

LETTERS

Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice

Takaki Komiyama¹, Takashi R. Sato¹, Daniel H. O'Connor¹, Ying-Xin Zhang^{1†}, Daniel Huber¹, Bryan M. Hooks¹, Mariano Gabitto^{2†} & Karel Svoboda¹

Cortical neurons form specific circuits¹, but the functional structure of this microarchitecture and its relation to behaviour are poorly understood. Two-photon calcium imaging can monitor activity of spatially defined neuronal ensembles in the mammalian cortex^{2–5}. Here we applied this technique to the motor cortex of mice performing a choice behaviour. Head-fixed mice were trained to lick in response to one of two odours, and to withhold licking for the other odour^{6,7}. Mice routinely showed significant learning within the first behavioural session and across sessions. Microstimulation^{8,9} and trans-synaptic tracing^{10,11} identified two non-overlapping candidate tongue motor cortical areas. Inactivating either area impaired voluntary licking. Imaging in layer 2/3 showed neurons with diverse response types in both areas. Activity in approximately half of the imaged neurons distinguished trial types associated with different actions. Many neurons showed modulation coinciding with or

preceding the action, consistent with their involvement in motor control. Neurons with different response types were spatially intermingled. Nearby neurons (within $\sim 150\mu\text{m}$) showed pronounced coincident activity. These temporal correlations increased with learning within and across behavioural sessions, specifically for neuron pairs with similar response types. We propose that correlated activity in specific ensembles of functionally related neurons is a signature of learning-related circuit plasticity. Our findings reveal a fine-scale and dynamic organization of the frontal cortex that probably underlies flexible behaviour.

We trained head-fixed mice in a lick/no-lick task^{6,7} (Fig. 1a, b). Mice were rewarded with water for licking in response to odour A (hits). Correct withholding of licking to odour B (correct rejections, CR) was not rewarded. Licking to odour B (false alarms) and withholding of licking to odour A (misses) were punished with a longer

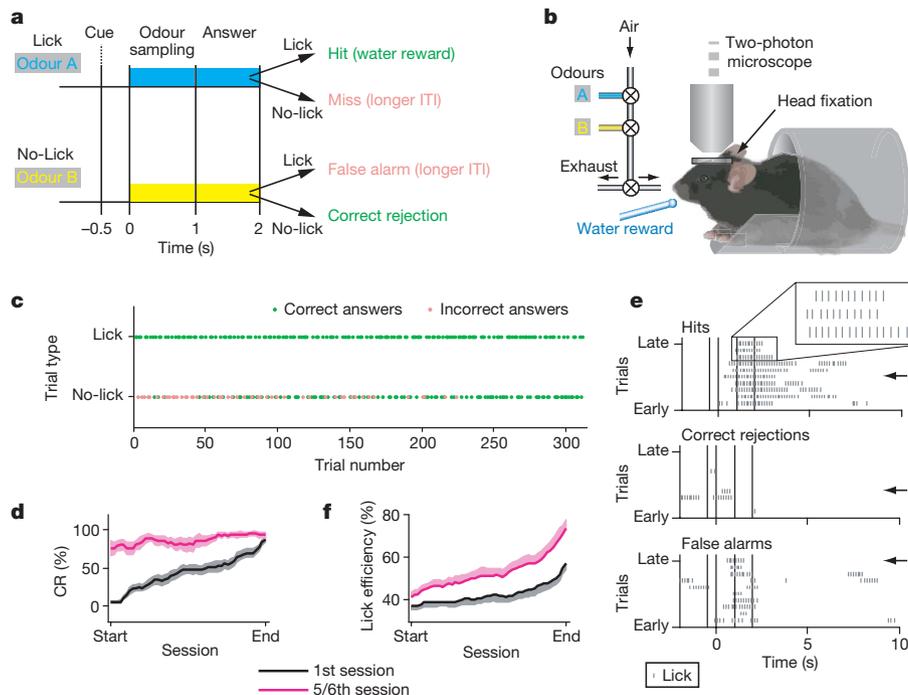


Figure 1 | Lick/no-lick task. **a**, Task structure. Odour was applied 0.5 s after a cue (clicking of the final valve, see Methods). After 1 s of odour sampling, licking was scored within a 1-s-long answer period. ITI, inter-trial interval. **b**, Schematic of a head-fixed mouse under a two-photon microscope performing the task. **c**, Performance during a first session. **d**, Average learning curves for the first ($n = 27$) and fifth/sixth ($n = 10$) sessions. CR

(%) is the percentage of correct rejections in no-lick trials, over 21 trials. **e**, Licking behaviour in ten sample trials of each trial type, evenly selected across the first session. Arrows, learning criterion (60% CR). Each tick represents a lick. Vertical lines correspond to **a**. Note the regularity in each lick bout (inset). **f**, Lick efficiency, the percentage of all licks during the answer period for hit trials, averaged over 21 trials.

¹Janelia Farm Research Campus, HHMI, Ashburn, Virginia 20147, USA. ²HHMI and Departments of Neurobiology and Neurosciences, University of California at San Diego, La Jolla, California 92093, USA. †Present addresses: The Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, USA (Y.-X.Z.); HHMI and Department of Biochemistry and Molecular Biophysics, Columbia College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA (M.G.).

inter-trial interval. All mice ($n = 27$) initially licked in response to both odours (Methods), but they learned to suppress licking in no-lick trials during the first session (Fig. 1c). Because misses were rare ($<0.4\%$), the fraction of correct rejections in no-lick trials accurately represented behavioural performance (Fig. 1d, black and Supplementary Fig. 1a). Mice reached a learning criterion (60% CR) on average within 195 trials (range 61–359, out of 200–450 trials per session; Supplementary Fig. 1b). This rapid learning allowed studying task-related activity before overtraining, in contrast to most other studies of choice-based behaviour^{12,13}. In subsequent behavioural sessions mice reached the learning criterion significantly faster (fifth/sixth session: mean, 45 trials; range, 1–118; $n = 10$) (Fig. 1d, magenta), demonstrating learning within and across sessions.

Licking consisted of stereotyped bouts of tongue protractions at 7.1 ± 0.7 Hz (Fig. 1e and Supplementary Fig. 2), similar to other species¹⁴. Rhythmic licking is generated by a central pattern generator in the brainstem, which is under descending control from the motor cortex¹⁴. With improving performance, licking became increasingly limited to the answer period (Fig. 1e, f and Supplementary Fig. 2a).

We identified putative motor cortex areas controlling licking using several approaches. First, optical⁹ and electrical microstimulation revealed an anterior-lateral area (centred 2.0 mm lateral and 2.4 mm anterior to bregma) where movement of the tongue, jaw and lip can be reliably evoked (Supplementary Fig. 3a, b)⁸.

However, the relationship between movements evoked by microstimulation and connectivity can be complex¹⁵, especially at stronger stimulus intensities¹⁶ (tongue movement required stronger stimuli than whiskers and forelimbs; Supplementary Fig. 3c). As a second method, we used Bartha strain pseudorabies virus, a retrograde trans-synaptic tracer^{10,11}. Four days after injecting green fluorescent protein (GFP)-expressing virus¹⁷ into the tongue, we found clusters of GFP-positive layer 5 pyramidal neurons bilaterally in the motor cortex, 1.2 mm lateral and 0.3 mm anterior to bregma (Supplementary Fig. 4). Stimulating this area typically caused movement of whiskers and/or forelimb; this does not preclude a role in the control of licking, because the effects of motor cortex activity on targets in the brainstem could be controlled, for example, by gating¹². Labelled cells were never found in the area identified by microstimulation, indicating that the areas defined by microstimulation (ALM, for anterior-lateral motor) and tracing (PMM, for posterior-medial motor) are distinct (Fig. 2a).

Two series of experiments confirmed the involvement of the ALM and PMM in the control of licking. First, anterograde tracing showed that both areas project to the brainstem reticular formation, where rhythmic licking is generated¹⁴ (Supplementary Fig. 5). Second, we measured the effects of local inactivation¹⁸ by injections of muscimol, a potent agonist of GABA (γ -aminobutyric acid) receptors, on licking in an odour detection task. Inactivation of either area, but not the somatosensory cortex ('1') or an anterior-medial area ('2'), impaired licking (Fig. 2b, Supplementary Fig. 6 and Supplementary Movies). Activity in both areas is therefore required for voluntary licking.

We recorded the ensemble activity (ALM, $n = 5$; PMM, $n = 5$) during the first learning session of the lick/no-lick task. One of the two areas was loaded with the Ca^{2+} -sensitive dye OGB-1 AM, allowing us to image the activity of ensembles of layer 2/3 cells using two-photon microscopy². Astrocytes were distinguished with the astrocytic marker SR101 (ref. 19) (Fig. 2c). On average we imaged 55 neurons and 2 astrocytes per experiment (Fig. 2d and Supplementary Fig. 7). Astrocytes showed occasional long calcium transients (for example, Fig. 2d, grey trace) without obvious relation to the task and were not analysed further. Of the 548 imaged neurons, 395 showed significant calcium transients and were analysed in detail (Supplementary Fig. 8 and Methods). For some analyses, we extracted from the fluorescence data a series of events (a measure of instantaneous activity), with magnitudes equal to the amplitudes of the calcium transients (Supplementary Fig. 9 and Methods). The amplitudes (relative change in fluorescence ($\Delta F/F_0$), mean = 17%, range = 4–99%) and decay times (median = 0.86 s) of neuronal calcium transients were comparable

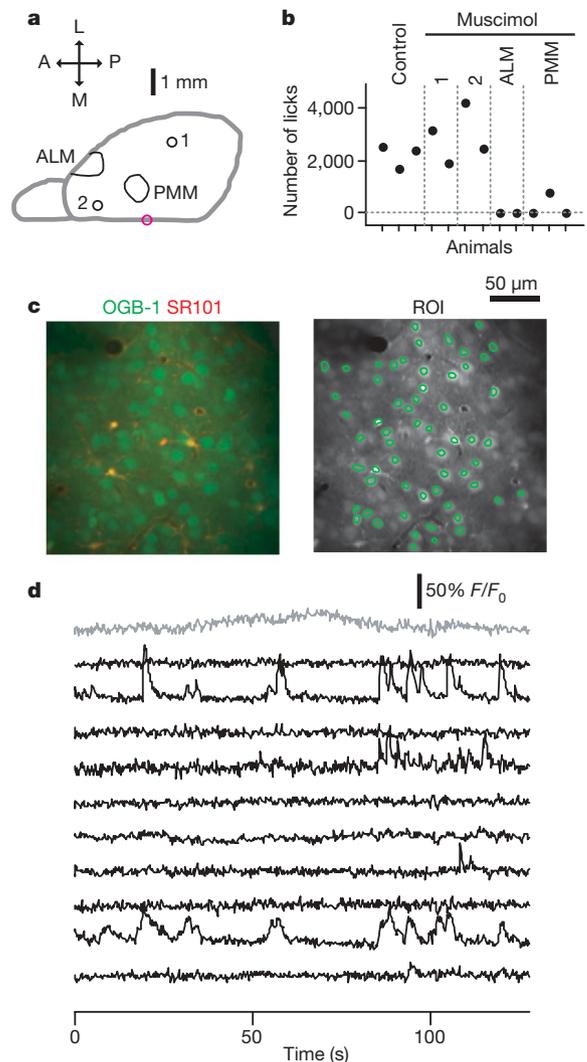


Figure 2 | Imaging motor cortex ensemble activity. **a**, Dorsal view of the mouse brain (right hemisphere). The imaged areas (ALM and PMM) are shown. '1' (somatosensory cortex) and '2' (anterior-medial cortex) denote control regions for muscimol injections. Magenta circle denotes the bregma. A, anterior; L, lateral; M, medial; P, posterior. **b**, Performance in an odour-detection task after muscimol injections in indicated areas. Each dot represents an animal. **c**, Layer 2/3 cells imaged with two-photon microscopy. Left, overlay of SR101 (red) and OGB-1 (green). Astrocytes are labelled by both dyes and thus appear yellow, whereas neurons are green. Right, regions of interest (ROI, green) overlaid on the OGB-1 channel. **d**, Example calcium traces from ten neurons (black) and one astrocyte (grey) imaged over 128 s.

to previous imaging studies^{5,20}. Neurons displayed sparse activity (events min^{-1} , median, range: ALM, 2.4, 0.7–38.1; PMM, 6.1, 0.9–38.4).

We first studied task-related activity after the learning criterion. A large fraction of neurons was active during specific phases of the task, but the response types were diverse. Some neurons were active in both hit and correct-rejection trials (task neurons). However, most were selectively excited either during hit (hit neurons) or correct rejection (CR neurons) trials (Fig. 3a; see Supplementary Fig. 10 for criteria²¹). Most of the hit neurons were also active after licking during inter-trial intervals in the absence of olfactory stimuli, supporting their roles in the control of licking (Supplementary Fig. 11).

The time course of task modulation was also diverse. Some neurons were modulated before or during the answer period (40% in PMM and 18% in ALM), whereas other neurons showed modulation only after the answer period (22% in PMM and 25% in ALM, for example, Fig. 3a, ALM CR-2). The task-related activity in both areas was qualitatively similar. However, hit neurons were more abundant in the PMM than in the ALM (Fig. 3b, left, $P < 0.001$, χ^2 test).

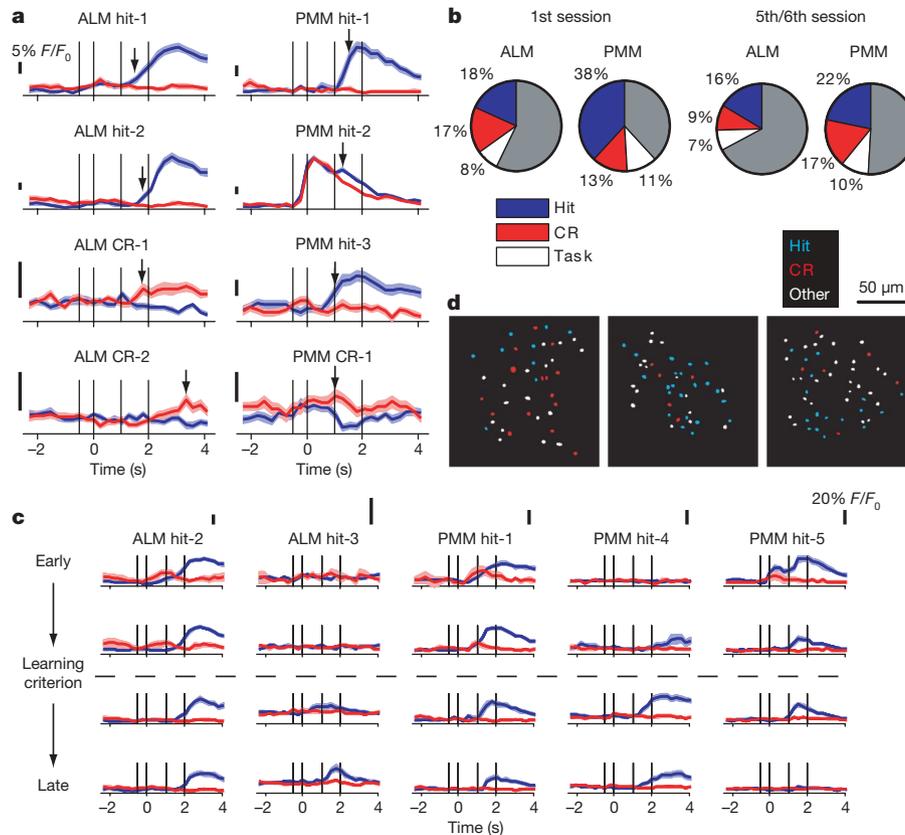


Figure 3 | Task-related activity. **a**, Average activity of eight task-related neurons during hit (blue) and correct rejection (red) trials (mean \pm s.e.m.). Arrows indicate the time in which the activity between trial types begins to diverge (TUD). Vertical lines correspond to Fig. 1a. **b**, Proportions of different response types during the first session and in the fifth/sixth session.

Task-related activity was dynamic (Fig. 3c). Although the activity of some hit neurons was stable during the first session (for example, ALM hit-2 and PMM hit-1 in Fig. 3c), other neurons showed dynamic activity patterns (for example, ALM hit-3, PMM hit-4 and PMM hit-5 in Fig. 3c). As a population, 28% (31 out of 112) of hit neurons increased their activity during the first session, and 14% (16 out of 112) decreased their activity (Supplementary Fig. 12).

We next analysed the spatial distributions of neurons with different response types. The probability that nearest neighbour neurons shared the same response type (hit–hit or CR–CR) was not significantly different from that in randomized data ($P = 0.56$). Other analyses also failed to show functional clustering (Methods). Thus, as a first approximation, the response types were spatially intermingled (Fig. 3d).

Temporal correlations in neuronal activity can shape the information capacity and robustness of neural codes (for example, ref. 13). Significant inter-neuronal correlations have been observed in mammalian cortices *in vivo* over tens of milliseconds^{22–26} to seconds^{13,22}, but the spatial scales of these correlations are poorly defined^{25,26}, partly due to the limited resolution of multi-electrode recording. In our data set, obvious correlations between pairs of neurons were occasionally apparent in the fluorescence traces (Fig. 4a).

We analysed the temporal correlations of pairs of neurons by computing a coincidence index, which measures the fraction of events that are coincident beyond chance expectation (Methods). Although the time resolution of our measurements is relatively low (250 ms), correlations at this time scale are known to reflect shorter time-scale interactions²². To minimize correlations imposed by the task, we calculated coincidence index during inter-trial intervals. Twenty-seven per cent (2,128 out of 7,937) of neuron pairs showed significantly positive correlation (Fig. 4b, at $P < 0.01$). The number

of negatively correlated pairs (1%, 41 out of 7,937) was at chance level; however, negative correlations in networks with low event rates are expected to be difficult to detect²⁷. Correlations were higher in the PMM than in the ALM (means, 0.035 versus 0.012, $P < 0.001$, Wilcoxon rank sum test). Similar results were obtained when we excluded periods during which the mouse was licking (Supplementary Fig. 13), or when correlations were calculated during trials (Supplementary Fig. 14). Data from the ALM and PMM showed similar trends and were pooled.

Correlations decreased with distance between neurons (length constant $\sim 154 \mu\text{m}$, Fig. 4c), indicating that the inter-neuronal distance is a key determinant of the strength of coupling between neurons^{3,28}. This length scale is similar to the sizes of typical dendritic and local axonal arbors²⁹ and the spatial extent of monosynaptic connectivity observed in layers 2/3 *in vitro*^{29,30}. Correlations therefore probably reflect local synaptic connectivity.

We looked for relationships between response type and correlations^{22–24,26}. Correlations were higher for neuron pairs of the same response type (hit–hit and CR–CR) compared to other pairs (Fig. 4d, $P < 0.001$, Wilcoxon rank sum test). Among neuron pairs of the same response type, pairs with the same (within our time resolution, 256 ms) time until divergence (TUD, arrows in Fig. 3a) had higher correlations than those with different TUD values (Fig. 4d, $P < 0.001$, Wilcoxon rank sum test). We quantified the similarity in response types for each pair (Methods). Correlations between pairs of neurons increased monotonically with similarity of their response types (Fig. 4e), across all distances (Fig. 4f). The strengths of temporal correlations, even during inter-trial intervals, are predictive of similarities in task-related responses. Thus, temporal correlations, probably imposed by specific local connectivity, depend on both inter-neuronal distance and task-related response types (Fig. 5b).

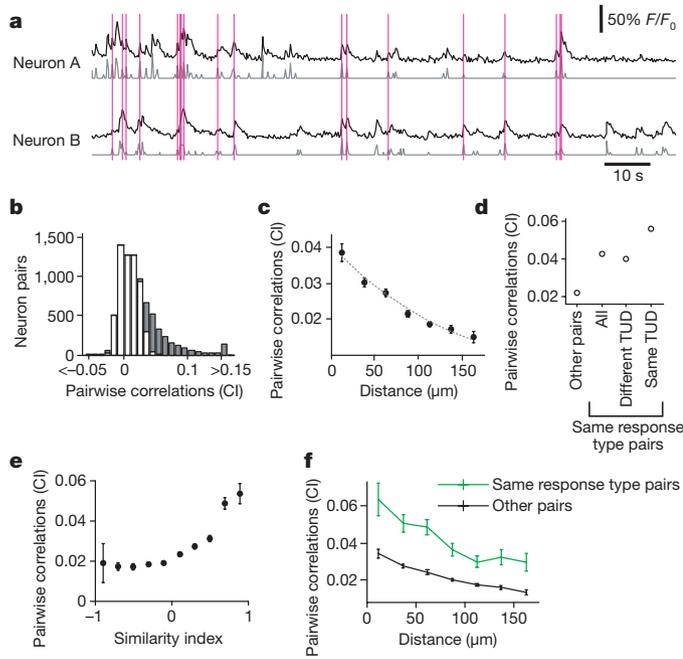


Figure 4 | Temporal structure of ensemble activity. **a**, Example of a neuron pair showing highly correlated activity. Black, raw fluorescent traces; grey, activity events (see Supplementary Fig. 9 and Methods); magenta lines, coincident events. **b**, Distribution of the coincidence index (CI) for 7,858 neuron pairs measured during inter-trial intervals in the first session. Grey bars represent significant pairs (permutation test by shuffling image frames, $P < 0.01$). **c**, Correlations decrease with distance between neuron pairs. Dotted line, exponential fit. **d**, Correlations are higher between neuron pairs with the same response type (hit–hit or CR–CR), especially when the time until divergence is the same. Error bars (s.e.m.) are contained in the symbols. **e**, Correlations between neuron pairs (during inter-trial intervals) are a monotonic function of similarity of the response types (during the task). **f**, Neuron pairs with the same response type show higher levels of correlations over all distances.

Correlations increased gradually with learning, specifically among neuron pairs of the same response type (Fig. 5a, ‘first session’, $P < 0.001$, Wilcoxon signed-rank test). We further characterized learning-related activity by imaging motor cortex activity in mice trained in the lick/no-lick task for several (5–6) sessions (ALM, $n = 4$; PMM, $n = 5$). The fraction of active neurons and the rate and size of their calcium transients were comparable to the data from the first session. Spatial and temporal aspects of task-related activity were qualitatively similar (Supplementary Figs 12, 15 and 16), but the proportion of task-related neurons was smaller (Fig. 3b, ALM, $P < 0.05$; PMM, $P < 0.01$, χ^2 test), indicating a sparsening of task-related activity with learning.

The correlations among neuron pairs of the same response type in the fifth/sixth session were significantly higher than in the first session (Fig. 5a, $P < 0.001$, Wilcoxon rank sum test). Correlations continued to increase during the session (Fig. 5a, $P < 0.001$, Wilcoxon signed-rank test), without an increase in event rates (Supplementary Fig. 17a). Increasing correlations within and across sessions suggest that the strengthening of coupling among neurons of the same response type is a hallmark of learning-related circuit reorganization.

Odour-selective input is relayed to spatially intermingled motor cortex neurons. Learning could induce biases in connection strength between particular odours and motor cortex neurons (odour A \rightarrow lick neurons, odour B \rightarrow no-lick neurons). Our data indicate that plasticity in local circuits then create ‘subnetworks’ of functionally related neurons (Fig. 5b). Mechanisms may include potentiation of excitatory connections between neurons of similar response types, and perhaps also enhanced inhibition between neurons of different response types. This local circuit plasticity would prevent simultaneous activation of

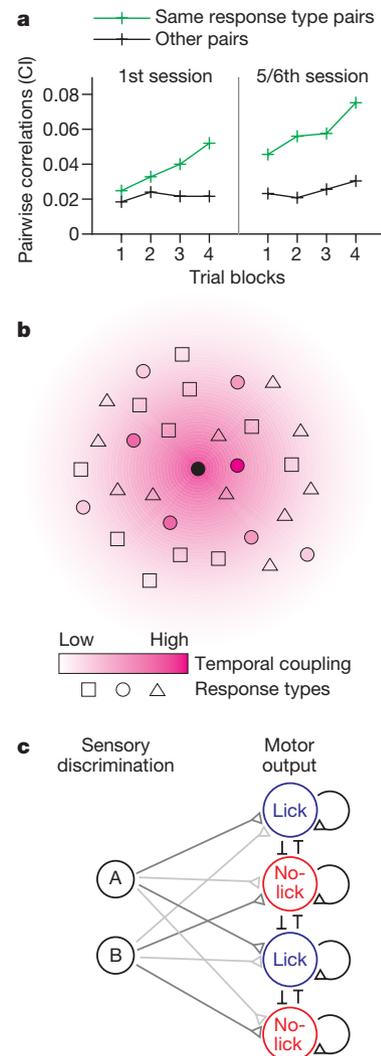


Figure 5 | Learning-related coupling of specific neuronal ensembles.

a, Correlations increase with learning, specifically for neuron pairs with the same response types. Trial blocks in the first session are defined in Fig. 3c. The fifth/sixth session was quartered. Error bars (s.e.m.) are contained in the symbols. **b**, Distance and response type affect inter-neuronal correlations. Different response types (symbols) are spatially intermingled. A given neuron (filled circle) is more tightly coupled to its close neighbours than distant ones (gradient), and to functionally related neurons (circles) than other intermingled neurons with different response types. **c**, Model. Sensory information (from odour-selective neurons A and B) is relayed to hit and CR neuron populations in the motor cortex. Learning leads to strengthening of connections between sensory and motor neurons (for example, A \rightarrow hit) and/or increased local recurrent excitation (black lines). Local inhibition might also be involved. Triangles, excitatory; lines, inhibitory.

different functional subnetworks that may code for inconsistent motor behaviours. Learning-related changes in local cortical circuits might thus underlie binary actions in our lick/no-lick task (Fig. 5c).

METHODS SUMMARY

C57BL/6 male mice (Charles River Laboratory) were water restricted, and trained daily (typically 5 days) for 15–30 min until they licked in response to two odours for a water reward ($\sim 5 \mu\text{l}$ per trial). Licking was monitored using a photodiode. The discrimination paradigm was then introduced by having one of the odour signal no-lick trials. Because of this training procedure the learning consisted mostly of withholding licking to the no-lick stimulus. Cortical ensemble activity was imaged during the first discrimination session or the fifth/sixth session. On the day of imaging, mice were anaesthetized with isoflurane ($< 1\%$) and a craniotomy ($\sim 2\text{--}2.5 \text{ mm}$) was made over the imaging area, leaving the dura intact. Dye solution was injected stereotaxically using a custom-built volumetric injector. After dye injection, a window was placed over the craniotomy and mice were released in their

home cage to recover from anaesthesia. About 75–90 min after dye loading, mice were mounted under a custom-built two-photon microscope where mice performed the task while cortical ensemble activity was imaged. The experimental session was terminated after the animal stopped responding to both odours.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 September 2009; accepted 8 February 2010.

Published online 7 April 2010.

- Song, S., Sjöström, P. J., Reigl, M., Nelson, S. & Chklovskii, D. B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, 1–13 (2005).
- Stosiek, C., Garaschuk, O., Holthoff, K. & Konnerth, A. *In vivo* two-photon calcium imaging of neuronal networks. *Proc. Natl Acad. Sci. USA* **100**, 7319–7324 (2003).
- Kerr, J. N. *et al.* Spatial organization of neuronal population responses in layer 2/3 of rat barrel cortex. *J. Neurosci.* **27**, 13316–13328 (2007).
- Sato, T. R., Gray, N. W., Mainen, Z. F. & Svoboda, K. The functional microarchitecture of the mouse barrel cortex. *PLoS Biol.* **5**, e189 (2007).
- Dombeck, D. A., Khabbazi, A. N., Collman, F., Adelman, T. L. & Tank, D. W. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* **56**, 43–57 (2007).
- Slotnick, B. & Restrepo, D. Olfactometry with mice. *Curr. Protoc. Neurosci.* **Chapter 8**, Unit–8.20 (2005).
- Verhagen, J. V., Wesson, D. W., Netoff, T. I., White, J. A. & Wachowiak, M. Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nature Neurosci.* **10**, 631–639 (2007).
- Li, C. X. & Waters, R. S. Organization of the mouse motor cortex studied by retrograde tracing and intracortical microstimulation (ICMS) mapping. *Can. J. Neurol. Sci.* **18**, 28–38 (1991).
- Ayling, O. G., Harrison, T. C., Boyd, J. D., Goroshkov, A. & Murphy, T. H. Automated light-based mapping of motor cortex by photoactivation of channelrhodopsin-2 transgenic mice. *Nature Methods* **6**, 219–224 (2009).
- Song, C. K., Enquist, L. W. & Bartness, T. J. New developments in tracing neural circuits with herpesviruses. *Virus Res.* **111**, 235–249 (2005).
- Fay, R. A. & Norgren, R. Identification of rat brainstem multisynaptic connections to the oral motor nuclei using pseudorabies virus. III. Lingual muscle motor systems. *Brain Res. Brain Res. Rev.* **25**, 291–311 (1997).
- Tanji, J. & Evarts, E. V. Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *J. Neurophysiol.* **39**, 1062–1068 (1976).
- Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).
- Travers, J. B., Dinardo, L. A. & Karimnamazi, H. Motor and premotor mechanisms of licking. *Neurosci. Biobehav. Rev.* **21**, 631–647 (1997).
- Rathelot, J. A. & Strick, P. L. Muscle representation in the macaque motor cortex: an anatomical perspective. *Proc. Natl Acad. Sci. USA* **103**, 8257–8262 (2006).
- Brecht, M. *et al.* Organization of rat vibrissa motor cortex and adjacent areas according to cytoarchitectonics, microstimulation, and intracellular stimulation of identified cells. *J. Comp. Neurol.* **479**, 360–373 (2004).
- Smith, B. N. *et al.* Pseudorabies virus expressing enhanced green fluorescent protein: A tool for *in vitro* electrophysiological analysis of transsynaptically labeled neurons in identified central nervous system circuits. *Proc. Natl Acad. Sci. USA* **97**, 9264–9269 (2000).
- O'Connor, D. H. *et al.* Vibrissa-based object localization in head-fixed mice. *J. Neurosci.* **30**, 1947–1967 (2010).
- Nimmerjahn, A., Kirchhoff, F., Kerr, J. N. & Helmchen, F. Sulforhodamine 101 as a specific marker of astroglia in the neocortex *in vivo*. *Nature Methods* **1**, 31–37 (2004).
- Kerr, J. N., Greenberg, D. & Helmchen, F. Imaging input and output of neocortical networks *in vivo*. *Proc. Natl Acad. Sci. USA* **102**, 14063–14068 (2005).
- Doucette, W. & Restrepo, D. Profound context-dependent plasticity of mitral cell responses in olfactory bulb. *PLoS Biol.* **6**, e258 (2008).
- Bair, W., Zohary, E. & Newsome, W. T. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* **21**, 1676–1697 (2001).
- Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *J. Neurosci.* **6**, 1160–1170 (1986).
- Georgopoulos, A. P., Taira, M. & Lukashin, A. Cognitive neurophysiology of the motor cortex. *Science* **260**, 47–52 (1993).
- Fujisawa, S., Amarasingham, A., Harrison, M. T. & Buzsáki, G. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nature Neurosci.* **11**, 823–833 (2008).
- Constantinidis, C., Franowicz, M. N. & Goldman-Rakic, P. S. Coding specificity in cortical microcircuits: a multiple-electrode analysis of primate prefrontal cortex. *J. Neurosci.* **21**, 3646–3655 (2001).
- Aertsen, A. M. & Gerstein, G. L. Evaluation of neuronal connectivity: sensitivity of cross-correlation. *Brain Res.* **340**, 341–354 (1985).
- Dombeck, D. A., Graziano, M. S. & Tank, D. W. Functional clustering of neurons in motor cortex determined by cellular resolution imaging in awake behaving mice. *J. Neurosci.* **29**, 13751–13760 (2009).
- Shepherd, G. M. G., Stepanyants, A., Bureau, I., Chklovskii, D. B. & Svoboda, K. Geometric and functional organization of cortical circuits. *Nature Neurosci.* **8**, 782–790 (2005).
- Holmgren, C., Harkany, T., Svennenfors, B. & Zilberter, Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol. (Lond.)* **551**, 139–153 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Rinberg for help with experiments; F. Collman, D. Tank, C. Zuker, T. O'Connor and V. Iyer for help with analysis and imaging software; L. W. Enquist for pseudorabies vectors; W. Denk for help with mechanical design; D. Dombeck, M. Andermann, A. Kerlin and C. Reid for discussions about imaging awake mice; B. Shields, A. Hu and S. Michael for help with histology; A. Arnold for help with imaging; J. Osborne and S. Bassin for machining; L. Luo, Z. Mainen and D. Rinberg for comments on the manuscript; A. C. Gontang for illustration. Supported by Howard Hughes Medical Institute. T.K. is a Helen Hay Whitney Foundation postdoctoral fellow.

Author Contributions T.K. and K.S. conceived the project. T.K. developed and performed most of the experiments. D.H.O. helped to develop head-fixed behaviour. D.H. developed the glass-plug imaging window. Y.-X.Z. and D.H. performed optical stimulation mapping. T.K. and T.R.S. performed electrical stimulation mapping. T.K., B.M.H. and T.R.S. performed PRV tracing. T.K., T.R.S. and K.S. analysed data. M.G. provided a software module for image segmentation. T.K. and K.S. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.K. (komyiamat@janelia.hhmi.org).

METHODS

Viral tracing. All procedures were in accordance with protocols approved by the Janelia Farm Institutional Animal Care and Use Committee. For retrograde trans-synaptic tracing, 10 μ l PRV-152 (pseudorabies virus encoding enhanced GFP, gift from L. Enquist) was injected bilaterally into the tongue of 4–10-week-old mice with a syringe (Hamilton). Virus injection into the tongue has been shown to label hypoglossal motoneurons and their presynaptic interneurons in the reticular formation of the brainstem^{11,31}. Four to five days later, mice were perfused and the brain was sectioned coronally (50- μ m thickness). Every other section (100- μ m intervals) was stained with an anti-GFP antibody and imaged (Axioimager, Zeiss). Layer 5 cell bodies were scored manually with NeuroLucida. For each section, the neuronal density along the medial-lateral axis was fit with a Gaussian function.

For anterograde tracing, adeno-associated virus encoding eGFP (AAV2/1-CAG-eGFP, T. Mao, personal communications) was stereotactically injected into the ALM or PMM using a custom-built volumetric injector. Thirty nanolitres of viral suspension was injected at depths of 400 and 800 μ m. Two weeks later, mice were perfused and the brain was sectioned (50- μ m thickness). Every other section (100- μ m intervals) was stained with an anti-GFP antibody and imaged (Axioimager, Zeiss or Nanoscope, Hamamatsu).

Microstimulation. Electrical microstimulation was performed on mice lightly anaesthetized with ketamine/xylazine. Intracortical microstimuli were delivered using a Pt/Ir (FHC) or tungsten (FHC) electrode at 300 Hz, 0.4 ms per pulse, 50 ms per train, typically at 50–100 μ A at the depth of approximately 800 μ m. Jaw and tongue protrusions were monitored visually.

Optical stimulation was performed on Thy-1::ChR2 mice (line 18)³² as described^{9,33}. The skull was thinned over the right hemisphere of the frontal cortex and covered with Kwik-sil (WPI) and a coverslip. Mice were lightly anaesthetized with ketamine/xylazine and blue light (473 nm, Crystalaser) was applied over a range of powers (\leq 35 mW) in a grid pattern with 500- μ m spacing. Jaw was held down by cotton thread to expose the tongue, and tongue movement was monitored visually.

Behavioural training. C57BL/6 male mice (Charles River Laboratory) were housed individually and water-restricted at 1 ml per day, starting at ages 5–7 weeks. After 10–14 days, custom-made headplates were implanted. The skull over the imaging area was covered by Kwik-Cast (WPI) and dental acrylic. For behavioural training, mice were head-fixed with their body placed in an acrylic tube (2.9 cm inside diameter; McMaster) and trained for 15–30 min per day until they consistently licked in response to two odours for a water reward (\sim 5 μ l per trial) (typically 5 days). Water was delivered by gravity controlled by a solenoid valve (NResearch) and unconsumed water was removed from the lickport using a pump (Fisher Scientific). Licking was monitored by breaking an infrared beam (between an 860-nm LED (Panasonic) and a phototransistor (Fairchild Semiconductor)). The setup was controlled by software (Z. Mainen and C. Brody) running on MATLAB (Mathworks) communicating with a real-time system (TDT and RTLinux). On the day of the first learning session, after approximately ten trials with the contingency as above, the discrimination paradigm was introduced in which one of the odours signalled no-lick trials. The experimental session was terminated after the animal stopped responding to both odours. The typical parameters are: ITI, 8.5 s; extra ITI as punishment, 4.5 s; odour sampling period, 1 s; answer period, 1 s; and cue period, 0.5 s. The trial type for each trial was chosen pseudorandomly.

Olfactory stimuli. A custom built olfactometer³⁴ delivered a 1:4 mixture of air saturated with odour (citril (lick) and isoamyl acetate (no-lick), Sigma) and clean air. The olfactometer constantly delivered clean air during inter-trial intervals. The final valve placed before the animal diverted the air flow away from the animal for 500 ms before odour onset. This served to create a sharp onset of odour delivery and also served as a trial initiation signal ('cue' in Fig. 1a).

Inactivation experiments. Muscimol hydrobromide (Sigma, 5 μ g μ l⁻¹ in HEPES-buffered artificial cerebral spine fluid (CSF)) was injected stereotactically using a custom-built volumetric injector. Under light anaesthesia (\sim 0.8% isoflurane), \sim 70 nl volumes were injected slowly (4 min per injection) at depths of 300 and 700 μ m. Coordinates were (in mm) lateral 2.0 and anterior 2.4 for ALM, lateral 1.2 and anterior 0.3 for PMM, lateral 3.3 and anterior -1 for the control area '1' (the barrel cortex) and lateral 0.3 and anterior 2 for the control area '2'. After bilateral injections, the skull was covered by Kwik-Cast (WPI) and a thin layer of dental acrylic. Mice were released in the home cage and allowed to recover from anaesthesia for at least 1 h.

Behavioural experiments were performed 2–2.5 h after the beginning of injections. At this time, animals were active in their home cages. In some cases, voluntary licking was assayed under head-fixation before the behavioural session, by bringing a small amount of water close to their mouths with a pipette (Supplementary Fig. 6 and Supplementary Movies). The odour-detection licking task was the same as the lick/no-lick task except that both odours signalled

reward. Even though the control injection areas set the upper bound for the effective spread of muscimol, we note that we did not define the exact size of inactivated areas.

Dye loading and imaging. On the day of imaging, mice were anaesthetized with isoflurane ($<$ 1%) and a craniotomy (\sim 2–2.5 mm) was made over the imaging area, leaving the dura intact. Dye solution (0.8 mM Oregon Green 488 BAPTA-1 (OGB-1) AM, 6% dimethylsulphoxide (DMSO), 1.2% (w/v) pluronic acid and 30 μ M SR101 in HEPES-buffered artificial CSF (160 mM NaCl, 6 mM KCl, 13 mM glucose, 13 mM HEPES, 2.5 mM CaCl₂, 2.5 mM MgSO₄) was injected stereotactically. Approximately 18 nl of dye solution was injected over 5 min. This typically resulted in loading of a cluster of cells with diameter of 400 μ m.

After dye injection, a window (a coverglass with a \sim 300- μ m-thick glass plug, inspired by ref. 5) was placed over the craniotomy and the edges were sealed with 1.5% agarose. The window was glued to the skull and headplate. Mice were released in their home cage to recover from anaesthesia. Between 75 and 90 min after dye loading, mice were mounted under a custom-built two-photon microscope. Images (512 \times 128 pixels) were acquired at 2 ms per line, 256 ms per frame. Five-hundred frames were acquired in a series, followed by a short rest interval (\sim 4 s). Trial signals from the behaviour control system, which were acquired simultaneously with imaging, were used to assign images to particular trials.

Fluorescence data analysis. Lateral motion was corrected in two steps (Supplementary Fig. 8a). A line-by-line correction was performed using a hidden-Markov model-based algorithm⁵, followed by cross-correlation-based image alignment (Turboreg, ImageJ plugin³⁵). ROIs containing the centres of OGB-1-labelled neurons were drawn semi-automatically. The pixels in each ROI were averaged to estimate the fluorescence of a single cell at a particular time point (that is, an image frame). For each ROI, 500 frames yielded 128-s long traces. Because fluorescence responses were sparse, the mode of the distribution of the fluorescent signal gave a good estimate of the baseline fluorescence (F_0 , Supplementary Fig. 8).

On hit trials, some cells exhibited consistent shifts of the baseline owing to motion artefacts produced by the animal's water consumption. We corrected this by dividing the fluorescence signal with the mode calculated for each time point over trials of each type (smoothed by averaging over three frames, with the central frame weighted twice as much as the surrounding two frames). This correction assumes that the activity is relatively sparse and the mode correctly represents the baseline. Visual inspection revealed that this was true for all but four neurons in the first session and 19 in the fifth/sixth session that had very reliable responses in most trials (that is, the mode did not represent the baseline correctly), and we did not apply the correction to them (for example, Supplementary Fig. 8l). For other neurons, this method corrected the baseline shift without affecting transients (Supplementary Fig. 8b–k).

The vast majority of calcium influx in cortical neurons is caused by action potentials^{5,20,36}. We thus identified calcium-dependent fluorescence transients as a proxy for neural activity. Under our experimental conditions, the decay of action-potential-evoked fluorescence transients are relatively slow², decaying with an exponential time-course with median time-constant $\tau = 860$ ms²⁰. For event detection (Supplementary Fig. 9), we detected segments of the trace that exceeded the baseline by twice the standard deviation of the baseline (σ) for at least two successive frames. σ was calculated from selecting values below baseline and appending values of the flipped sign, to construct an estimate of the baseline fluctuations without contaminations from true calcium signals. Events were scored if the trace exceeded baseline by 3σ for at least one frame, or if the detection criterion by template matching was above three (ref. 37). The template used for template matching was six image frames of exponential decay. After detecting the transients we estimated the instantaneous activity by subtracting the bleed through of signal from previous transients at earlier time points. For this, we simply subtracted 73% ($e^{-256\text{ms}/\tau} = 73\%$, where 256 ms is the frame period) of the value above baseline of the previous frame from each frame within the regions identified as transients.

Using the same methods we identified transients with negative amplitudes, which are probably caused by noise, including any remaining motion artefacts⁵. We excluded neurons from further analysis if the ratio of the number of positive to negative transients was below ten. We thereby rejected \sim 28% of the neurons (153 of 548) in the first session and \sim 29% (222 of 757) in the fifth/sixth session. This rejects neurons that were imaged close to their dorsal or ventral poles and thus had a poorer signal-to-noise ratio⁴, but this also excludes neurons with very low levels of activity. Thus, we analysed 395 neurons (193 from ALM and 202 from PMM) in the first session, and 535 neurons (207 from ALM and 328 from PMM) in the fifth/sixth session. Of these neurons, the ratio of positive to negative transients for the population was 34.6, indicating a false positive rate of \sim 2.9%.

Analysis of dynamic changes of activity levels. Dynamism of hit neurons was quantified as follows. For each hit neuron, the peak fluorescence for each hit trial was measured and earlier and later trials were compared by bootstrap. Neurons

that showed significantly different ($P < 0.01$) activity levels in earlier and later trials were classified as dynamic (Supplementary Fig. 12). Earlier and later trials were defined as before and after the learning criterion for the first session, and first and second halves for the fifth/sixth session.

Identification of lick bouts. Lick bouts were defined as three licks or more, with all inter-lick intervals < 300 ms. Lick bouts were scored as 'trial-licks' if the bout overlapped with the answer period of a trial, and as 'ITI-licks' if the bout did not overlap with the 4-s period after the cue presentation of any trial.

Analysis of spatial clustering. We examined the spatial clustering of response types in eight experiments from the first session and seven experiments from the fifth/sixth session that showed three or more neurons of each of the two response types (hit neurons and CR neurons).

First, we calculated the probability that the nearest neighbour of each hit neuron and CR neuron was of the same response type. The P -value was calculated by randomly shuffling labels. The mean P -value was 0.56 (range: 0.13–0.93) for the first session and 0.30 (range: 0.03–0.76) for the fifth/sixth session.

Second, we measured the distance from each hit neuron and CR neuron to the nearest neuron of the same response type (hit-to-hit, or CR-to-CR). The P -value was calculated by randomly shuffling labels. The mean P -value was 0.48 (range: 0.11–0.88) for the first session and 0.33 (range: 0.03–0.64) for the fifth/sixth session. We also measured the distance from each hit neuron and CR neuron to the second nearest neuron of the same response type. The mean P -value was 0.33 (range: 0.12–0.61) for the first session and 0.39 (range: 0.003–0.97) for the fifth/sixth session.

Third, we calculated the mean distance between all pairs of hit or CR neurons (D_{hit} and D_{CR}). This measure is similar to a previously used measure of clustering termed 'focality'³⁸. P -values were calculated by randomly shuffling labels. For the first session, the mean P -value for D_{hit} was 0.37 (range: 0.04–0.92) and for D_{CR} was 0.41 (range: 0.11–0.88). For the fifth/sixth session, the mean P -value for D_{hit} was 0.20 (range: 0.11–0.33) and for D_{CR} was 0.46 (range: 0.01–0.99).

All of these analyses therefore failed to reject the null hypothesis that the response types are intermingled.

Correlations and the coincidence index. To measure correlations between pairs of neurons we counted coincident events. Measurements of coincident activity are expected to capture correlations with shorter time-constants than the time resolution of our measurements (256 ms per time point), including monosynaptically coupled pairs, polysynaptically coupled pairs, and pairs of neurons with common input. For each neuron i , the vector of binarized events is given by $E_i = (E_i(0), E_i(\Delta t), E_i(2\Delta t), \dots, E_i((N-1)\Delta t))$, where N is the number of image frames, and Δt is the frame interval. Trials (6 s starting with the trial initiation cue) or inter-trial intervals were concatenated. The coincidence index (CI) for the pair i, j is given by

$$CI_{ij} = \frac{E_i \cdot E_j - N \langle E_i \rangle \langle E_j \rangle}{N \sqrt{\langle E_i \rangle \langle E_j \rangle}}$$

$E_i \cdot E_j$ is the number of coincident events between neurons i, j (events occurring in the same image frame), and $N \langle E_i \rangle \langle E_j \rangle$ is the expected number of coincident events if neurons i and j are independent. The denominator normalizes the coincidence index by the geometric mean of the event number of the two neurons. CI thus measures the fraction of coincident events between two neurons above chance expectation. CI is ~ 0 for independent pairs, positive for coincident pairs and negative for anti-correlated pairs. Using events, rather than fluorescence, reduces sensitivity to small co-fluctuations caused by laser power or motion artefacts.

CI calculation during ITI excluding when the mouse is licking. From fluorescence traces from ITI, image frames containing a lick and their immediate adjacent frames were excluded before coincidence index calculation. This excludes data from 750 ms surrounding each lick, much larger than the lick interval within lick bouts.

Similarity index. We used the fluorescence traces $F_i(t)$ to calculate similarities between task-related responses. To quantify the similarity of response types for each neuron pair i, j , we calculated the average selectivity of each neuron i for hit and correct rejection trials

$$\Delta F_i(t) = \langle F_{i,\text{hit}}(t) \rangle - \langle F_{i,\text{CR}}(t) \rangle$$

where $\langle \rangle$ denotes the average over trials. We define the similarity between neurons i, j as the normalized dot product of their selectivities.

$$SI = \frac{\Delta F_i(t) \bullet \Delta F_j(t)}{|\Delta F_i(t)| |\Delta F_j(t)|}$$

SI measures the similarity of task-related activity. Unless both neurons are modulated by the task, SI is expected to be near zero. SI diverges from zero with task modulation of neurons i, j and also with temporal overlap of the modulation.

31. Dobbins, E. G. & Feldman, J. L. Differential innervation of protractor and retractor muscles of the tongue in rat. *J. Comp. Neurol.* **357**, 376–394 (1995).
32. Arenkiel, B. R. *et al.* *In vivo* light-induced activation of neural circuitry in transgenic mice expressing channelrhodopsin-2. *Neuron* **54**, 205–218 (2007).
33. Hira, R. *et al.* Transcranial optogenetic stimulation for functional mapping of the motor cortex. *J. Neurosci. Methods* **179**, 258–263 (2009).
34. Bodyak, N. & Slotnick, B. Performance of mice in an automated olfactometer: odor detection, discrimination and odor memory. *Chem. Senses* **24**, 637–645 (1999).
35. Thévenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).
36. Svoboda, K., Denk, W., Kleinfeld, D. & Tank, D. W. *In vivo* dendritic calcium dynamics in neocortical pyramidal neurons. *Nature* **385**, 161–165 (1997).
37. Clements, J. D. & Bekkers, J. M. Detection of spontaneous synaptic events with an optimally scaled template. *Biophys. J.* **73**, 220–229 (1997).
38. Yaksi, E., Judkewitz, B. & Friedrich, R. W. Topological reorganization of odor representations in the olfactory bulb. *PLoS Biol.* **5**, e178 (2007).