevmatikakis and Giovannucci, 2017). Residual bidirectional scanning artifact vectors were calculated using a highest-pixel-line signal correlation between the two scanning directions on the entire frame. Inter-trial registration was calculated using a custom-written cross-correlation algorithm based on the rigid image stack registration plugin in ImageJ. All calculated lateral motion corrections were applied on both the mRuby2 and GCaMP6s signals. Second, axial motion correction was performed using cross-correlation on linearly interpolated volumes (with a factor 3). The image planes with the highest correlation to a reference image, defined as the center image plane of the first volume, were selected. For an unbiased extraction of the GCaMP6s fluorescence signals from individual neurons, regions of interest (ROIs) were drawn manually for each session based on neuronal shape using the mRuby2 signal. The fluorescence time-course of each neuron ($F_{measured}$) was measured as the average of all pixel values of the GCaMP6s signal within the ROI. Local neuropil signal ($F_{neuropil}$) was measured for each ROI as the average of pixel values within an automatically defined ring of 15 µm width, 2 µm away from the ROI (excluding overlap with surrounding ROIs) (Kerlin et al., 2010). The fluorescence signal of a cell body was then estimated as $F(t) = F_{measured}(t) - r \times F_{neuropil}(t)$ with $r = 0.7$ (Chen et al., 2013c). Residual trends were removed by subtracting the 8th percentile of each trial (Dombeck et al., 2007). Normalized calcium traces $\Delta F/F_0$ were calculated as $(F - F_0)/F_0$, where F_0 is the median of the individual mean baseline fluorescence signal of each trial over a 1-s period before the start of the stimulation. For individual stimulation sessions (see **Individual stimulation session and neuron categorization**) and spontaneous activity recordings, F_0 is the 30th percentile of each trial trace. For display, traces were additionally filtered with a Savitzky-Golay function (2nd order, 500-ms span).

Activity onset analysis

Normalized calcium traces ($\Delta F/F_0$) were aligned to either the onset of the texture presentation or to the first lick during the texture presentation for each neuron across all hit trials of an expert session. For both realignments, the onset of the neuronal response was calculated as the time, relative to the texture or first lick onset, at which the average of the response reached half of its maximum amplitude.

Individual stimulation session and neuron categorisation

Prior to the start of the training, 9 mice were imaged in the experimental training configuration, where task stimulation features were presented independently in a pseudo-random fashion. Data acquisition was organized in trials of 10 s, each starting with a 3-s baseline after which one of the following conditions was presented at a random time within a 4-s window: 2-s texture, 0.2-s sound (auditory cue) or water valve opening, and finishing with another 3-s of recording. In 20% of the trials, no stimulation was applied. Whisking and licking events were recorded during the course of the session.

To determine if neuronal activity was significantly modulated by texture or sound stimulations, we compared, for each neuron across trials, the average normalized fluorescence during 1 s before and after stimulus onset using a paired-sample t-test at a significance threshold of 5%. To account for noise in our data due to possible stimulation-induced movement artefacts, we performed the same test using the mRuby2 signal. None of the neuron showed a significant change in mRuby2 signal upon texture and sound stimulations.

We used a random forests machine-learning algorithm to decode behavioural features (licking and whisking rates) from the activity of single neurons. This procedure allowed us to categorize single neurons as either decoding whisking, licking or both. Given the slow kinetics of calcium transients captured by the GCaMP6s sensor, spiking rates were inferred from the $\Delta F/F_0$ trace and used as input to the algorithm, which allowed to temporally match behavioral event variations (i.e. whisking or licking rates) to neuronal activity. Firing rates at each imaging frame were inferred from normalized calcium traces ($\Delta F/F_0$) using a fast nonnegative deconvolution method (https://github.com/jovo/oopsi) (Vogelstein et al., 2010) with variable background fluorescence estimation and a K_d of 144 nM (Lock et al., 2015). In

order for the algorithm to capture differences in activity levels between neurons, all trial traces of all neurons recorded per mouse were concatenated before inferring spikes. To account for putatively preceding pre-motor and/or following sensory activity relative to behavioural events in S1, neuronal activity was shifted negatively and positively in time with a maximum shift of 500 ms. Eleven time bins of inferred firing rates (discretized in time bins of 100 ms) centred on zero time-shift were used to predict instantaneous behavioural features and composed a vector $X_i(t) = [x_i(t - 500\ ms), \dots\ x_i(t), \dots, x_i(t + 500\ ms)]$ where $x_i(t)$ represents the inferred firing rates of the *i*th neuron at zero time-shift. Licking and whisking rates were down sampled to 11.5 Hz in order to temporally match calcium imaging data. The *ranger* function of the ranger R package version 0.10.1 was used to construct regression forests, with each behavioural feature as dependent variable and the binned inferred firing rates of a given neuron as predictors. For each neuron, two regression forests were constructed, one to decode whisking and the other licking. Most arguments of the function were kept at default settings, except the following: the number of trees was set to 128, the minimum size of terminal nodes was set to 2, the number of predictor variables randomly sampled at each node split was set to the maximum between 1 or the third of the number of predictors, and the variable importance mode was set to “impurity”. To obtain a prediction for all trials, 5-fold cross-validation was applied by training the algorithm on 80% of the trials (i.e. training set) and evaluating it on the remaining 20% of the trials (i.e. test set). Since data acquisition was discretized by trial, for each cross-validation the training and test set trials were concatenated for training and prediction, respectively. For each neuron and for each behavioural feature, the decoding accuracy was assessed by computing the Pearson’s product-moment correlation coefficient between the observed and predicted behavioural event fluctuations. In order to get an estimate of the noise in the prediction levels, the same analysis was performed using the mRuby2 signal as a control. Neurons were classified as decoding a given behavioural feature if their Pearson’s correlation coefficient computed on the GCaMP6s signal was five standard deviations away from the mean of the Pearson’s correlation coefficients for all neurons computed on the mRuby2 signal. Neurons meeting these criteria for both whisking and licking were classified as decoding both behavioural features (Figure S4D,E).

Spontaneous activity correlation

Spontaneous calcium transients were recorded for 10 min after mice reached the expert level before and after texture reversal. Pairwise Pearson's correlation coefficients were calculated on the normalized calcium traces.

Discriminability index

The selectivity of each neuron was expressed by a discriminability index (DI) that was calculated based on neurometric functions using a receiver operating characteristic (ROC) analysis (Britten et al., 1992; Stuttgen and Schwarz, 2008; Takahashi et al., 2016). Normalized mean calcium responses ($\Delta F/F_0$) during the 2 s stimulus presentations in the P120 texture trials were compared to the P280 texture trials. ROC curves were generated by plotting, for all threshold levels, the fraction of P120 trials against the fraction of P280 trials for which the response exceeded threshold. Threshold levels were defined as a linear function from the minimal to the maximal calcium responses. DI was computed from the area under the ROC curve (AUC) as follows: $DI = (AUC - 0.5) \times 2$. DI values vary between -1 and 1. Positive or negative values indicate larger or smaller responses to P120 than to P280 texture presentations, respectively. Statistical significance of the measured DI value was assessed by performing a permutation test, from which a sampling distribution was obtained by shuffling the texture labels of the trials 10,000 times. The measured DI was considered significant when it was outside of the 2.5th – 97.5th percentiles interval of the sampling distribution. For the choice index, the same calculation was performed, with the difference that trials where the animal licked during the answer period were compared to trials with no lick.

Error history

Error history for each neuron was calculated as the normalized difference between the average calcium response during 'hits' ($\bar{R}$) over a sliding window of 200 trials as follows:

$$Error\ history(t) = \frac{\bar{R}_{hit(post\ FA)}(t-100{:}\,t+100) - \bar{R}_{hit(post\ hit)}(t-100{:}\,t+100)}{\bar{R}_{hit}(t-100{:}\,t+100)}$$

where $R_{hit(post\ FA)}$ is a 'hit' response preceded by a 'FA', $R_{hit(post\ hit)}$ is a 'hit' response preceded by another 'hit', R_{hit} is any 'hit' response and t is the trial position along the learning sequence. To estimate the fraction of neurons with an error history above chance over reversal learning, for each neuron, all 'hit' responses within each window of 200 trials were randomly permuted, replacing $R_{hit(post\ FA)}$, $R_{hit(post\ hit)}$ and R_{hit} in their respective trial positions and an error history value was calculated. This process was repeated 1000 times to obtain 95% confidence intervals for each observed error history value.

Immunohistochemistry

Post-hoc immunohistochemistry of GABA was performed on mRuby2/GCaMP6s expressing neurons. 100-µm thick tangential sections were produced using a vibratome (Leica VT 1000). The sections were washed 3 × 3 minutes in 500 µl Tris-buffered saline (0.1 M Tris, 150 mM NaCl) containing 0.1% Tween (TBST), then pre-treated with TBST and 0.1% Triton-X for 20 minutes followed by a 3 × 3 minute TBST wash. They were blocked in 300 µl TBST containing 10% normal donkey serum (ab7475, Abcam) for 1 hour and incubated with mouse anti-GABA antibody (ab86186, Abcam) diluted 1:500 for 72 hours at 4°C. After another 5 × 3 minutes wash in TBST they were incubated in 300 µl of donkey anti-mouse antibody coupled to Alexa Fluor 647 (A32787, Thermo Fisher Scientific) diluted 1:200 in TBST for one hour at room temperature. Finally, they were washed 10 × 3 minutes in TBST and then for one hour in PBS before being mounted onto glass slides. We applied Fluoroshield mounting medium with DAPI (Abcam) before applying the coverslip. The sections were imaged using a Zeiss Confocal LSM800 Airyscan.

S1 inactivation and whisker trimming

To inactivate S1, the GABA (G-aminobutyric acid) agonist muscimol was injected in expert mice. During the test session, high baseline performance (>70%) was first recorded for 100 trials before the injection was performed. Under light anesthesia (4% isoflurane at 0.4 l/min), a small hole was drilled through the imaging window above the previously mapped C2 barrel column to provide access to a glass pipette through which 300 nl of Muscimol (Bodipy-TMR-

X, 5mM in cortex buffer with 5% DMSO, Thermo Fisher Scientific) was injected at 300 and 500 µm below the pia. Mice were left to recover for 45 min and their behavioural performance was then assessed for another 100 trials. For the whisker trimming experiment, a similar baseline performance was first recorded for 100 trials before trimming the whiskers on the side of the snout contralateral to the texture presentations, and tested the performance for 50 trials. This ensured that trimming itself did not alter performance. Then, the whiskers that were in contact with the textures (ipsilateral to the texture presentation side) were trimmed, and the effect on task performance was measured for another 50 trials.

Statistics

All statistics were performed using MATLAB. For all figures, significance levels were denoted as $*P<0.05$, $**P<0.01$, $***P<0.001$. No statistical methods were used to estimate sample sizes. Non-parametric tests were used for sample sizes smaller than 15.

Supplemental Information

Dynamic encoding of perceptual features in primary somatosensory cortex upon reversal learning

Ronan Chéreau, Tanika Bawa, Leon Fodoulian, Alan Carleton, Stéphane Pagès, Anthony Holtmaat

Figure S1 related to Figure 1

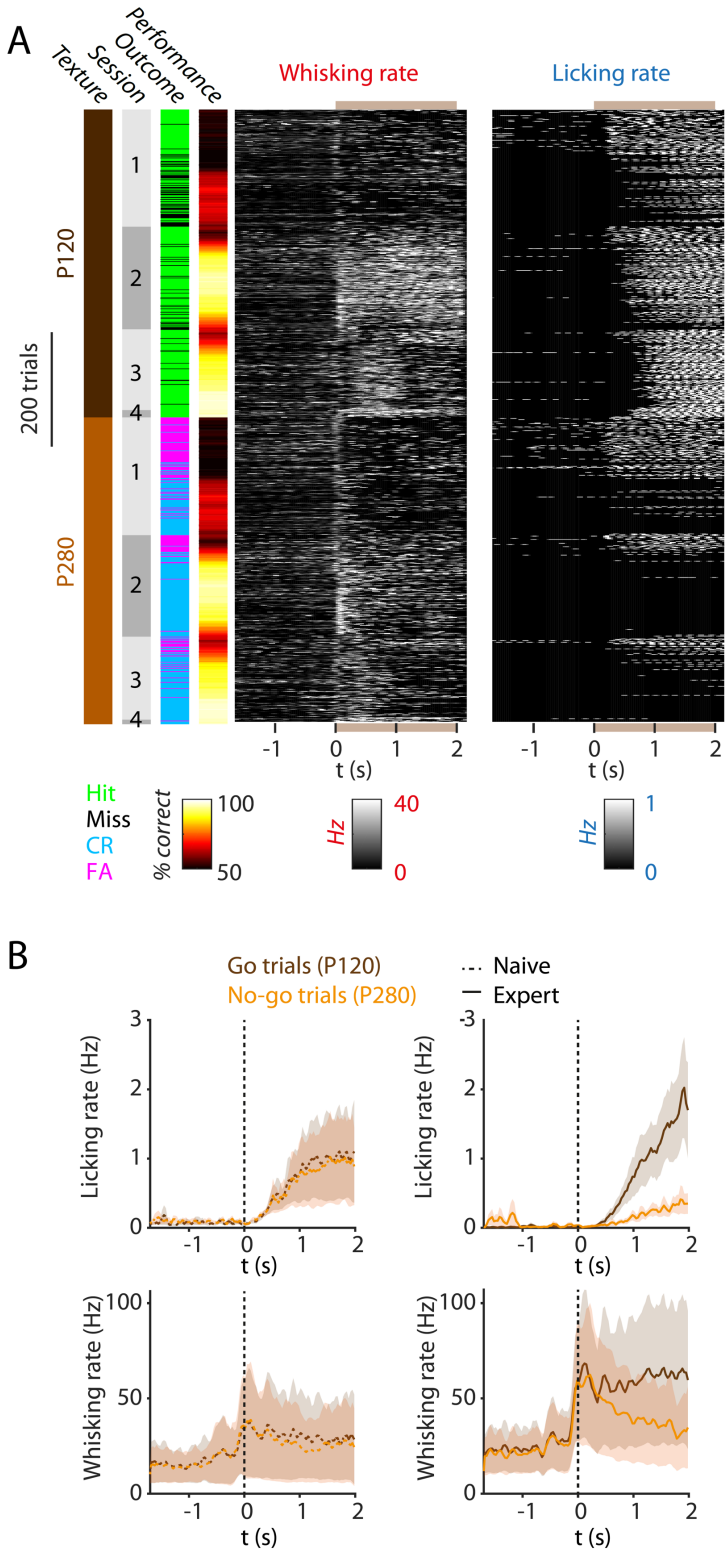

Figure S1. Evolution of whisking and licking strategies across learning.

(A) Single trial whisking and licking rates for a mouse during training, aligned to texture presentation onset (t=0s). Trials were sorted by texture (P120, Go stimulus and P280, No-go

stimulus). For each trial, the corresponding session number, trial outcome and performance are shown. The shaded brown area indicates the time during which the texture is presented (0-2 s).

(B) Average time course of licking (top) and whisking (bottom) rates aligned to texture presentation onset (vertical dotted line, t=0s) in a naïve session (dashed lines) vs. an expert session (solid lines) for 8 mice. While the licking rate is similar between textures when mice are naïve to the task, it increases for the Go texture and decreases for the No-go texture when mice become experts. The sampling strategy increases with learning for both textures indicated by the increase in whisking rate around the onset of texture presentation. The whisking rates remained elevated upon presentation of the Go texture, which may indicate an increase in arousal state and/or whisker movement induced by increased licking. Shaded areas represent s.e.m.

Figure S2 related to Figure 1

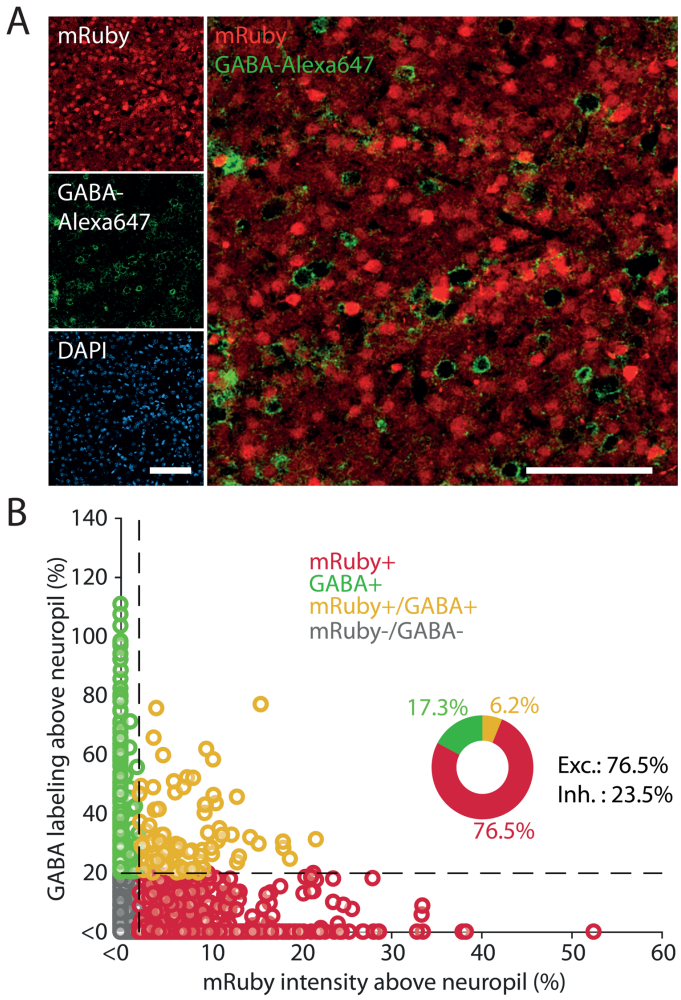

Figure S2. Fraction of mRuby2-P2A-GCaMP6s-expressing neurons producing GABA.

(A) Example of mRuby2-expression (red), GABA immunostaining (green), and DAPI labeling (blue) of L2/3 neurons in S1 on the left and a merge of the mRuby2 expression and GABA immunostaining. Note that the vast majority of GABA-positive neurons is negative for GABA and vice versa. Scale bars: 100 μm.

(B) Scatter plot comparing mRuby2 fluorescence intensity compared to GABA labeling fluorescence intensity both normalized to neuropil (each data point represents a neuron, *N*=2276 neurons measured from 5 mice). Normalized signal intensity difference above 20% was considered as labeled. Pie chart shows the percentage of neurons that were mRuby2 positive only (red), GABA positive only (green), and mRuby2+GABA positive (yellow). Only 6.2% of mRuby2-positive neurons were GABA positive, indicating that the vast majority of mRuby2/GCaMP6s-expressing neurons are pyramidal neurons.

Figure S3 related to Figure 1

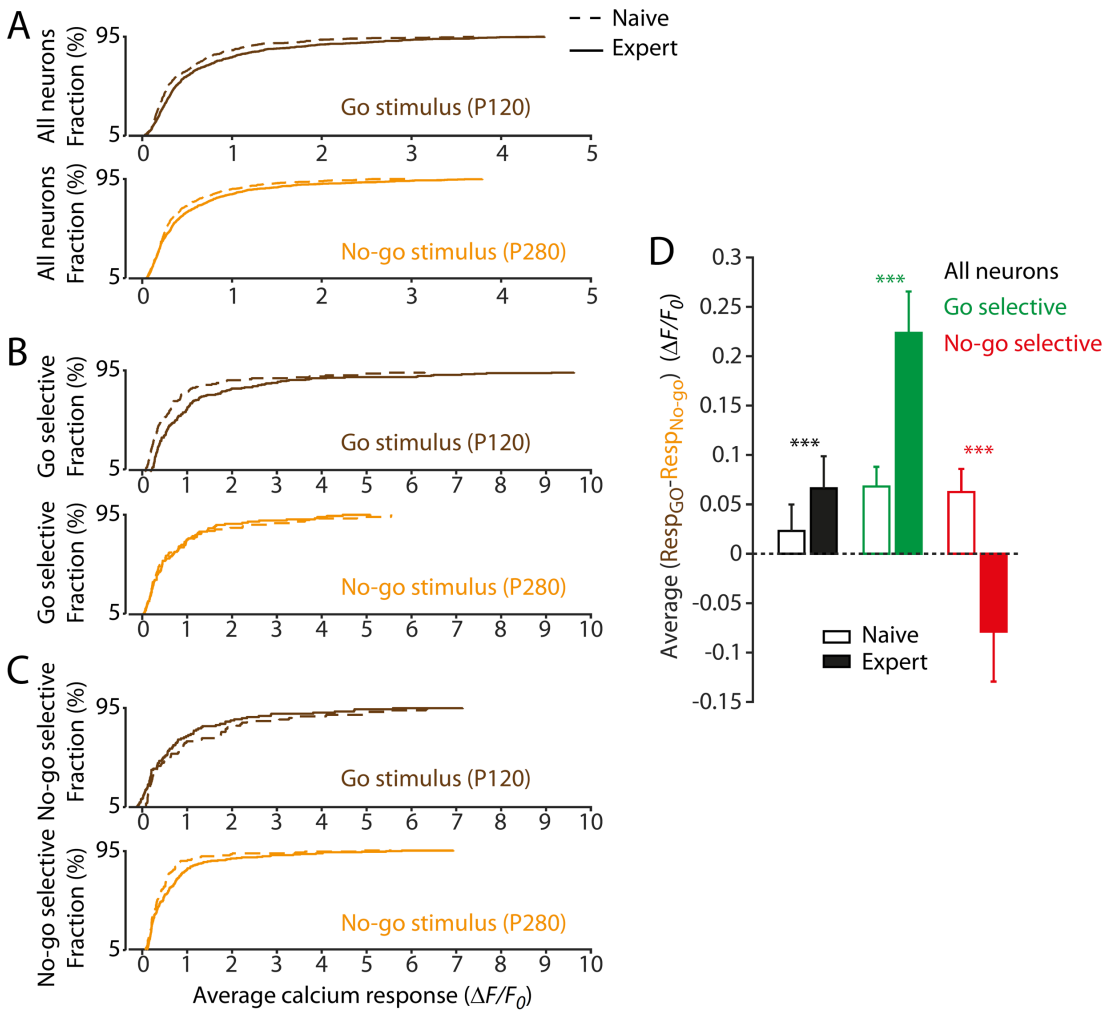

Figure S3. Evolution of calcium responses across learning.

(A) Cumulative distributions of the mean calcium responses per neuron for the Go stimulus (P120) and No-go stimulus (P280) in naïve (dashed line) and expert sessions (continuous line). For presentation purposes, the top and bottom 5% of the population measurements are not shown.

(B) Same for Go selective neurons only.

(C) Same for No-go selective neurons only.

(D) Average difference between normalized calcium responses of the Go and No-go stimuli for all neurons (black), Go selective neurons (green), No-go selective neurons (green) in naïve and expert sessions (all neurons: *N*=600 neurons in naïve and *N*=931 neurons in expert, Wilcoxon rank sum test naïve vs. expert, ****P*=9e-5; Go selective neurons: *N*=116 neurons in naïve and *N*=193 neurons in expert, Wilcoxon rank sum test naïve vs. expert, ****P*=5.8e-10; No-go

selective neurons: *N*=71 neurons in naïve and *N*=98 neurons in expert, Wilcoxon rank sum test naïve vs. expert, ****P*=3.6e-4).

Figure S4 related to Figure 1

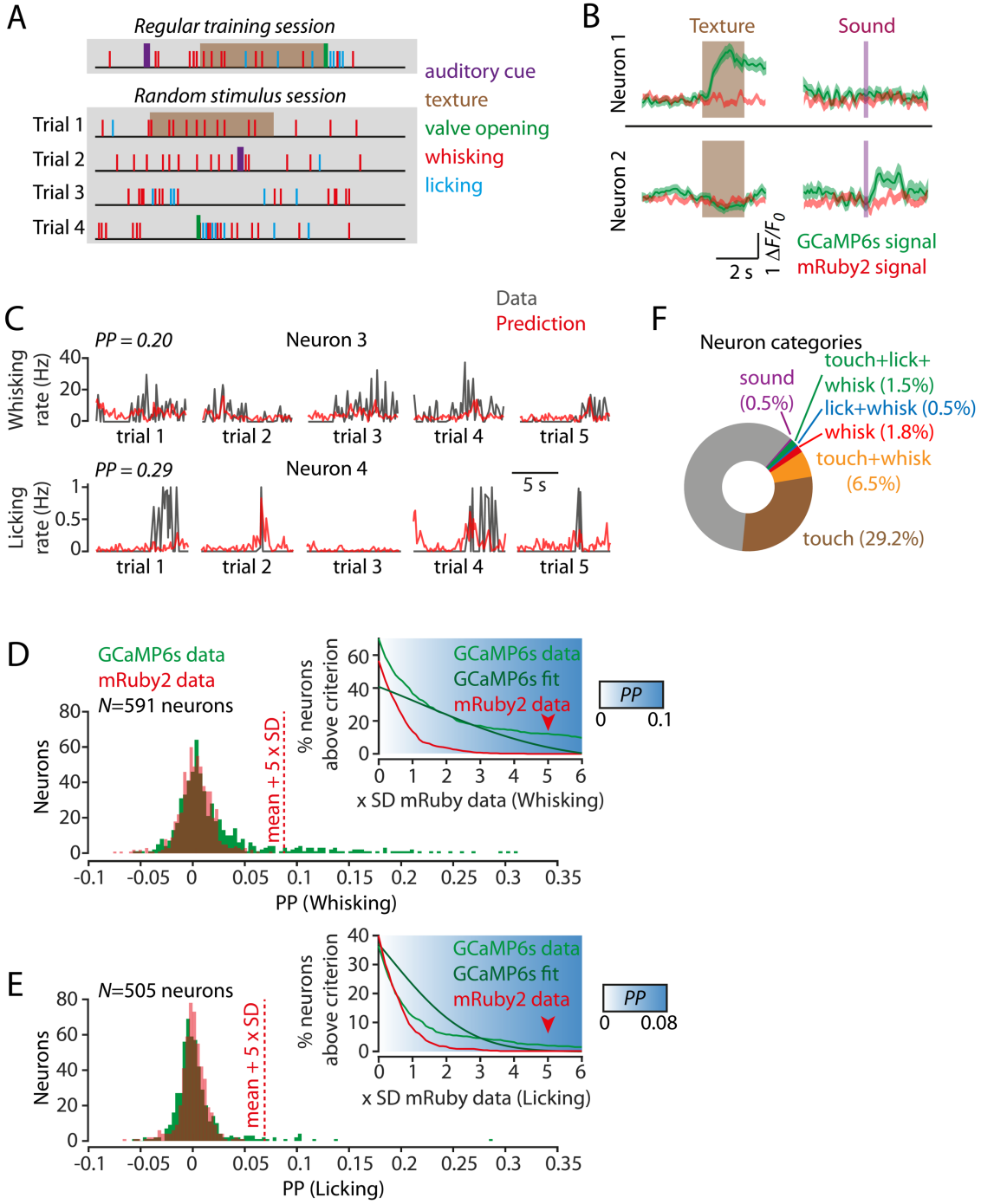

Figure S4. Distributions of prediction powers for decoding whisking and licking rates.

(A) In a regular training session, stimuli and behavioral are tightly coupled, complicating the interpretation of the exact cause of the neuronal response unclear. Therefore, prior to training, a random stimulus session was performed during which the 2-s texture presentation, the sound (auditory cue) or the water reward were delivered in a pseudo-random fashion while recording spontaneous whisking and licking rates. Here shown, various trial examples in a random stimulus session as compared to a regular training session.

(B) Examples of two neurons significantly responding to texture or sound (black: GCaMP6s, red: mRuby2). Shaded areas represent s.e.m.

(C) Examples of two neurons decoding behavioral features as predicted by a Random-Forest decoder that was trained on inferred neuronal spike rates, and whisking and licking rates. Plots of whisking and licking rates (Data) superimposed with the scaled model prediction (Prediction). PP, prediction power.

(D and E) Prediction power (PP) distributions for whisking rates (D) and licking rates (E) reflecting the similarity between the behavioral data and the Random Forest decoding model based on the activity of individual neurons (inferred spikes; STAR Methods). The signal intensity fluctuation recorded from the mRuby2 channel was used as control. A neuron with a PP above 5 standard deviations (SD) of the control distribution was considered significantly predictive of the behavioral feature (*N*=591 neurons, 55 neurons predicted whisking rate and 10 neurons predicted licking behavior). The inset shows the percentage of neurons above a variable threshold criterion from the measured GCaMP6s PP values (light green) compared to normally distributed GCaMP6s PP values (dark green) and the measured mRuby2 PP values (red). This analysis shows that the fraction of neurons detected as predictive for whisking and licking is higher than a fraction obtained by chance for criterion values above 5 SD of the mRuby2 distribution. The variation of predictive fractions for whisking and licking around the chosen criterion of 5 SD (from 4 to 6 SD) of the mRuby2 distribution (indicated by arrows) is negligible.

(F) Percentage of neuron categories determined by combining outcomes of the response analysis in (B) and the behavioral decoding analysis in (C) (*N*=505 neurons, 8 mice).

Figure S5 related to Figure 2

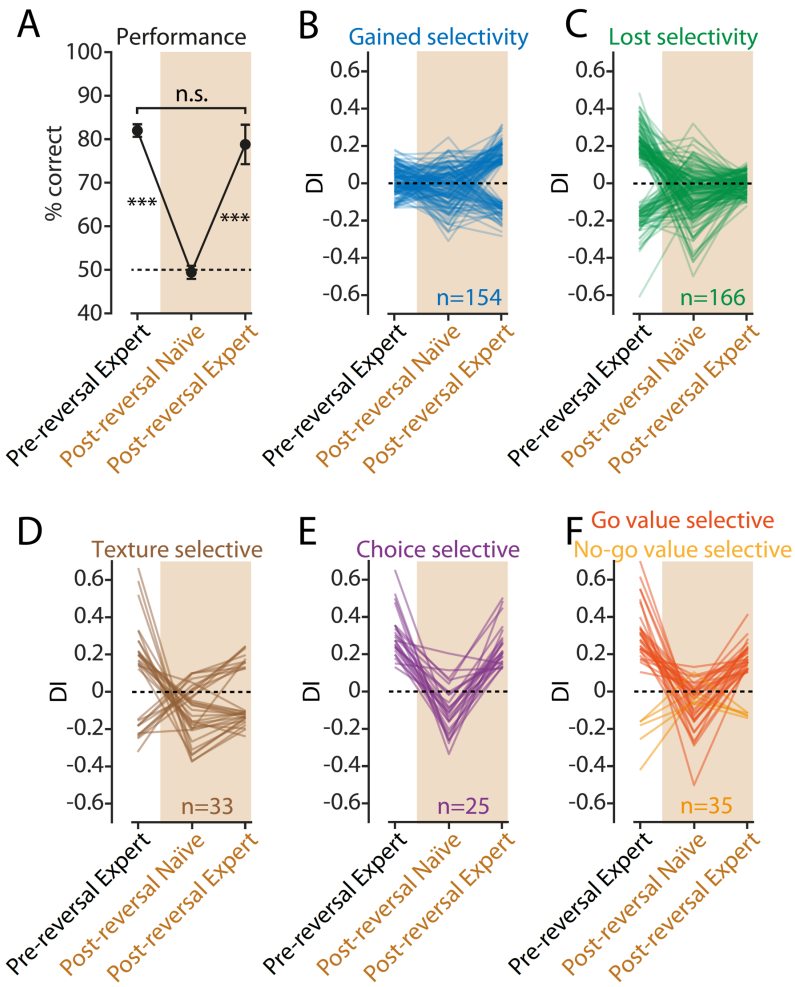

Figure S5. Performance pre and post-reversal learning and selectivity of the neuronal classes.

(A) Average performance of mice in pre-reversal expert, post-reversal naïve and post-reversal expert sessions (*N*=12 mice, rank sum test, pre-reversal expert vs. post-reversal naïve: ***P=1.8e-4, post-reversal naïve vs. post-reversal expert: ***P=3.3e-4, pre-reversal expert vs. post-reversal expert: P=0.62).

(B-F) Evolution of DI of all neuronal classes in the corresponding pre-reversal expert, post-reversal naïve and post-reversal expert sessions.

Figure S6 related to Figure 2

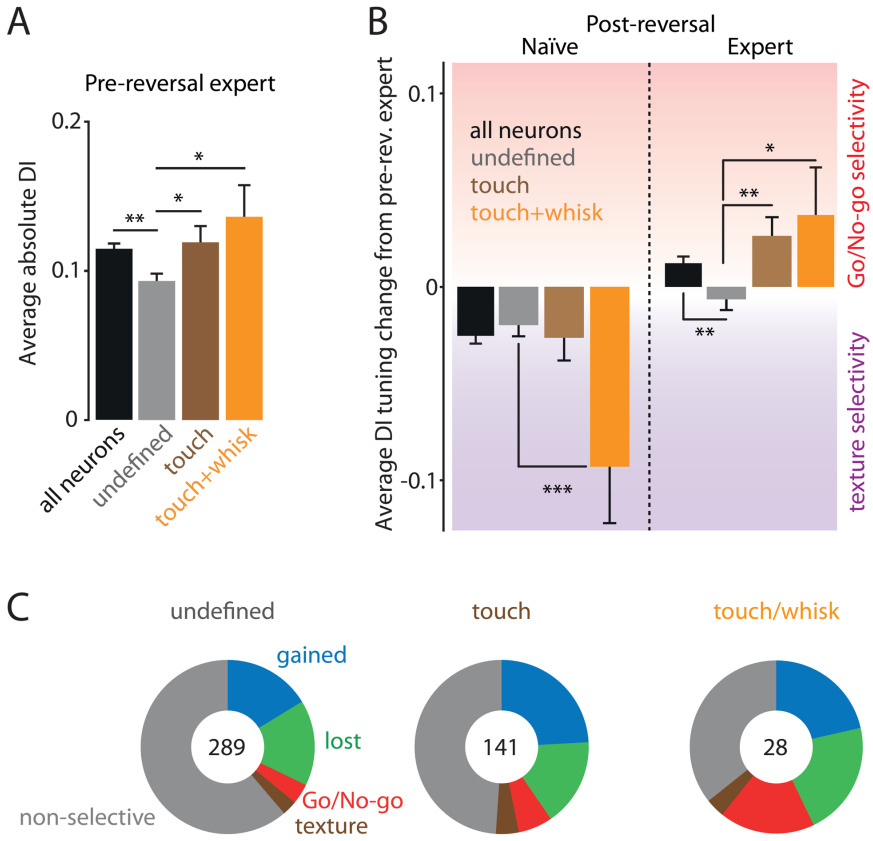

Figure S6. Changes in selectivity across neuronal categories.

(A) Average absolute DI in the pre-reversal expert session for all neurons pooled and the main neuronal categories. The selectivity is highest for touch and touch/whisk categories (two-sample t-test vs. undefined category, *N*=363: all neurons, *N*=863, ***P*=0.0014; touch neurons, *N*=174, **P*=0.0132; touch/whisk neurons, *N*=35, **P*=0.0128).

(B) Average DI change in post-reversal naïve and expert mice relative to their absolute pre-reversal DI. After texture inversion, selectivity of the touch/whisk category significantly reverts its sign, indicating that neurons preferentially respond to the previous Go texture P120 (two-sample t-test vs. undefined category, *N*=289: touch/whisk neurons, *N*=28, ****P*<0.001). When mice become experts again after texture inversion, the selectivity of touch and touch/whisk neurons tends to revert back to the new Go stimulus (P280 texture; two-sample t-test vs. undefined category, *N*=289: all neurons, *N*=682, ***P*=0.007; touch neurons, *N*=141, ***P*=0.002; touch/whisk neurons, *N*=28, **P*=0.025). Other categories of neurons are not shown and no tests were performed due to their small sample size (touch/lick/whisk, *N*=6; whisk, *N*=8; lick/whisk, *N*=2; sound, *N*=3).

(C) Pie charts showing the fractions of non-selective neurons, neurons that gained or lost selectivity, and neurons that were stimulus or texture selective, for the three main categories. The number of neurons that are Go/No-go selective is higher in the touch and touch/whisk categories.

Figure S7 related to Figure 2

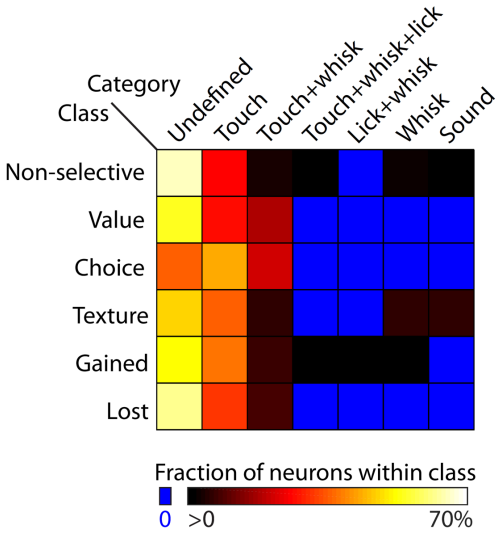

Figure S7. Fraction of neuronal categories within all identified functional classes.

Table showing the percentage of neurons of each categories within the different classes of neurons. Notably, none of the choice and value neurons was categorized as predictive for licking (number of neurons per classes for which corresponding categories were defined prior training: non-selective 271, value neurons 11, choice 14, texture 17, gained 90, lost 75). Fractions of 0% are represented in blue in the look up table.

Figure S8 related to Figure 3

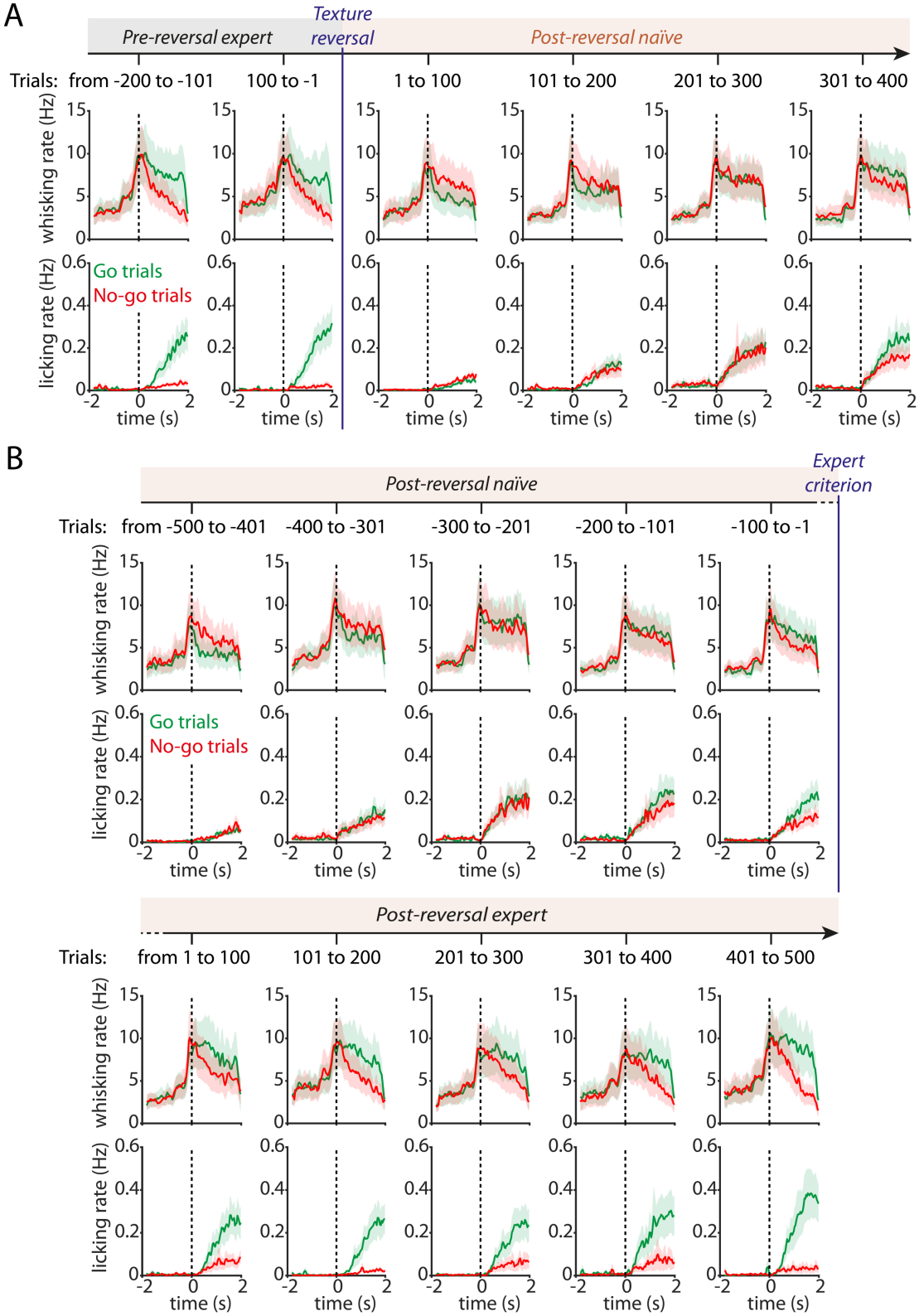

Figure S8. Whisking and licking rates pre and post texture reversal.

(A) Time course of whisking (top) and licking (bottom) rates aligned to texture presentation onset (dotted line) across two intervals of 100 trials before texture reversal (P120 texture was

presented on the Go trials in green, P280 on the No-go trials in red) and the first four intervals after reversal (P280 texture was presented on the Go trials in green, P120 on the No-go trials in red). We observed a decrease in whisking and licking rates for the Go trials immediately after texture reversal. This may reflect a loss of commitment to the task due a decline in the number of water rewards due to increased FAs and Misses. Whisking and licking rates gradually increase 200 trials after texture reversal as mice start to randomly lick again upon presentation of the Go and No-go stimulus.

(B) Average time course of whisking (top) and licking (bottom) rates aligned to texture presentation (dotted line) from five bins of 100 trials before and after the expert criterion (70% performance). The whisking strategy only moderately changed over the 300 trials prior to reaching the expert criterion. When mice became experts, the whisking rates remained elevated upon the Go stimulus presentations, which may reflect whisker movements induced by the motor action of licking, similar to Figure S1. Shaded areas represent s.e.m.

Figure S9 related to Figure 3

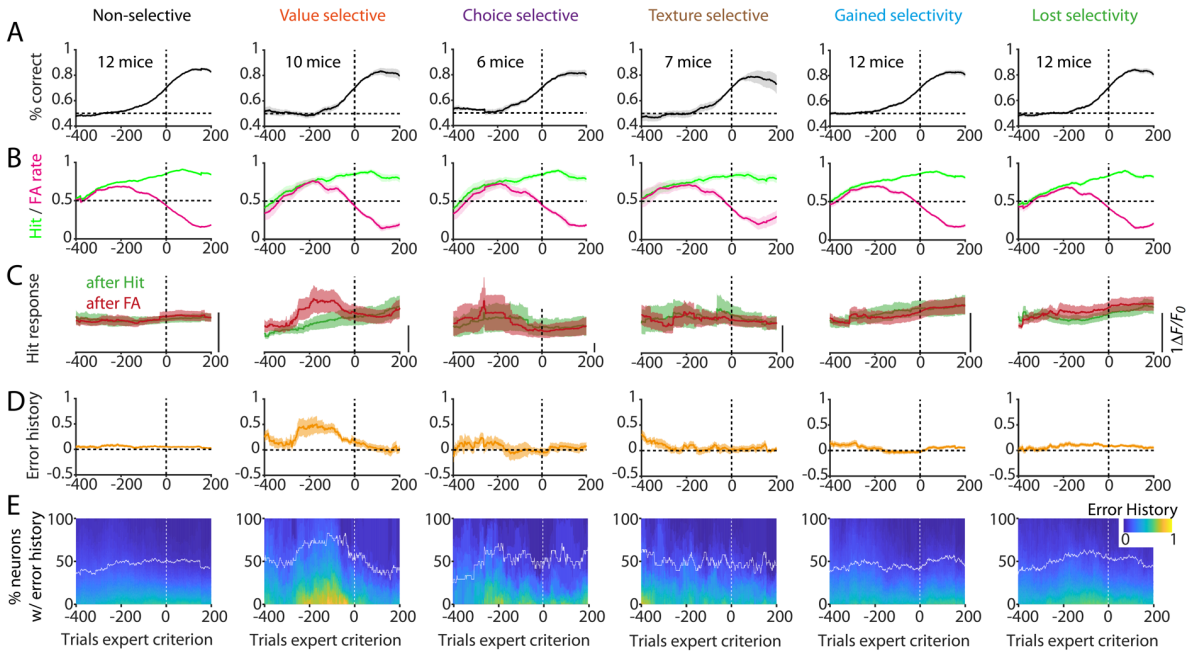

Figure S9. Evolution of error history for all identified neuronal classes during reversal learning.

(A) Average performance during reversal learning. Performance curves were realigned to expert criterion (Non-selective neurons, *N*=412, 12 mice; Value neurons, *N*=35, 10 mice; Choice neurons, *N*=24, 6 mice; Texture neurons *N*=33, 7 mice, Neurons gaining selectivity after reversal *N*=144, 12 mice; Neurons losing selectivity after reversal *N*=152, 12 mice).

(B) Corresponding 'Hit' and 'FA' rates.

(C) Average 'Hit' response after a 'Hit' trial (green) and after a 'FA' trial (red).

(D) Calculated mean error history. (A to D) Shaded areas represent 95% CI.

(E) Percentage of neurons per class with significant error history rate (white line) above the cumulative distributions of error history values.

Figure S10 related to Figure 3

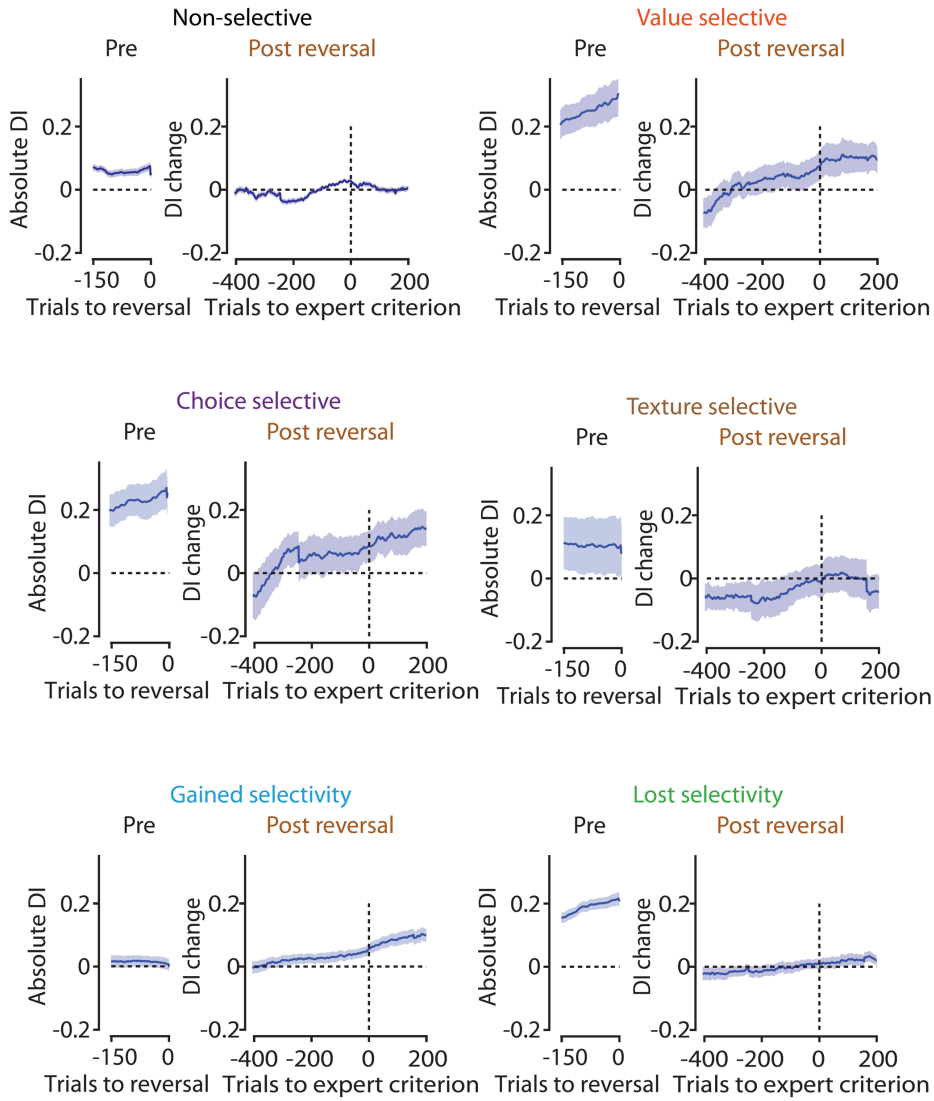

Figure S10. Evolution of selectivity for all identified neuronal classes.

DI profiles for the various classes of neurons. For each panel, the graph on the left shows the average absolute DI across mice over the last 150 trials pre-reversal. The right graph shows the average change in DI from the absolute DI on the left, post-reversal, and realigned across mice to the expert criterion (Non-selective neurons, *N*=412, 12 mice; Value neurons, *N*=35, 10 mice; Choice neurons, *N*=24, 6 mice; Texture neurons *N*=33, 7 mice, Neurons gaining selectivity after reversal *N*=144, 12 mice; Neurons losing selectivity after reversal *N*=152, 12 mice).

REFERENCES

Akrami, A., Kopec, C.D., Diamond, M.E., and Brody, C.D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature 554, 368-372.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12, 4745-4765.

Chen, J.L., Carta, S., Soldado-Magraner, J., Schneider, B.L., and Helmchen, F. (2013a). Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. Nature 499, 336-340.

Chen, J.L., Margolis, D.J., Stankov, A., Sumanovski, L.T., Schneider, B.L., and Helmchen, F. (2015). Pathway-specific reorganization of projection neurons in somatosensory cortex during learning. Nat Neurosci 18, 1101-1108.

Chen, J.L., Pfaffli, O.A., Voigt, F.F., Margolis, D.J., and Helmchen, F. (2013b). Online correction of licking-induced brain motion during two-photon imaging with a tunable lens. J Physiol 591, 4689-4698.

Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V.*, et al.* (2013c). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295-300.

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron 56, 43-57.

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1, 1-47.

Francis, N.A., Winkowski, D.E., Sheikhattar, A., Armengol, K., Babadi, B., and Kanold, P.O. (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. Neuron 97, 885-897 e886.

Garion, L., Dubin, U., Rubin, Y., Khateb, M., Schiller, Y., Azouz, R., and Schiller, J. (2014). Texture coarseness responsive neurons and their mapping in layer 2-3 of the rat barrel cortex in vivo. Elife 3, e03405.

Goltstein, P.M., Coffey, E.B., Roelfsema, P.R., and Pennartz, C.M. (2013). In vivo two-photon Ca2+ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. J Neurosci 33, 11540-11555.

Guo, Z.V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J.T., Feng, G., and Svoboda, K. (2014). Flow of cortical activity underlying a tactile decision in mice. Neuron 81, 179-194.

Hattori, R., Danskin, B., Babic, Z., Mlynaryk, N., and Komiyama, T. (2019). Area-Specificity and Plasticity of History-Dependent Value Coding During Learning. Cell.

Holtmaat, A., Bonhoeffer, T., Chow, D.K., Chuckowree, J., De Paola, V., Hofer, S.B., Hubener, M., Keck, T., Knott, G., Lee, W.C.*, et al.* (2009). Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. Nat Protoc 4, 1128-1144.

Holtmaat, A., and Caroni, P. (2016). Functional and structural underpinnings of neuronal assembly formation in learning. Nat Neurosci 19, 1553-1562.

Hubel, D.H., and Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. J Physiol 148, 574-591.

Hwang, E.J., Dahlen, J.E., Mukundan, M., and Komiyama, T. (2017). History-based action selection bias in posterior parietal cortex. Nat Commun 8, 1242.

Kerlin, A.M., Andermann, M.L., Berezovskii, V.K., and Reid, R.C. (2010). Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. Neuron 67, 858-871.

Lacefield, C.O., Pnevmatikakis, E.A., Paninski, L., and Bruno, R.M. (2019). Reinforcement Learning Recruits Somata and Apical Dendrites across Layers of Primary Sensory Cortex. Cell Rep 26, 2000-2008 e2002.

Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. Nature 532, 459-464.
Lock, J.T., Parker, I., and Smith, I.F. (2015). A comparison of fluorescent Ca(2)(+) indicators for imaging local Ca(2)(+) signals in cultured cells. Cell Calcium 58, 638-648.
Makino, H., Hwang, E.J., Hedrick, N.G., and Komiyama, T. (2016). Circuit Mechanisms of Sensorimotor Learning. Neuron 92, 705-721.
Pecka, M., Han, Y., Sader, E., and Mrsic-Flogel, T.D. (2014). Experience-dependent specialization of receptive field surround for selective coding of natural scenes. Neuron 84, 457-469.
Peron, S.P., Freeman, J., Iyer, V., Guo, C., and Svoboda, K. (2015). A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. Neuron 86, 783-799.
Pho, G.N., Goard, M.J., Woodson, J., Crawford, B., and Sur, M. (2018). Task-dependent representations of stimulus and choice in mouse parietal cortex. Nat Commun 9, 2596.
Pi, H.J., Hangya, B., Kvitsiani, D., Sanders, J.I., Huang, Z.J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. Nature 503, 521-524.
Pnevmatikakis, E.A., and Giovannucci, A. (2017). NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. J Neurosci Methods 291, 83-94.
Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. Biomed Eng Online 2, 13.
Pooresmaeili, A., Poort, J., and Roelfsema, P.R. (2014). Simultaneous selection by object-based attention in visual and frontal cortex. Proc Natl Acad Sci U S A 111, 6467-6472.
Poort, J., Khan, A.G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G.B., Mrsic-Flogel, T.D., and Hofer, S.B. (2015). Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. Neuron 86, 1478-1490.
Rose, T., Jaepel, J., Hubener, M., and Bonhoeffer, T. (2016). Cell-specific restoration of stimulus preference after monocular deprivation in the visual cortex. Science 352, 1319-1322.
Safaai, H., von Heimendahl, M., Sorando, J.M., Diamond, M.E., and Maravall, M. (2013). Coordinated population activity underlying texture discrimination in rat barrel cortex. J Neurosci 33, 5843-5855.
Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature 412, 549-553.
Stiebler, I., Neulist, R., Fichtel, I., and Ehret, G. (1997). The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation. J Comp Physiol A 181, 559-571.
Stuttgen, M.C., and Schwarz, C. (2008). Psychophysical and neurometric detection performance under stimulus uncertainty. Nat Neurosci 11, 1091-1099.
Sul, J.H., Kim, H., Huh, N., Lee, D., and Jung, M.W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron 66, 449-460.
Takahashi, N., Oertner, T.G., Hegemann, P., and Larkum, M.E. (2016). Active cortical dendrites modulate perception. Science 354, 1587-1590.
Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., and Paninski, L. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. J Neurophysiol 104, 3691-3704.
Watabe-Uchida, M., Eshel, N., and Uchida, N. (2017). Neural Circuitry of Reward Prediction Error. Annu Rev Neurosci 40, 373-394.
Wimmer, R.D., Schmitt, L.I., Davidson, T.J., Nakajima, M., Deisseroth, K., and Halassa, M.M. (2015). Thalamic control of sensory selection in divided attention. Nature 526, 705-709.
Yang, H., Kwon, S.E., Severson, K.S., and O'Connor, D.H. (2016). Origins of choice-related activity in mouse somatosensory cortex. Nat Neurosci 19, 127-134.