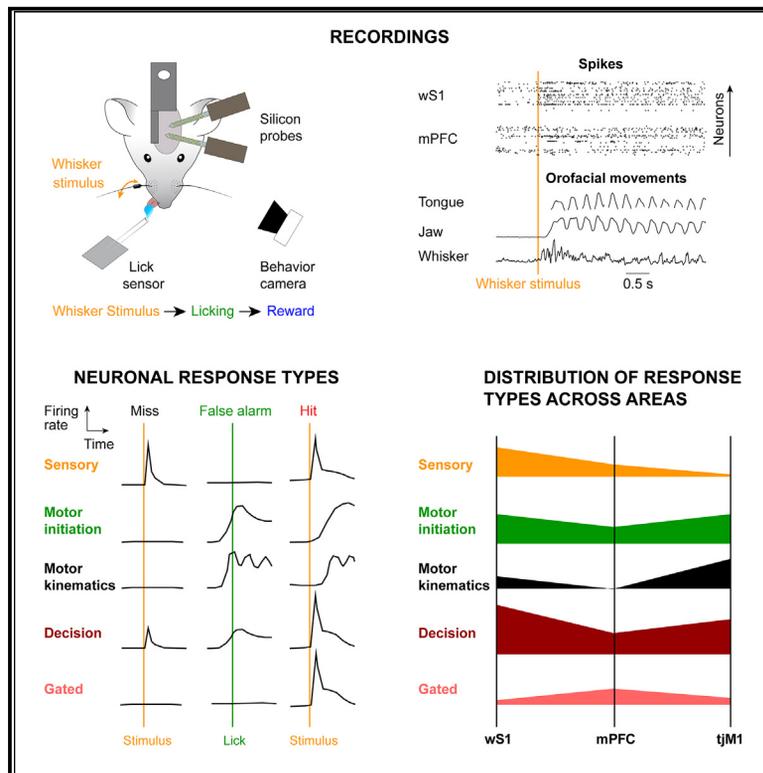


# Distributed and specific encoding of sensory, motor, and decision information in the mouse neocortex during goal-directed behavior

## Graphical abstract



## Authors

Anastasiia Oryshchuk, Christos Sourmpis, Julie Weverbergh, ..., Wulfram Gerstner, Carl C.H. Petersen, Sylvain Crochet

## Correspondence

carl.petersen@epfl.ch (C.C.H.P.), sylvain.crochet@epfl.ch (S.C.)

## In brief

Combining high-density extracellular electrophysiological recordings with high-speed videography of orofacial movements of mice performing a psychometric whisker sensory detection task reported by licking, Oryshchuk et al. found overall highly distributed task-related neuronal activity, but sensory, motor, and decision information were most selectively represented, respectively, in sensory, motor, and decision information were most selectively represented, respectively, in sensory, motor, and decision information were most selectively represented, respectively, in sensory, motor, and prefrontal cortices.

## Highlights

- Highly distributed neocortical activity as mice lick in response to whisker stimuli
- Task-independent encoding of whisker stimulus properties only in sensory cortex
- Task-independent coding of precise tongue movement kinematics only in motor cortex
- Learned highly selective representation of decision in prefrontal cortex



## Report

# Distributed and specific encoding of sensory, motor, and decision information in the mouse neocortex during goal-directed behavior

Anastasiia Oryshchuk,<sup>1</sup> Christos Sourmpis,<sup>1,2</sup> Julie Weverbergh,<sup>1</sup> Reza Asri,<sup>1</sup> Vahid Esmaeili,<sup>1</sup> Alireza Modirshanechi,<sup>2</sup> Wulfram Gerstner,<sup>2</sup> Carl C.H. Petersen,<sup>1,4,\*</sup> and Sylvain Crochet<sup>1,3,4,5,\*</sup>

<sup>1</sup>Laboratory of Sensory Processing, Brain Mind Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

<sup>2</sup>School of Life Sciences and School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

<sup>3</sup>Institut National de la Santé et de la Recherche Médicale (INSERM), 6900 Lyon, France

<sup>4</sup>These authors contributed equally

<sup>5</sup>Lead contact

\*Correspondence: [carl.petersen@epfl.ch](mailto:carl.petersen@epfl.ch) (C.C.H.P.), [sylvain.crochet@epfl.ch](mailto:sylvain.crochet@epfl.ch) (S.C.)

<https://doi.org/10.1016/j.celrep.2023.113618>

## SUMMARY

Goal-directed behaviors involve coordinated activity in many cortical areas, but whether the encoding of task variables is distributed across areas or is more specifically represented in distinct areas remains unclear. Here, we compared representations of sensory, motor, and decision information in the whisker primary somatosensory cortex, medial prefrontal cortex, and tongue-jaw primary motor cortex in mice trained to lick in response to a whisker stimulus with mice that were not taught this association. Irrespective of learning, properties of the sensory stimulus were best encoded in the sensory cortex, whereas fine movement kinematics were best represented in the motor cortex. However, movement initiation and the decision to lick in response to the whisker stimulus were represented in all three areas, with decision neurons in the medial prefrontal cortex being more selective, showing minimal sensory responses in miss trials and motor responses during spontaneous licks. Our results reconcile previous studies indicating highly specific vs. highly distributed sensorimotor processing.

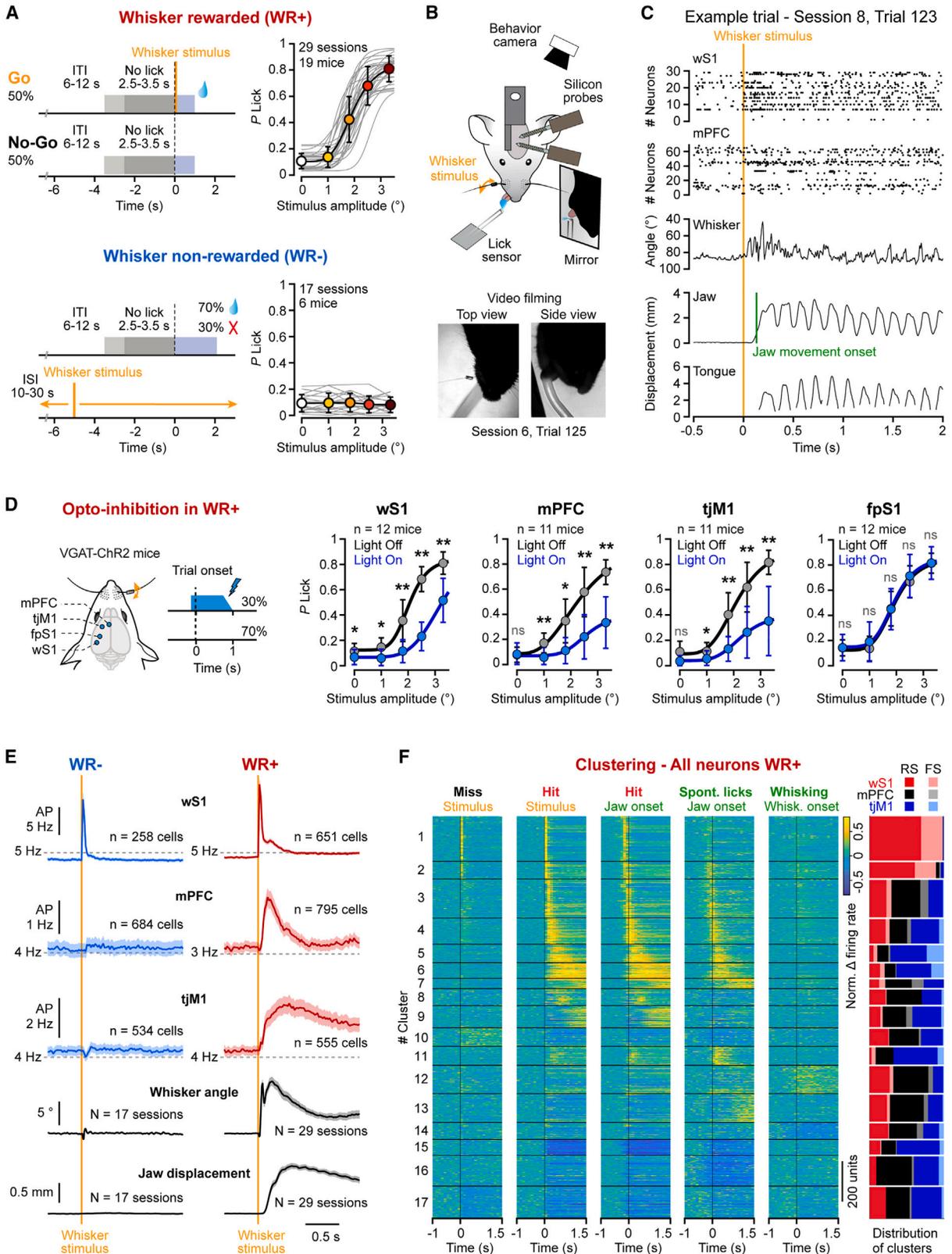
## INTRODUCTION

Learning to generate the appropriate action in response to a given external stimulus is a key function for the survival of any animal. Sensorimotor transformations require the succession of several processes including sensory perception, decision-making, and motor execution. In mammals, even simple sensorimotor transformations involve a large and distributed network of cortical and subcortical areas, including sensory, motor, associative, and high-order cortical areas, as well as striatum, superior-colliculus, cerebellum, and various midbrain regions.<sup>1–8</sup> Early work in non-human primates suggested a sequential transformation of sensory signals into decisions and then motor commands as the information flows from sensory to frontal and motor cortical areas, with sensory areas mostly encoding the physical qualities of the sensory stimulus and frontal regions encoding the decision to initiate the motor response during correct trials.<sup>1,9,10</sup> However, this idea of a segregated representation of different task dimensions across brain areas has recently been challenged by studies in mice showing the presence of decision information and reward-related signals at early stages of sensorimotor transformations, including sensory cortical areas,<sup>11–14</sup> as

well as brain-wide motor-related activity that could constitute a confounding factor with decision-related activity.<sup>15–19</sup> Therefore, it remains unclear whether task variables are encoded in a distributed manner across cortical areas or whether some variables are more specifically represented in distinct areas.

To address this question, we compared the neuronal representation of sensory, motor, and decision information in a sensory, a motor, and a higher-order cortical area in mice trained to lick for water reward in response to a brief whisker stimulus and other mice exposed to the same whisker stimuli that did not predict reward availability. We performed high-density extracellular recordings from the primary whisker sensory cortex (wS1), the first cortical area involved in the processing of the whisker stimulus<sup>6,20–22</sup>; from the primary tongue and jaw motor area (tjM1), a motor area involved in the control of directional licking<sup>23,24</sup>; and from the medial prefrontal cortex (mPFC), a higher-order area potentially involved in decision-making.<sup>25,26</sup> In order to disentangle sensory, motor, and decision information and following the pioneering work in non-human primates,<sup>1</sup> we used a psychophysical version of the whisker detection task in which whisker stimuli of variable amplitude were randomly presented.<sup>12,27,28</sup> To carefully take into account movement-related





(legend on next page)

neuronal activity, we simultaneously monitored orofacial movements using high-speed video filming.

## RESULTS

### High-density extracellular electrophysiological recordings in mice performing a psychometric whisker detection task

Head-restrained, water-restricted mice were trained to respond to a brief whisker stimulus of variable amplitude by licking a water spout in order to receive a reward (whisker rewarded [WR+] mice) (Figure 1A; Table S1A). Once trained, we performed high-density extracellular recordings simultaneously from two areas among wS1, mPFC, and tJM1. High-speed video filming during the recordings allowed precise assessment of the contribution of movements to the task-related neuronal activity in each cortical area (Figures 1B and 1C). To test to what extent the encoding of sensory and motor-related information might emerge through learning, we also recorded from the same areas in other mice for which the whisker stimulus was not associated with reward (whisker non-rewarded [WR−] mice) (Figure 1A; Table S1A).

We first tested the involvement of the three areas in the execution of the task in WR+ mice (Figure 1D; Table S1B). Focal optogenetic inhibition during the presentation of the whisker stimulus and response window induced a decrease in licking probability for all stimulus amplitudes, thus indicating that neuronal activity in wS1, mPFC, and tJM1 contributes to task execution, in good agreement with previous studies.<sup>23,26,29–31</sup> Importantly, optoinhibition of the forepaw S1 (fpS1), located ~1 mm from wS1, had no impact upon task performance.

We recorded 1,476 neurons across 17 sessions in 6 WR− mice and 2,001 neurons across 29 sessions in 19 WR+ mice (Figures S1A–S1C). In WR+ mice, the whisker stimuli evoked a fast and high-amplitude whisker protraction followed by the opening of the jaw and licking, whereas the same stimuli evoked little whisker movement and no licking in WR− mice. On average,

neuronal activity increased prominently after the whisker stimuli in all three areas in WR+ mice but only in wS1 in WR− mice (Figure 1E). Clustering analysis of the neuronal responses for different types of trials and aligned to the stimulus or movement onset times revealed a wide variety of responses across the three cortical areas in WR+ mice, ranging from a brief increase in firing predominant in wS1 to longer-lasting dynamics typical of tJM1 (Figures 1F and S1D). To better understand what information is encoded in each area, we next analyzed different trial types separately.

### Task-independent representations of pure sensory and motor information

We first analyzed trials in which mice did not lick in response to the whisker stimuli to assess the encoding of purely sensory information (Figures 2A–2D and S2A). We used receiver operating characteristic (ROC) analysis to split the neurons into positive and negative responses to the whisker stimulus and to compute the proportion of significantly modulated neurons in each area (Table S2A). We found a fast and brief evoked response in wS1 in both WR+ and WR− mice in the first 50 ms following the whisker stimulus, with 34% and 43% significantly positively modulated cells, respectively (Figure 2A). Little sensory-evoked response was observed in mPFC and tJM1 in both WR+ and WR− mice. We could decode significantly the presence of the stimulus (Miss vs. Correct rejection trials) from the population activity in wS1 in both WR+ and WR− mice (accuracy: WR+ =  $0.73 \pm 0.10$ ,  $p = 3.1 \times 10^{-5}$ ; WR− =  $0.76 \pm 0.10$ ,  $p = 7.8 \times 10^{-3}$ ; mean  $\pm$  SD, Wilcoxon signed-rank test). It was also possible to significantly decode the whisker stimulus, albeit with low accuracy, from mPFC and tJM1 activity in WR+ mice but not from WR− mice (Figure 2B; Table S2B). However, we did not find any significant difference in decoding accuracy between WR+ and WR− mice (Table S2B). Neurons in wS1 excited by whisker deflection showed a significantly higher correlation between the evoked response (change in firing rate) and the whisker stimulus amplitude, compared to mPFC and tJM1, in

### Figure 1. High-density extracellular recording and monitoring of orofacial movements during a psychometric whisker detection task

(A) Whisker-rewarded (WR+) mice (top) were trained to lick within 1 s following the whisker stimulus (go trials) but not in the absence of the stimulus (no-go trials), whereas for whisker non-rewarded (WR−) mice (bottom), the whisker stimulus was decorrelated to reward delivery. Lick probability ( $P_{lick}$ ; right) was determined during the 1 s following whisker stimulus as a function of the stimulus amplitude for behavioral sessions with electrophysiological recordings. The response probability was fitted with a sigmoid function for WR+ mice. Gray lines show individual recording sessions. Filled circles with error bars indicate mean response probability across sessions (mean  $\pm$  SD).

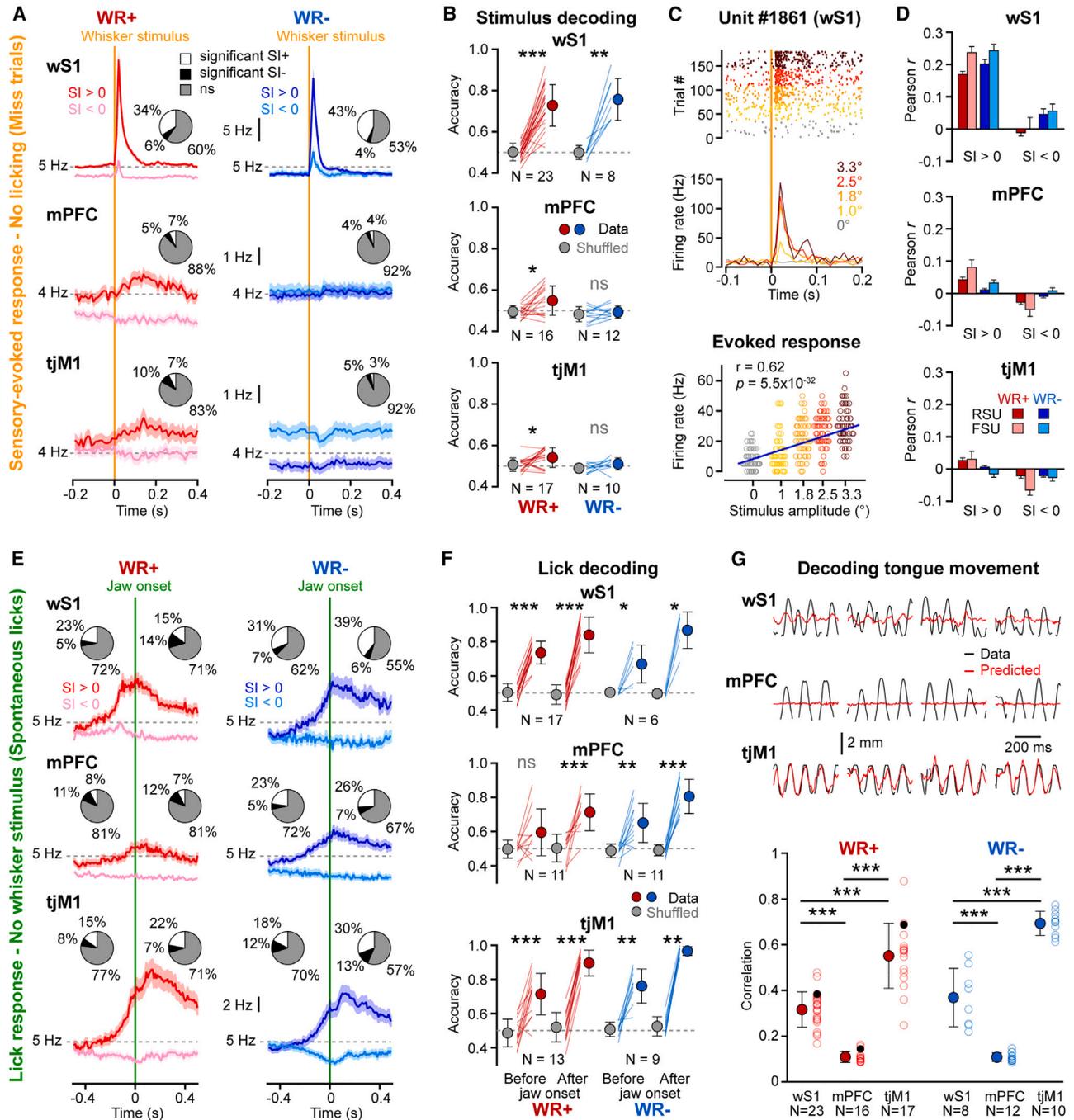
(B) Neuronal activity of wS1, tJM1, or mPFC was recorded with two silicon probes in head-restrained mice performing one of the two behavioral tasks. Facial movements were monitored using high-speed video filming.

(C) Example single Hit trial in a WR+ mouse showing (from top to bottom) raster plots of simultaneously recorded neuronal activity in wS1 and mPFC (neurons were sorted according to recording depth), whisker angular position, and jaw and tongue displacement. Upward deflections indicate, respectively, whisker protraction, jaw opening, and tongue protrusion. The green line on the jaw movement trace indicates the onset of jaw opening.

(D) Focal optogenetic inactivation during the response window was performed in VGAT-ChR2 mice for wS1, tJM1, mPFC, and fpS1 (30% trials) (left). Lick probability was determined as a function of the whisker stimulus amplitude for control (light off, black) and test (light on, blue) trials for the 4 cortical areas (right). Filled circles with error bars show mean  $\pm$  SD. Statistical significance between light-off and light-on trials was assessed using Wilcoxon signed-rank tests with Bonferroni correction for each area (5 tests): non-significant [ns]  $p > 0.05$ , \* $p < 0.05$ , and \*\* $p < 0.01$ .

(E) Grand-averaged peristimulus time histograms (PSTHs) of action potential (AP) firing (mean  $\pm$  SEM, 20 ms bins) for wS1, mPFC, and tJM1, as well as whisker and jaw displacement aligned to the whisker stimulus onset, for all stimulus trials of amplitudes 1.8°, 2.5°, and 3.3° (WR− mice, blue, left; WR+ mice, red, right).

(F) Clustering analysis of the neuronal responses for different trial types and alignments revealed 17 clusters in WR+ mice. Heatmaps showing the mean normalized activity for each neuron, grouped by clusters (left) according to Miss and Hit trials aligned to stimulus onset; Hit and Spontaneous licks (Spont. licks) aligned to jaw movement onset; and whisking aligned to whisking onset. Composition of each cluster was quantified according to cortical areas (wS1, red; mPFC, black; tJM1, blue) and neuron types (RS [regular spiking] units, dark colors; FS [fast-spiking] units, light colors) (right). Proportions are normalized according to the total number of recorded neurons in each area.



**Figure 2. Task-independent representation of sensory and motor information**

(A) Grand average sensory-evoked responses in Miss trials (mean  $\pm$  SEM, 10 ms bins) for WR+ (left) and WR- (right) mice. Neurons were divided into positive (dark colors) and negative (light colors) responses based on their maximum selectivity index (SI; Miss vs. Correct rejection trials) within the first 200 ms after the stimulus onset. Pie charts show the proportion of neurons with significant positive (SI+, white), negative (SI-, black), and ns (gray) SIs.

(B) Population decoding accuracy (Miss vs. Correct rejection trials, decoded in the 200 ms period after stimulus onset) for WR+ (red) and WR- (blue) mice. The performance of the decoder was assessed by comparing the results to label-shuffled (Shuffled) versions (gray filled circles) using Wilcoxon signed-rank tests: ns  $p > 0.05$ , \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$ . Lines indicate individual recording sessions. Filled circles with error bars show mean accuracy (mean  $\pm$  SD).

(C) Raster plot (top) and PSTHs (middle) from an example wS1 neuron in a WR+ mouse for the sensory-evoked responses in Miss trials sorted according to stimulus strength. For the same neuron, correlation between the number of evoked APs (200 ms after stimulus onset) and stimulus strength (bottom) is shown. Blue line shows linear fit. Pearson's coefficient of correlation ( $r$ ) and the  $p$  value are indicated.

(D) Mean  $r$  for the neurons recorded in wS1, mPFC, and tJM1. Neurons with a positive SI are on the left, and neurons with a negative SI are on the right. WR+ mice are in red, and WR- mice are in blue. RS units (RSUs) are in dark colors, and FS units (FSUs) are in light colors. Data are represented as mean  $\pm$  SEM.

(legend continued on next page)

both groups of mice (WR+ mice: wS1 vs. mPFC,  $p = 5.6 \times 10^{-27}$ ; wS1 vs. tJM1,  $p = 1.8 \times 10^{-31}$ ; WR– mice: wS1 vs. mPFC,  $p = 3.5 \times 10^{-46}$ ; wS1 vs. tJM1,  $p = 2.2 \times 10^{-45}$ ; Wilcoxon rank-sum test with Bonferroni correction) (Figures 2C and 2D; Table S2C). It was possible to decode the stimulus amplitude above chance level from the population activity in wS1 but not in mPFC or tJM1 (Figure S3A). Thus, the encoding of the physical properties of the whisker stimulus in the absence of licking appears to be restricted to wS1 and does not depend upon the learning of the whisker detection task.

We then investigated the encoding of pure motor information during non-rewarded Spontaneous lick bouts (Figures 2E–2G and S2A). Aligning the neuronal activity to the onset of the jaw movement revealed important changes in neuronal activity that started before jaw opening in all three areas in both WR+ and WR– mice (Figure 2E), with the highest proportion of significantly modulated cells found in wS1 (28% in WR+ mice and 38% in WR– mice) and the lowest in mPFC (19% in WR+ and 28% in WR– mice). After the jaw-opening onset, the proportion of significantly modulated cells increased further in all three areas but was highest in tJM1 and wS1 (tJM1-WR+ = 29%; tJM1-WR– = 43%; wS1-WR+ = 29%; wS1-WR– = 45%) and lowest in mPFC (19% in WR+ mice and 33% in WR– mice) (Table S2D). Accordingly, it was largely possible to decode the jaw opening before and after movement onset from the population activity of the three areas (Figure 2F; Table S2E). However, tJM1 achieved much better accuracy than mPFC or wS1 when decoding continuous tongue movements from the population activity (WR+ mice: tJM1 vs. wS1,  $p = 1.7 \times 10^{-5}$ ; tJM1 vs. mPFC,  $p = 3.2 \times 10^{-6}$ ; WR– mice: tJM1 vs. wS1,  $p = 1.4 \times 10^{-4}$ ; tJM1 vs. mPFC,  $p = 2.6 \times 10^{-4}$ ; Wilcoxon rank-sum test with Bonferroni correction) (Figures 2G and S3B–S3D; Table S2F). Thus, although movement initiation was broadly represented across cortical areas, as previously reported,<sup>16–18</sup> fine tongue movement kinematics were best represented in tJM1, in agreement with recent studies.<sup>23,24</sup>

### Sensory signals in wS1 and mPFC, but not tJM1, in Hit trials

We next examined Hit trials in which WR+ mice successfully responded to the whisker stimulus by licking (Figures 3 and S2B). Neuronal responses aligned to the stimulus onset showed a graded evoked response that increased with stimulus amplitude for positively responding (selectivity index [SI] > 0) neurons in all three areas (Figure 3A; Table S3A). It was possible to significantly decode the stimulus (Hit vs. False alarm) after the stimulus onset from all three areas (accuracy: wS1 =  $0.87 \pm 0.06$ ,  $p = 2.9 \times 10^{-4}$ ; mPFC =  $0.68 \pm 0.10$ ,  $p = 2.0 \times 10^{-3}$ ; tJM1 =  $0.74 \pm 0.12$ ,  $p = 7.3 \times 10^{-4}$ ; mean  $\pm$  SD, Wilcoxon signed-rank tests) (Figure 3B; Table S3B). However, when realigning the evoked responses

to the onset of jaw opening, responses to all whisker stimulus amplitudes as well as to Spontaneous licks overlapped in tJM1, whereas graded responses before the jaw onset were still observed in wS1 and mPFC. In good agreement, it was possible to decode significantly the stimulus (Hit vs. Spontaneous licks) before jaw-opening onset only from the population activity of wS1 and mPFC but not from the activity of tJM1 (accuracy: wS1 =  $0.73 \pm 0.09$ ,  $p = 2.9 \times 10^{-4}$ ; mPFC =  $0.64 \pm 0.11$ ,  $p = 4.2 \times 10^{-2}$ ; tJM1 =  $0.58 \pm 0.14$ ,  $p = 0.22$ ; mean  $\pm$  SD, Wilcoxon signed-rank tests) (Figure 3B; Table S3B).

The graded responses to whisker stimuli observed in tJM1 could be due to a decrease in both the mean and variability of reaction time with increasing stimulus amplitude (Figure 3C). Indeed, the evoked activity in tJM1 appeared to correlate with jaw-opening onset rather than stimulus onset time, whereas in wS1 and mPFC, the evoked activity was time locked to the whisker stimulus (Figure 3D). Comparing the evoked response for different reaction times for a given stimulus amplitude ( $2.5^\circ$ ) showed that the evoked response in wS1 and mPFC was locked to the stimulus onset time, whereas the response in tJM1 drifted in time as the reaction time increased (Figure 3E). Thus, in Hit trials, tJM1 mostly encoded motor-related information, whereas both wS1 and mPFC represented a mixture of sensory and motor-related information.

Realigning the neuronal activity to the jaw-opening onset time also revealed a sequence of activation in Hit trials with the peak of activity occurring first in wS1, then in mPFC, and finally in tJM1 (Figure S2D). Interestingly, in Spontaneous lick trials, the activation of wS1 also preceded the activation of mPFC and tJM1 in WR+ mice, whereas all three areas were simultaneously activated in WR– mice, suggesting that some Spontaneous licks in WR+ might be driven by false sensory percepts.

### Gated decision neurons in mPFC

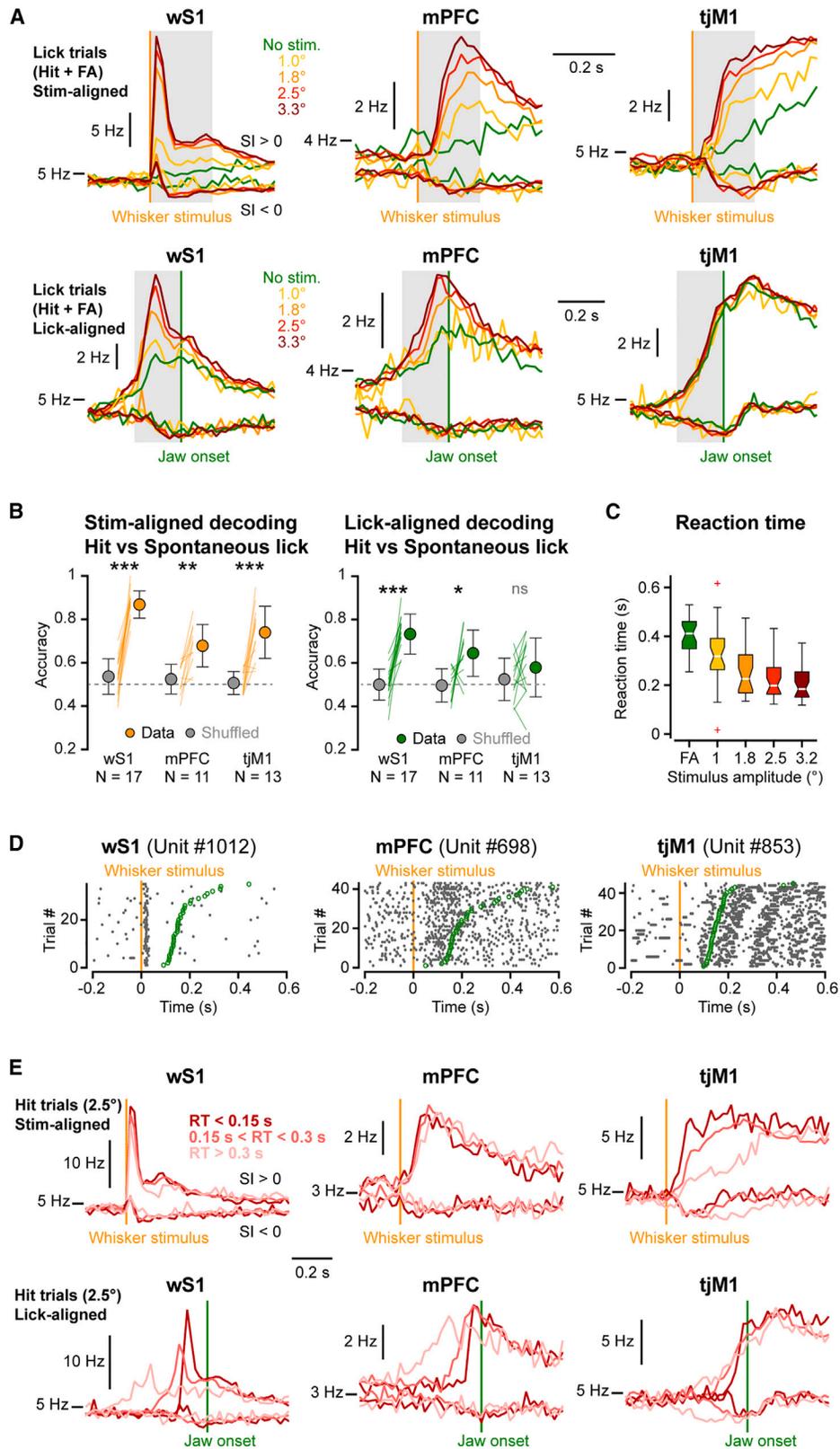
Finally, we investigated the decision to lick in response to the whisker stimulus (Figure 4). Neurons representing the decision to lick should distinguish between Hit and Miss trials. Clear differences in the neuronal evoked responses between Hit and Miss trials were observed for all stimulus amplitudes, in all three areas, including during the first 50 ms in wS1 (Figures 4A–4C; Table S4B). It was possible to significantly decode Hit from Miss trials for the near-threshold stimulus amplitude from the population activity in all three areas (accuracy: wS1 =  $0.76 \pm 0.11$ ,  $p = 4.6 \times 10^{-5}$ ; mPFC =  $0.60 \pm 0.10$ ,  $p = 2.6 \times 10^{-2}$ ; tJM1 =  $0.71 \pm 0.09$ ,  $p = 6.1 \times 10^{-5}$ ; mean  $\pm$  SD, Wilcoxon signed-rank tests) (Figure 4B; Table S4A).

However, neurons representing the decision to lick in response to the whisker stimulus should also distinguish between Hit and Spontaneous licks (Figures 4D–4G). We thus defined decision neurons as neurons with significant selectivity

(E) Similar to (A) but for averaged neuronal activity aligned to jaw-opening onset for Spontaneous licks in the absence of a sensory stimulus. The proportion of neurons with significant SI+ or SI– were computed 200 ms before (left) and after (right) jaw onset time.

(F) Similar to (B) but for decoding accuracy of licking computed 200 ms before (left) and after (right) jaw onset time.

(G) Example continuous decoding of tongue movements for long lick bouts showing predicted (red) and real (black) tongue movements from one example session for each area (4 licking bouts of 500 ms) (top). Correlation between real and predicted tongue movements for individual recording sessions (open circles, except black coloring to indicate the examples shown on top) and mean  $\pm$  SD (filled circles with error bars) for each area in WR+ (red) and WR– (blue) mice (bottom) are shown. Differences between areas were tested using Wilcoxon rank-sum tests with Bonferroni correction (3 tests): \*\*\* $p < 0.001$ .



(legend on next page)

for Hit vs. Miss trials just after the whisker stimulus for near-threshold stimulus amplitude and with significant selectivity for Hit vs. Spontaneous licks just before jaw-opening onset. Plotting one SI against the other revealed a small group of decision neurons with significant SIs for both Hit vs. Miss and Hit vs. Spontaneous licks (Figure 4D; Table S4C) that were found in all areas but with the highest proportion in wS1 (wS1 vs. tJM1:  $p = 1.9 \times 10^{-7}$ ; wS1 vs. mPFC:  $p = 1.9 \times 10^{-9}$ ; chi-squared test with Bonferroni correction) (Figure 4G; Table S4D) and were more abundant in deep layers (5 and 6) in all three areas (Figures S4A and S4B). Plotting the peristimulus time histograms (PSTHs) for the decision neurons with significant positive SIs for each area, aligned to the stimulus and the jaw onset times, revealed that decision neurons in wS1 still showed a clear sensory-evoked response in Miss trials and that neurons in tJM1 showed clear motor-related activity for Spontaneous licks. In contrast, decision neurons in mPFC showed the highest selectivity for Hit trials (Figure 4E). We thus subdivided decision neurons into sensory, motor, sensorimotor, and gated neurons depending on whether they also showed a significant sensory-evoked response in Miss trials (sensory neurons), a significant motor response during Spontaneous licks (motor neurons), both (sensorimotor neurons), or none (gated neurons) (Figures 4F and S4). Interestingly, decision neurons in mPFC had the highest proportion of gated neurons responding selectively in Hit trials but not in Miss trials or Spontaneous licks (mPFC vs. wS1:  $p = 1.1 \times 10^{-6}$ ; mPFC vs. tJM1:  $p = 0.026$ ; chi-squared test with Bonferroni correction) (Figure 4G; Table S4D). These neurons were predominantly found in deep layers of the ventral mPFC (infralimbic and medial orbitofrontal areas) (Figures S4C and S4D).

## DISCUSSION

### Representation of sensory and motor information irrespective of learning

In good agreement with previous studies, we found that task-evoked activity is highly distributed across areas when mice lick in response to a sensory stimulus, due both to sensory and motor-related activity.<sup>8,16,18,31</sup> Over the last decade, many studies have established that motor activity has a global impact on nearly all brain regions and that task-driven computations are embedded within a broader context of movement-related information.<sup>8,15–17,32–34</sup> However, by looking more specifically at sensory activity in the absence of licking, we found that pure sensory

information is best represented in the primary sensory cortex (Figures 2A–2D). Similarly, fine movement kinematics were best represented in the primary motor cortex (Figures 2E–2G), in good agreement with recent studies showing that tJM1 neurons encode both the direction and the amplitude of tongue movements during licking.<sup>23,24</sup> The representation of the physical properties of the sensory stimulus in the primary sensory cortex and of movement kinematics in the primary motor cortex appeared to be a stable property of the cortical circuits in these areas, although it does not rule out a possible refinement of the representation through learning.<sup>35–38</sup>

### Representation of sensory information develops with learning in mPFC

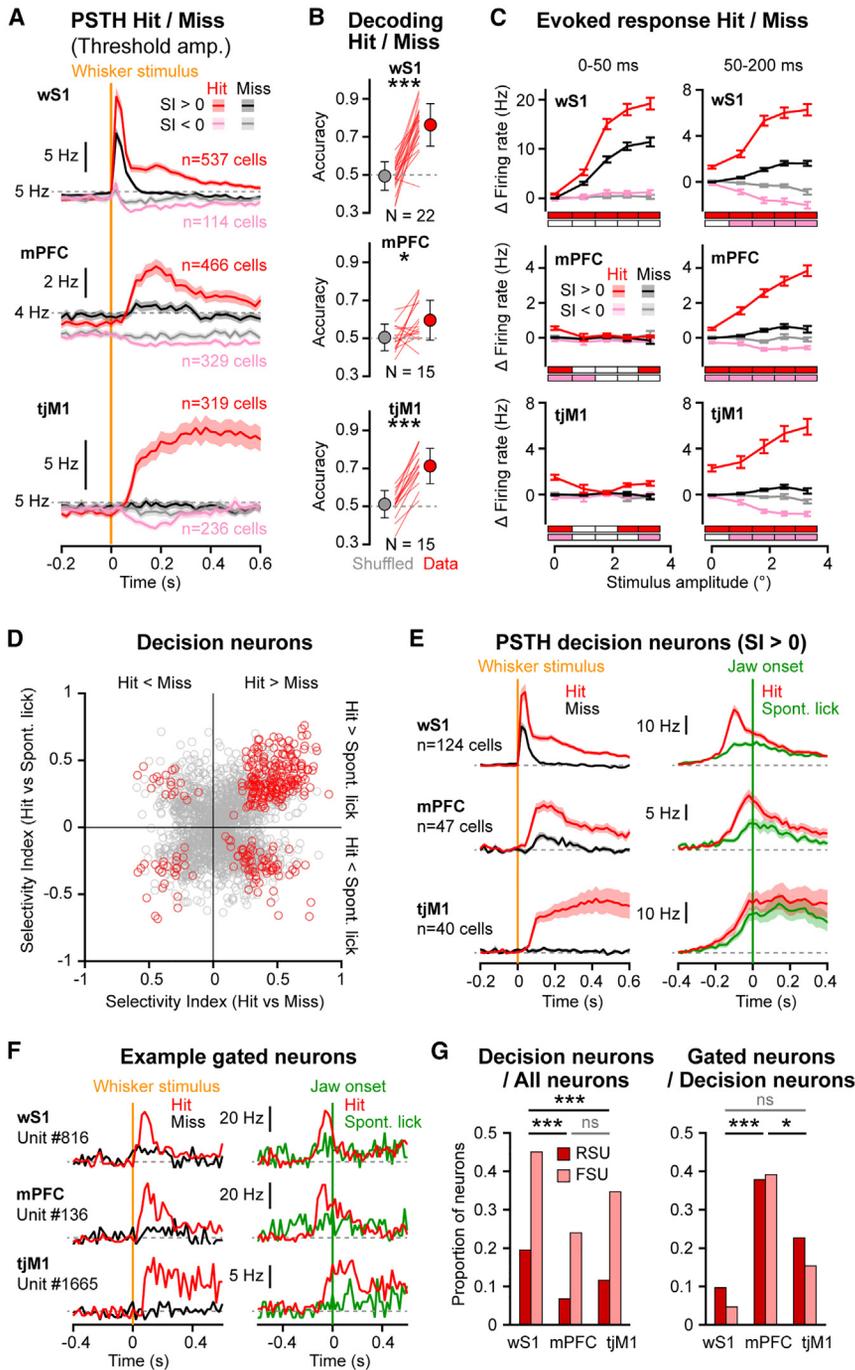
Previous investigations have reported evoked responses for behaviorally relevant sensory stimuli that developed during learning in mPFC.<sup>26,31,39,40</sup> In good agreement, we found that neurons in mPFC had prominent responses in Hit trials of WR+ mice (Figure 3) but not in Miss trials and not in mice that were passively exposed to unrewarded whisker stimuli (WR– mice). Since mPFC activity is modulated by licking in both WR+ and WR– mice, the Hit-related activity that develops across learning in mPFC could simply reflect the acquired motor response. However, when aligning neuronal activity to jaw movement onset, we found that it was still possible to decode Hit trials from Spontaneous licks before movement onset in mPFC. Furthermore, a graded response that correlated with stimulus amplitude, before the movement onset, was observed in mPFC. Thus, mPFC activity cannot be fully explained by motor activity and a sensory-related activity developed in WR+ mice that had learned to associate the sensory stimulus with the reward. The graded response before movement onset could represent either encoding of physical properties of the stimulus (stimulus amplitude) or an increase in reward expectation for stronger stimuli.<sup>41</sup>

### Decision information is distributed but more selective in mPFC

Hit vs. Miss differences are commonly used to characterize perceptual decisions in detection tasks. However, decision and motor responses are often highly correlated and thus difficult to disentangle. To address this issue, we proposed a more stringent definition of decision information coding, where a neuron is considered as representing decision if it can distinguish between Hit and Miss trials just after the stimulus presentation as well as distinguishing Hit trials and Spontaneous licks just before

### Figure 3. Sensory stimulus representation in Hit trials in wS1 and mPFC but not tJM1

(A) Stimulus-aligned (top) and jaw-onset-aligned (bottom) grand average neuronal responses for different stimulus amplitudes in Hit trials, and False alarm (FA) trials, in wS1 (left), mPFC (middle), and tJM1 (right). Neurons were divided into positive ( $SI > 0$ ) and negative ( $SI < 0$ ) responses.  
 (B) Decoding accuracy of Hit vs. Spontaneous lick trials 200 ms after stimulus onset time (left) or 200 ms before lick onset time (right). Real data: orange and green filled circles; Shuffled data: gray filled circles (mean  $\pm$  SD). Orange and green lines show accuracies for individual recording sessions. Significant differences between shuffled and real data were tested using Wilcoxon signed-rank tests: ns  $p > 0.05$ , \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$ .  
 (C) Boxplot showing the mean reaction time (jaw-opening onset time) for each stimulus amplitude across all the recording sessions. The box plots show the median (white bar), the upper and lower quartiles, the minimum and maximum (whiskers), and the outliers (crosses).  
 (D) Example raster plots of neurons recorded in wS1 (left), mPFC (middle), and tJM1 (right) for Hit trials (stimulus amplitude  $2.5^\circ$ ) aligned to the stimulus onset and sorted by reaction times (green circles).  
 (E) Grand average PSTHs for Hit trials with stimulus amplitude  $2.5^\circ$  with different reaction times (RTs) and aligned to the stimulus onset (top) or the jaw onset (bottom). Trials with early ( $<0.15$  s), intermediate ( $>0.15$  and  $<0.3$  s), and late ( $>0.3$  s) RTs are displayed with different shades of red. Neurons are divided into positive and negative SIs.



**Figure 4. Decision-related information is encoded in all three areas**

(A) Comparison of grand average neuronal responses for Hit (red and pink) and Miss (black and gray) trials for near-threshold stimulus amplitude aligned to the whisker stimulus (mean  $\pm$  SEM). Neurons were divided based on their SI for Hit vs. Correct rejection trials as in Figure 3.

(B) Population decoding accuracy of Hit vs. Miss trials for the near-threshold stimulus amplitude 200 ms after stimulus onset. Significant differences between Shuffled (gray) and real data (red) accuracies were tested using Wilcoxon signed-rank test: \* $p < 0.05$  and \*\*\* $p < 0.001$ . Lines indicate decoding from individual recording sessions. Filled circles with error bars show mean accuracy (mean  $\pm$  SD).

(C) Grand average mean evoked firing rate (mean  $\pm$  SEM) as a function of the stimulus amplitude for Hit (red and pink) and Miss (black and gray) trials computed within the first 50 ms (left) and 50 to 200 ms (right) after the whisker stimulus. Red and black, neurons with positive Hit vs. Correct rejection SIs; pink and gray, neurons with negative SIs. Colored bars below each graph indicate significant differences between Hit and Miss mean evoked firing rates for each amplitude (Wilcoxon signed-rank tests with Bonferroni correction across amplitudes): top bars, red  $p < 0.05$  for  $SI > 0$ ; bottom bars, pink  $p < 0.05$ ; for  $SI < 0$ ; white, ns.

(D) SI for Hit vs. Spontaneous licks (200 ms before jaw onset) plotted against the SI for Hit vs. Miss trials (200 ms after stimulus onset). Each circle represents one neuron (all neurons from all areas are plotted together). Red circles indicate neurons for which both SIs were significant.

(E) Averaged PSTHs (mean  $\pm$  SEM, 20 ms bins) of neurons for which both SIs were significant and positive for each area. Hit (red) and Miss (black) trials (near-threshold amplitude) aligned to the stimulus onset (left). Hit (strongest stimulus amplitude, red) and Spontaneous lick (green) trials aligned to the jaw onset (right).

(F) PSTHs for example gated neurons: Hit and Miss trials, near-threshold amplitude, and stimulus aligned (left) and Hit and Spontaneous lick trials, strongest amplitude, and jaw onset aligned (right). (G) Proportion of RSU and FSU decision neurons (for which both SIs were significant) relative to the total number of RSUs and FSUs recorded in each area (left). Proportion of gated neurons among the decision neurons in each area (right).

movement initiation. By doing so, we identified a small proportion of decision neurons in all three areas but with the highest proportion in wS1 (Figure 4G). In line with this finding, a recent study has identified decision neurons in wS1 as mice performed a two-alternative forced-choice texture discrimination task and provided causal evidence of the role of these neurons in driving behavior and enhancing performance.<sup>13</sup> Neurons encoding behavioral decisions have also been found in the primary auditory cortex as mice performed an auditory discrimination task.<sup>14</sup> The impor-

tance of wS1 in the decision process in our task likely reflects the success or failure of sensory perception (detection), the first essential step in goal-directed sensorimotor transformation. Overall, our results are in good agreement with recent studies showing a broad representation of decision information across cortical areas.<sup>7,16</sup> Importantly, it should be noted that more complex tasks requiring integration of multimodal sensory information or short-term memory would probably involve additional frontal cortical areas in the decision process.<sup>6,7,18</sup>

In contrast to the expected role of mPFC in decision-making, the proportion of decision neurons was not the highest in mPFC.

However, mPFC decision neurons appeared to have the highest selectivity when Hit vs. Miss responses and Hit vs. Spontaneous lick activity were jointly considered. Gated neurons—which we defined as not responding to sensory stimulus alone or to licking alone—were found in the highest proportion among decision neurons in mPFC (Figure 4G). Thus, altogether, our results show that mPFC activity encodes a mixture of task-related sensory, motor, and decision information, in good agreement with previous studies<sup>26,31,39–42</sup> (Figure S4E). Optogenetic inactivation of mPFC strongly impairs mouse performance in sensory detection (Figure 1D) and discrimination tasks,<sup>18,26,31</sup> suggesting it plays a direct role in governing the sensory-to-motor transformation, perhaps through enforcing specific learned task rules and instructing other brain areas about the appropriate behavioral response.<sup>43</sup> In addition, mPFC likely plays many other roles such as keeping track of trial history, outcome, and context.<sup>41,44–46</sup> In our study, we focused our analysis on the neuronal activity immediately after the sensory stimulus or before movement, but we also observed some clusters with a high proportion in mPFC that exhibit late and selective activity on Hit trials or Spontaneous lick trials (Figures 1F and S1D). Such neurons could play an important role in encoding trial feedback that may be used for learning or updating behavior on a trial-by-trial basis.

### Limitations of the study

A critical aspect of our study is the definition of decision neurons. Here, we defined sensory-guided decision-making as the specific learned response of the mouse to the whisker stimulus preceding the initiation of licking in Hit trials. By doing so, we identified gated neurons whose activity cannot be explained by the sensory or motor response but can only reflect this decision process that links the sensory stimulus to the appropriate action. In future experiments, it will also be important to investigate two-alternative forced-choice (2AFC) tasks, which are considered to be better suited to address the decision process because, regardless of the choice, the animal has to make a motor response and because they allow us to distinguish between an error in the decision (wrong choice) and an absence of response due to disengagement.<sup>47,48</sup>

Another important aspect of the sensorimotor transformation that we could not address in the current study is how information flows across the different cortical areas. Future experiments combining simultaneous high-density recordings from multiple brain regions and optogenetic manipulations of targeted neuronal populations will help address causal mechanisms of interareal communication.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Experimental design
  - Implantation of metal head-post
  - Skull preparation and craniotomies
  - Behavioral paradigm
  - Electrophysiological recordings
  - Optogenetic manipulations
  - Histology and localization of electrode/optical fiber tracks
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral data analysis
  - Quantification of orofacial movements
  - Electrophysiology data
  - Clustering neuronal responses
  - Receiver Operating Characteristic analysis
  - Pearson correlations
  - Population decoding
  - Statistics

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2023.113618>.

### ACKNOWLEDGMENTS

We thank Candice Stoudmann for help with histology and mouse breeding, the members of the Petersen laboratory for helpful discussions, and James Poulet for constructive comments on the manuscript. This work was supported by the Swiss National Science Foundation: grants 31003A\_182010 (C.C.H.P.), 310030\_146252 (C.C.H.P.), and TMAG-3\_209271 (C.C.H.P.).

### AUTHOR CONTRIBUTIONS

A.O., C.C.H.P., and S.C. conceptualized the study; A.O. performed the experiments; A.O. obtained histological data; A.O., C.S., J.W., and R.A. analyzed the data; V.E., A.M., and W.G. advised on clustering and decoding of neuronal data; A.O., C.C.H.P., and S.C. wrote the manuscript; all authors discussed and edited the manuscript; and C.C.H.P. and S.C. provided overall supervision.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 13, 2023

Revised: October 27, 2023

Accepted: December 8, 2023

Published: December 26, 2023

### REFERENCES

1. Romo, R., and de Lafuente, V. (2013). Conversion of sensory signals into perceptual decisions. *Prog. Neurobiol.* 103, 41–75.
2. Svoboda, K., and Li, N. (2018). Neural mechanisms of movement planning: motor cortex and beyond. *Curr. Opin. Neurobiol.* 49, 33–41.
3. Crochet, S., Lee, S.-H., and Petersen, C.C.H. (2019). Neural circuits for goal-directed sensorimotor transformations. *Trends Neurosci.* 42, 66–77.
4. Gallero-Salas, Y., Han, S., Sych, Y., Voigt, F.F., Laurenczy, B., Gilad, A., and Helmchen, F. (2021). Sensory and behavioral components of

- neocortical signal flow in discrimination tasks with short-term memory. *Neuron* 109, 135–148.e6.
5. Cruz, K.G., Leow, Y.N., Le, N.M., Adam, E., Huda, R., and Sur, M. (2023). Cortical-subcortical interactions in goal-directed behavior. *Physiol. Rev.* 103, 347–389.
  6. 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.
  7. Pinto, L., Rajan, K., DePasquale, B., Thiberge, S.Y., Tank, D.W., and Brody, C.D. (2019). Task-dependent changes in the large-scale dynamics and necessity of cortical regions. *Neuron* 104, 810–824.e9.
  8. Zatzka-Haas, P., Steinmetz, N.A., Carandini, M., and Harris, K.D. (2021). Sensory coding and the causal impact of mouse cortex in a visual decision. *eLife* 10, e63163.
  9. de Lafuente, V., and Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* 8, 1698–1703.
  10. Alvarez, M., Zainos, A., and Romo, R. (2015). Decoding stimulus features in primate somatosensory cortex during perceptual categorization. *Proc. Natl. Acad. Sci. USA* 112, 4773–4778.
  11. 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.e2.
  12. Lee, C.C.Y., Kheradpezhoh, E., Diamond, M.E., and Arabzadeh, E. (2020). State-dependent changes in perception and coding in the mouse somatosensory cortex. *Cell Rep.* 32, 108197.
  13. Buetfering, C., Zhang, Z., Pitsiani, M., Smallridge, J., Boven, E., McElligott, S., and Häusser, M. (2022). Behaviorally relevant decision coding in primary somatosensory cortex neurons. *Nat. Neurosci.* 25, 1225–1236.
  14. Francis, N.A., Mukherjee, S., Koçillari, L., Panzeri, S., Babadi, B., and Kanoold, P.O. (2022). Sequential transmission of task-relevant information in cortical neuronal networks. *Cell Rep.* 39, 110878.
  15. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* 22, 1677–1686.
  16. Steinmetz, N.A., Zatzka-Haas, P., Carandini, M., and Harris, K.D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature* 576, 266–273.
  17. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brain-wide activity. *Science* 364, 255.
  18. Esmaeili, V., Tamura, K., Muscinelli, S.P., Modirshanechi, A., Boscaglia, M., Lee, A.B., Oryshchuk, A., Foustoukos, G., Liu, Y., Crochet, S., et al. (2021). Rapid suppression and sustained activation of distinct cortical regions for a delayed sensory-triggered motor response. *Neuron* 109, 2183–2201.e9.
  19. Avitan, L., and Stringer, C. (2022). Not so spontaneous: Multi-dimensional representations of behaviors and context in sensory areas. *Neuron* 110, 3064–3075.
  20. Esmaeili, V., Tamura, K., Foustoukos, G., Oryshchuk, A., Crochet, S., and Petersen, C.C.H. (2020). Cortical circuits for transforming whisker sensation into goal-directed licking. *Curr. Opin. Neurobiol.* 65, 38–48.
  21. Diamond, M.E., and Arabzadeh, E. (2013). Whisker sensory system - from receptor to decision. *Prog. Neurobiol.* 103, 28–40.
  22. Kwon, S.E., Yang, H., Minamisawa, G., and O'Connor, D.H. (2016). Sensory and decision-related activity propagate in a cortical feedback loop during touch perception. *Nat. Neurosci.* 19, 1243–1249.
  23. Mayrhofer, J.M., El-Boustani, S., Foustoukos, G., Auffret, M., Tamura, K., and Petersen, C.C.H. (2019). Distinct contributions of whisker sensory cortex and tongue-jaw motor cortex in a goal-directed sensorimotor transformation. *Neuron* 103, 1034–1043.e5.
  24. Xu, D., Dong, M., Chen, Y., Delgado, A.M., Hughes, N.C., Zhang, L., and O'Connor, D.H. (2022). Cortical processing of flexible and context-dependent sensorimotor sequences. *Nature* 603, 464–469.
  25. Le Merre, P., Åhrlund-Richter, S., and Carlén, M. (2021). The mouse prefrontal cortex: Unity in diversity. *Neuron* 109, 1925–1944.
  26. Pinto, L., and Dan, Y. (2015). Cell-type-specific activity in prefrontal cortex during goal-directed behavior. *Neuron* 87, 437–450.
  27. Stüttgen, M.C., and Schwarz, C. (2008). Psychophysical and neurometric detection performance under stimulus uncertainty. *Nat. Neurosci.* 11, 1091–1099.
  28. Takahashi, N., Oertner, T.G., Hegemann, P., and Larkum, M.E. (2016). Active cortical dendrites modulate perception. *Science* 354, 1587–1590.
  29. Sachidhanandam, S., Sreenivasan, V., Kyriakatos, A., Kremer, Y., and Petersen, C.C.H. (2013). Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat. Neurosci.* 16, 1671–1677.
  30. 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.
  31. Le Merre, P., Esmaeili, V., Charrière, E., Galan, K., Salin, P.-A., Petersen, C.C.H., and Crochet, S. (2018). Reward-based learning drives rapid sensory signals in medial prefrontal cortex and dorsal hippocampus necessary for goal-directed behavior. *Neuron* 97, 83–91.e5.
  32. Poulet, J.F.A., and Crochet, S. (2018). The cortical states of wakefulness. *Front. Syst. Neurosci.* 12, 64.
  33. Ayaz, A., Stäuble, A., Hamada, M., Wulf, M.A., Saleem, A.B., and Helmchen, F. (2019). Layer-specific integration of locomotion and sensory information in mouse barrel cortex. *Nat. Commun.* 10, 2585.
  34. Minderer, M., Brown, K.D., and Harvey, C.D. (2019). The spatial structure of neural encoding in mouse posterior cortex during navigation. *Neuron* 102, 232–248.e11.
  35. 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.
  36. Poort, J., Wilmes, K.A., Blot, A., Chadwick, A., Sahani, M., Clopath, C., Msrisc-Fogel, T.D., Hofer, S.B., and Khan, A.G. (2022). Learning and attention increase visual response selectivity through distinct mechanisms. *Neuron* 110, 686–697.e6.
  37. Jurjut, O., Georgieva, P., Busse, L., and Katzner, S. (2017). Learning enhances sensory processing in mouse V1 before improving behavior. *J. Neurosci.* 37, 6460–6474.
  38. Peters, A.J., Chen, S.X., and Komiyama, T. (2014). Emergence of reproducible spatiotemporal activity during motor learning. *Nature* 510, 263–267.
  39. Otis, J.M., Namboodiri, V.M.K., Matan, A.M., Voets, E.S., Mohorn, E.P., Kosyk, O., McHenry, J.A., Robinson, J.E., Resendez, S.L., Rossi, M.A., and Stuber, G.D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature* 543, 103–107.
  40. Peters, A.J., Marica, A.-M., Fabre, J.M.J., Harris, K.D., and Carandini, M. (2022). Visuomotor learning promotes visually evoked activity in the medial prefrontal cortex. *Cell Rep.* 41, 111487.
  41. Lak, A., Okun, M., Moss, M.M., Gurnani, H., Farrell, K., Wells, M.J., Reddy, C.B., Kepecs, A., Harris, K.D., and Carandini, M. (2020). Dopaminergic and prefrontal basis of learning from sensory confidence and reward value. *Neuron* 105, 700–711.e6.
  42. Kim, E., Bari, B.A., and Cohen, J.Y. (2021). Subthreshold basis for reward-predictive persistent activity in mouse prefrontal cortex. *Cell Rep.* 35, 109082.
  43. Brockett, A.T., Tennyson, S.S., deBettencourt, C.A., Kallmyer, M., and Roesch, M.R. (2022). Medial prefrontal cortex lesions disrupt prepotent action selection signals in dorsomedial striatum. *Curr. Biol.* 32, 3276–3287.e3.

44. Moorman, D.E., and Aston-Jones, G. (2015). Prefrontal neurons encode context-based response execution and inhibition in reward seeking and extinction. *Proc. Natl. Acad. Sci. USA* *112*, 9472–9477.
45. Lui, J.H., Nguyen, N.D., Grutzner, S.M., Darmanis, S., Peixoto, D., Wagner, M.J., Allen, W.E., Kebschull, J.M., Richman, E.B., Ren, J., et al. (2021). Differential encoding in prefrontal cortex projection neuron classes across cognitive tasks. *Cell* *184*, 489–506.e26.
46. Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., and Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell* *184*, 2750–2766.e17.
47. Carandini, M., and Churchland, A.K. (2013). Probing perceptual decisions in rodents. *Nat. Neurosci.* *16*, 824–831.
48. Aguillon-Rodriguez, V., Angelaki, D., Bayer, H., Bonacchi, N., Carandini, M., Cazettes, F., Chapuis, G., Churchland, A.K., Dan, Y., et al.; International Brain Laboratory (2021). Standardized and reproducible measurement of decision-making in mice. *eLife* *10*, e63711.
49. Zhao, S., Ting, J.T., Atallah, H.E., Qiu, L., Tan, J., Gloss, B., Augustine, G.J., Deisseroth, K., Luo, M., Graybiel, A.M., and Feng, G. (2011). Cell type-specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nat. Methods* *8*, 745–752.
50. Yamashita, T., and Petersen, C.C.H. (2016). Target-specific membrane potential dynamics of neocortical projection neurons during goal-directed behavior. *eLife* *5*, e15798.
51. Matteucci, G., Guyoton, M., Mayrhofer, J.M., Auffret, M., Foustoukos, G., Petersen, C.C.H., and El-Boustani, S. (2022). Cortical sensory processing across motivational states during goal-directed behavior. *Neuron* *110*, 4176–4193.e10.
52. Shamash, P., Carandini, M., Harris, K.D., and Steinmetz, N.A. (2018). A tool for analyzing electrode tracks from slice histology.. Preprint at bioRxiv.
53. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* *21*, 1281–1289.
54. Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M., and Kenneth D, H. (2016). Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. Preprint at bioRxiv.
55. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., and Redish, A.D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* *131*, 1–11.
56. Hill, D.N., Mehta, S.B., and Kleinfeld, D. (2011). Quality metrics to accompany spike sorting of extracellular signals. *J. Neurosci.* *31*, 8699–8705.
57. Esmaeili, V., Oryshchuk, A., Asri, R., Tamura, K., Foustoukos, G., Liu, Y., Guiet, R., Crochet, S., and Petersen, C.C.H. (2022). Learning-related congruent and incongruent changes of excitation and inhibition in distinct cortical areas. *PLoS Biol.* *20*, e3001667.
58. McCormick, D.A., Connors, B.W., Lighthall, J.W., and Prince, D.A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J. Neurophysiol.* *54*, 782–806.
59. Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Omelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W., and Witten, I.B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* *570*, 509–513.
60. Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature* *497*, 585–590.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Chemicals, peptides, and recombinant proteins</b>		
Dil (1,1'-Dioctadecyl-3,3,3',3'-Tetramethylindocarbocyanine Perchlorate)	Invitrogen, USA	Cat# D282
Loctite super glue	Henkel, Germany	401
Silicone elastomer	World Precision Instruments, USA	Kwik-Cast
32% paraformaldehyde (PFA)	Electron Microscopy Science, USA	32% PFA solution
Self-curing denture acrylic	Kulzer, Germany	Paladur
<b>Deposited data</b>		
Dataset and MATLAB analysis code	This study	<a href="https://doi.org/10.5281/zenodo.10115924">https://doi.org/10.5281/zenodo.10115924</a>
Allen Mouse Common Coordinate Framework	Allen Mouse Brain Atlas	<a href="https://alleninstitute.github.io/AllenSDK/reference_space.html">https://alleninstitute.github.io/AllenSDK/reference_space.html</a>
<b>Experimental models: Organisms/strains</b>		
Mouse: B6.Cg-Tg(Slc32a1-COP4*H134R/EYFP)8Gfng/J	The Jackson Laboratory	JAX: 014548
Mouse: C57BL/6 wild type	Janvier (France)	C57BL/6Jrj
<b>Software and algorithms</b>		
MATLAB R2021a	MathWorks, USA	SCR_001622
Kilosort2	Pachitariu et al., 2016 <sup>54</sup> <a href="https://doi.org/10.1101/061481">https://doi.org/10.1101/061481</a>	<a href="https://github.com/jamesjun/Kilosort2">https://github.com/jamesjun/Kilosort2</a>
Allen CCF tools	Shamash et al., 2018 <sup>52</sup> <a href="https://doi.org/10.1101/447995">https://doi.org/10.1101/447995</a>	<a href="https://github.com/cortex-lab/allenCCF">https://github.com/cortex-lab/allenCCF</a>
DeepLabCut 2.2b7	Mathis et al., 2018 <sup>53</sup> <a href="https://doi.org/10.1038/s41593-018-0209-y">https://doi.org/10.1038/s41593-018-0209-y</a>	<a href="https://github.com/DeepLabCut/DeepLabCut">https://github.com/DeepLabCut/DeepLabCut</a>
<b>Other</b>		
32-channel silicon probe	NeuroNexus, USA	A1x32-Poly2-10mm-50 s-177
Optrode	NeuroNexus, USA	A1x32-Poly3-10mm-50 s-177-OA32
Digital headstage	Blackrock Microsystems, USA	CerePlex™ M32
Data acquisition system	Blackrock Microsystems, USA	CerePlex™ Direct
Blue Laser	GMP SA, Switzerland	MBL-F-473/200mW
470-nm high power LED	Thorlabs, USA	M470F3
200-μm diameter 0.22NA patch cable	Thorlabs, USA	M84L01
200-μm diameter 0.22NA implantable	Thorlabs, USA	CFMXA05
Fiber Optic Cannulae		
High speed camera	Optronis, Germany	CL 600 X 2/M

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sylvain Crochet ([sylvain.crochet@epfl.ch](mailto:sylvain.crochet@epfl.ch)).

#### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- The processed data (spike times, probe locations, behavior, orofacial movements) have been deposited at Zenodo and are publicly available at <https://doi.org/10.5281/zenodo.10115924>.
- All original code used to analyze the data has been deposited at Zenodo and is publicly available at <https://doi.org/10.5281/zenodo.10115924>.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All procedures were approved by the Swiss Federal Veterinary Office (License number VD-1628) and were conducted in accordance with the Swiss guidelines for the use of research animals. For optogenetic inactivation, we used VGAT-ChR2 mice [B6.Cg-Tg(Slc32a1-COP4\*H134R/EYFP)8Gfng/J, JAX: 014548].<sup>49</sup> For electrophysiological recordings, we used C57BL/6 wild-type mice and VGAT-ChR2 mice. Mice (male and female) were at least 6 weeks old at the time of head-post implantation (see below) and were kept in a reverse light/dark cycle (light from 7 p.m. to 7 a.m.), in ventilated cages at a temperature of  $22 \pm 2^\circ\text{C}$  with food available *ad libitum*. During behavioral training, water was restricted to 1 mL a day with at least 2 days of free-access to water in the home cage every 2 weeks. All mice were weighed and inspected daily during behavioral training.

### METHOD DETAILS

#### Experimental design

This study did not involve randomization or blinding. We did not estimate sample size before carrying out the study.

#### Implantation of metal head-post

Mice were first implanted with a metal head-post under anesthesia using a mixture of ketamine and xylazine (ketamine: 125 mg/kg, xylazine: 10 mg/kg, i.p.). Carprofen was injected (100  $\mu\text{L}$  at 0.5 mg/mL, i.p.) for analgesia before the start of surgery. Body temperature was kept at  $37^\circ\text{C}$  throughout the surgery with a heating pad. An ocular ointment (Vita-Pos, Pharma Medica AG, Switzerland) was applied over the eyes to prevent them from drying. A mix of lidocaine and bupivacaine was injected below the scalp as local analgesic and a solution of povidone-iodine (Betadine, Mundipharma Medical Company, Bermuda) was applied for skin disinfection before skin incision. A part of the scalp was removed with surgical scissors to expose the skull which was then cleaned with cotton buds and a scalpel blade to remove the periosteal tissue. After disinfection with Betadine, the skull was dried with cotton buds and a thin layer of super glue (Loctite super glue 401, Henkel, Germany) was applied across the dorsal part of the skull. A custom-made head fixation implant was glued to the right hemisphere. The head implant was then further secured with self-curing denture acrylic (Paladur, Kulzer, Germany or Ortho-Jet, Lang, USA). A chamber was built with denture acrylic along the edge of the bone covering the left hemisphere. This intact, transparent skull preparation was used to perform intrinsic optical signal (IOS) imaging experiments. Mice were returned to their home cages and ibuprofen (Algifor Dolo Junior, Verfora SA, Switzerland) was added to the drinking water for three days after surgery.

#### Skull preparation and craniotomies

For electrophysiological recordings, up to 3 small craniotomies (<0.5 mm in diameter) were made over the regions of interest in the left hemisphere using a dental drill under isoflurane anesthesia (2–3% in  $\text{O}_2$ ). The craniotomies were protected using a silicone elastomer (Kwik-Cast, World Precision Instruments, Sarasota, FL, USA). Regions of interest were the whisker primary somatosensory cortex (wS1), the medial prefrontal cortex (mPFC) and the tongue-jaw primary motor cortex (tjM1), selected based on previous studies.<sup>18,23,31</sup> The right C2 whisker representation in the left wS1 was mapped using IOS imaging under isoflurane anesthesia (1–1.5% in  $\text{O}_2$ ). For the other regions, stereotaxic coordinates relative to bregma were used: tjM1: AP 2.0 mm; Lat 2.5 mm; and mPFC: AP 2 mm; Lat 0.5 mm. For optogenetic inactivation experiments, the bone over the regions of interest was thinned and a thin layer of superglue was applied to protect the skull for stable optical access over days. For the inactivation of mPFC, a small craniotomy was made for the insertion of an optical fiber.

#### Behavioral paradigm

Head-restrained, water-restricted mice were trained in either of two different behavioral tasks. Whisker-rewarded mice (WR+) were trained in a psychometric version of a whisker-detection task to lick for water reward in response to a brief single-whisker deflection of variable amplitude. Whisker non-rewarded mice (WR-) were trained to lick for reward at random times while being exposed to the same range of whisker stimuli that were not predictive of the reward delivery. During all sessions, an ambient white noise was played continuously at 80 dB to mask distracting noises as well as the sound artifacts produced by the whisker stimulus.

The task structure closely resembled that previously used in simple whisker detection tasks<sup>29,31,50</sup> but with varying strength of the whisker stimulation. New trials started without any preceding cue following a variable 6–12-s inter-trial interval that included a final variable 2.5–3.5-s no-lick period, during which mice were required not to lick, otherwise trial initiation was aborted. Stimulus trials and

No stimulus trials (Catch trials) were presented randomly with equal probabilities. Stimulus trials included four whisker stimulus amplitudes of 1°, 1.8°, 2.5°, and 3.3° deflection of the right C2 whisker, also delivered with equal probabilities. All trial types were presented in a pseudo-randomized manner. Mice were required to lick the water spout within a 1-s response window following the whisker stimulus in order to receive a drop of sweet water as a reward. Contacts of the tongue with the water spout were detected online by a piezo-electric sensor attached to the spout. Trials when the mouse licked the water spout within the 1-s response window after a whisker stimulus, were considered as Hit trials, and were rewarded with a drop (4  $\mu$ L) of sweet water; if the mouse did not lick within the response window after a whisker stimulus, no reward was delivered, and the trial was classified as a Miss; if mice licked within the response window when no whisker stimulation was delivered (Catch trials), the trial was classified as a False alarm; and if no licking occurred during catch trials, they were considered as Correct rejection trials.

Mice were trained daily with one training session a day and a break of 2 days with free access to water every 2 weeks. Training of naive animals started with two days of free-licking and handling, during which mice were gradually accustomed to head-fixation and learned to associate licking of the reward spout with water delivery. Following the training on the second day, all whiskers apart from C2 were trimmed. On the third day, mice began to be trained in the whisker detection task. At the beginning of the session, mice were presented with a few associative trials during which they received automatic water reward paired with a strong whisker stimulation (amplitude 3.3°) to build an association between whisker deflection and reward. Mice rapidly learned the task after a few associative trials and licked in response to the whisker stimulus, after which the associative trials were discontinued. The consecutive daily training sessions aimed to improve the performance of the mice by increasing the Hit rate and decreasing the False alarm rate. The proportion of Catch trials was progressively increased with training days from 30% at the beginning of training to 50%. The inter-trial interval was progressively increased with training days from 5–6 s to 6–12 s for final experiments, making the mouse calmer (decreased spontaneous licking). During the first few training days (3–5 days), the strength of the magnetic pulse was kept at the maximum (3.3° whisker deflection), which can easily be detected by mice. When mice displayed a good level of performance (Hit rate >70%; False alarm rate <30%) the strength of the stimulation was gradually decreased to 2.5°, making the whisker stimulus more difficult to detect. During these sessions, to facilitate task performance, a short block of 10–40 trials with the highest whisker stimulus amplitude (3.3° whisker deflection) was always presented at the beginning of the session. When performance became stable with a False alarm rate below 30% and Hit rate above 70% with the 2.5° stimulus amplitude, the full psychometric task was introduced and the mice were further trained for a few additional days in the final version of the task. For recording sessions, four whisker stimulation trials with different amplitudes were presented with equal probability. Each recording session started with a few (5–10) trials of strong whisker stimulations (3.3°) followed by all trial types. These few starting trials were not included in the analysis and were only used to reduce the effect of over-motivation at the beginning of behavioral sessions when animals were most thirsty.<sup>51</sup>

In this study, the stimulus strengths used were calibrated across the physiological range of whisker deflections and the ability of mice to detect them. All stimuli for the psychophysical detection task were specifically designed to be very precise in duration, shape, and magnitude. In our detection task, the whisker stimulus is delivered through a brief magnetic pulse acting upon a small metal particle attached to the right C2 whisker. Magnetic pulses of 1-ms duration and varying in amplitude were achieved by applying biphasic voltage pulses (a 0.5-ms duration positive step followed by a 0.5-ms duration negative step) with different amplitudes to the coil. The evoked whisker deflections were measured using a high-precision infrared displacement sensor in a deeply anesthetized mouse to avoid spontaneous whisker movements.

Whisker-non-rewarded mice (WR-) were trained to lick for a reward at random times while being exposed to the same range of whisker stimuli that were not predictive of the reward delivery. WR- mice were exposed to whisker stimulation for a similar number of days as the WR+ mice. The trial structure was very similar to that used for WR+ mice. A new trial started after 8–12 s of intertrial interval plus 2.5–3.5 s of no-lick window, followed by a 2-s response window. If by chance the mouse licked during the response window it received a reward with a 70% probability. Whisker stimulation was completely independent of the trial structure and could occur at any time, with the interstimulus interval calculated to match the distribution of stimuli in WR+ mice. For later analysis of Miss trials, only whisker stimulation with no lick 2.5 s before (no-lick quiet window) and 1 s after the stimulation were considered. For Spontaneous licks, only non-rewarded licks preceded by 2.5 s without any lick or whisker stimulus, and no stimulus 1 s before or after the lick onset were considered.

### Electrophysiological recordings

Acute recordings were performed on both groups of mice after training. One or two recording sessions (over two consecutive days) were carried out for each mouse. Extracellular spikes were recorded using single-shank silicon probes (A1x32-Poly2-10mm-50 s-177, NeuroNexus, MI, USA) with 32 recording sites covering 775  $\mu$ m of the cortical depth. In most sessions, two probes were inserted in two different brain areas. Before insertion, the probes were coated with Dil (1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate, Invitrogen, USA) for post hoc recovery of the recording location (see below). The neural data were bandpass filtered (0.3 Hz - 7.5 kHz), amplified and digitized (30 kHz) using a digital headstage (CerePlex M32, Blackrock Microsystems, UT, USA), and recorded using the data acquisition system (CerePlex Direct, Blackrock Microsystems, UT, USA). The headstage was connected to an external reference chloridized silver wire that was placed in the recording chamber filled with Ringer solution. After probe insertion, mice were left for ~30 min for the probes and brain to stabilize after which the behavioral session and electrophysiological recordings started.

### Optogenetic manipulations

To achieve potent and focal inhibition of the cortical areas of interest, we used VGAT-ChR2 mice expressing ChR2 in all cortical GABAergic neurons.<sup>6,49</sup> Optogenetic stimulation of three superficial cortical areas wS1, tJM1, and fpS1 was achieved through a transparent skull preparation with skull thinning above the area of interest to facilitate the penetration of the light into the cortex. We used a 200- $\mu$ m diameter 0.22NA patch cable (M84L01, Thorlabs, USA) at the end of which a cannula (CFMXA05, Thorlabs USA) was attached and positioned right above the area of interest. During the experiments, the rest of the exposed skull was covered with Kwik-cast to avoid off-target light delivery to nearby cortical areas. For mPFC inhibition a cannula (CFMXA05, Thorlabs USA) of 200- $\mu$ m diameter was implanted at a 1500- $\mu$ m depth below the pia, just above mPFC, or acutely lowered to a similar depth just before the inactivation session. The optic fiber was coupled to a 470 nm high-power LED (M470F3, Thorlabs, USA).

Optogenetic inactivation was performed in 13 WR+ VGAT-ChR2 mice, inhibiting one area of interest per session. The order for the areas was randomized across mice. Testing sessions started when mice reached expert levels of performance ( $d$ -prime  $>1$  for the strong whisker stimulus amplitude). An ambient blue masking light was used in the training sessions as well as on testing days. Light trials (30%) were randomly interleaved with light-off control trials (70%) in both Stimulus and Catch trials. In light trials, a 100 Hz train of blue light pulses (50% duty-cycle, mean power 5 mW) was applied 100 ms before and 1000 ms after trial onset (with the final 100 ms having a ramp down to prevent rebound excitation).

### Histology and localization of electrode/optical fiber tracks

At the end of the experiments, mice were perfused with phosphate buffered saline (PBS) followed by 4% paraformaldehyde (Electron Microscopy Science, USA) in PBS. Brains were post-fixed overnight at room temperature. Serial 100- $\mu$ m coronal sections were cut by a conventional vibratome (VT 1000S; Leica, Wetzlar, Germany). The slices were then mounted and the fluorescent Dil electrode tracks were imaged using an epifluorescence microscope (Leica DM5500). Matlab-based software (Allen CCF tools, <https://github.com/cortex-lab/allenCCF>) was used to register brain slices and probe locations to the Allen mouse brain atlas.<sup>52</sup>

## QUANTIFICATION AND STATISTICAL ANALYSIS

All analyses were performed using custom-written codes in MATLAB and Python (population decoding analysis).

### Behavioral data analysis

To quantify the performance of the animals during the psychophysical detection task, we computed the Hit rate (the number of Hit trials divided by the number of Stimulus trials) for each stimulus amplitude and the False alarm rate (the number of False alarm trials divided by the number of Catch trials). To construct psychometric curves, we fitted the detection performance across stimulus intensities with a sigmoidal function by using the “fit” function in MATLAB:

$$\text{Fitting model: } f(x) = ((\alpha - \beta) / (1 + (x/\gamma)^\lambda)) + \beta$$

where  $x$  is stimulus intensity;  $f(x)$  is the detection (lick) probability;  $\alpha$  and  $\beta$  define the lapse rates of the curve;  $\gamma$  defines the middle point; and  $\lambda$  the slope of the curve.  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\lambda$  are free parameters that were fitted using a maximum likelihood method. The detection threshold was estimated by fitting, and defined as the middle point of the psychometric curve. The near-threshold stimulus for each session was defined as the stimulus amplitude nearest to the detection threshold. Individual session psychometric functions were constructed after cutting the end of each behavioral session to exclude any trials after the mouse failed to lick in response to the strongest whisker stimulus 4 times in a row, thus only including engaged trials in our analysis.

For Miss trial analysis (Figures 2A–2D), all trials were considered, including Miss trials at the end when mice were disengaged and stopped licking because we did not find differences between the evoked activity during Miss trials occurring when the mice were engaged vs. when the mice were disengaged. Similarly, for Hit trial analysis (Figure 3) all Hit trials were considered. However, for the direct comparison of Hit vs. Miss trials (Figure 4), we selected only engaged trials for the near-threshold stimulus amplitude, to limit distribution bias across the session. We also defined Spontaneous licks as all non-rewarded licks with at least 2.5 s of no-lick baseline. Thus, Spontaneous licks included False alarm licks during Catch trials, as well as spontaneous licks during intertrial intervals. For the optogenetic inactivation experiments, the effect of inactivation was assessed only for engaged trials.

### Quantification of orofacial movements

We monitored the orofacial movements (right C2 whisker, tongue, and jaw) of mice during each behavioral session using a high-speed camera (CL 600 X 2/M, Optronics, Germany; 200 or 500 Hz frame rate, 0.5 or 1 ms exposure, and 512x512-pixel resolution) under blue and infrared light illumination. The high-speed camera was positioned above the mouse and focused on the C2 whisker. A side mirror was used to capture the motion of the tongue and the jaw. To extract the position of each body part across time we used the Python-based toolbox DeepLabCut 2.2b7 (DLC), a software that uses deep learning networks for markerless pose estimation.<sup>53</sup> To train the network and to improve its ability to generalize, we first created a joint video that consisted of a sample of random frames of Hit trials from all recorded sessions. We trained the network on around 1,000 labeled frames using the k-means automatic extraction algorithm for the maximum number of iterations (1,030,000 iterations). Afterward, we used the trained network to extract the position of several body parts over time (C2 whisker base and middle point, tongue tip, jaw tip, and nose tip) for all the video filming data,

for all the sessions, and all mice. For most sessions, we observed high-quality performance of the network, but for a small subset of sessions, we further improved the model by refining the labels of 20 outlier frames and retraining the network. We filtered out all body part position estimations below the 60% likelihood. Using the X and Y extracted coordinates for each body part, we then computed the whisker angular position as the angle between the whisker (whisker base to whisker middle point) and the mouse head midline; the jaw/tongue displacement as the distance between the tip of the jaw/tongue, and the position of the resting (closed) position of the jaw. The jaw and tongue traces were then mean-filtered (10 ms) and multiplied by pixel size to obtain the real displacement in mm. In our analyses of free whisking, we used epochs that were defined as periods of time with power of whisker angular velocity higher than  $40,000 \text{ deg}^2 \cdot \text{s}^{-2}$  for at least 300 ms, preceded by at least 500 ms without whisker movements, in the absence of jaw movement. The lick onset time was defined as the time at which the jaw displacement crossed a threshold (Jaw displacement SD/7) just before the first peak of the jaw trace.

### Electrophysiology data

Extracellular spike detection and clustering was performed using Kilosort2<sup>54</sup> (<https://github.com/MouseLand/Kilosort>). Each cluster was then inspected manually and refined when needed. Only "well-isolated" units based on the sorting quality matrix<sup>55,56</sup> (<https://github.com/cortex-lab/sortingQuality>) were selected for the analyses. Single units were categorized as regular spiking units (RSUs) or fast-spiking units (FSUs) based on the duration of the spike waveform as described previously.<sup>57,58</sup> RSUs had spike peak-to-baseline duration  $>0.34$  ms and FSUs  $<0.26$  ms. Neurons with intermediate spike duration were considered as undefined units. In most of our analyses we included all the neurons (RSU, FSU and undefined).

### Clustering neuronal responses

We used a clustering method to define groups of neurons with similar response patterns across different trial types in the three recorded areas of WR+ mice.<sup>31</sup> All 2,001 neurons were included in the analysis. We considered five trial types: Miss trials aligned to the stimulus onset; Hit trials aligned to the stimulus onset; Hit trials aligned to the jaw movement onsets; Spontaneous licks aligned to the jaw onset; and free whisking (in the absence of jaw movements) aligned to the onset of whisking. For stimulus trials (Hit and Miss) we used the stimulus amplitudes  $1.8^\circ$ ,  $2.5^\circ$  and  $3.3^\circ$ , with the number of Hit and Miss trials adjusted by downsampling for each stimulus amplitude. For each neuron, we computed the mean PSTHs for each of the trial types, with 1 s of baseline and 1.5 s of response time for both stimulus and lick alignment, with a 100 ms resolution. We subtracted the baseline (mean response from  $-0.5$  to 0 s for stimulus-aligned trials, and from  $-1$  to  $-0.5$  s for jaw and whisker onset alignments). We then concatenated the PSTHs for all five trial types for each neuron resulting in a vector of 125 firing rate values and normalized to the range of values across these 125 bins. We subsequently combined PSTH vectors for all neurons resulting in an activity matrix  $X \in \mathbb{R}^{2,001 \times 125}$  where each row  $i$  corresponds to the concatenated normalized firing rate of the neuron  $i$  across the 5 trial types. Principal Component Analysis (PCA) and spectral embedding were used to reduce redundancy and detect non-convex clusters, as described in Esmaeili et al. 2021.<sup>31</sup> Neurons were then clustered using a Gaussian Mixture Model and the number of clusters was selected based on the minimum Bayesian Information Criterion.<sup>59</sup>

### Receiver Operating Characteristic analysis

We used Receiver Operating Characteristic (ROC) analysis to classify neurons as positively or negatively modulated and to assess whether single neurons significantly encoded information about different task variables by computing the Selectivity Index (SI) in bins of 50 ms after single-trial baseline subtraction. The SI was computed by shifting and scaling the area under the ROC curve (AUC) between  $-1$  and 1:

$$\text{Selectivity index} = 2 * (\text{AUC} - 0.5),$$

To examine the significance of SI, the AUC was compared to a label-shuffled distribution (AUC shuffled), obtained by shuffling the labels of the trials 100 times and performing non-parametric permutation tests. When SI was computed across time, the p values were corrected using False Discovery Rate correction (FDR) to account for all the bins tested. Neurons were classified as positively (SI  $> 0$ ) or negatively (SI  $< 0$ ) responding based on the maximum of the absolute value of the SI in the time window under consideration. Grand-average PSTHs were computed separately for neurons with positive and negative responses. The number of significantly modulated neurons was computed as the number of neurons with significant SI (positive or negative) in at least one 50-ms time bin in the time window under consideration (usually 200 ms).

To assess pure sensory information coding, we computed the SI for each neuron comparing stimulus-evoked activity (stimulus amplitudes  $1.8^\circ$ ,  $2.5^\circ$  and  $3.3^\circ$ ) in the absence of licking (Miss trials in WR+ mice and almost all whisker trials in WR-mice) to no-stimulus trials (Correct rejection trials in WR+ mice and no-lick no-stim trials for WR-mice). The proportion of significant neurons was measured by counting the number of neurons with significant SI within 200 ms after the stimulus onset. To assess pure motor information, we computed the SI comparing the activity in Spontaneous lick trials to Correct rejection trials. The proportion of significant neurons was measured by counting the number of neurons with significant SI in the 200 ms before, or 200 ms after, the jaw opening onset. To assess the response of the neurons in Hit trials, we computed the SI for Hit trials compared to Correct rejection trials, aligned to the stimulus (or trial) onset. Finally, to define decision neurons, we computed the SI of each neuron comparing Hit trials and Miss trials (for near-threshold stimulus amplitude) aligned to stimulus onset, and the SI for Hit trials (highest stimulus amplitude,

3.3°) vs. Spontaneous lick trials, aligned to lick onset. Neurons were classified as decision neurons if they had both a significant SI for Hit vs. Miss within the 200 ms after the stimulus onset, and a significant SI for Hit vs. Spontaneous lick within the 200 ms before jaw opening onset. Decision neurons were classified as gated if they had neither significant SI for sensory response (Miss vs. Correct rejection), nor for motor response (Spontaneous lick vs. Correct rejection).

### Pearson correlations

To assess the linear correlation between neuronal responses and the amplitude of the whisker stimuli, we computed the Pearson correlation for each neuron between the evoked change in firing rate over the 200-ms period after stimulus onset (relative to the pre-stimulus baseline) and the stimulus amplitude. The significance of the correlation was assessed using a t test ( $p < 0.05$ ). The statistical difference in the Pearson's coefficients between areas was assessed using the Wilcoxon rank-sum test with Bonferroni correction. It should be noted that this analysis is a very conservative measurement of the correlation between single-neuron activity and the stimulus amplitude that could not account for potential encoding of the stimulus amplitude by population activity through non-linear mixed selectivity in mPFC.<sup>60</sup> Therefore, we also used a population decoding approach of the whisker stimulus amplitude (see below).

### Population decoding

Trial-based decoding of the sensory stimulus (stimulus vs. no stimulus, or stimulus amplitude) and motor response (lick vs. no lick) was performed from all the neurons (RSUs, FSUs and undefined units) recorded in one brain area during one session for non-overlapping 100-ms time windows. We used logistic regression for binary signal classification (e.g., stimulus vs. no-stimulus) and multinomial logistic regression for the stimulus amplitude classification. We used a randomized downsampling of trials to select the same number of trials from each category. For example, when decoding lick (Hit) vs. no lick (Miss) for near-threshold stimulus amplitude, the number of Hit and Miss trials used was the minimum between the number of Hit and Miss trials ( $\min[N_{\text{Hit}}, N_{\text{Miss}}]$ ). For trial-based prediction of the stimulus amplitudes, an equal number of trials for each amplitude was selected.

We also used linear regression for continuous signal regression to perform continuous decoding of the tongue movement. For continuous decoding, we used 5 ms kernels that extended causally for 50 ms and acausally for 50 ms for every neuron. Based on the activity of these 100 ms, we predicted the behavioral signal at each timestep (one timestep corresponds to 5 ms duration). For the continuous prediction of the tongue trace, we used 500 ms-long licking bouts occurring when the mouse was collecting a reward, starting at least 500 ms after the whisker stimulus.

We estimated the accuracy or the explained variance using 5-fold stratified cross-validation in order to ensure that our labels/conditions had the same data distribution in the training and test set. In order to avoid overfitting, we used L2-regularization where the strength of the regularization was determined with a small grid search with nested cross-validation. To estimate the significance of the decoding results, we compared our results with the average measure of 50 label-shuffled splits using the Wilcoxon signed-rank test (for each session, 50 shuffled versions were made and one averaged measure was computed; then the decoding measures of all sessions were compared to these corresponding averaged measures). Recordings that had less than 8 neurons per area or less than 26 trials per label/condition were not used for our decoding analyses.

### Statistics

Data are represented as mean  $\pm$  SD unless otherwise noted. The Wilcoxon signed-rank test was used to assess significance in paired comparisons; and the Wilcoxon rank-sum test was used for unpaired comparisons (MATLAB implementations). Kruskal-Wallis test followed by Tukey-Kramer HSD test when appropriate were used for comparison between multiple (>3) groups.

The statistical tests used and n numbers are reported explicitly in the main text or figure legends. When necessary, p values were corrected for multiple comparisons using either Bonferroni correction or FDR correction, and the method used is indicated in the main text or figure legends.