Representational drift influences learning of a sensory cortical microstimulation task

Pancholi, Ravi; Ryan, Lauren; Peron, Simon

Center for Neural Science
New York University
4 Washington Place Rm. 621
New York, NY 10003

Lead Contact: Simon Peron; speron@nyu.edu

SUMMARY

Sensory cortical representations can be highly dynamic, raising the question of how representational stability impacts perception. We used volumetric two-photon calcium imaging to track neural activity as mice learned an optical microstimulation task. Mice were required to report the number of photostimulation pulses delivered to opsin-expressing neurons in layer 2/3 of primary vibrissal somatosensory cortex. We found that the amount of photostimulus-evoked activity could predict animal choice. Over the course of training, this photostimulus-evoked activity rapidly declined, particularly among highly photoresponsive neurons. Population activity levels across training were similar between animals that learned and those that did not. However, the photoresponsive population exhibited turnover both within and across behavioral sessions that was elevated among animals that failed to learn. Animals that failed to learn also exhibited a faster deterioration in stimulus decoding. Our results show that the rate at which sensory cortical representations change constrains learning and impacts behavior.

INTRODUCTION

Cortical activity can be highly dynamic, even in the context of identical sensory input. In olfactory (Schoonover et al., 2021), visual (Deitch et al., 2021; Marks and Goard, 2021; Perez-Ortega et al., 2021), somatosensory (Peron et al., 2015b), multimodal (Driscoll et al., 2017), and motor cortical areas (Rokni et al., 2007), as well as hippocampus (Ziv et al., 2013), the population of neurons representing a particular sensory or motor feature changes over time. This turnover in the stimulus-responsive population is known as representational drift (Lutcke et al., 2013; Rule et al., 2019). Because variance in neural activity within sensory cortices can interfere with readout in downstream areas, representational drift has been proposed as a fundamental constraint on perception (Clopath et al., 2017). Its ubiquity in both neocortex and allocortex suggests that it is a fundamental cortical feature. Despite this, the link between representational drift and behavior remains poorly understood.

Representational drift is typically studied in animals receiving input at the sensory epithelium. Under such conditions, dynamics in the sensory stream prior to cortex can contribute to drift. Moreover, fanout at pre-cortical processing stages distributes information across multiple cortical and subcortical targets, reducing the relative perceptual importance of the cortical area under study (Diamond et al., 2008). Microstimulation of cortex circumvents pre-cortical processing and simplifies interpretation of behavioral effects by explicitly requiring the animal to use specific features of the evoked activity. Microstimulation has been used extensively in the study of perception (Afraz et al., 2006; Cohen and Newsome, 2004; Histed et al., 2012; O'Connor et al., 2013; Salzman et al., 1990; Yang and Zador, 2012), and animals can readily be trained to report microstimulation (Butovas and Schwarz, 2007; Dalgleish et al., 2020; Huber et al., 2008; Leal-Campanario et al., 2006). However, electrical microstimulation of cortex evokes activity among a sparse and spatially distributed population (Histed et al., 2009), suffers from instability due to slowly emerging gliosis (Polikov et al., 2005), and produces a prominent stimulus artifact that complicates simultaneous recording (Weiss et al., 2019). Optical microstimulation of cortex (Dalgleish *et al.*, 2020; Finkelstein et al., 2021; Histed and Maunsell, 2014; Huber *et al.*, 2008) overcomes these constraints and allows for refined spatial and genetic control of evoked activity as well as chronic concurrent recording via large-scale calcium imaging (Peron et al., 2015a). Consequently, optical microstimulation of cortex in conjunction with population imaging is well suited to the study of representational drift and its impact on behavior.

Here, we train mice to report the intensity of optical microstimulation in a subset of opsin-expressing pyramidal neurons in layer (L) 2/3 of primary vibrissal somatosensory cortex (vS1). Mice must discriminate between a high and low number of optical microstimulation pulses delivered to vS1. Over the course of training, we track neural dynamics in the opsin-expressing volume using volumetric two-photon calcium imaging (Peron *et al.*, 2015a). Photostimulus-evoked activity predicts animal choice and declines over the course of learning. Mice exhibit different rates of learning, with the amount of evoked activity failing to account for this difference. Throughout learning, the neurons

comprising the photoresponsive population change at rates that vary across animals. Animals where this turnover is greater, both within and across behavioral sessions, are less likely to learn the task. Moreover, these animals exhibit a faster decline in the efficacy of stimulus decoding using neural activity. Together, these results show that representational drift constrains learning and impacts perception.

RESULTS

Animals can learn an optical microstimulation pulse count discrimination task

We monitored cortical dynamics as mice were trained on an optical microstimulation task. Transgenic mice expressing GCaMP6s (Chen et al., 2013) in cortical excitatory neurons (Slc17a7-Cre X Ai162) (Daigle et al., 2018) were virally transfected with the soma-restricted opsin ChRmine (AAV-8-CaMKIIa-ChRmine-mScarlet-Kv2.1-WPRE) (Marshel et al., 2019), resulting in opsin expression in pyramidal neurons in L2/3 of vS1. To allow for optical access, a cranial window was implanted over vS1 (**Figure 1A**). Following recovery, we used widefield imaging during whisker deflection to confirm that opsin expression was restricted to vS1 (typically the C row of whiskers; **Figure S1**). A miniature light-emitting diode (LED) was affixed to the cranial window adjacent to the site of opsin expression for optogenetic stimulation.

Light pulses (5 ms long; minimum inter-pulse onset interval, 50 ms) were delivered to the opsin-expressing area and animals were required to report whether the pulse count was high or low. Mice were rewarded for licking the left of two lickports on low pulse count trials (1 or 3 pulses; 0 pulses in early training) and the right lickport on high pulse count trials (7 or 9 pulses), with 5 pulse count trials rewarded randomly between left and right (**Figure 1B**). LED output was stable within and across training sessions (**Figure S2**). Increasing the pulse count, the pulse amplitude, or the inter-pulse interval all resulted in greater photostimulation-evoked activity in awake animals not performing a task (Histed and Maunsell, 2014) (**Figure S3**). On each trial, photostimulus pulses were presented during the stimulus epoch (500 ms). This was followed by a short delay period (500 ms), after which mice indicated their response by licking during the response epoch (<2 s) (**Figure 1C**). Stimulus pulses were randomly permuted among 9 time bins to prevent animals from solving the task using timing cues (**Figure 1D**, **S4**). A spectrally matched LED was placed near the mouse's eyes and illuminated during each time bin to obscure potential visual cues from photostimulation ('masking flash', **Figure 1B**).

Mice (**Table S1**) were first trained on a photostimulation detection task (9 vs. 0 pulses; **Figure 1E**). Mice were classified as 'learners' (n = 6) if their performance improved over the first week of training. Mice that failed to improve were classified as 'non-learners' (n = 5) and were removed from the training cohort within 1-2 weeks. Learners reached criterion performance (d-prime > 1.5) in 9.8 ± 3.7 days (mean ± SD) after which additional pulse counts were added until mice were proficient at the full intensity discrimination task. Learners achieved high performance at this final stage (78.3% ± 3.4% correct; **Figure 1F**, **G**). To confirm that animals required opsin expression to perform the task, we trained a separate cohort of mice not infected with opsin ('no opsin'; **Figure 1F**). These mice could not learn the task after 20 training sessions (performance on final day: 48.6% ± 1.4% correct; n = 5 mice). Lesioning the opsin-expressing area in mice that had progressed to the final task degraded performance (average performance of 3 sessions before lesion: 78.8% correct, 3 sessions after lesion: 60.9%, p-value = 0.014, paired t-test; **Figure 1H**). Thus, mice can learn to

discriminate the number of photostimulus pulses delivered to a subset of neurons in L2/3 of vS1 via optical microstimulation.

Optical microstimulation response increases with pulse count

To assess neural activity during optical microstimulation, we recorded from L2/3 of the opsin-expressing area using volumetric two-photon calcium imaging (Peron *et al.*, 2015a; Peron *et al.*, 2015b). Neurons distributed among three 800-by-800 µm planes spaced 60 µm apart were tracked for 29.8 ± 6.9 days (mean ± SD; n = 6 mice; **Figure 2A**). Imaged neurons were separated into opsin-expressing (692.6 ± 123.2 neurons per mouse) and opsin non-expressing (2,287.3 ± 466.9 neurons per mouse) populations based on the presence of mScarlet (Methods). Based on estimates of L2/3 pyramidal neuron density in vS1 (Lefort et al., 2009; Peron *et al.*, 2015b), we imaged 15-20% of infected L2/3 neurons, for a total of ~3,500 to ~4,500 opsin-expressing neurons per mouse. Neurons exhibited diverse responses to photostimulation (**Figure 2B**, **C**). We found both opsin-expressing and opsin non-expressing excitatory neurons that responded to photostimulation, with photoresponsive neurons showing increasing responses with increased pulse count (**Figure 2D**). Photoresponsive neurons were broadly spatially distributed (**Figure 2E**). Higher photostimulus pulse count consistently increased the fraction of neurons responsive to stimulation (**Figure 2F**), as well as the response amplitude (**Figure 2G**) and response probability (**Figure 2H**) of individual neurons. This was true for both opsin-expressing and opsin non-expressing neurons. In all cases, the opsin-expressing population responded more reliably and strongly to photostimulation than the opsin non-expressing population.

The amount of microstimulation-evoked activity predicts animal choice

Higher microstimulation pulse counts drive larger responses in photoresponsive neurons. Thus, mice could solve the task by licking in one direction if activity fell below a threshold, and the other if activity exceeded a threshold. If mice employ this strategy, variability in evoked activity for a given pulse count should influence choice. Specifically, for a given pulse count, trials where the animal reported more stimulation (right lick) should exhibit more evoked activity than trials where the animal reported less stimulation (left lick). We found that a subset of individual neurons indeed showed this pattern across all pulse counts, whereas other neurons exhibited a difference on only a subset of pulse counts or no difference at all (**Figure 3A, B**).

Sensory decisions are believed to result from pooling across many neurons (Britten et al., 1996; Shadlen et al., 1996), so we next asked whether the predicted difference in evoked activity was also present at the scale of the photoresponsive population rather than just among individual neurons. Among strongly photoresponsive opsin-expressing neurons (response probability > 0.25) during an example session, we found that the mean evoked ΔF/F across trials was no different for left and right lick trials (**Figure 3C**). Among responsive opsin non-expressing neurons, however, the response was higher on trials where the animal indicated a stronger stimulus for all pulse counts. Across mice, the peak-normalized response on lick-left and lick-right trials did not differ across pulse counts for opsin expressing neurons (**Figure 3D, E**; P = 0.060, t-test comparing

right and left lick trials, n=6 mice). However, for opsin non-expressing neurons, the difference was significant at all pulse counts (P = 0.017), with greater levels of evoked activity on trials where the animal reported higher activity. Thus, the amount of evoked activity in a subset of neurons immediately downstream from the opsin-expressing population can predict animal choice, suggesting that these neurons may influence the animal's behavior.

Activity levels decline over time, but overall activity is not predictive of learning

Having observed a link between neural activity and perceptual choice in our task, we next sought to determine whether L2/3 dynamics during training impacted learning rates. Sensory cortical representations in L2/3 exhibit diverse dynamics during learning, with changes in aggregate responsiveness as well as in the responses of single neurons (Poort et al., 2015). We first asked whether aggregate responsiveness impacted learning by tracking neural activity across training (**Figure 4A, B**). Restricting our analysis to trials with 9 photostimulus pulses, which were present in all sessions, we identified neurons with stable, increasing, or decreasing responses to photostimulation (**Figure 4C**). The most responsive neurons exhibited the largest declines in responsiveness (**Figure 4D, E**). Though responsiveness did not change for the bottom 95% of neurons (percentile of response amplitude; Methods), we observed a significant decline in response amplitude and probability for the top percentiles of responsive neurons (**Figure 4F**). Thus, training results in a sparsification of the photostimulation response, primarily due to declining responses among the most responsive neurons.

To control for changes in responsiveness due to changes in opsin expression, we tracked mScarlet fluorescence ('redness') over the course of training (**Figure S5**; Methods). We found that overall redness across all neurons did not change over time (P = 0.490, paired t-test, n = 6 mice; **Figure S5B**). Relative redness was also stable across days (**Figure S5C**, **D**). Thus, observed changes in neural responsiveness are unlikely to be due to changes in opsin expression.

Stronger optical microstimulation is more detectable (Dalgleish *et al.*, 2020; Huber *et al.*, 2008), so we next asked whether the amplitude of the evoked response was greater among learners. Specifically, we looked for differences in evoked responses across the days common to mice that learned and those that did not learn (days 1-8). Restricting our analysis to the top 5th percentile of responsive neurons in each population, we found that opsin-expressing neurons showed comparable response amplitudes, quantified using z-scored $\Delta F/F$, among both learners (3.27 ± 1.11) and non-learners (4.01 ± 1.15; P = 0.305, t-test comparing learners and non-learners). Response probability among both learners (**Figure 4G**, mean response probability, days 1-8: 0.69 ± 0.11) and non-learners (0.79 ± 0.11) also showed no difference (P = 0.180). Opsin non-expressing neurons also showed no difference in response amplitude (learners: 1.39 ± 0.38; non-learners: 1.33 ± 0.31; P = 0.772) or response probability between learners and non-learners (**Figure 4H**; learners: 0.35 ± 0.12; non-learners: 0.33 ± 0.07; P = 0.698). Thus, activity declined in both learners and non-learners, and evoked activity was comparable among both cohorts.

Greater representational drift in the photoresponsive population impairs learning

Even though aggregate responsiveness was similar among both learners and non-learners, it is possible that the turnover in the photoresponsive population accounts for the discrepancy in learning across animals. Such 'representational drift' has been observed in many cortical areas (Rule *et al.*, 2019), and is thought to influence perception and learning (Clopath *et al.*, 2017). Within a single training session, we observed neurons whose responses were stable across the session and others whose responses increased or decreased (**Figure 5A**). At the population level, we computed the evoked ΔF/F (Methods) for individual 9-pulse trials (**Figure 5B**) and obtained a vector of values (one per neuron) for a given trial. To quantify representational drift, we computed the correlation between evoked ΔF/F vectors across trial pairs (**Figure 5C**). Periods of elevated trial-to-trial correlation were interspersed with periods of lower correlation, though the population response could be stable over many trials (**Figure 5D**).

Because evoked activity in our optical microstimulation task is by design perceptually relevant, we next asked whether representational drift impacted perception (Clopath *et al.*, 2017). To obtain an overall measure of drift for a given session, we computed the mean trial-to-trial correlation for a given session in all mice. Among opsin-expressing neurons, there was no difference in the mean trial-to-trial correlation across the training days common to all mice (first 8 days) between learners and non-learners (**Figure 5E**; learners: 0.58 ± 0.05, n = 6 mice; non-learners: 0.58 ± 0.11, n = 5; P = 0.835, t-test). Among opsin non-expressing neurons, however, learners had higher mean trial-to-trial correlation than non-learners (learners: 0.47 ± 0.02, n = 6 mice; non-learners: 0.34 ± 0.03, n = 5; P < 0.001, t-test). Thus, among opsin non-expressing neurons, within-session representational drift is higher among non-learners.

We next examined representational drift across sessions. For each session, we generated a vector with the mean evoked ΔF/F across 9-pulse trials for each neuron. We then compared days by computing the correlation between one day's vector and analogous vectors obtained for subsequent days. For sessions farther apart in time, correlations between these vectors declined (**Figure 6A**, **B**). Aggregated across animals, the opsin-expressing population showed a decline in the rate of turnover over the course of training, whereas opsin non-expressing neurons exhibited steady drift (**Figure 6C**). To quantify this, we measured the number of days it took for R^2 between two population vectors to fall below 0.5. In learners, this increased from 3.9 ± 0.3 days in early training (days 1-3) to 6.2 ± 1.2 days in middle training (days 8-10) for opsin-expressing neurons (**Figure 6D**; P = 0.035, t-test, n = 6 mice), but remained unchanged among opsin non-expressing neurons (early: 4.6 ± 0.2 days; late: 5.4 ± 0.8 days; P = 0.163, t-test). Thus, inter-day representational drift slowed for opsin-expressing neurons as training progressed.

To determine whether the rate of inter-day drift impacted learning, we compared learners with non-learners using the days for which we had data across all mice (days 1-8). Correlations between single day population response vectors were high for both learners and non-learners among opsin-expressing neurons but were lower for non-

learners among opsin non-expressing neurons (**Figure 6E**). We measured the mean number of days until R^2 between two population vectors fell below 0.5. For days with only a few subsequent days, this condition was often never met; we therefore restricted our analysis to the days for which all animals yielded a value for this metric (days 1-3). We found no difference among learners and non-learners for opsin-expressing neurons (**Figure 6F**; days until $R^2 < 0.5$, learners: 3.9 ± 0.3, n = 6 mice; non-learners: 3.8 ± 0.7 days, n = 5; p = 0.759, t-test). For opsin non-expressing neurons, however, R^2 remained above 0.5 longer for learners (4.6 ± 0.1 days, n = 6 mice) than for non-leaners (2.4 ± 0.3 days, n = 5; p = 0.048, t-test). Thus, mice that fail to learn the task exhibit higher levels of both intra- and inter-day representational drift among opsin non-expressing neurons.

Stimulus decoding degrades more rapidly in mice that fail to learn the task

Representational drift is thought to impact downstream decoding of sensory cortical activity (Clopath *et al.*, 2017), suggesting that the ability to decode the stimulus from neural activity may be different among learners and non-learners. We therefore compared decoding between learners and non-learners over the course of training. We evaluated decodability using receiver operating characteristic (ROC) analysis, discriminating between evoked ΔF/F on high (7, 9) and low (0, 1, and 3) pulse count trials. We quantified decoding efficacy by computing the area under the ROC curve (AUC; Methods) for each neuron on a given session. Stimulus decoding declined over the course of training in individual mice (**Figure 7A**). Though transitions to more difficult task variants reduced decodability, decoding declined even between these transitions. In learners, the AUC among the top 5% of neurons by decoding ability on a given day (Methods) declined from 0.84 ± 0.02 (days 1-3) to 0.70 ± 0.01 (days 20-22) among opsin-expressing neurons (**Figure 7B**; P = 0.004, n = 6 mice, t-test; comparing days 1-3 vs. days 20-22). Simultaneously, the fraction of opsin-expressing neurons with an AUC > 0.75 declined from 0.12 ± 0.02 to 0.02 ± 0.01 (**Figure 7C**; P = 0.003). Smaller declines were observed among opsin non-expressing neurons, with the AUC among the top 5% of neurons dropping from 0.69 ± 0.02 to 0.61 ± 0.01 (P = 0.044, days 1-3 vs. days 20-22) and the fraction of neurons with AUC > 0.75 declining from 0.02 ± 0.02 to 0.01 ± 0.00 (P = 0.006). Thus, in mice that learned the task, stimulus decoding slowly declines over the course of training, despite steady or even improving performance.

We next asked whether decoder performance differed among learners and non-learners. We examined decoding on the first and last three common days (days 1-3 and 6-8, respectively). Initially (days 1-3), decoding among the top 5% of opsin-expressing neurons was higher for non-learners than learners (**Figure 7D**; learners: 0.84 ± 0.02, n = 6; non-learners: 0.91 ± 0.01, n = 5; P = 0.006, t-test comparing learners with non-learners), with no difference among opsin non-expressing neurons (learners: 0.69 ± 0.02; non-learners: 0.65 ± 0.02, P = 0.422). The fraction of neurons with AUC > 0.75 was also comparable among learners and non-learners for both opsin-expressing (**Figure 7E**; learners: 0.12 ± 0.02, n = 6; non-learners: 0.15 ± 0.04, n = 5; P = 0.414) and opsin non-expressing neurons (learners: 0.02 ± 0.02, n = 6; non-learners: 0.01 ± 0.00, n = 5; P = 0.148). For the last three common days (days 6-8), however, decoding for the top 5% of neurons was no longer significantly different among opsin-expressing

neurons (learners: 0.82 ± 0.02; non-learners: 0.75 ± 0.07, P = 0.168), but became so among opsin non-expressing neurons, with non-learners exhibiting worse decoding performance (learners: 0.66 ± 0.01; non-learners: 0.59 ± 0.02, P = 0.028). Similarly, the fraction of neurons with AUC > 0.75 was comparable among opsin-expressing neurons in learners vs. non-learners (learners: 0.07 ± 0.0, n = 6; non-learners: 0.04 ± 0.05, n = 5; P = 0.266), but was lower among opsin non-expressing neurons (learners: 0.02 ± 0.00, n = 6; non-learners: 0.00 ± 0.00, n = 5; P = 0.027). Thus, in later training, opsin non-expressing neurons exhibited worse stimulus decoding among non-learners.

DISCUSSION

We tracked neural activity in vS1 as mice learned to perform an optical microstimulation pulse count discrimination task (**Figure 1**). Evoked activity in this task was proportional to the number of photostimulation pulses delivered (**Figure 2**). Evoked activity was predictive of animal choice, though only among opsin non-expressing neurons indirectly driven by photostimulation (**Figure 3**). Individual mice learned at different rates, with some mice failing to learn the task in the time provided. Over the course of training, evoked activity among the photoresponsive population declined, though aggregate evoked activity failed to account for the differences in learning rate (**Figure 4**). Throughout training, the photoresponsive population exhibited a steady rate of turnover, or representational drift. Learning was fastest in animals that exhibited low levels of drift among opsin non-expressing neurons, both within and across sessions (**Figures 5**, **6**). Increased drift coincided with greater declines in stimulus decoding, and mice that failed to learn exhibited poorer stimulus decoding in later training (**Figure 7**).

Increased rates of representational drift and degradation in stimulus decoding were only predictive of learning among opsin non-expressing neurons. Thus, the activity directly evoked in cortex by photostimulation did not account for the rate of learning. However, the efficacy with which that activity drove consistent responses in downstream neural populations did impact perception (Clopath *et al.*, 2017; Lutcke *et al.*, 2013; Rule *et al.*, 2019). Photostimulation produced more consistent responses among the directly stimulated population, as demonstrated by the higher correlations among the population response for opsin-expressing neurons both within and across sessions relative to opsin non-expressing neurons (**Figure 5E, 6F**). Factors that could contribute to varied recruitment of downstream neurons include top-down modulation (Lee et al., 2013; Sadeh and Clopath, 2022) as well as variable excitatory recurrence in downstream populations (Carrillo-Reid et al., 2019; London et al., 2010; Marshel *et al.*, 2019; Peron et al., 2020).

Representational drift has been observed in many sensory (Deitch *et al.*, 2021; Marks and Goard, 2021; Perez-Ortega *et al.*, 2021; Peron *et al.*, 2015b; Schoonover *et al.*, 2021) and motor cortices (Carmena et al., 2005; Huber et al., 2012; Masamizu et al., 2014; Rokni *et al.*, 2007), as well as hippocampus (Kentros et al., 2004; Manns et al., 2007; Muzzio et al., 2009; Ziv *et al.*, 2013). Under some conditions, however, sensory cortical (Cohen et al., 2013; Mank et al., 2008; McMahon et al., 2014), motor cortical (Chestek et al., 2007; Dhawale et al., 2017; Ganguly and Carmena, 2009), and hippocampal (Thompson and Best, 1990) representations can be stable for extended periods at the single neuron level. Factors that contribute to greater response stability include extended periods of task training (Peron *et al.*, 2015b; Peters et al., 2014; Poort *et al.*, 2015), task engagement (Danielson et al., 2016; Kentros *et al.*, 2004; Muzzio *et al.*, 2009), and artificial versus naturalistic stimuli (Marks and Goard, 2021). We observe representational drift in the context of direct cortical stimulation, even in highly trained animals engaged in the task. This supports the view that drift is an intrinsic feature of cortex, though it is does not rule out the possibility that other structures contribute to drift (Sadeh and Clopath, 2022).

We observe sparsification of the photoresponsive population over the course of the first week of training, both in terms of the size of the responsive population and the typical response for a given neuron. With novel natural stimuli, both increases and decreases in responsive populations have been observed. For example, novel, behaviorally-relevant tones become overrepresented across learning in auditory cortex, (Kato et al., 2015; Weinberger, 2004) whereas whisker representations can shrink following environmental enrichment in vS1 (Polley et al., 2004). Candidate mechanisms for the latter decline in responsiveness include increased inhibitory input onto pyramidal neurons (Knott et al., 2002) and reduced excitability (Barth et al., 2004; Desai et al., 1999). Given that the most active neurons experience disproportionately larger response declines, sparsification is unlikely to be due to global mechanisms, such as a general increase in inhibition. Instead, cell-specific changes likely account for the observed dynamics.

Microstimulation has been used extensively in the study of perception (Cohen and Newsome, 2004; Histed *et al.*, 2012), yet tracking the dynamics of microstimulation-evoked activity over learning has proven difficult. Chronic electrical microstimulation poses numerous challenges: it is invasive, can inadvertently activate neurons up to millimeters away from the stimulation site (Histed *et al.*, 2009), and can cause tissue degradation at the site of electrode insertion (Polikov *et al.*, 2005). Furthermore, simultaneous recording is complicated by a prominent stimulation artifact (Weiss *et al.*, 2019), and even studies with concurrent recording rarely examine the activity evoked at the stimulation site (Allison-Walker et al., 2021; Butovas and Schwarz, 2003; Chen et al., 2020; Hao et al., 2016) (but see (Sombeck et al., 2022)). Finally, electrical microstimulation cannot target genetically separable neural populations. Optical microstimulation overcomes these issues. Photostimulation using an external LED does not require an intracortical implant, the stimulation is spatially and genetically confined, and calcium imaging allows us to observe evoked responses in thousands of neurons at single cell resolution (Peron *et al.*, 2015a). In contrast to objective-based optogenetic illumination, our LED-based approach also allows the objective to move while maintaining consistent stimulation at the target site. This allows for imaging at multiple depths and other brain areas and permits training away from the microscope.

We show that over the course of optical microstimulation task training, activity in the stimulated region of vS1 exhibits representational drift both within and across sessions that is predictive of learning speed. Thus, representational drift fundamentally constrains perception.

ACKNOWLEDGMENTS

We thank Bettina Voelcker, Keelin O'Neil, and Aaron Lanz for discussion and technical assistance. We thank Sadrah Sadeh, Claudia Clopath, Robert Froemke, Anthony Movshon, and Michael Long for discussion. This work was supported by the Whitehall Foundation and the National Institutes of Health (R01NS117536; F31NS120483; T32GM007308).

AUTHOR CONTRIBUTIONS

R.P. and S.P. designed the study. R.P. carried out the experiments. L.R. performed histology and developed the lesion method. R.P. and S.P. performed data analysis and wrote the paper.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

FIGURES

Pancholi et al., Figure 1

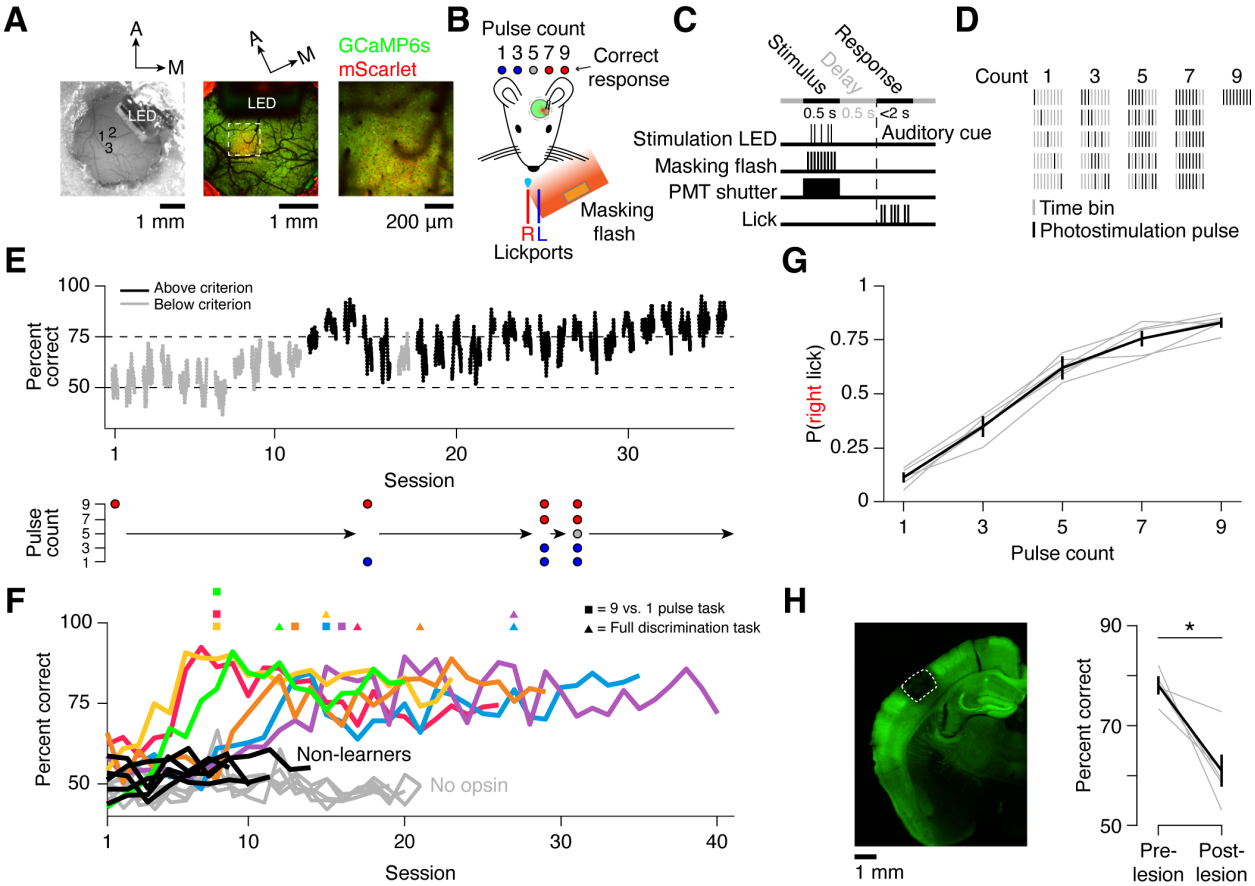

Figure 1. Mice learn to discriminate photostimulus pulse count using evoked vS1 activity. (A) Opsin expression in barrel cortex. Left, widefield view of cranial window showing viral injection sites and LED; center, two-photon image of cranial window (green, GCaMP6s fluorescence; red, mScarlet fluorescence); right, higher magnification two-photon image. **(B)** Mice must lick the right (red) lickport following presentation of 7 or 9 LED pulses or the left (blue) lickport following presentation of 0, 1, or 3 pulses to obtain reward. 5 pulses are randomly rewarded. A spectrally matched masking flash LED was used throughout. **(C)** Task timing during a single trial (Methods). **(D)** Example photostimulation pulse timing permutations. **(E)** Performance during learning. Top, task performance as a function of training day for one animal, with each dot indicating the rolling average of 61 trials centered on the shown trial (Black, days where peak d-prime ≥ 1.5; grey, d-prime < 1.5); bottom, pulse count progression. **(F)** Task performance as a function of training day for individual animals. Colored, animals expressing opsin that learned the task; black, animals expressing opsin that failed to learn the task; grey, animals with no opsin. **(G)** Right lick probability as a function of pulse count in final version of the task. Grey, individual animals; black, mean ± standard error of the mean (SEM). **(H)** Lesion of opsin-expressing tissue. Left, coronal section following laser lesion (box, lesion extent; green, GCaMP6s fluorescence); right, task performance before and

after lesion. Grey, average performance over 3 sessions pre- and post-lesion for individual animals; black: mean ± SEM. *$P < 0.05$ for paired t-test.

Pancholi et al., Figure 2

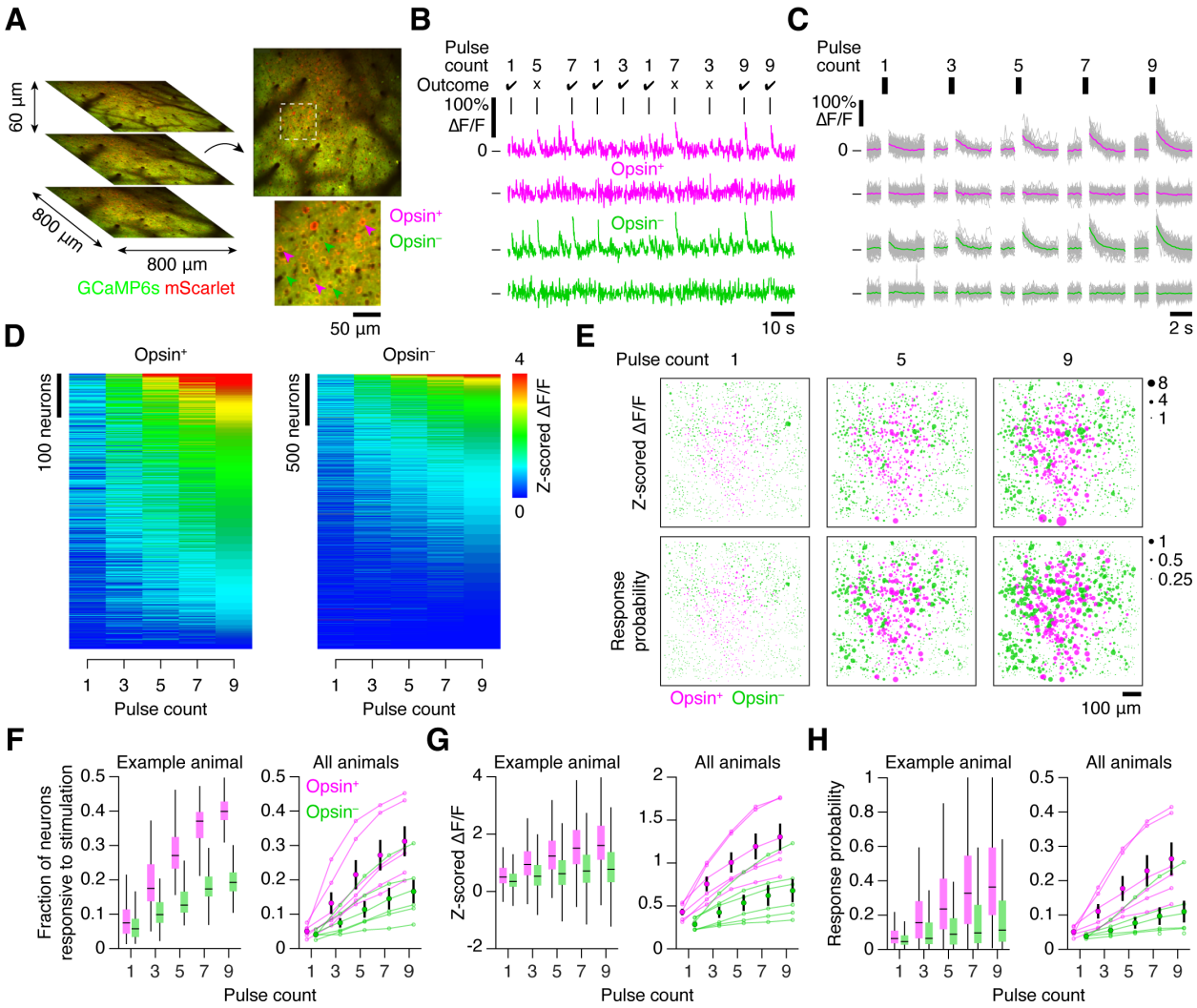

Figure 2. Photostimulation response increases with pulse count. (A) Volumetric two-photon imaging. Left, three planes (800-by-800 µm, 60 µm inter-plane distance) imaged simultaneously at 7 Hz; right, example plane with example opsin-expressing (magenta) and opsin non-expressing (green) neurons. **(B)** Photostimulation-evoked ΔF/F traces for 4 neurons across 10 consecutive trials. Vertical bars indicate stimulation epochs. Trial outcome indicates whether the animal made a correct (check) or incorrect (cross) behavioral response. **(C)** ΔF/F traces for the 4 neurons in (B) showing mean response to each pulse count over a single session. Grey, individual trials; color; mean. **(D)** Response amplitudes (z-scored ΔF/F) of all opsin-expressing and opsin non-expressing neurons from one animal sorted by responses to the 9 pulse stimulus. **(E)** Spatial map of neural responses from one imaging plane. Size of dot indicates either response amplitude (top) or response probability (bottom) to 1, 5, or 9 stimulus pulses. **(F)** Fraction of neurons responsive to stimulation as a function of pulse count. Left, example animal (box plot denotes quartiles, whiskers denote 1st and 99th percentiles); right, all animals (small dot, median; large dot, grand mean; error bars, SEM). **(G)** Same

as (F) for response amplitude as a function of pulse count. **(H)** Same as (F) for response probability as a function of pulse count.

Pancholi et al., Figure 3

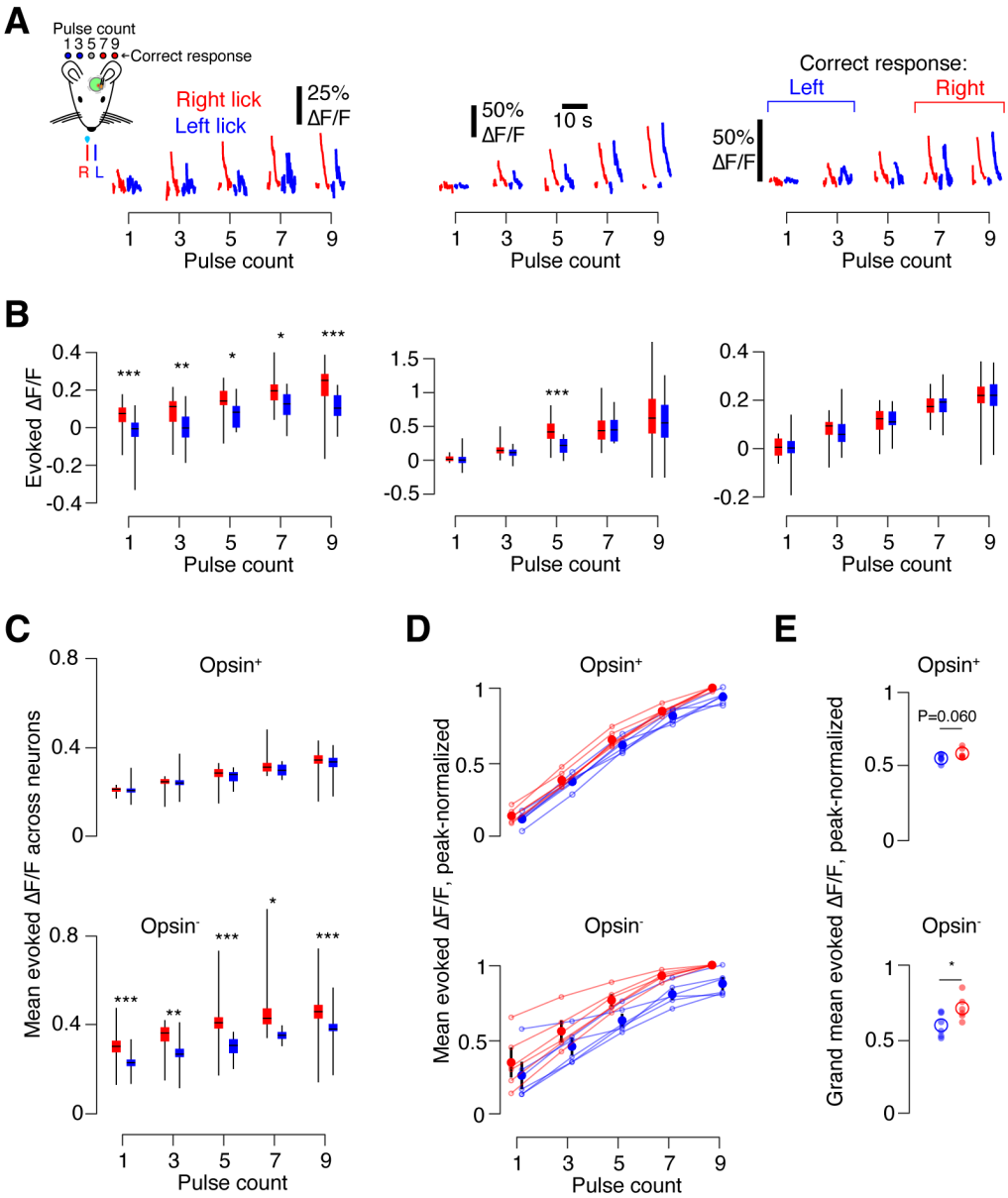

Figure 3. Neural activity in vS1 predicts behavioral choice. (A) Example neuron ΔF/F traces for all pulse counts, grouped by response. Red, animal licked right; blue, animal licked left. Left, neuron with choice-predictive activity for all pulse counts. Middle, choice-predictive activity on 5 pulse trials only. Right, neuron without choice-predictive activity. **(B)**. Evoked ΔF/F across all trials in an example session for neurons in (A). Box plot denotes quartiles, whiskers denote 1st and 99th percentiles, horizontal line denotes median. *P < 0.05, **P < 0.01, ***P < 0.001, t-test; unlabeled, P ≥ 0.05. **(C)** Mean photostimulation-evoked ΔF/F across all neurons during an example session. Box plots as in (B). Top, opsin-expressing neurons; bottom, opsin non-expressing. **(D)** Peak-normalized photostimulation-evoked ΔF/F across all mice (n = 6). Thin line, individual

mouse. Dot, mean across mice. Vertical line, SEM. **(E)** Peak-normalized photostimulation-evoked ΔF/F across all mice, averaged across all pulse counts. Small circle, single animal mean. Large circle, cross-animal mean.

Pancholi et al., Figure 4

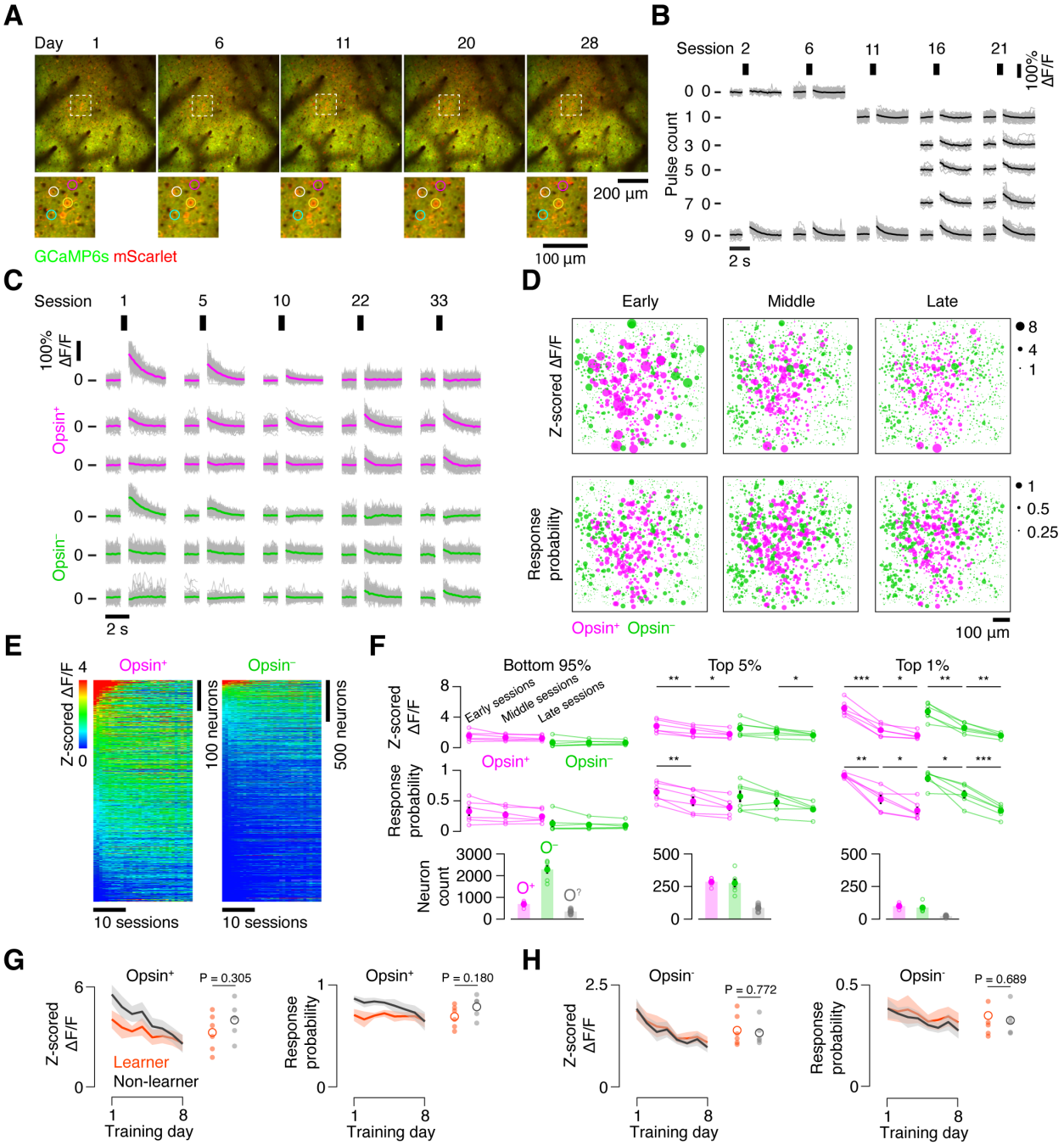

Figure 4. Population dynamics over the course of training and their impact on learning. (A) Example imaging plane across multiple days. **(B)** Responses of an example neuron across training for each pulse count. Vertical bars indicate stimulation epoch. Grey, individual trials; black, mean. As training progressed, the pulse counts used changed. **(C)** Responses of example neurons on 9 pulse trials across training. Top three rows: opsin-expressing neurons (magenta) with decreasing (top), stable (middle), and increasing (bottom) responses across sessions. Bottom three rows: same for three

opsin non-expressing neurons (green). Grey, individual trials; color, mean. **(D)** Spatial maps of neurons from one imaging plane at three time points along the training progression. Size of dot indicates either response amplitude (top) or response probability (bottom). **(E)** Response amplitude across sessions for all neurons from one animal. Neurons are sorted by the mean response amplitude over the first three training sessions. **(F)** Aggregate responsiveness across training. Left, bottom 95% of neurons by responsiveness (magenta, opsin-expressing; green, opsin non-expressing). Each point indicates an animal's median response amplitude (top; z-scored ΔF/F) or probability of response (middle) across three early, middle, or late sessions. Large dot, grand mean across animals. Error bars, SEM. Bottom, the total number of opsin-expressing (O^+), opsin non-expressing (O^-), and expression-ambiguous ($O^?$) neurons (Methods). Center, same as left panel but for neurons in the top 5% of response amplitudes. Right, same as left panel but for neurons in the top percentile of response amplitudes. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, paired t-test; unlabeled, $p \geq 0.05$. **(G)** Change in response amplitude (left, z-scored ΔF/F) and response probability (right) for the training days common to both learners (orange) and non-learners (black) among the top 5% of most responsive opsin-expressing neurons. Dark line, mean; envelope, SEM. Dots reflect the mean across all days for individual mice, with hollow circle indicating the grand mean. P-value is given for t-test comparing learners and non-learners. **(H)** As in (G), but for opsin non-expressing neurons.

Pancholi et al., Figure 5

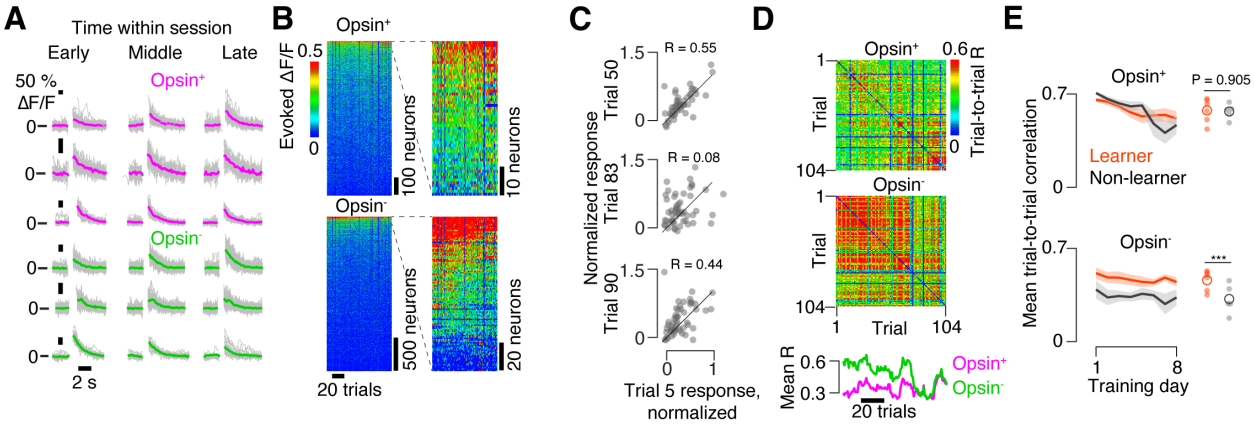

Figure 5. Representational drift within a session. (A) Example response to 9-pulse stimulus for several neurons at various points during the session. Magenta, opsin-expressing; green, opsin non-expressing. Top neuron, increasing response; middle, stable response; bottom, declining response. **(B)** Photostimulus-evoked ΔF/F for opsin-expressing (top) and opsin non-expressing (bottom) neurons for an example animal across all trials from a single session. **(C)** Pearson correlation between evoked ΔF/F response on an early trial and various subsequent trials. Each dot represents a neuron that was responsive at some point during the session. **(D)** Correlation between evoked ΔF/F vectors for all pairs of trials on an example imaging day. Top, matrix for opsin-expressing neurons; middle, matrix for opsin non-expressing neurons; bottom, average correlation with all other trials across individual trials. **(E)** Mean within-day representational drift during early sessions (1-8) in opsin expressing (top) and opsin non-expressing (bottom) neurons. Black, non-learners; orange, learners. Thin lines, individual animals; thick line, mean across animals (n = 6 mice, learners; n = 5 non-learners). Small circle, individual animal mean; large circle, cross-animal mean. ***P < 0.001, t-test comparing learners and non-learners.

Pancholi et al., Figure 6

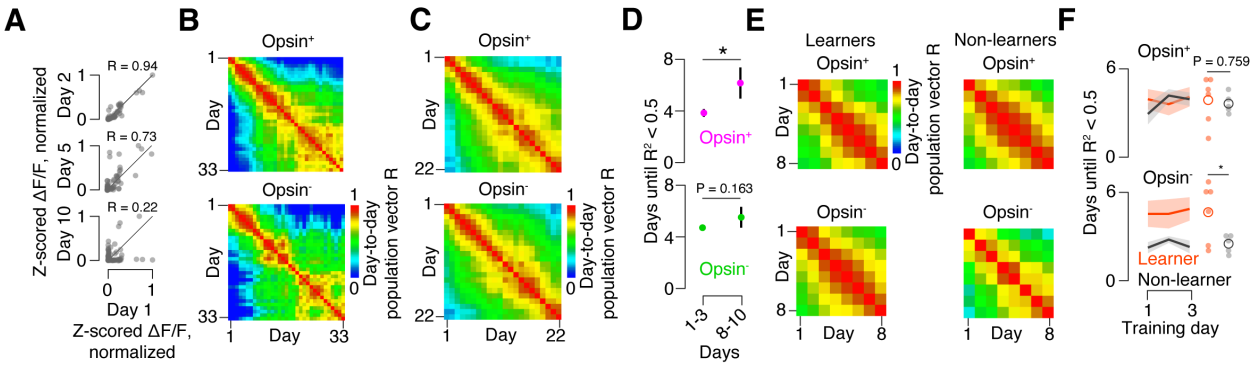

Figure 6. Representational drift across sessions. (A) Pearson correlation between mean evoked ΔF/F response on day 1 and example subsequent days. Each dot represents a neuron that was responsive in at least one session. **(B)** Cross-day correlation matrix for an example mouse. Top, opsin-expressing neurons; bottom, opsin non-expressing neurons. **(C)** Average cross-day correlation matrix across all mice that learned the task for the first 22 days of imaging. **(D)** Number of days until R^2 drops below 0.5. **(E)** Average cross-day correlation matrix across all mice for the first 8 days of imaging. Left, mice that learned; right, mice that did not learn. **(F)** Number of days from given starting point until cross-day R^2 drops below 0.5. Top, opsin-expressing neurons; bottom, opsin non-expressing neurons. Orange, learners; black, non-learners. Dark line, mean across mice. Envelope, SEM. Small circle, individual animal mean for days 1-3; large circle, cross-animal mean. *$P < 0.05$, t-test comparing learners and non-learners.

Pancholi et al., Figure 7

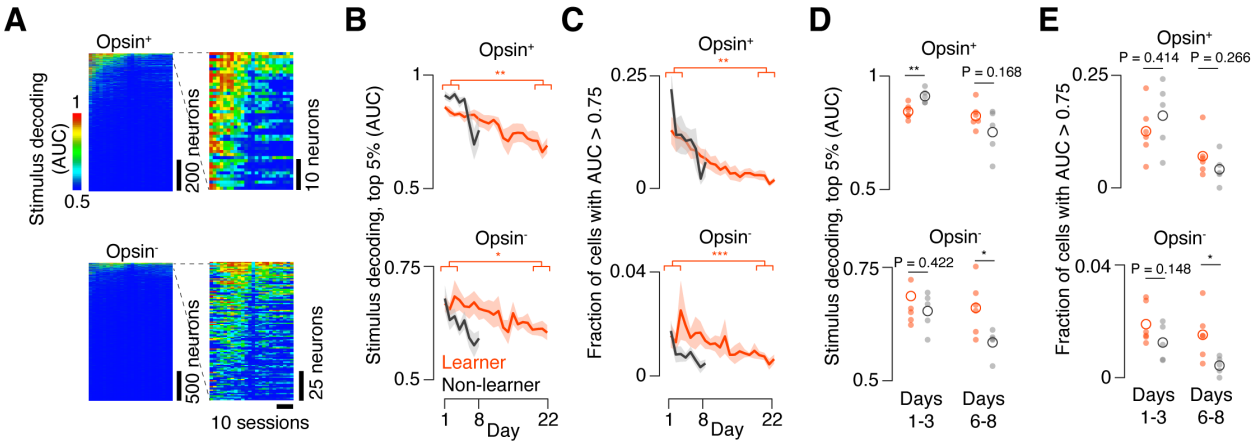

Figure 7. Stimulus decoding across sessions. (A) Decoding of stimulus across all neurons and sessions for an example mouse. Stimulus decoding for individual neurons was calculated as the area under the curve (AUC) for a linear decoder that must indicate if the stimulus contained a high (> 5) or low (< 5) pulse count (Methods). Top, opsin-expressing neurons. Bottom, opsin non-expressing neurons. Left, entire population, sorted by mean AUC across days. Right, top 5% by AUC. **(B)** Stimulus decoding (AUC) among best-decoding 5% of cells up to the last common session for learners (orange, n=6) and non-learners (black, n=5). Dark line, mean; envelope, SEM. Top, opsin-expressing population; bottom, opsin non-expressing. *P < 0.05, **P < 0.01, ***P < 0.001, paired t-test comparing mean across days 1-3 with mean of days 6-8. **(C)** As in (B), but for the number of neurons with AUC > 0.75. **(D)** Stimulus decoding (AUC) among top 5% of decoding cells up to the last common session for learners (orange, n = 6) and non-learners (black, n = 5) averaged across the first three common days (1-3; left) or last three common days (6-8; right). Small circle, single animal mean; large circle, cross-animal mean. P-value is given for t-test comparing learners and non-learners. **(E)** As in (D), but for the number of neurons with AUC > 0.75.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact. Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Simon Peron (speron@nyu.edu).

Materials Availability. This study did not generate new unique reagents.

Data and Code Availability. Source code used in this paper will be made available at http://github.com/peronlab upon publication. Data from this paper will be made available in a public repository (DANDI archive) upon publication.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse lines

Adult Ai162 (JAX 031562) X Slc17a7-Cre (JAX X 023527) (Daigle *et al.*, 2018) mice (15 male, 1 female; **Table S1**) were used throughout. These mice express GCaMP6s exclusively in cortical excitatory neurons in a tetracycline transactivator-dependent manner. To suppress expression during development, breeders were fed a diet that included doxycycline (625 mg/kg doxycycline; Teklad). Mice received doxycycline until weaning. All animal procedures were in compliance with protocols approved by New York University's University Animal Welfare Committee.

METHODS DETAILS

Surgical preparation

Mice (6-10 weeks old) were anesthetized with isoflurane during viral injections, surgical implantation, and LED placement (3% induction, 1.5% maintenance). A titanium headbar was attached to the skull with cyanoacrylate (Vetbond). A circular craniotomy (3.5 mm diameter) was made in the left hemisphere over vS1 (center: 3.3 mm lateral, 1.7 mm posterior from bregma) using a dental drill (Midwest Tradition, FG 1/4 drill bit).

Following the craniotomy, virus encoding the soma-localized opsin ChRmine and the red fluorophore mScarlet (AAV-8-CaMKIIa-ChRmine-mScarlet-Kv2.1-WPRE, 2.48×10^{13} vg/mL, diluted 1:500 in 1X PBS; generously provided by Dr. Karl Deisseroth) was injected into vS1. A glass capillary (Wiretrol II, Drummond) was pulled using a micropipette puller (P-97, Sutter Instrument) and the tip beveled to 25° with a tip diameter of 25 µm. The pipette was back-filled with mineral oil (M5904, Sigma-Aldrich) and 2 µL of viral solution was drawn into the tip. Three 100 nL injections were made 300 µm below the dura and spaced 400 µm apart in a triangle centered on the typical anatomical location of the C2 barrel. For each injection, the pipette was lowered into the brain at a rate of 300 µm/min, followed by a 1-minute pause after which virus was injected at a rate of 20 nL/min using a hydraulic micromanipulator (Narishige MO-10). Following a 2 minute pause, the pipette was withdrawn at a rate of 300 µm/min with an additional 1 minute pause at a depth of 150 µm below the dura. Finally, the dura was

removed using a pair of fine forceps (Fine Science Tools) and a double-layer cranial window (4.5 mm external diameter, 3.5 mm inner diameter; #1.5 coverslip; adhered with Norland 61 UV glue) was placed over the craniotomy. The cranial window and headbar were affixed to the skull with dental acrylic (Orthojet, Lang Dental).

Following surgical recovery, mice were trimmed to whiskers C1-3 and placed on water restriction, with regular trimming thereafter. To confirm that the area of opsin-expression fell within vS1, the location of the barrels corresponding to the C1-3 whiskers was identified by measuring touch-evoked GCaMP6s ΔF/F at coarse resolution during an imaging session during which the mouse was awake but not performing any task (4X objective; 3 x 3 mm field of view; **Figure S1A-C**).

LED placement was performed 2-4 weeks after injection. LEDs (590 nm, LXZ1-PL01, Lumileds) were connected to the LED driver by soldering 3 cm of polyurethane enameled copper wire (34 AWG) to each pad of the LED. A 2-pin, flat flex cable connector (Digikey) was then soldered to the free ends of the copper wire and secured with epoxy (Devcon). To waterproof the LED, a thin layer of clear nail polish (Sally Hansen) was applied to all surfaces of the LED. LEDs were then affixed to the animals' cranial windows. Following anesthesia, a small arc (60°) of dental cement was drilled away on the anterior-medial edge of the craniotomy to expose the edge of the cranial window. Cyanoacrylate was applied to the drilled area to ensure that the craniotomy remained sealed. The LED was placed at a 30° angle on the edge of the cranial window with the most lateral edge of the LED 0.5-1 mm away from the center of the opsin-expressing area (**Figure S1**). The medial edge of the LED and the copper wires were secured to the dental cement and the LED connector secured to the posterior edge of the headbar using cyanoacrylate. Waterproofing was confirmed by placing water on the cranial window and ensuring that no current passed between the water and the LED.

Light propagation model

The LED was modeled as an 8x8 grid of point sources positioned in water above the cranial window (**Figure S1E**). A light ray from each point source thus passed through the following: water, interface between water and window glass, glass, interface between glass and brain, and brain tissue. At each interface, Fresnel equations were used to calculate the fractional transmission of light. Within the brain, scattering of light was accounted for using empirical measurements (Li et al., 2019). The power density at different depths within the brain was computed by scaling the max power at the LED face (measured using a photodiode) by the fractional intensity, fractional transmission at each interface, and fraction of light remaining after scattering. This quantity was divided by the surface area of a half-sphere to yield the predicted power density at a given location in the brain. Light radiation from each point source was treated independently and averaged to compute the final predicted power density.

Photostimulation system

Optogenetic stimulus delivery was controlled by a LabJack T7, with commands originating from a MATLAB user interface on a separate computer. The stimulation and masking flash LEDs were controlled using LED drivers (T-cube, ThorLabs) whose

output currents were controlled by the LabJack. The signal to the stimulation LED was terminated using a 4-pin flat flex cable connector (Digikey) that was mated to the LED connector on the animal during behavior. The masking flash consisted of 3 LEDs (595 nm, XPEBAM-L1-0000-00A01, Cree LED) that were spectrally matched to the stimulation LED and placed near the animal's face to illuminate the eyes.

Behavior

Mice were trained on an optogenetic pulse count discrimination task in which the number of light pulses from the stimulation LED was associated with a water reward from one of two lickports (**Figure 1B**). Each trial consisted of three epochs: stimulus (500 ms), delay (500 ms), and response (< 2 s) (**Figure 1C**). During the stimulus epoch, between 0 and 9 light pulses (5 ms, 20 Hz) were presented from the stimulation LED. Pulses were randomly distributed among 9 time bins (5 ms, 20 Hz) to prevent the animal from using timing cues to solve the task (**Figure 1D**, **S4**). Masking flash pulses (15 ms, 20 Hz, starting 5 ms prior to the optogenetic stimulus pulse and ending 5 ms after) were presented at each time bin on all trials to prevent the animal from using visual cues to solve the task. The PMT shutters were closed 50 ms before the start of the first photostimulus time bin and opened 50 ms after the end of the last bin (total closure time: 505 ms). Thus, the masking flash provided a visual cue for the onset of the stimulus epoch, whereas the shutter provided an auditory cue. Following the stimulus epoch, there was a 500 ms delay after which an auditory cue (3 kHz, 50 ms) signaled the beginning of the response epoch. At this point, mice could indicate their decision by licking one of two lickports spaced 3 mm apart along the medial-lateral axis. The lickports were moved into reach during the response epoch by electrical motors. Well-trained mice withheld licking until the lickport came into reach. Correct responses were rewarded with 3-7 µL of water which the animal could collect for 1-2 s. Incorrect responses were punished with immediate withdrawal of the lickport and a timeout. After the response epoch, there was a 7 s inter-trial interval.

Training proceeded in a standard sequence. Following LED placement, water-restricted mice were handled and head-fixed to habituate animals to the behavioral apparatus. Animals were first trained on a photostimulus detection task in which 9 stimulation pulses predicted reward from the right lickport and 0 pulses predicted reward from the left lickport (**Figure 1B**, **E**). As mice attained our performance criterion (peak d-prime over a 61 trial window > 1.5) for more than 2 consecutive sessions for a given stage, the task progressed as follows: 0 pulses (left) vs. 9 pulses (right); 1 pulse (left) vs. 9 pulses (right); 1 and 3 pulses (left) vs. 7 and 9 pulses (right); 1, 3, and 5 pulses (left) vs. 5, 7, and 9 pulses (right). 5 pulse trials were randomly and equally rewarded between left and right. To maintain high performance on the final task and reduce bias, we modified the probability of presentation of each pulse count such that 1 and 9 pulse trials occurred up to 8x more frequently than intermediate pulse counts.

The behavioral task was controlled by a BPod state machine (Sanworks) and custom MATLAB software. The auditory response tone was controlled by a low-latency audio board (Bela). Lickport motion was controlled by a set of 3 motorized actuators (Zaber) and an Arduino. Licks were sensed using a custom electrical detection circuit (Janelia).

Two-photon microscopy

Calcium imaging was performed using a custom two-photon microscope (http://openwiki.janelia.org/wiki/display/shareddesigns/MIMMS). The microscope consisted of a 940 nm laser (Chameleon Ultra 2, Coherent), a Pockels cell (350-80-02, Conoptics), two galvanometric scanners (6SD11268, Cambridge Technology), a resonant scanner (6SC08KA040-02Y, Cambridge Technology), a 16x objective (N16XLWD-PF, Nikon), an emission filter for green fluorescence (FF01-510/84-30, Semrock), an emission filter for red fluorescence (FF01-650/60, Semrock), and two GaAsP PMTs (H10770PB-40, Hamamatsu). Each PMT had an associated shutter (VS14S1T1, Vincent Associates) that was controlled by a voltage signal from the LabJack.

Imaging data was acquired using Scanimage (Vidrio Technologies). Three 800-by-800 µm imaging planes axially spaced 60 µm apart were acquired at a rate of ~7 Hz. The objective was moved axially (total depth, 180 µm) with a piezo (P-725KHDS, Physik Instrumente). Power was depth-adjusted by the acquisition software with an exponential length constant of ~250 µm.

Imaging data were processed on the NYU High Performance Computing cluster using a semi-automated software pipeline that included image registration, segmentation, neuropil subtraction, ΔF/F computation, and calcium event detection (Peron *et al.*, 2015b). After the first imaging day, motion-corrected mean images were collected for each plane and used as references during imaging on subsequent days. Alignment across days was performed as previously described (Huber *et al.*, 2012).

Opsin expression was measured using mScarlet fluorescence. We defined opsin-expressing, opsin non-expressing, and ambiguous neurons using imaging data collected in the penultimate session for each animal. The green fluorescence signal was linearly demixed from the red fluorescence signal. Using the demixed pixel values, we calculated a 'redness score' as the mean pixel intensity across all pixels used for that neuron. We also computed the fraction of pixels with redness exceeding a noise threshold for each neuron following de-mixing. For each mouse, we manually selected a combined mean pixel value and pixel fraction threshold above which neurons were considered opsin-expressing. Neurons with a mean redness and fraction of red pixels below a second, lower set of thresholds were considered opsin non-expressing. Neurons with an intermediate redness score and fraction of red pixels were considered ambiguous and were excluded from analysis.

Cortical lesions and histology

Cortical lesions were performed with either an 800 nm (Chameleon Ultra 2, Coherent) or 1040 nm (Fidelity HP, Coherent) laser by focusing the laser at 200-300 µm depth for 12-40 s at 100% power (1.6 W; **Figure 1H**) at the infection site. Between 2 and 8 sites spaced 200-400 µm apart were targeted within the opsin-expressing area. Lesion success was visually confirmed by an increase in GCaMP6s fluorescence in the targeted area. In 2 animals, a second round of lesions was conducted the subsequent day because opsin-expression and evoked activity was still evident. Animals were

awake but not performing the task during lesioning and were monitored for signs of distress or discomfort. Imaging data was not collected following lesioning. Animals were perfused after completion of training. Brains were sectioned using a microtome (Leica), mounted on glass slides, and imaged on a fluorescent light microscope (VS120, Olympus). Histology revealed that typical lesions removed most, but not all, opsin-expressing neurons.

Quantifying responsiveness

For analyses of responsiveness (**Figures 2, 4**), neurons were classified as responsive or non-responsive in every trial by comparing the post-stimulation ΔF/F to the baseline ΔF/F. Baseline ΔF/F was selected to be the ΔF/F for the ~5.5 s (39 frames) preceding shutter closure on each trial. The evoked ΔF/F was calculated as the mean ΔF/F of the two frames immediately following shutter reopening. Neurons were considered photoresponsive on a given trial if the evoked ΔF/F exceeded the 97.5th percentile or fell below the 2.5th percentile of the distribution of baseline ΔF/F values across all trials in a session.

To compare responses between neurons and across sessions, we z-scored the post-stimulation ΔF/F, using the mean and standard deviation of the baseline ΔF/F over all trials for a neuron in a given session. Highly responsive neurons were identified by first computing the mean z-scored ΔF/F values over 9-pulse trials for all neurons and sessions, yielding one value per neuron per session. Neurons were identified as highly responsive if they exceeded either the 95th or 99th percentiles of this distribution on at least one session of the sessions being considered. Where not specified, the 95th percentile was used.

Within- and cross-day representational stability analyses

To quantify within-day representational stability (**Figure 5**), we first measured the evoked ΔF/F across all neurons for 9-pulse trials. For each 9-pulse trial, we constructed a population response vector consisting of the evoked ΔF/F for each neuron. The Pearson correlation between all pairs of population response vectors was then computed, yielding an inter-trial correlation matrix. Analysis was performed separately for opsin-expressing and opsin non-expressing populations. To compare mice that learned with those that did not, we first computed the mean of this matrix, using the upper triangular portion of the matrix excluding the diagonal. The comparison was made using the mean of this value across the training days that were common to all mice in the dataset (days 1-8).

To assess cross-day representational stability (**Figure 6**), we first computed the mean evoked ΔF/F across all neurons on 9-pulse trials. This resulted in a single population response vector consisting of the mean evoked ΔF/F for each neuron for that session. The Pearson correlation between all pairs of population response vectors was then computed, yielding a cross-day correlation matrix. To compare mice that learned with those that did not, we used this matrix to calculate the number of days from each day until R^2 fell below 0.5 (i.e., the day beyond which a given day's population response vector could no longer account for at least half of the variance for subsequent days'

population response vector). We only examined the days for which all mice had a valid value for this number (days 1-3).

Stimulus decoding analysis

Stimulus decoding was assessed using receiver operating characteristic (ROC) analysis. For each day, we partitioned the trials into high (7, 9 pulse) and low (0, 1 and 3 pulse) stimulus intensity trials, excluding 5 pulse trials on days where they were present. For each trial, we obtained the evoked $\Delta F/F$. ROC analysis was performed by sliding a criterion threshold through the range of evoked $\Delta F/F$ values across the two trial types, thereby classifying responses as false alarms ($\Delta F/F >$ threshold on low pulse count trial) or hits ($\Delta F/F >$ threshold on high pulse count trial). We report the area under the curve (AUC) resulting from this analysis (Green and Swets, 1966). Two metrics were employed to assess decoding efficacy: first, we computed the mean AUC across the top 5% of AUC values for a given day; second, we examined the number of neurons having an AUC exceeding 0.75 on any given day. In all cases, statistical comparisons were made by taking the average of one of the aforementioned metrics across three days. For comparing early to late decoding among mice that learned the task, we compared days 1-3 to the last three days common to all ($n = 6$) mice that learned: days 20-22. For comparisons between learners and non-learners, we used days 1-3 and the last three common days, days 6-8.

Statistical analysis

For comparisons between samples, two-tailed paired and unpaired t-tests were used. Where appropriate, the Bonferroni adjustment was applied to paired and unpaired t-tests to correct for multiple comparisons. For correlation tests, Pearson's correlation was used to identify a linear correlation coefficient (R) and test for significance.

SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

Animal ID	Sex	Age at training start (days)	Opsin$^+$ cell count	Opsin$^-$ cell count	Learning status	Opsin expression
293010	M	75	775	2593	Learner	Yes
293157	M	96	718	2679	Learner	Yes
293167	M	66	486	2413	Learner	Yes
295578	M	77	624	2640	Learner	Yes
295678	M	83	836	1621	Learner	Yes
010852	M	87	717	1778	Learner	Yes
293156	M	107	725	2530	Non-learner	Yes
294755	M	90	766	2360	Non-learner	Yes
295137	M	74	863	2183	Non-learner	Yes
295331	M	83	693	2178	Non-learner	Yes
010580	M	92	681	1609	Non-learner	Yes
297108	M	67	N/A	N/A	N/A	No
297109	M	67	N/A	N/A	N/A	No
297670	M	58	N/A	N/A	N/A	No
297023	M	81	N/A	N/A	N/A	No
297027	F	82	N/A	N/A	N/A	No
007417	M	N/A	829	2731	N/A	Yes

Table S1. **List of animals**. All mice were transgenic Ai162 X Slc17a7-Cre (Daigle et al., 2018), expressing GCaMP6s exclusively in excitatory neurons. Opsin non-expressing controls were only used in Figure 1. One animal (0074147) was never trained and used for calibration of photostimulation parameters in **Figure S3**.

Pancholi et al., Figure S1

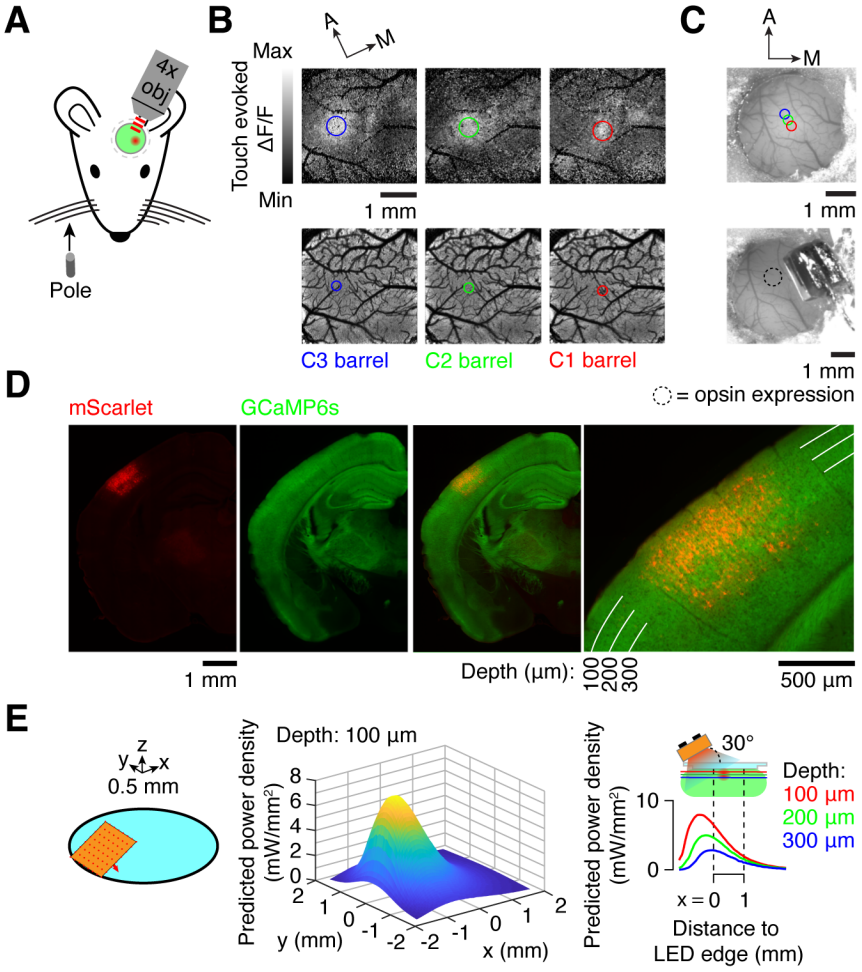

Figure S1. Targeting single barrel for superficial opsin expression. (A) Barrel mapping was conducted by driving a pole into whiskers C1-3 one at a time during widefield (4x) two-photon calcium imaging. **(B)** Touch-evoked ΔF/F following whisker stimulation. Top, ΔF/F following touch by indicated whisker; bottom, image showing barrel locations and corresponding vasculature. **(C)** Cranial window prior to LED placement (top) and following LED placement (bottom). Colored circles indicate the identified barrel locations from (B), and the dotted circle indicates the site of opsin expression. **(D)** Opsin expression. Left to right: expression of mScarlet, the fluorophore co-expressed with ChRmine; expression of GCaMP6s in a coronal section; overlay of GCaMP6s and mScarlet; magnified overlay. **€** Simulation of light propagation in the opsin-expressing tissue (Methods). Left, geometric model of the LED and cranial window; center, predicted power density surface at a depth of 100 µm in the brain; right, predicted power density at various depths in the brain and distances from the LED edge.

Pancholi et al., Figure S2

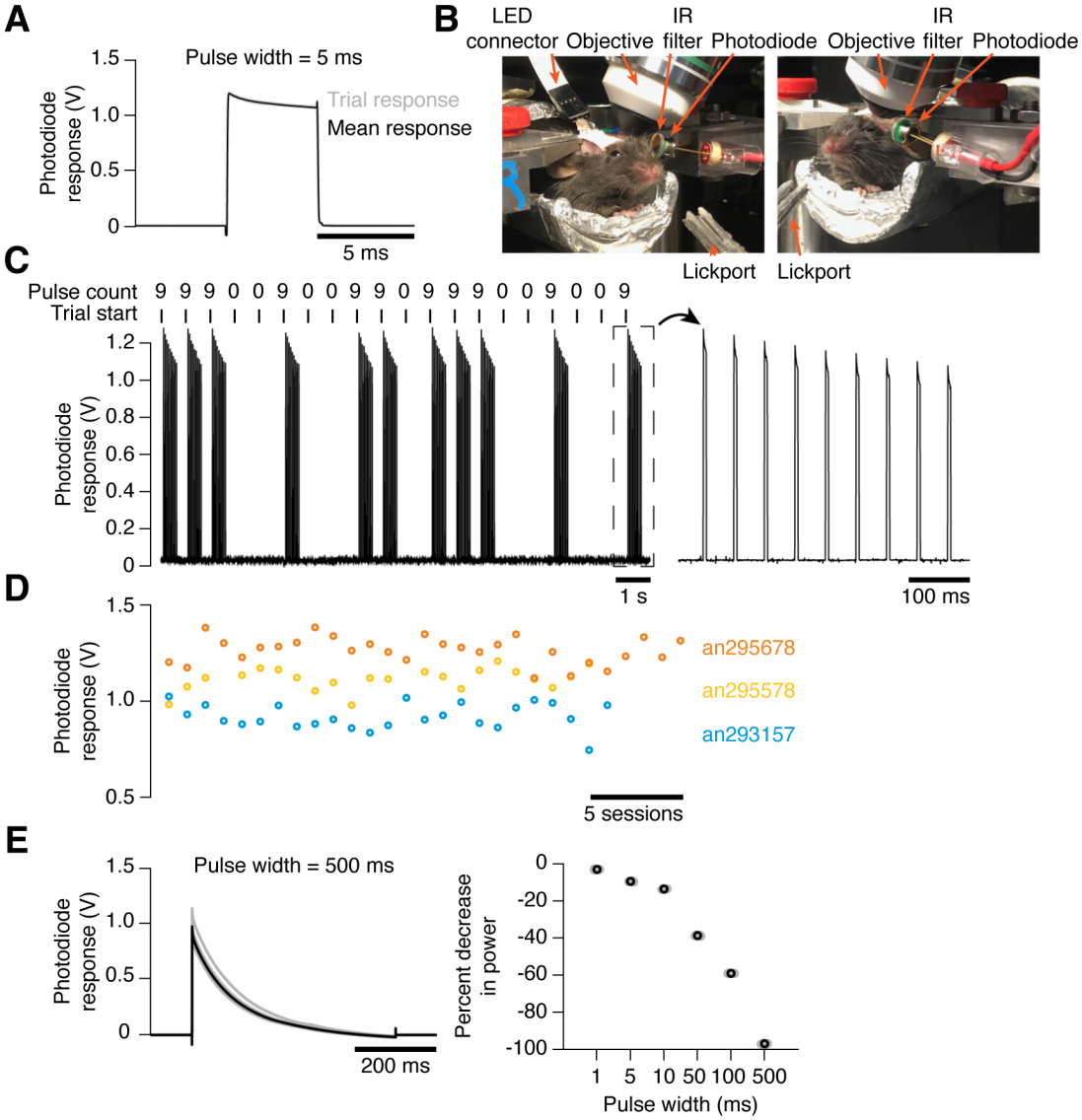

Figure S2. LED stimulation is robust and reproducible. (A) LED responses as measured by a standalone photodiode. 10 repetitions of a single 5 ms pulse at 100% power. Grey, individual repetitions; black, mean. **(B)** Placement of photodiode used to measure LED output. **(C)** Photodiode responses to LED stimulation across 20 consecutive trials in one behavioral session. **(D)** Photodiode responses to 9 pulse trials across sessions for 3 animals. LED was tested for 10-20 trials at the conclusion of each behavioral session following removal of water for imaging objective. **(E)** Photodiode responses to long duration pulses. Left, same as (A) for a 500 ms pulse; right, percent decrease in power between pulse start and pulse end for various pulse widths. Grey, individual repetitions; black, mean across repetitions.

Pancholi et al., Figure S3

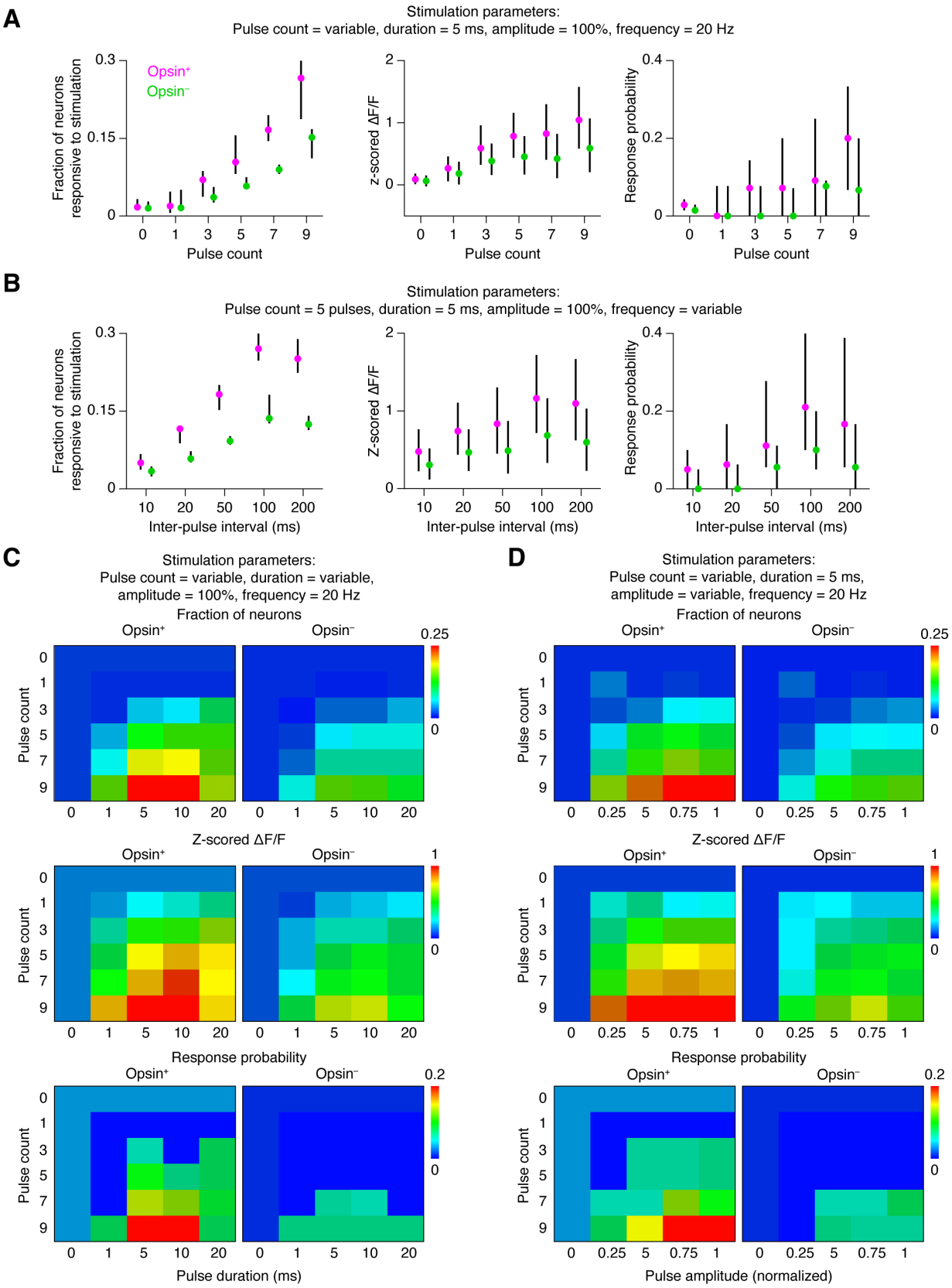

Figure S3. Neural responsiveness varies as a function of stimulation parameters. **(A)** Fraction of neurons responding (left), response amplitude (center), and response probability (right) as a function of pulse count for opsin-expressing (red) and opsin non-expressing (blue) neurons. Data are shown as median ± inter-quartile range for a single mouse that was awake but not performing a task (an007417, **Table S1**). **(B)** Same as (A) except as a function of inter-pulse interval. **(C)** Fraction of neurons responding (top), response amplitude (middle), and response probability (bottom) as function of both pulse count and pulse duration. **(D)** Same as (C) except for pulse count and pulse amplitude.

Pancholi et al., Figure S4

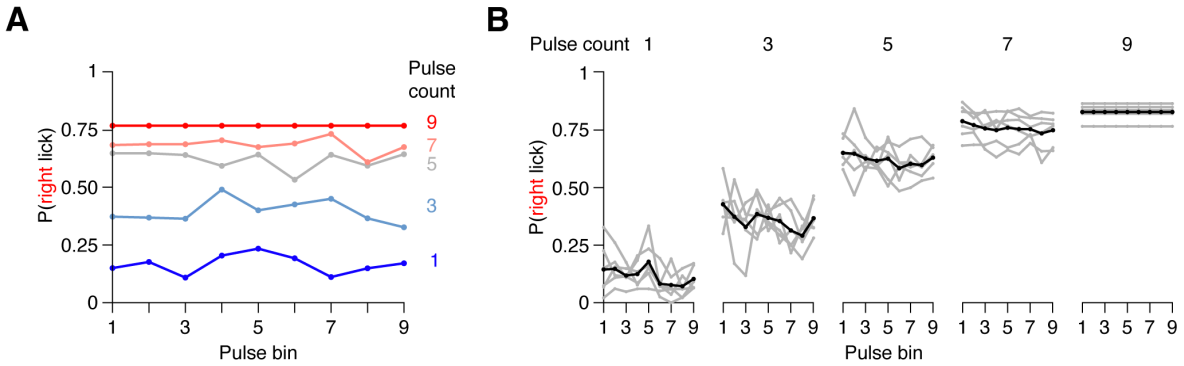

Figure S4. Mice do not use pulse timing to solve the task. (A) Right lick probability in each of the 9 possible pulse time bins for each pulse count in the final discrimination task for one animal. **(B)** Same as (A) for all animals. Grey, individual animals; black, mean.

Pancholi et al., Figure S5

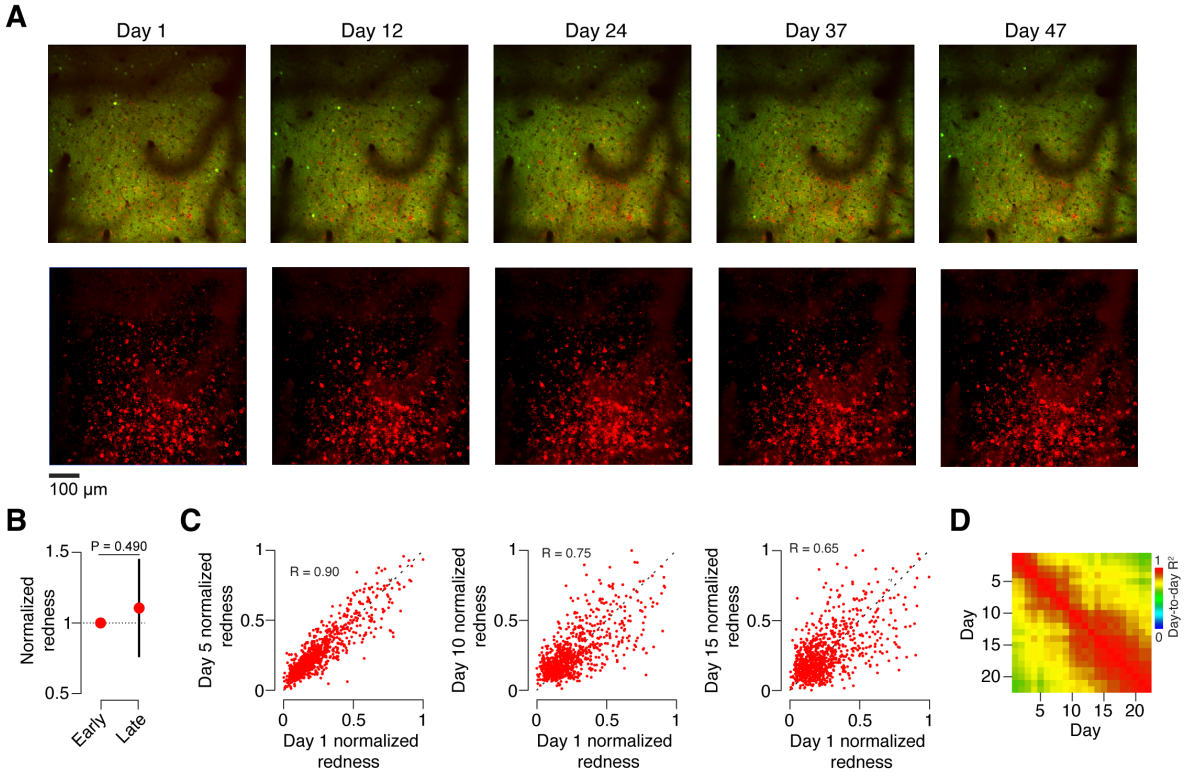

Figure S5. Opsin expression over the course of learning. (A) Example imaging plane across nearly 7 weeks. Top, GCaMP6s (green) and opsin (red) expression; bottom, linearly demixed red channel signal showing opsin expression. **(B)** Mean redness for the earliest and latest sessions. Error bars, SEM. P-value given for paired t-test comparing early to late sessions (n = 6 mice). **(C)** Relationship between redness on day 1 and redness on subsequent days for example animal. **(D)** Day-to-day correlation of redness across all opsin-expressing neurons averaged across animals (n = 6 mice).

REFERENCES

Afraz, S.R., Kiani, R., and Esteky, H. (2006). Microstimulation of inferotemporal cortex influences face categorization. Nature *442*, 692-695.

Allison-Walker, T., Hagan, M.A., Price, N.S.C., and Wong, Y.T. (2021). Microstimulation-evoked neural responses in visual cortex are depth dependent. Brain Stimul *14*, 741-750. 10.1016/j.brs.2021.04.020.

Barth, A.L., Gerkin, R.C., and Dean, K.L. (2004). Alteration of neuronal firing properties after in vivo experience in a FosGFP transgenic mouse. J Neurosci *24*, 6466-6475.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci *13*, 87-100.

Butovas, S., and Schwarz, C. (2003). Spatiotemporal effects of microstimulation in rat neocortex: a parametric study using multielectrode recordings. J Neurophysiol *90*, 3024-3039.

Butovas, S., and Schwarz, C. (2007). Detection psychophysics of intracortical microstimulation in rat primary somatosensory cortex. Eur J Neurosci *25*, 2161-2169.

Carmena, J.M., Lebedev, M.A., Henriquez, C.S., and Nicolelis, M.A. (2005). Stable ensemble performance with single-neuron variability during reaching movements in primates. J Neurosci *25*, 10712-10716. 10.1523/JNEUROSCI.2772-05.2005.

Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A., and Yuste, R. (2019). Controlling Visually Guided Behavior by Holographic Recalling of Cortical Ensembles. Cell *178*, 447-457 e445. 10.1016/j.cell.2019.05.045.

Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature *499*, 295-300. 10.1038/nature12354.

Chen, X., Wang, F., Fernandez, E., and Roelfsema, P.R. (2020). Shape perception via a high-channel-count neuroprosthesis in monkey visual cortex. Science *370*, 1191-1196. 10.1126/science.abd7435.

Chestek, C.A., Batista, A.P., Santhanam, G., Yu, B.M., Afshar, A., Cunningham, J.P., Gilja, V., Ryu, S.I., Churchland, M.M., and Shenoy, K.V. (2007). Single-neuron stability during repeated reaching in macaque premotor cortex. J Neurosci *27*, 10742-10750. 10.1523/JNEUROSCI.0959-07.2007.

Clopath, C., Bonhoeffer, T., Hubener, M., and Rose, T. (2017). Variance and invariance of neuronal long-term representations. Philos Trans R Soc Lond B Biol Sci *372*. 10.1098/rstb.2016.0161.

Cohen, L., Koffman, N., Meiri, H., Yarom, Y., Lampl, I., and Mizrahi, A. (2013). Time-lapse electrical recordings of single neurons from the mouse neocortex. Proc Natl Acad Sci U S A *110*, 5665-5670. 10.1073/pnas.1214434110.

Cohen, M.R., and Newsome, W.T. (2004). What electrical microstimulation has revealed about the neural basis of cognition. Curr Opin Neurobiol *14*, 169-177.

Daigle, T.L., Madisen, L., Hage, T.A., Valley, M.T., Knoblich, U., Larsen, R.S., Takeno, M.M., Huang, L., Gu, H., Larsen, R., et al. (2018). A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. Cell *174*, 465-480 e422. 10.1016/j.cell.2018.06.035.

Dalgleish, H.W., Russell, L.E., Packer, A.M., Roth, A., Gauld, O.M., Greenstreet, F., Thompson, E.J., and Hausser, M. (2020). How many neurons are sufficient for perception of cortical activity? Elife *9*. 10.7554/eLife.58889.

Danielson, N.B., Zaremba, J.D., Kaifosh, P., Bowler, J., Ladow, M., and Losonczy, A. (2016). Sublayer-Specific Coding Dynamics during Spatial Navigation and Learning in Hippocampal Area CA1. Neuron *91*, 652-665. 10.1016/j.neuron.2016.06.020.

Deitch, D., Rubin, A., and Ziv, Y. (2021). Representational drift in the mouse visual cortex. Curr Biol *31*, 4327-4339 e4326. 10.1016/j.cub.2021.07.062.

Desai, N.S., Rutherford, L.C., and Turrigiano, G.G. (1999). Plasticity in the intrinsic excitability of cortical pyramidal neurons. Nature Neuroscience *2*, 515-520.

Dhawale, A.K., Poddar, R., Wolff, S.B., Normand, V.A., Kopelowitz, E., and Olveczky, B.P. (2017). Automated long-term recording and analysis of neural activity in behaving animals. Elife *6*. 10.7554/eLife.27702.

Diamond, M.E., von Heimendahl, M., Knutsen, P.M., Kleinfeld, D., and Ahissar, E. (2008). 'Where' and 'what' in the whisker sensorimotor system. Nat Rev Neurosci *9*, 601-612. 10.1038/nrn2411.

Driscoll, L.N., Pettit, N.L., Minderer, M., Chettih, S.N., and Harvey, C.D. (2017). Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. Cell *170*, 986-999 e916. 10.1016/j.cell.2017.07.021.

Finkelstein, A., Fontolan, L., Economo, M.N., Li, N., Romani, S., and Svoboda, K. (2021). Attractor dynamics gate cortical information flow during decision-making. Nat Neurosci *24*, 843-850. 10.1038/s41593-021-00840-6.

Ganguly, K., and Carmena, J.M. (2009). Emergence of a stable cortical map for neuroprosthetic control. PLoS biology *7*, e1000153. 10.1371/journal.pbio.1000153.

Green, D.M., and Swets, J.A. (1966). Signal detection theory and psychophysics (John Wiley & Sons Ltd).

Hao, Y., Riehle, A., and Brochier, T.G. (2016). Mapping Horizontal Spread of Activity in Monkey Motor Cortex Using Single Pulse Microstimulation. Front Neural Circuits *10*, 104. 10.3389/fncir.2016.00104.

Histed, M.H., Bonin, V., and Reid, R.C. (2009). Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. Neuron *63*, 508-522. S0896-6273(09)00545-5 [pii]
10.1016/j.neuron.2009.07.016.

Histed, M.H., and Maunsell, J.H. (2014). Cortical neural populations can guide behavior by integrating inputs linearly, independent of synchrony. Proc Natl Acad Sci U S A *111*, E178-187. 10.1073/pnas.1318750111.

Histed, M.H., Ni, A.M., and Maunsell, J.H. (2012). Insights into cortical mechanisms of behavior from microstimulation experiments. Prog Neurobiol. 10.1016/j.pneurobio.2012.01.006.

Huber, D., Gutnisky, D.A., Peron, S., O'Connor, D.H., Wiegert, J.S., Tian, L., Oertner, T.G., Looger, L.L., and Svoboda, K. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. Nature *484*, 473-478. 10.1038/nature11039.

Huber, D., Petreanu, L., Ghitani, N., Ranade, S., Hromadka, T., Mainen, Z., and Svoboda, K. (2008). Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. Nature *451*, 61-64. 10.1038/nature06445.

Kato, H.K., Gillet, S.N., and Isaacson, J.S. (2015). Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance. Neuron *88*, 1027-1039. 10.1016/j.neuron.2015.10.024.

Kentros, C.G., Agnihotri, N.T., Streater, S., Hawkins, R.D., and Kandel, E.R. (2004). Increased attention to spatial context increases both place field stability and spatial memory. Neuron *42*, 283-295. 10.1016/s0896-6273(04)00192-8.

Knott, G.W., Quairiaux, C., Genoud, C., and Welker, E. (2002). Formation of dendritic spines with GABAergic synapses induced by whisker stimulation in adult mice. Neuron *34*, 265-273.

Leal-Campanario, R., Delgado-Garcia, J.M., and Gruart, A. (2006). Microstimulation of the somatosensory cortex can substitute for vibrissa stimulation during Pavlovian conditioning. Proc Natl Acad Sci U S A *103*, 10052-10057.

Lee, S., Kruglikov, I., Huang, Z.J., Fishell, G., and Rudy, B. (2013). A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nat Neurosci *16*, 1662-1670. 10.1038/nn.3544.

Lefort, S., Tomm, C., Floyd Sarria, J.C., and Petersen, C.C. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. Neuron *61*, 301-316. 10.1016/j.neuron.2008.12.020.

Li, N., Chen, S., Guo, Z.V., Chen, H., Huo, Y., Inagaki, H.K., Chen, G., Davis, C., Hansel, D., Guo, C., and Svoboda, K. (2019). Spatiotemporal constraints on optogenetic inactivation in cortical circuits. Elife *8*. 10.7554/eLife.48622.

London, M., Roth, A., Beeren, L., Hausser, M., and Latham, P.E. (2010). Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. Nature *466*, 123-127. 10.1038/nature09086.

Lutcke, H., Margolis, D.J., and Helmchen, F. (2013). Steady or changing? Long-term monitoring of neuronal population activity. Trends Neurosci *36*, 375-384. 10.1016/j.tins.2013.03.008.

Mank, M., Santos, A.F., Direnberger, S., Mrsic-Flogel, T.D., Hofer, S.B., Stein, V., Hendel, T., Reiff, D.F., Levelt, C., Borst, A., et al. (2008). A genetically encoded calcium indicator for chronic in vivo two-photon imaging. Nat Methods.

Manns, J.R., Howard, M.W., and Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. Neuron *56*, 530-540. 10.1016/j.neuron.2007.08.017.

Marks, T.D., and Goard, M.J. (2021). Stimulus-dependent representational drift in primary visual cortex. Nat Commun *12*, 5169. 10.1038/s41467-021-25436-3.

Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., et al. (2019). Cortical layer-specific critical dynamics triggering perception. Science *365*. 10.1126/science.aaw5202.

Masamizu, Y., Tanaka, Y.R., Tanaka, Y.H., Hira, R., Ohkubo, F., Kitamura, K., Isomura, Y., Okada, T., and Matsuzaki, M. (2014). Two distinct layer-specific dynamics of cortical ensembles during learning of a motor task. Nat Neurosci *17*, 987-994. 10.1038/nn.3739.

McMahon, D.B., Jones, A.P., Bondar, I.V., and Leopold, D.A. (2014). Face-selective neurons maintain consistent visual responses across months. Proc Natl Acad Sci U S A *111*, 8251-8256. 10.1073/pnas.1318331111.

Muzzio, I.A., Levita, L., Kulkarni, J., Monaco, J., Kentros, C., Stead, M., Abbott, L.F., and Kandel, E.R. (2009). Attention enhances the retrieval and stability of visuospatial and

olfactory representations in the dorsal hippocampus. PLoS Biol *7*, e1000140. 10.1371/journal.pbio.1000140.

O'Connor, D.H., Hires, S.A., Guo, Z.V., Li, N., Yu, J., Sun, Q.Q., Huber, D., and Svoboda, K. (2013). Neural coding during active somatosensation revealed using illusory touch. Nat Neurosci *16*, 958-965. 10.1038/nn.3419.

Perez-Ortega, J., Alejandre-Garcia, T., and Yuste, R. (2021). Long-term stability of cortical ensembles. Elife *10*. 10.7554/eLife.64449.

Peron, S., Chen, T.W., and Svoboda, K. (2015a). Comprehensive imaging of cortical networks. Curr Opin Neurobiol *32*, 115-123. 10.1016/j.conb.2015.03.016.

Peron, S., Pancholi, R., Voelcker, B., Wittenbach, J.D., Olafsdottir, H.F., Freeman, J., and Svoboda, K. (2020). Recurrent interactions in local cortical circuits. Nature *579*, 256-259. 10.1038/s41586-020-2062-x.

Peron, S.P., Freeman, J., Iyer, V., Guo, C., and Svoboda, K. (2015b). A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. Neuron *86*, 783-799. 10.1016/j.neuron.2015.03.027.

Peters, A.J., Chen, S.X., and Komiyama, T. (2014). Emergence of reproducible spatiotemporal activity during motor learning. Nature *510*, 263–267.

Polikov, V.S., Tresco, P.A., and Reichert, W.M. (2005). Response of brain tissue to chronically implanted neural electrodes. J Neurosci Methods *148*, 1-18. 10.1016/j.jneumeth.2005.08.015.

Polley, D.B., Kvasnak, E., and Frostig, R.D. (2004). Naturalistic experience transforms sensory maps in the adult cortex of caged animals. Nature *429*, 67-71.

Poort, J., Khan, A.G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G.B., Mrsic-Flogel, T.D., and Hofer, S.B. (2015). Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. Neuron *86*, 1478-1490. 10.1016/j.neuron.2015.05.037.

Rokni, U., Richardson, A.G., Bizzi, E., and Seung, H.S. (2007). Motor learning with unstable neural representations. Neuron *54*, 653-666. 10.1016/j.neuron.2007.04.030.

Rule, M.E., O'Leary, T., and Harvey, C.D. (2019). Causes and consequences of representational drift. Curr Opin Neurobiol *58*, 141-147. 10.1016/j.conb.2019.08.005.

Sadeh, S., and Clopath, C. (2022). Contribution of behavioural variability to representational drift. bioRxiv, 2022.2001.2002.474731. 10.1101/2022.01.02.474731.

Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. Nature *346*, 174-177. 10.1038/346174a0.

Schoonover, C.E., Ohashi, S.N., Axel, R., and Fink, A.J.P. (2021). Representational drift in primary olfactory cortex. Nature *594*, 541-546. 10.1038/s41586-021-03628-7.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neurosci *16*, 1486-1510.

Sombeck, J.T., Heye, J., Kumaravelu, K., Goetz, S.M., Peterchev, A.V., Grill, W.M., Bensmaia, S., and Miller, L.E. (2022). Characterizing the short-latency evoked response to intracortical microstimulation across a multi-electrode array. J Neural Eng *19*. 10.1088/1741-2552/ac63e8.

Thompson, L.T., and Best, P.J. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. Brain Res *509*, 299-308. 10.1016/0006-8993(90)90555-p.

Weinberger, N.M. (2004). Specific long-term memory traces in primary auditory cortex. Nat Rev Neurosci *5*, 279-290. 10.1038/nrn1366.

Weiss, J.M., Flesher, S.N., Franklin, R., Collinger, J.L., and Gaunt, R.A. (2019). Artifact-free recordings in human bidirectional brain-computer interfaces. J Neural Eng *16*, 016002. 10.1088/1741-2552/aae748.

Yang, Y., and Zador, A.M. (2012). Differences in sensitivity to neural timing among cortical areas. J Neurosci *32*, 15142-15147. 10.1523/JNEUROSCI.1411-12.2012.

Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nat Neurosci *16*, 264-266. 10.1038/nn.3329.