Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex

Highlights
- V1 neurons increasingly discriminate task-relevant stimuli with learning
- Chronic imaging reveals single cell changes underlying this population effect
- Learning-related changes are reduced when animals ignore task-relevant stimuli
- Anticipatory and behavioral choice-related signals emerge in reward-predicting cells

Authors
Jasper Poort, Adil G. Khan, Marius Pachitariu, ..., Georg B. Keller, Thomas D. Mrsic-Flogel, Sonja B. Hofer

Correspondence
sonja.hofer@unibas.ch

In Brief
By tracking the same visual cortex neurons across days, Poort et al. demonstrate how learning a visual task leads to increasingly distinguishable representations of relevant stimuli. These changes parallel the emergence of diverse non-sensory signals in specific neuronal subsets.
Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex

Jasper Poort,1,2,5 Adil G. Khan,1,2,5 Marius Pachitariu,3 Abdellatif Nemri,1,2,5 Ivana Orsolic,1 Julija Krupic,2 Marius Bauza,2 Maneesh Sahani,3 Georg B. Keller,1 Thomas D. Mrsic-Flogel,1,2 and Sonja B. Hofer1,2,∗

1Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland
2University College London, 21 University Street, London WC1E 6DE, UK
3Gatsby Computational Neuroscience Unit, University College London, 17 Queen Square, London WC1N 3AR, UK
4Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
5Co-first author
*Correspondence: sonja.hofer@unibas.ch
http://dx.doi.org/10.1016/j.neuron.2015.05.037

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

We determined how learning modifies neural representations in primary visual cortex (V1) during acquisition of a visually guided behavioral task. We imaged the activity of the same layer 2/3 neuronal populations as mice learned to discriminate two visual patterns while running through a virtual corridor, where one pattern was rewarded. Improvements in behavioral performance were closely associated with increasingly distinguishable population-level representations of task-relevant stimuli, as a result of stabilization of existing and recruitment of new neurons selective for these stimuli. These effects correlated with the appearance of multiple task-dependent signals during learning: those that increased neuronal selectivity across the population when expert animals engaged in the task, and those reflecting anticipation or behavioral choices specifically in neuronal subsets preferring the rewarded stimulus. Therefore, learning engages diverse mechanisms that modify sensory and non-sensory representations in V1 to adjust its processing to task requirements and the behavioral relevance of visual stimuli.

INTRODUCTION

Primary areas of the sensory neocortex are thought to faithfully represent the identity of stimuli in the external environment. Yet as animals learn the association between a sensory stimulus and its behavioral relevance, or improve their perceptual capabilities with training, stimulus representations in sensory cortical areas can change (Schoups et al., 2001; Yang and Maunsell, 2004; Rutkowski and Weinberger, 2005; Blake et al., 2006; Li et al., 2008; Wiest et al., 2010; Gdalyahu et al., 2012; Goltstein et al., 2013; Yan et al., 2014). Such changes may lead to enhanced and more distinct representations of task-relevant stimuli, and therefore improve the salience of information relayed to downstream areas.

The nature and effect sizes of learning-related changes to neural representations vary strongly between different studies, potentially depending on modality, sensory cortical area, and the behavioral task (Schoups et al., 2001; Yang and Maunsell, 2004; Rutkowski and Weinberger, 2005; Li et al., 2008; Ghose et al., 2002; Law and Gold, 2008). The repeated association between a stimulus and reward can lead to lasting, task-independent changes in cortical representations of that stimulus (Schoups et al., 2001; Rutkowski and Weinberger, 2005; Goltstein et al., 2013). Alternatively, the expression of learning-related changes to sensory responses can also depend on the animals being engaged in the task (Li et al., 2004; 2008; Polley et al., 2006), consistent with observations that even in primary sensory cortex neuronal responses can be influenced by non-sensory, task-dependent signals reflecting the animal’s attentive state, expectations, or behavior (see, for example, Ress and Heeger, 2003; Shuler and Bear, 2006; Li et al., 2008; Niell and Stryker, 2010; Keller et al., 2012; David et al., 2012; Stănioiu et al., 2013; Nienborg and Cumming, 2014). Therefore, the strategies by which learning can modify cortical sensory processing are diverse but remain poorly understood. Specifically, how do individual neurons change their response properties as stimuli acquire behavioral relevance? To what extent do these changes persist when the animals are not engaged in the task? How do learning-induced response changes relate to the appearance of non-sensory, task-dependent signals? Do these non-sensory signals act globally, or do they target specific neuronal subsets encoding behaviorally relevant sensory features?

To address these questions, it is crucial to track the activity of the same cells over the course of learning. We therefore used chronic two-photon calcium imaging (Huber et al., 2012; Chen et al., 2013) in mouse V1 while the animals learned to perform a visual discrimination task in virtual reality. We observed a robust and progressive population-wide increase in neural selectivity in cortical layer 2/3 (L2/3) during learning—an effect related to greater day-to-day stability of single cell response preferences as well as to an increase in the number of cells selective for task-relevant stimuli. Improvements in V1 selectivity were reduced when animals disengaged from the task. Task acquisition additionally led to the appearance of both anticipatory and behavioral choice-related signals in a specific subpopulation of neurons whose firing predicted the reward. Therefore,
learning the relationship between visual cues and their behavioral relevance leads to concerted changes in the representation of both sensory and non-sensory task-related information in a primary sensory cortical area.

RESULTS

Behavioral Task

Mice can perform complex visually guided behaviors, but they often require weeks of training to achieve high performance levels when head restrained (Andermann et al., 2010; Glickfeld et al., 2013; Pinto et al., 2013). Virtual reality environments offer an advantage for training head-fixed animals because they allow active engagement with the sensory world, for example, when the animal’s locomotion on a treadmill is directly coupled to optic flow changes in the visual scene (Hölscher et al., 2005; Dombeck et al., 2007). We hypothesized that this type of active visuomotor engagement approximates ethological situations when mice encounter behaviorally relevant stimuli during navigation, exploration, or foraging. Indeed, we found that this enabled rapid visually guided learning (see below).

We trained head-fixed mice to discriminate two grating patterns of different orientations in a virtual reality environment (see Movie S1 available online) from random starting points, mice were abruptly presented with a corridor containing either vertical or angled (40° relative to vertical) gratings on both walls. The abrupt appearance of the grating corridors provided precise control of stimulus timing. Mice were rewarded for licking in response to the vertical grating corridor with a drop of soya milk delivered through a reward spout (hit trial; reward was given if a lick was detected in a region a short distance into the grating corridor, referred to as the reward zone). No punishment was given for licking in response to the non-rewarded, angled grating corridor (false-alarm trial). Most mice progressed rapidly from indiscriminate licking (example lick raster plots in Figure 1C, top) to licking only within the grating corridors in response to both gratings (Figure 1C, middle), and finally to nearly exclusive licking in response to the rewarded, vertical grating (Figure 1C, bottom) and withholding licking in the non-rewarded angled grating corridor (correct rejection trials). Mice typically slowed down while licking in the rewarded grating corridor and learned to accelerate upon seeing the non-rewarded grating (Figure 1C, right panels). We quantified task performance by calculating the behavioral d’-prime for each training session, which is a measure of the difference in the proportions of hit and false-alarm trials (Figure 1D; see Experimental Procedures). Mice usually learnt the task within 3–6 days (Figure 1D) and eventually reached high behavioral accuracies (behavioral d’-prime in last session 3.2 ± 0.7, corresponding to 89% ± 8% correct responses, mean ± SD; see Figure S1).

Figure 1. Rapid Learning of a V1-Dependent Visual Discrimination Task in Virtual Reality

(A) Schematic of the virtual reality setup.

(B) Task schematic with virtual corridor wall patterns. CR, correct rejection; FA, false alarm.

(C) Changes in licking over learning in an example mouse. Licks (dots) aligned to grating onset in vertical grating (left, blue shading) and angled grating (middle, pink shading) trials. Red dots, reward delivery; yellow dots, licking after reward delivery. Right, average running speed for session shown on left, aligned to grating onset for vertical (blue) and angled (red) trials. Shading, SEM.

(D) Behavioral performance (behavioral d’-prime; see Experimental Procedures) of five mice imaged on consecutive training sessions. See also Figure S1.

(E) Behavioral performance in the visual and an equivalent odor discrimination task (see Experimental Procedures, average across sessions) as a function of light intensity during bilateral optogenetic silencing of visual cortex. PV-ChR2, transgenic mice expressing Channelrhodopsin-2 in parvalbumin-positive interneurons (n = 4 mice, 10 visual and 4 odor discrimination sessions), WT, wild-type mice (n = 3 mice, 7 sessions). *p < 0.05, **p < 0.001 after Bonferroni correction, Wilcoxon rank-sum test comparing PV-ChR2 to WT in the visual task.
We tested whether V1 activity was required for visual discrimination in this task by optogenetically silencing V1 in both hemispheres of fully trained animals in a random subset of trials. We silenced the cortex during grating corridor presentation by photostimulation of parvalbumin-positive inhibitory interneurons expressing Channelrhodopsin-2 in transgenic mice (Boyden et al., 2005; Lien and Scanziani, 2013; Glickfeld et al., 2013). Visual discrimination performance decreased progressively when increasing the intensity of blue light directed to V1 in transgenic mice, (Figure 1E; Friedman test, $\chi^2[4] = 32.44, p < 10^{-5}$), but not in wild-type control mice (Figure 1E; Friedman test, $\chi^2[4] = 5.76, p = 0.22$). The same transgenic mice were additionally trained in an analogous odor discrimination task in the same virtual corridor (see Experimental Procedures), which they continued to perform normally even when illuminating V1 with high light intensities (Figure 1E; Friedman test, $\chi^2[3] = 0.20, p = 0.98$), demonstrating that only visual processing was affected by this optogenetic manipulation.

Response Dynamics Underlying Increase in Neuronal Selectivity during Learning
Having established the necessity of V1 for this visual discrimination task (see also Glickfeld et al., 2013), we examined how the activity of neuronal populations in V1 changed during learning. For this purpose, we expressed the calcium indicator GCaMP6 (Chen et al., 2013) in V1 using AAV vectors, and chronically recorded calcium signals (32 Hz frame rate) in L2/3 using two-photon microscopy (Denk et al., 1990) while the animals performed the task (Figures 2A and 2B; on average 199 trials per session, range 31–342 trials). We imaged the same populations of neurons (75 ± 27 cells per mouse; mean ± SD) either in each training session over the entire time course of learning (five mice, Figure 1D), before and after learning (three mice), or only after learning (three mice). Neurons exhibited diverse response profiles during the task (Figures 2A, 2B, and S2). While some neurons responded to features in the approach corridor (Figure 2A, cell 1; Figure S7), many cells responded to both the vertical and angled grating corridors, and their responses were often stronger to one grating than the other (Figures 2B and S2). In other neurons, the calcium signal decreased during grating presentation (Figure 2B, cell 8; Figure S2). Despite variability in response amplitudes and in the degree of response selectivity from session to session (see below and Figures 2E–2G), the majority of neurons maintained their response profiles over time (Figures 2B, S3A, and S3B).

To quantify how the preference and selectivity of individual neurons for the two grating corridors changed during learning, we derived an index of neuronal selectivity for each neuron in each training session (defined as the difference between the average responses to vertical and angled gratings in a time window 0–1 s after grating onset, normalized by the pooled standard deviation of responses across trials). By binning sessions with similar behavioral performance (Figure 2C), we observed a gradual broadening of the distribution of neuronal selectivity over learning, resulting in both more positive values (higher preference for the rewarded, vertical grating) and more negative values (higher preference for the non-rewarded, angled grating). Consequently, the fraction of selective neurons rose significantly over learning (Figure 2D), including an increase in the percentage of cells preferring the non-rewarded grating corridor (12% to 19%, $p = 0.02$, bootstrap test), and a larger increase in the percentage of cells preferring the rewarded grating corridor (12% to 32%, $p < 10^{-4}$; Figure 2D; see Figure S4 for individual mice). Restricting the analysis only to neurons with a significant response increase after grating corridor onset ($p < 0.01$, Wilcoxon signed-rank test) yielded similar results (Figure S5).

The increase in neuronal selectivity was caused by an increase in reliability of responses (mean standard deviation of responses within 0–1 s window from grating onset, pre learning = 0.088, post learning = 0.063, $p = 0.001$, bootstrap test, 27 sessions before and S2 after learning), as well as an increased difference in response amplitude to the two gratings with learning (mean absolute response difference; pre learning = 0.017, post learning = 0.024, $p = 0.016$, bootstrap test). However, there was no consistent strategy by which individual neurons changed their response amplitudes to the two gratings (Figure S3C).

We next determined whether the increase in selectivity for grating stimuli was restricted to neurons with specific response properties. Neurons preferentially responding to either grating before learning were no more likely to increase their selectivity during learning than non-selective neurons ($R = -0.06, p = 0.20$; Figure S3E). Moreover, individual cells showed relatively large variability in how they changed their selectivity over learning (Figures S3D and S3F). The increase in selectivity, therefore, involved diverse modes of response change distributed over many neurons across the L2/3 population in V1.

Neuronal responses can show considerable variability from one day to the next (Huber et al., 2012; Peters et al., 2014; Ziv et al., 2013). We quantified day-to-day fluctuations of stimulus preferences of individual cells and how they changed during learning (Figures 2E–2G). We computed the likelihood of neurons maintaining their grating selectivity from one day to the next (persistence of response preference, Figure 2F) within different stages of learning: before animals showed improvements in their behavioral performance (pre learning), during learning, and after the behavioral performance had stabilized (post learning; see Experimental Procedures). While neurons were relatively more likely to lose their selectivity from one day to the next before learning, it was rare for neurons to completely switch from preferring one grating to the other (on average 3% before learning). Over learning, the persistence of selective responses increased, and cells preferring either the rewarded or the non-rewarded stimulus became more stable in their stimulus preference (Figure 2F; rewarded-grating-preferring cells, pre = 49% to post = 70%, $p < 10^{-3}$; non-rewarded-grating-preferring cells, pre = 17% to post = 55%, $p < 10^{-4}$, bootstrap tests). We additionally determined the probability of non-selective neurons becoming selective from one day to the next (Figure 2G). As learning progressed, non-selective neurons became more likely to acquire a preference for the rewarded, vertical grating, but not for the non-rewarded, angled grating (Figure 2G, rewarded-stimulus-preferring cells, $p < 10^{-4}$; non-rewarded-stimulus-preferring cells, $p = 0.29$, bootstrap tests). Therefore, the increasing preference for task-relevant stimuli in L2/3 of V1 during learning was a result of a stabilization of response selectivity to both gratings as well as an increased...
conversion of unselective neurons into those more selective for the rewarded grating.

**Progressive Increase of Population-wide Stimulus Discriminability in V1 with Learning**

We next determined how these learning-related changes in single-neuron selectivity influenced the ability of neuronal populations to discriminate the grating stimuli. As a composite measure of selectivity in a population with both positive and negative selectivity indices, we computed the root-mean-square of grating selectivity of all neurons imaged simultaneously (population selectivity) over the time course of stimulus presentation (200 ms sliding window; see Experimental Procedures) for different behavioral discrimination performance levels. Colors denote bins of behavioral d-prime from chance performance (blue) to expert performance (orange).

**Figure 2. Chronic Two-Photon Imaging of Single Cells across Learning**

(A) Example calcium traces of four V1 neurons during the task in an expert mouse, aligned to running speed (gray trace on top), licking (black lines), and reward delivery (red lines). Blue and red shading indicate time spent in the vertical and angled grating corridor, respectively.

(B) Average responses and corresponding images of four additional example cells in four training sessions aligned to grating onset (dashed vertical line). Values above each trace on day 6 denote neuronal selectivity for grating corridors, computed from responses 0–1 s after grating onset (see Experimental Procedures).

(C) Histograms of neuronal selectivity (positive values: cells prefer vertical, rewarded gratings; negative values: cells prefer angled, non-rewarded gratings) for different behavioral discrimination performance levels. Colors denote bins of behavioral d-prime from chance performance (blue) to expert performance (orange).

(D) Grating selectivity of the same neurons (rows) across sessions (columns) in the first three and last three sessions; cells were ordered based on the selectivity averaged across the middle four sessions; n = 8 mice.

(E) Persistence of response selectivity across consecutive training sessions during different stages of learning. Values are the probability of a neuron with a grating preference on one day to maintain this preference on the next day within each learning stage (response 0–1 s after grating onset; vertical grating, N_vertical = 51, N_angle = 121, N_non = 79; angled grating, N_vertical = 90, N_angle = 95, N_non = 200 cells). Errorbars depict SEM (determined by bootstrapping with replacement). Pre learning, behavioral d-prime (d̃) of both sessions < 1, and Δd̃ < 0.5 (14 session pairs); during learning: d̃ first session < 2, d̃ second session > 0.5, Δd̃ > 0.5 (14 session pairs); after learning: d̃ both sessions > 2, absolute Δd̃ < 0.5 (19 session pairs).

(F) The fraction of non-selective cells becoming selective for task-relevant stimuli across consecutive training sessions during different stages of learning (as in F). Values are the probability cells non-selective on one day (N_non = 549, N_angle = 417, N_non = 422) to develop a preference for one of the two gratings the next day within each learning stage.

n = 11 mice for all panels, except where indicated. See also Figures S2–S5.
0.26, post learning = 0.46, p < 10^{-4}, bootstrap test, comparison within 0–1 s window post grating onset), and rose sharply after grating onset only in well-trained mice. Additionally, we trained a linear decoder to predict which stimulus the mouse had encountered in each trial (vertical versus angled grating corridor) from calcium responses of all cells imaged simultaneously (see Experimental Procedures). The ability of the decoder to classify trials correctly increased strongly with improved behavioral performance during learning, such that classification accuracy exceeded 90% in expert mice (Figure 3B). Therefore, as mice got better at discriminating the two gratings, population-level representations of these task-relevant stimuli became increasingly distinguishable. In individual animals, neuronal population selectivity closely tracked the session-by-session changes in behavioral performance (Figure S6); there was a high positive correlation between the average population selectivity (0–1 s post grating onset) and the behavioral d-prime for individual sessions (Figure 3C; R = 0.64, p < 10^{-4}, n = 78 sessions).

These results suggest that the increased selectivity of V1 neurons during training is a specific effect of learning the discrimination task. Indeed, neither response amplitude nor response selectivity for stimulus features in the approach corridor increased during learning (p = 0.38, pre- versus post learning, Wilcoxon signed-rank test), even though those features did evoke reliable responses in subsets of cells (Figure S7). Therefore, learning-related changes in V1 activity were specific to task-relevant grating stimuli and were not a consequence of repeated exposure to the same visual environment over multiple sessions (Frenkel et al., 2006).

**Task Dependence of Learning-Induced Increases in Neuronal Selectivity**

To address this question, we trained expert mice to switch between blocks of the visual discrimination task and an analogous olfactory discrimination task. Mice learned to lick to obtain a reward in response to one of two different odors while running through the virtual corridor where they occasionally encountered the grating stimuli used in the visual discrimination task (see Experimental Procedures). Mice learned to switch rapidly between the two tasks within the same session, such that they successfully discriminated the grating stimuli in the visual task but ignored the same grating stimuli (while successfully discriminating odors) during the intervening olfactory blocks (Figure 4A; see Experimental Procedures, Movie S1, and Figures S8A–S8F). Although the average response amplitudes to the grating stimuli did not change in the olfactory blocks (Figure S9; p > 0.32), most neurons became less selective (Figure 4B), as the fractions of neurons preferring both the rewarded and non-rewarded stimuli decreased (Figure 4C; all p values < 10^{-4}, bootstrap test). Consequently, population selectivity for the same grating stimuli decreased significantly in the olfactory blocks compared to the visual blocks (Figure 4D, p = 0.014, bootstrap test), but remained above the pre learning level (p = 10^{-4}). Moreover, when the same visual stimuli were played back to fully trained but anesthetized mice, the selectivity of V1 populations was further reduced compared to the olfactory blocks (p = 0.002) but still higher than before learning (Figure 4D, p = 0.04). These results indicate that there may be two causes underlying the learning-related increase in stimulus selectivity in V1: a more lasting, task-independent change in the visual circuits, and a task-dependent modulation that depended on the animals being engaged in visual discrimination. The fact that the selectivity of most neurons increased during visual discrimination (Figure 4B) suggests that the task-dependent signals mediating these effects have a widespread influence on neuronal populations in V1.
Changes in Motor Behavior with Training Cannot Account for the Increase in Neuronal Selectivity

Several possible causes may underlie the task-dependent changes of stimulus selectivity in V1 during learning. Responses to task-relevant stimuli could be specifically modified to give rise to more distinguishable representations at the population level, thus allowing for easier perceptual discrimination. In addition, changes in V1 activity could also reflect signals associated with the behavioral outcome of the task, including signals related to the animals’ motor behavior, which are known to modulate the activity of V1 neurons (Neill and Stryker, 2010; Keller et al., 2012; Saleem et al., 2013). Neither the average running speed nor activity of V1 neurons (Niell and Stryker, 2010; Keller et al., 2012) to the animals’ motor behavior, which are known to modulate the activity of V1 neurons (Neill and Stryker, 2010; Keller et al., 2012; Saleem et al., 2013), there was no positive relationship between how strongly cells were modulated by running and/or optic flow speed and their change in grating selectivity over the course of learning (Figures S11C–S11G). Note that in these tasks, gratings were always preceded by a gray corridor to ensure also that the visual input preceding the task-relevant stimuli was uniform across conditions. Thus, when the speed profiles of task-relevant visual stimuli were identical in all conditions, we again found that V1 neurons increased their grating selectivity over the course of learning, as well as when the animals engaged in the visual compared to the olfactory discrimination task (Figures S11B and S11G, respectively).

Second, we tested if locomotion-related response modulation in V1 influenced the learning-related changes in neuronal selectivity. We did not observe any speed-related differences in neuronal population selectivity computed from trials with matched running speed profiles in all conditions (Figures S5A and S5B). Specifically, V1 neurons showed increased selectivity after learning independent of running speed (Figure 5A; population selectivity within 0–0.5 s from grating onset, pre- versus post learning, slow: p = 0.02; fast: p = 0.01, bootstrap test). Moreover, while some neurons showed a correlation between their calcium signal and running speed, as expected from previous studies (Neill and Stryker, 2010; Keller et al., 2012; Saleem et al., 2013), there was no positive relationship between how strongly cells were modulated by running and/or optic flow speed and their change in grating selectivity over learning (Figures S12A and S12D; see Experimental Procedures). Indeed, the exclusion of neurons whose responses were modulated by running did not alter the increase in V1 population selectivity over learning (Figures S12B and S12C, and S12E and S12F).
Figure 5. Neuronal Changes during Learning Cannot Be Explained by Changes in Running Behavior

(A) Population selectivity for different running speeds (thick traces, fast running trials; thin traces, slow running trials) matched across sessions before (behavioral d-prime < 1, gray traces) and after learning (behavioral d-prime > 2, black traces). Data are from 27 sessions pre learning and 52 sessions post learning. See also Figures S10–S13.

(B) Average running speeds corresponding to conditions in (A). Solid lines indicate vertical grating trials and dashed lines angled grating trials. There was no difference in running speeds within the same speed bin across stimuli, nor before/after learning (all comparisons p > 0.08).

(C) Time course of decoding performance from grating onset (probability of correct classification of grating corridor type, vertical versus angled). Decoder was based either on cumulative neuronal activity (solid lines) or cumulative running speed (dashed lines) for different behavioral discrimination performance levels during learning. n = 11 mice for (A)–(C).

(D) Decoding performance as in (C), before learning (behavioral d-prime < 1, gray lines, 23 sessions, 11 mice), and after learning (behavioral d-prime > 2) for sessions with delayed divergence of running behavior in vertical and angled grating trials (see Experimental Procedures; purple, 8 sessions, 7 mice), and sessions with matched behavioral d-prime but early divergence of running behavior (black lines, 8 sessions, 6 mice). In all panels, 0 s = grating corridor onset.

Third, while there was some modulation of V1 activity by signals related to the animals’ licking, excluding neurons modulated by licking did not change the learning effect (Figures S12G–S12L). Fourth, we found that any signals related to eye position, eye movements, and pupil size could not account for the increased neuronal selectivity after learning (see Supplemental Information and Figures S13A–S13F). Furthermore, we conducted similar analyses to control for any differences in motor behavior during the visual and olfactory discrimination task (Figures S8G–S8J and S13G–S13J), and found that variations in locomotion, licking, eye movements, or pupil size could not explain the task-dependent improvements of neuronal selectivity in V1.

Finally, we trained the linear decoder introduced above on either the population activity of V1 neurons or the running speed of the mouse to predict trial type (vertical versus angled grating corridor; see Experimental Procedures; Figure 5C). Due to the systematic divergence of running speed after mice had entered the grating corridors (see above), the ability of the decoder to classify trials correctly based on running speed strongly improved over learning. However, the decoder trained on V1 activity allowed for earlier classification of the stimulus than the decoder trained on running speed (Figure 5C, top behavioral d-prime bin V1 activity versus running speed at 150 ms, p < 10^{-4}, bootstrap test). Indeed, even the short-latency V1 activity before running speed divergence (typical divergence > 220 ms after stimulus onset) allowed for a significant improvement in grating classification during learning (bottom versus top behavioral d-prime bin at 220 ms, p = 0.001, bootstrap test). Importantly, in post learning sessions (behavioral d-prime > 2), during which mice showed a delayed divergence in their running speeds in response to the rewarded and non-rewarded gratings (running divergence > 400 ms after grating onset), neuronal activity allowed for an equally early and accurate classification of the grating stimuli compared to sessions with matched behavioral d-prime but with earlier running speed divergence (neuronal decoding performance early versus late running divergence: p > 0.1 for all time bins 0–0.5 s from grating onset; Figure 5D). Therefore, learning led to improvements in the ability of V1 populations to discriminate task-relevant stimuli before the animal acted on its decision either to slow down and lick for reward, or to speed up and suppress licking. Taken together, the increase of neuronal selectivity in V1 with training cannot be explained by the modulation of V1 activity by any of the measured motor parameters (running, licking, eye movements, pupil dilation) nor by any differences in optic flow before and after learning.

The Emergence of Signals Reflecting Behavioral Outcome during Learning

The information related to the animal’s own action is not the only non-sensory signal that can influence V1 activity. Other task-related signals relaying information about the attentional state, expectations, or behavioral choice have also been observed in visual cortical areas (Moran and Desimone, 1985; Britten et al., 1996; Shuler and Bear, 2006; Stänjor et al., 2013; Nienborg and Cunningham, 2014). To identify such signals in V1 activity, we compared responses to the non-rewarded, angled grating during correct rejection trials (CR, mouse withheld licking and accelerated) and false-alarm trials (FA, mouse incorrectly licked and slowed down). Because the visual stimulus identity during CR and FA
trials was the same but the behavior of the animal was different (i.e., stopping and licking versus running; see also Figure 6A), we could identify neurons whose responses were not behaviorally modulated (no significant response difference between CR and FA trials despite a strong difference in behavior; Figure 6B) and those that were (significantly different responses between CR and FA trials; Figure 6B). When we excluded all behaviorally modulated cells from the analysis, we still found that the proportion of neurons selective for the rewarded and non-rewarded gratings significantly increased over learning (Figure 6C; all p values < 0.04, bootstrap test), similar to the effects for the entire population (Figure 2D). These results again demonstrate that the improvement in V1 selectivity for both task-relevant stimuli after learning is not caused by signals related to the change in the animals’ behavior during learning, associated changes in optic flow speed, or task-related signals such as reward expectation.

Importantly, however, visually evoked activity of many cells was modulated by the behavioral response (up to 40% of selective neurons; Figure 6B). This difference was apparent at the population level because a decoder trained on predicting the behavioral choice in response to the non-rewarded grating (CR versus FA trials) from neuronal activity of all cells performed above chance and improved with learning (Figure 6D; highest versus lowest behavioral d-prime, p = 0.01, bootstrap test). Interestingly, on average, neurons preferentially responding to the rewarded grating showed significantly different responses between CR and FA trials, while neurons preferring the non-rewarded grating did not (Figure 6E; rewarded-stimulus-preferring cells, p < 10^-4, n = 336; non-rewarded-stimulus-preferring cells, p = 0.31, n = 194, Wilcoxon rank-sum test). Therefore, signals related to the behavioral outcome developed over learning and mainly influenced a specific subgroup of neurons preferring the rewarded stimulus.

The Emergence of Anticipatory Signals during Learning
Analysis of neuronal activity just before the onset of the grating corridors revealed another task-dependent signal that developed during training, presumably related to the animals’ anticipation. While mice started each new trial at a different, random position in the approach corridor, the abrupt onset of the grating corridors was always preceded by the same pattern of black and white circles on the corridor walls (see Figure S7A). Some neurons increased their activity just before grating onset with learning (Figure 7A), suggesting that they had developed anticipatory signals (Jaramillo and Zador, 2011; Totah et al., 2013), which might reflect the animals’ ability to eventually predict and anticipate the time point of appearance (but not the identity) of the grating corridors from the preceding corridor wall pattern.
Importantly, only the neurons preferring the rewarded stimulus, and not the neurons preferring the non-rewarded stimulus, developed this pre-stimulus activity increase during learning (Figures 7B and 7C, pre- versus post learning, rewarded-stimulus-preferring cells: p = 0.001; non-rewarded-stimulus-preferring cells: p = 0.14, bootstrap test). The existence of these specific, putative anticipation signals was supported by a significant decrease in pre-stimulus activity during anesthesia after learning only in cells preferring the rewarded grating (Figure 7C, rewarded-stimulus-preferring cells: p < 10^{-3}; non-rewarded-stimulus-preferring cells: p = 0.06, trend in the opposite direction, bootstrap test). Taken together, non-sensory signals, both before and after appearance of the task-relevant stimuli, seem to influence primarily a specific ensemble of cells that preferentially responded to the stimulus that predicts the reward.

**DISCUSSION**

We show that learning leads to concerted changes in how L2/3 neurons in V1 process visual and non-visual signals related to the behavioral task. By tracking individual neurons during learning, we observed a net recruitment and stabilization of neurons selective for task-relevant stimuli, resulting in improved stimulus discriminability at the population level, which closely correlated with the behavioral performance of the animals. These learning-induced enhancements of stimulus representation in V1 diminished substantially when animals did not engage in the visual discrimination task, suggesting that putative top-down signals contribute to increased population-level discriminability. In parallel, we observed the emergence of additional task-dependent signals in a specific subpopulation of cells—neurons preferentially responding to the rewarded stimulus developed anticipatory responses prior to the appearance of task-relevant stimuli and additional activity related to the animal’s behavioral choice after stimulus onset.

**Learning-Related Changes in Mouse V1**

We developed a visually guided task in which head-fixed mice learned to discriminate two grating patterns in a virtual reality environment in which the animals’ running controlled their position in a corridor (Hölscher et al., 2005; Dombeck et al., 2007). Most mice learned to perform this task with high behavioral accuracy within 1 week (behavioral d-prime > 3, corresponding to accuracy levels of > 90%). We speculate that task acquisition was facilitated by the fact that mice had active control over their visual environment (locomotion coupled to visual feedback), resulting in a more naturalistic visual experience (Gibson, 1979) that seemed to promote engagement in the task. We showed that task performance was dependent on visual cortex activity and, importantly, that responses of V1 neurons to task-relevant stimuli became progressively more distinguishable, leading to more selective task-relevant information in V1 circuits.

The closed-loop nature of behavioral tasks in virtual reality makes it necessary to separate sensory and motor influences on neuronal responses. Specifically, it was important to control for the changes in running speed and the resultant changes in the optic flow speed over the course of training in relation to the observed changes in V1 activity. The learning-related increase in neuronal selectivity did not decrease (1) when comparing responses only in running speed-matched conditions before and after learning, (2) when comparing responses to identical optic flow before and after learning, (3) when excluding neurons from the analysis whose responses were modulated by running and visual flow speed, and (4) when only including neurons with similar responses to the same grating in FA and CR trials even though the animals’ behavior (running speed and licking) and the optic flow differed. Moreover, learning-induced increase in V1 selectivity did not diminish when controlling for licking, eye position, eye movements, and pupil size. Therefore, the improvement in V1 stimulus discriminability during training could not be accounted for by any changes in the animals’ motor behavior we could measure or by associated changes in visual input.

Finally, even though somatic GCaMP6 signals the occurrence of spiking with a slight delay (time to peak for one action potential > ~40 ms; Chen et al., 2013), we found improved discriminability of task-relevant stimuli in V1 within approximately 200 ms after stimulus onset, which preceded the animal’s behavioral response and changes in locomotion. This suggests that learning may increase the salience of information relayed to...
downstream areas to better inform behavioral decisions. Importantly, these results are comparable with those of a recent study of learning-related changes in V1 of macaque monkeys using multiunit recordings (Yan et al., 2014), suggesting that learning exerts similar effects on a primary sensory cortex in rodents and primates.

**Selectivity Changes in Individual Neurons during Learning**

Tracking the activity of the same identified cells throughout learning allowed us to investigate which changes in single cells underlie population-wide improvements in stimulus selectivity. Previous studies in visual cortex have shown differences in orientation tuning at or close to task-relevant grating orientations in animals trained in visually guided tasks compared to control conditions (Schoups et al., 2001; Yang and Maunsell, 2004; Goltstein et al., 2013). These results suggest that increases in population selectivity might have been mainly due to an increase in response selectivity of neurons that already had shown some orientation tuning before learning. However, we did not find that learning-related changes are especially pronounced in or even restricted to neurons with particular visual response properties. Specifically, neurons already selectively responding to one of the two task-relevant grating stimuli before learning were not more likely to increase their selectivity than non-selective neurons during learning.

One change in single cell responses that led to increased stimulus discriminability at the population level was a learning-induced decrease in day-to-day fluctuations of selectivity for task-relevant stimuli in individual neurons, akin to response stabilization observed in the motor cortex (Huber et al., 2012; Peters et al., 2014). Neurons preferring either the rewarded or the non-rewarded stimulus became more likely to maintain their response selectivity across consecutive training sessions. In parallel, we found an increased recruitment of previously non-selective neurons to become selective for the rewarded grating stimulus during training, which may explain the larger proportion of neurons selective for this stimulus in expert mice.

**Task Engagement Enhances Neural Selectivity in V1**

We successfully trained mice to switch between a visual and an olfactory discrimination task several times within the same training session. Mice ignored the grating stimuli during the olfactory discrimination task, and this allowed us to test whether the learning-related enhancement in task-relevant visual stimulus processing was hardwired or task-dependent. Population-level discriminability for grating stimuli was reduced but not decreased to pre learning levels when expert animals were not engaged in the visual discrimination task. Therefore, learning led to both task-independent and task-dependent enhancements in the processing of relevant stimuli in V1. Task-independent changes likely reflect more persistent alterations to visual circuits, akin to those previously observed outside the task or under anesthesia in visual cortex after learning (Schoups et al., 2001; Yang and Maunsell, 2004; Goltstein et al., 2013). The existence of task-dependent changes, however, suggests that non-sensory signals directly contribute to the enhanced processing of behaviorally relevant stimuli (Li et al., 2004, 2008; Polley et al., 2006). Such modulatory signals, which depend on the animals’ behavioral context, could be relayed by excitatory projections of cortical or subcortical origin (Krauzlis et al., 2013; McAlonan et al., 2008; Zhang et al., 2014), or may additionally involve cholinergic input from the basal forebrain (Pinto et al., 2013). Importantly, we found that these signals seem to increase the selectivity of most neurons encoding both the rewarded and non-rewarded stimuli when animals actively engaged in visual discrimination.

**Emergence of Task-Specific Anticipatory and Behavioral-Choice-Related Signals in V1**

Coinciding with the changes in the representations of task-relevant stimuli, we observed the appearance of two additional types of task-dependent signals during learning. First, neurons preferring the rewarded stimulus developed anticipatory responses prior to the appearance of task-relevant stimuli (Jaramillo and Zador, 2011; Totah et al., 2013). These signals are unlikely to be visually evoked, as they are not visible in neurons preferring the vertical grating before learning or under anesthesia. Instead, they likely arise through the learned association between a specific corridor position and the appearance of a grating stimulus, suggesting that processing in V1 is influenced by stimulus expectation, perhaps to prime activity in those neurons whose firing best predicts a reward. These anticipatory signals may thus reflect reward expectation (the rewarded stimulus will appear with 50% likelihood). For example, they could be the neural signature of a type of “wishful thinking” by the animals—stimulus expectation that preferentially evokes the cortical representation of the rewarded and therefore preferred stimulus.

Over the course of training, some neurons also increasingly exhibited enhanced responses during error trials in which the animals incorrectly sought reward in response to angled gratings, suggesting their activity might be related to the animal’s behavioral choice (Britten et al., 1996; Ress and Heeger, 2003; Nienborg and Cumming, 2014), or reward expectation as previously observed in V1 (Shuler and Bear, 2006; Stånigor et al., 2013). Importantly, both the anticipatory and the behavioral choice-related signals emerged predominantly in neurons responding preferentially to the rewarded stimulus. We hypothesize that these signals may arise by strengthening of inputs from areas encoding reward expectation (e.g., orbitofrontal cortex; Tremblay and Schultz, 1999). Activity-dependent Hebbian mechanisms would permit this strengthening to occur specifically on V1 neurons preferring the rewarded stimulus, because these are consistently active before and during the time of reward delivery. With learning, as the animals increasingly develop an expectation of reward (i.e., just before and during the task-relevant stimulus appearance), neurons preferring the rewarded stimulus in V1 would be preferentially activated by projections conveying these putative top-down signals. This mechanism may act in concert with cholinergic signaling that has been proposed to explain reward timing-related plasticity in V1 (Chubykin et al., 2013).

The appearance of non-sensory signals in neuronal ensembles preferring the stimulus associated with a reward contrasts
with the modulation of sensory stimulus responses when mice were engaged in the visual discrimination task, which acted more generally by increasing the selectivity of neurons encoding both the rewarded and the non-rewarded stimuli. Identifying the sources of these diverse task-dependent signals is an important next step for clarifying their role in shaping early sensory processing. The sophisticated genetic tools available in mice will help elucidate the role of the many cortical and subcortical areas providing input to V1 during learned behaviors, as well as specific inhibitory cell types or different neuro-modulator systems in the emergence and expression of learning-related changes.

In summary, as a mouse learns the behavioral significance of a visual stimulus, the responses of L2/3 neurons in V1 become more selective for task-relevant stimuli, leading to enhanced stimulus discriminability at the population level. In parallel, multiple task-dependent signals emerge during learning and differentially influence the firing of neurons within the V1 circuit. This demonstrates the remarkable flexibility by which a primary sensory cortex can tailor its processing to the requirements of a task and to the behavioral relevance of sensory stimuli.

**EXPERIMENTAL PROCEDURES**

**Surgical Procedures and Imaging**

All experimental procedures were carried out in accordance with the institutional animal welfare guidelines and licensed by the UK Home Office and the Swiss cantonal veterinary office. A virus expressing GCaMP6f or GCaMP6m (AAV2/1-hsyn-GCaMP6-WPRE; Chen et al., 2013) was injected in the primary visual cortex (V1) in the right hemisphere of C57Bl/6J mice (P49–P57). Imaging and behavioral training started approximately 3 weeks after surgery. We imaged GCaMP6-labeled neurons in layer 2/3 in 93 training sessions and 12 recording sessions under isoflurane anesthesia in 11 mice with a custom-built resonant scanning two-photon microscope with a frame rate of 32 Hz. *Supplemental Experimental Procedures* contain further details about surgical and imaging procedures.

**Behavioral Tasks**

Mice were head-fixed and trained to run on a styrofoam cylinder. A reward delivery spout was positioned near the snout of the mouse, and licks were detected using a piezo disc sensor. Mice were then trained in a visual discrimination task in which the running speed on the cylinder was detected with an optical mouse and used to control the speed at which mice moved through a virtual environment presented on two screens in front of them. A trial started when the mouse was positioned at a random starting point in an approach corridor with walls showing black and white circles on a gray background. When the mouse reached a specific point in the corridor, it was randomly teleported to one of two grating corridors with either a vertical or an angled grating on the walls. In the vertical grating corridor, the mouse was rewarded with a drop of soya milk, for licking the spout after it had entered a “reward zone,” a short distance into the grating corridor. No punishment was given for licking in the angled grating corridor.

A subset of mice was trained to switch between blocks of an olfactory and visual discrimination task. In olfactory blocks, mice were rewarded with one of two odors. During this task, mice were also presented with the vertical and angled grating corridor at different positions in the approach corridor. Mice learnt to ignore these irrelevant grating stimuli while accurately discriminating the odors. On switching to the visual block, mice started licking selectively to the rewarded grating as before. See *Supplemental Experimental Procedures* for further details about the visual stimulus, behavioral tasks, and training.

**Bilateral Optogenetic Silencing of V1 Activity**

Bilateral silencing of V1 was carried out in four transgenic mice (three males, one female) expressing channelrhodopsin-2 in parvalbumin-expressing interneurons (Hippenmeyer et al., 2005; Madisen et al., 2012). Additionally, three male wild-type C57Bl/6J mice underwent identical surgical and experimental procedures. Mice were implanted with two cranial windows over both visual cortices. Intrinsic imaging was used to determine the extent of V1, and all regions excluding V1 were covered with black paint. In expert mice (>90% performance levels), V1 was silenced by illuminating both cranial windows with 470 nm light at one of four intensities shortly before and during the grating corridor. In 30% of trials no light stimulation was applied. The same mice were also trained on an olfactory discrimination task as described above (but without grating stimuli), V1 was silenced shortly before and during presentation of the odors. For further details, see *Supplemental Experimental Procedures*.

**Data Analysis**

Image stacks were corrected for motion, and regions of interest (ROIs) were selected for each cell in each session. Raw fluorescence time series $F(t)$ were obtained for each cell by averaging across pixels within each ROI. Baseline fluorescence $F_0(t)$ was computed by smoothing $F(t)$ (causal moving average of 0.75 s) and determining for each time point the minimum value in the preceding 60 s time window. The change in fluorescence relative to baseline, $\Delta F/F$, was computed by taking the difference between $F$ and $F_0$, and dividing by $F_0$.

To analyze responses to the vertical and angled grating corridors, neuronal activity was aligned to the onset of the grating corridor for each trial. A Wilcoxon rank-sum test was used to determine if responses—the average $\Delta F/F$ in a time window of 1 s after grating onset—in the two conditions were significantly different ($p < 0.05$), and the sign of the difference determined the response preference. The persistence of stimulus preference (Figure 2F) was defined as the probability that a cell that significantly preferred one of the two gratings on one day also preferred the same grating on the next day. Recruitment of non-selective cells (Figure 2G) was defined as the probability that a cell with no stimulus preference on one day became selective to one of the two gratings on the next day. We computed these measures for three stages of learning, based on the behavioral d-prime (bDP) of two consecutive sessions: before learning (bDP of both sessions < 1, and $\Delta$ bDP < 0.5, Nsession = 14), during learning (bDP of first session < 2, bDP second session > 0.5, and $\Delta$ bDP > 0.5, Nsession = 14), and after learning (both bDP > 2 and absolute change in bDP < 0.5, Nsession = 19). Varying the criteria to define different stages of learning led to similar results (data not shown).

To quantify the selectivity of neural responses we computed a response selectivity index (SI) for individual cells from the difference between the mean response in the first second after grating onset to the vertical and angled grating corridor, divided by the pooled standard deviation of the responses

$$SI = \frac{(R_v - R_a)}{SIAv}$$

where

$$SIAv = \frac{\sum n_i (n_i - 1)}{\sum n_i}$$

and $n_i$ is the number of trials in condition $i$ for $k$ conditions. Therefore, positive values indicate a preference for the vertical grating corridor and negative values a preference for the angled grating corridor. Please note that in the manuscript text the term selectivity substitutes for SI. To obtain a combined measure of grating discriminability for simultaneously imaged populations of neurons, population selectivity was computed by taking the average of the squared selectivity index across cells and taking the square root:

$$\sqrt{\frac{\sum_{i=1}^{k} SI^2}{Ncell}}$$

A bootstrap test (Efron, 1979) was used to test for significant differences between conditions that contained both dependent and independent data points. To test whether changes in the proportion of cells preferring the vertical or angled grating, or without preference across two conditions (typically before and after learning), were significant, we first computed for each session the
proportions of cells in each category. Next, we randomly picked the same number of sessions (the minimum across conditions) from both conditions, and repeated this 10,000 times. We then computed in both conditions the average cell proportion across sessions, and we also computed the proportion after randomly assigning sessions to one of the two conditions. The p value was given by the number of bootstraps in which the proportion change in the actual data was greater than the proportion change with randomly assigned condition labels. Similarly, bootstrapping was also used to assign significance to the differences in population selectivity, decoding performance, and pre-stimulus activity increase, by comparing the difference in the original data to the difference with randomly assigned condition labels.

To control for the effect of running speed and optic flow on neural responses and selectivity across learning, grating responses were compared specifically in trials that were matched for running speed across sessions and stimulus conditions (Figures 5A and 5B). First, the average running speed was determined in sliding 200 ms time windows from −0.5 to +0.5 s around the onset of the grating corridor (50 ms step size). Then responses in each time window of each trial were assigned to one of three groups, depending on running speed (three bins divided equally from the 2.5% percentile to the 97.5% percentile of the average running speed, across all sessions). Data for each time window were only included if it contained at least ten trials of both grating conditions. In the highest speed bin, not enough matched data were available across learning, thus restricting the analyses to the lowest speed bin (referred to as “slow”) and the intermediate speed bin (“fast”).

To quantify the accuracy with which two conditions (either trials with vertical and angled grating corridors (Figures 3B, 5C, and 5D) or FA and CR trials (Figure 5D) could be classified at time t relative to grating onset, a cumulative decoder was employed. From training data (30 trials of both conditions), the decoder constructed for each neuron n a model of the response using as parameters the mean response to the vertical (µ̂VA(t)) and angled grating corridor (µ̂BA(t)) and the variance of the noise σ̂n, to maximize the observed log-likelihood of the data under a Gaussian noise model. On test trials (the remaining trials that were not used as training data), the log-likelihood at time t that trial k belongs to condition C (where C was for instance the vertical (V) or angled grating corridor (A) condition) is proportional to

\[ L_C(t) = -\sum_{D_{test}} \sum_{C_0} (D_{test}(t - T_{start})) - \mu^C_{ VA}(t - T_{start})^2 / (2\sigma^2_n), \]

where \( D \) indicates deconvolved \( \Delta F/F \) (see Supplemental Experimental Procedures). If \( L_V > L_A \), the trial was assigned to the vertical condition, otherwise to the angled grating condition. To obtain at each time point t the cumulative likelihood \( L_C \), the summation only included time points starting from \( T_{start} \), which was the time of the grating onset, up until time t. Note that without the temporal accumulation of log-likelihood, the decoder would be equivalent to a linear discriminant analysis. To determine the time point at which there was a detectable divergence of running speed between vertical and angled grating trials, we performed a Wilcoxon rank-sum test on the average d-prime over 400 ms (n = 8 sessions in n = 7 mice, average d-prime 2.59). We paired each of these sessions to a unique session with the smallest difference in behavioral d-prime, but with time of divergence < 400 ms (n = 8 sessions in n = 6 mice, average d-prime 2.61).

To analyze responses during FA and CR trials, only sessions with at least 15 FA trials were included in the analysis (Figure 6). These were predominantly sessions at intermediate learning stages, as most expert mice made very few mistakes by the end of training (see Figure S1). Behaviorally modulated cells were defined as cells with significantly different activity for FA and CR trials in the first second after grating corridor onset (p < 0.05, Wilcoxon rank-sum tests). To obtain average responses for cells preferring the vertical or the angled grating corridor (Figure 6E), neurons were classified as vertical (or angled) preferring if they significantly preferred the vertical (or the angled) grating corridor in at least one session and never switched preference, and responses of such cells were averaged across the sessions in which they showed a significant preference.

The relative response increase before grating onset (Figure 7C) was calculated for each cell as the difference in the average \( \Delta F/F \) signal between two time windows, −0.25 s to 0 s and −1 s to −0.75 s, divided by the average \( \Delta F/F \) signal in the −1 to −0.75 s window, where t = 0 s is time of grating onset. To compare pre-stimulus responses before and after learning, responses were averaged on the first day of training and the first day post learning (behavioral d-prime > 2) for each cell. Neurons were classified as vertical (or angled) preferring if they significantly preferred the vertical (or the angled) grating corridor (p < 0.05, Wilcoxon rank-sum test).

SUPPLEMENTAL INFORMATION

Supplemental Information includes 13 figures, one movie, and Supplemental Experimental Procedures and can be found with this article at http://dx.doi.org/10.1016/j.neuron.2015.05.037.

AUTHOR CONTRIBUTIONS


ACKNOWLEDGMENTS

We thank Petr Znamenskiy, Pieter Roelfsema, Daniel Huber, Jerry Chen, and Frijof Heimchen for comments on earlier versions of the manuscript, and all members of the lab for comments and discussions. We thank John O’Keefe for support throughout the project, and Dinu Florin Albeanu for help with the intrinsic imaging setup. We thank the GENIE Program and Janelia Farm Research Campus of the Howard Hughes Medical Institute for making GCaMP6S material available, and the UCL Mechanical Workshop for building custom parts for the microscope and experimental setup. This work was supported by the Wellcome Trust (S.B.H., 095853, and T.D.M.-F., 095074, J.K., 50096 and M.B., 50096), the European Research Council (S.B.H., HigherVision 337797, T.D.M.-F., NeuroV1sion 616509), the Marie Curie Actions of the European Union’s FP7 program (J.P., 332141 and A.G.K., 301742), the Gatsby Charitable Foundation (M.P. and M.S.), and the Biozentrum core funds (University of Basel).

Received: December 8, 2014
Revised: March 7, 2015
Accepted: April 29, 2015
Published: June 4, 2015

REFERENCES


AUTHOR CONTRIBUTIONS

We thank Petr Znamenskiy, Pieter Roelfsema, Daniel Huber, Jerry Chen, and Frijof Heimchen for comments on earlier versions of the manuscript, and all members of the lab for comments and discussions. We thank John O’Keefe for support throughout the project, and Dinu Florin Albeanu for help with the intrinsic imaging setup. We thank the GENIE Program and Janelia Farm Research Campus of the Howard Hughes Medical Institute for making GCaMP6S material available, and the UCL Mechanical Workshop for building custom parts for the microscope and experimental setup. This work was supported by the Wellcome Trust (S.B.H., 095853, and T.D.M.-F., 095074, J.K., 50096 and M.B., 50096), the European Research Council (S.B.H., HigherVision 337797, T.D.M.-F., NeuroV1sion 616509), the Marie Curie Actions of the European Union’s FP7 program (J.P., 332141 and A.G.K., 301742), the Gatsby Charitable Foundation (M.P. and M.S.), and the Biozentrum core funds (University of Basel).

Received: December 8, 2014
Revised: March 7, 2015
Accepted: April 29, 2015
Published: June 4, 2015

REFERENCES


Supplemental Information

Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex

INVENTORY OF SUPPLEMENTAL INFORMATION:

Supplemental Figures S1 – S13

- **Figure S1.** Task performance of individual mice, related to Figure 1.
- **Figure S2.** Grating responses of all cells before and after learning, related to Figure 2.
- **Figure S3.** Changes in grating responses of individual neurons during learning, related to Figure 2.
- **Figure S4.** Neuronal changes during learning in individual mice, related to Figure 2.
- **Figure S5.** Learning-effect in cells with a response increase in grating corridors, related to Figure 2.
- **Figure S6.** Relationship between behavioral discrimination performance and neuronal population selectivity for individual mice, related to Figure 3.
- **Figure S7.** Neuronal responses during the approach corridor, related to Figure 3.
- **Figure S8.** Switching between visual and olfactory discrimination task: behavior and controls, related to Figure 4.
- **Figure S9.** Amplitudes of grating responses, related to Figure 4.
- **Figure S10:** Running and licking profiles of example mice during the visual discrimination task, related to Figure 5.
- **Figure S11.** Visual discrimination task with fixed visual flow, related to Figure 5.
- **Figure S12.** No effect of neuronal signals related to the animals’ running and licking behavior on the learning-related increase in neuronal selectivity, related to Figure 5.
- **Figure S13.** Effect of eye position, eye movement and pupil size, related to Figures 5, 7.

Supplemental Movie S1

Supplemental Experimental Procedures
**Figure S1: Task performance of individual mice**

Percentage of correct vertical trials (hit trials, blue) and correct angled trials (correct rejections, red) and behavioral d-prime (black) for all imaged mice and all imaging sessions. Mice M1 to M5 were imaged daily throughout learning, M6 to M8 were imaged before and after learning (separated by 6 training sessions without imaging), and M9 to M11 were imaged daily after learning (preceded by 6 training sessions without imaging).
Figure S2: Grating responses of all cells before and after learning

Average responses of individual cells to vertical (rewarded) and angled (non-rewarded) grating trials from the first training session (before learning, left) and the first training sessions after learning (right, behavioral d-prime > 2). Responses are aligned to grating corridor onset, baseline corrected (change in $\Delta F/F$, see Experimental Procedures), grouped by grating preference and ordered by the mean response 0 to 1 s after grating onset.
Figure S3 (related to Figure 2)

(A-B) Relationship between pre-learning (behavioral d-prime < 1) and post-learning (behavioral d-prime > 2) stimulus response to the vertical (A) or angled (B) grating. Each cell’s response is the difference in response during the second before and the second after grating onset divided by the pooled standard deviation (positive values indicate an increase in activity and negative values a decrease in activity after grating onset). Neurons with a significant response increase (I) or decrease (D) after grating onset (P < 0.05) mostly retained their stimulus response type (I: 75% of cells, D: 77% of cells) and only rarely completely changed their response type (I to D: 7%, D to I: 4%). (C) Proportions of cells with significant changes in response amplitude to the two gratings over learning (Wilcoxon rank-sum test, P < 0.05), for all neurons that increased their selectivity during learning (N = 308). Of all neurons that showed increased grating selectivity after learning, 35% increased their response to the preferred (pref) as well as to the non-preferred (non-pref) grating, 32% decreased their response to both gratings, and for 24% response amplitude changes did not reach significance. Only 8% increased their response to the preferred and decreased their response to the non-preferred grating, and less than 1% of cells showed other combinations of response strength changes such as an increased response to the preferred but no change in the response to the non-preferred grating. (D) Selectivity of the same neurons before and after learning (positive values indicate preference for vertical, negative values preference for angled gratings, see Experimental Procedures). (E) Relationship between the change in neuronal selectivity across learning and selectivity before learning. Values on x- and y-axis are calculated from different subsets of trials to prevent interdependencies. (F) Relationship between the change in selectivity across learning and the average pre-learning response to vertical and angled gratings. For each cell’s average grating response, positive values indicate an increase in activity and negative values a decrease in activity after grating onset.
Figure S4: Neuronal changes during learning in individual mice

(A) Proportions of neurons significantly preferring the vertical (Pref V) or the angled grating (Pref A), and those without preference (No Pref), before (sessions with behavioral d-prime < 1) and after learning (behavioral d-prime > 2) for individual mice; session mean ± SEM computed from responses 0 – 1 s after grating onset. (B-E) Average values for individual mice for neuronal selectivity (B), absolute difference in response amplitude to the vertical and angled grating (C), standard deviation of the grating response amplitudes over all trials in a session (D, pooled for vertical and angled grating), and mean response amplitudes to the gratings (E, pooled for vertical and angled grating); session mean ± SEM computed from responses 0 – 1 s after grating onset.
Figure S5: Learning-effect in cells with a response increase in grating corridors.

For all panels in this figure, only neurons showing a significant increase in their calcium signal after the onset of at least one of the grating corridors (P<0.01, Wilcoxon signed-rank test) have been included in the analysis. (A) Average responses of individual cells to vertical and angled grating trials from the first training session (before learning, left) and the first training sessions after learning (right, behavioral d-prime > 2). Responses are aligned to grating onset, baseline corrected (change in ∆F/F, see Experimental Procedures), grouped by grating preference, and ordered by the mean response 0 to 1 s after grating onset. (B) Histograms of single-cell selectivity (computed from responses 0 – 1 s after grating onset) for different behavioral discrimination performance levels during learning. Colors denote bins of behavioral d-prime, ranging from chance performance (blue) to expert performance (orange). (C) Proportions of neurons significantly preferring the vertical or the angled grating, and those without preference, before (sessions with behavioral d-prime < 1) and after learning (behavioral d-prime > 2) computed from responses 0 – 1 s after grating onset. Error bars depict SEM (determined by bootstrapping with replacement). N = 11 mice.
Figure S6: Relationship between behavioral discrimination performance and neuronal population selectivity for individual mice

Top: behavioral d-prime over learning for three example mice. Bottom: time course of neuronal population selectivity aligned to grating onset for the training sessions shown in the top panel. Mouse M7 was trained without two-photon calcium imaging for 6 sessions between imaging sessions 4 and 5.
Figure S7: Neuronal responses during the approach corridor

(A) Image of the approach corridor in the virtual environment. (B-D) Responses during the approach corridor of three example neurons before (left) and after learning (right). Top panel shows responses during single trials as a function of corridor position, bottom shows average across trials. Grating corridors start at position 0. Shading is SEM. (E) Average neuronal selectivity for approach corridor features (left, see Experimental Procedures) and the grating corridors (right) for individual mice before and after learning. Thick symbols and lines are mean values for all mice. Selectivity is enhanced for gratings after learning (Wilcoxon signed-rank test, P = 0.0078), but not for approach corridor features (P = 0.25). N = 8 mice. (F) Average peak response during the approach corridor for individual mice before and after learning. Thick symbols and lines are mean values for all mice. Response amplitude to features of the approach corridor does not significantly change during learning (P = 0.31).
Figure S8: Switching between visual and olfactory discrimination task: behavior and controls

(A-C) Lick rasters as in Figure 1C for an example mouse. Licks (black dots) are aligned to onset of the vertical grating (left, blue shading) and angled grating (right, pink shading) during the visual (A) and the olfactory task (C), or are shown aligned to the onset of the rewarded odor 1 (green shading) and unrewarded odor 2 (grey shading, B). (D,E) Running speed profiles aligned to grating onsets (D) and odor onsets (E) for the same mouse. (F) Average running speed aligned to the onset of the rewarded odor 1 (green shading) and unrewarded odor 2 (grey shading, B). (G) Left: correlation of ΔF/F difference in neuronal grating selectivity and running speed (coefficient of determination R²) and the difference in neuronal grating selectivity in the visual and olfactory block. Right: time course of population selectivity aligned to grating onset, as in Figure 4D, but excluding the top 50% of cells most strongly modulated by running. (H) Same as in G, but for ΔF/F variance explained by running (R²) as determined by a generalized linear model (GLM, see Suppl. Experimental Procedures). (I) Same as in G, but for the correlation between ΔF/F and licking rate. (J) Same as in H, but for variance explained by licking as determined by the GLM. Shading is SEM.
Figure S9: Amplitudes of grating responses

Average amplitudes of the calcium signal 0 - 1 s after grating onset of all cells in response to the vertical (V) and angled grating (A) before learning (blue bars, behavioral d-prime < 1, 27 sessions, 11 mice), in visual discrimination blocks after learning (orange bars, 11 sessions, 4 mice), in olfactory discrimination blocks in the same session (black bars, 11 sessions, 4 mice), and when the same visual stimuli were played back to fully trained, anesthetised mice (grey bars, 12 sessions, 11 mice). From data shown in Figure 4D. Grating responses were slightly but not significantly stronger before learning compared to the sessions in which mice switched between the visual and olfactory discrimination tasks (P > 0.10). Responses were similar during the visual and olfactory task in trained mice (P > 0.32) but weaker during anesthesia (P < 10^{-4}).
Figure S10: Running and licking profiles of example mice during the visual discrimination task

(A,B) Licking raster plots and average running profiles of two additional example mice during three training sessions, in the same format as Figure 1C. Licks (black dots) aligned to grating onset in vertical grating (left, blue shading) and angled grating (middle, pink shading) trials. Red dots: reward delivery, yellow dots: licking after reward delivery. Right, average running speed for sessions shown on left, aligned to grating onset for vertical (blue) and angled (red) trials. Shading, SEM.
Figure S11: Visual discrimination task with fixed visual flow

(A) Schematic of the modified task design: the grey approach corridor was of random length, drawn from an exponential distribution (flat hazard rate). After mice had learnt the task, they were presented with gratings uncoupled from running speed with identical speed profiles as observed by them before learning, trial by trial, 0 - 1 s from grating onset. Otherwise the visual flow was coupled to running speed. (B) Similar to Figure 2D. Proportions of cells preferring one of the gratings increased with learning when visual stimuli and their speed were identical before and after learning; computed from 0-1 s from grating onset; bootstrap test, \( N = 868 \) cells pre learning, 825 cells post learning, 3 mice. (C) Schematic of the modified switching task design for visual and olfactory discrimination blocks. Similar to A, but gratings were always presented with a fixed, identical speed profile (average of all conditions from the preceding training session) for 0 - 1 s from grating onset. In olfactory blocks, one of two odors was presented with a random delay after the gratings. (D) Average running speed profile of all mice during the tasks described in C, aligned to grating onset. \( N = 4 \) mice. (E) Visual (black) and olfactory (green) discrimination performance for individual mice, as in Figure 4A, showing that mice successfully switched their behavior between blocks. (F) Similar to Figure 4C. Proportions of cells preferring one of the gratings decreased in the olfactory blocks, while visual stimuli and their speed were identical in the two conditions. Bootstrap test, \( N = 1075 \) cells, 4 mice. (G) Grating selectivity of individual neurons during the visual and the olfactory task with fixed visual flow. The majority of neurons reduced their selectivity during the olfactory block, as in Figure 4B.
Figure S12: No effect of neuronal signals related to the animals’ running and licking behavior on the learning-related increase in neural selectivity

(A) Relationship between the correlation of $\Delta F/F$ and running speed (coefficient of determination $R^2$) and the change in neuronal selectivity during learning for individual cells before (black) and after (grey) learning. (B,C) Time course of population selectivity aligned to grating onset for different behavioral performance levels, as in Figure 3A, but excluding neurons strongly modulated by running. (D-F) Same as in A-C, but for $\Delta F/F$ variance explained by running ($R^2$) as determined by a generalized linear model (GLM, see Experimental Procedures). (G-I) Same as in A-C, but for the correlation between $\Delta F/F$ and licking rate. (J-L) Same as in D-F, but for variance explained by licking ($R^2$) as determined by the GLM. The $R^2$ threshold for exclusion was chosen to exclude either the top 25% or the top 50% cells most modulated by running or licking.

Figure S12: No effect of neuronal signals related to the animals’ running and licking behavior on the learning-related increase in neural selectivity

(A) Relationship between the correlation of $\Delta F/F$ and running speed (coefficient of determination $R^2$) and the change in neuronal selectivity during learning for individual cells before (black) and after (grey) learning. (B,C) Time course of population selectivity aligned to grating onset for different behavioral performance levels, as in Figure 3A, but excluding neurons strongly modulated by running. (D-F) Same as in A-C, but for $\Delta F/F$ variance explained by running ($R^2$) as determined by a generalized linear model (GLM, see Experimental Procedures). (G-I) Same as in A-C, but for the correlation between $\Delta F/F$ and licking rate. (J-L) Same as in D-F, but for variance explained by licking ($R^2$) as determined by the GLM. The $R^2$ threshold for exclusion was chosen to exclude either the top 25% or the top 50% cells most modulated by running or licking.
Figure S13 (related to Figures 5 and 7)

(A) Sample image of the eye contralateral to the imaged hemisphere. (B) Distributions of average horizontal (nasal-temporal axis) and vertical (ventral-dorsal axis) eye positions and pupil sizes of the contralateral eye in individual trials 0-1 s after onset of the vertical (V) and angled (A) grating, before (pre, behavioral d-prime <1) and after (post, behavioral d-prime >2) learning. (C) Saccade rate 0-1 s after grating onset before and after learning. (D-F) Proportions of neurons significantly preferring the vertical grating, the angled grating or without preference 0-1 s after grating onset, after equalizing the distributions of horizontal and vertical eye positions in all conditions (D), after equalizing the distributions of pupil sizes (E), and when excluding all trials with eye movements (F). (G-J) Similar to B-F but during sessions in which mice switched between a visual discrimination task and an olfactory discrimination task during which the same gratings were shown but were irrelevant. (K) Pre-stimulus activity increase similar to Figure 7C, but after excluding all trials with eye movements. Error bars are SEM.
Movie S1: Mouse performing the visual and olfactory discrimination task
(related to Figures 1 and 4)

Movie of a well-trained mouse performing the visual discrimination task in a virtual corridor as illustrated in Figure 1. The mouse is rewarded with a drop of soy milk when it licks a spout only when encountering vertical but not angled gratings on the corridor walls. In the second part of the movie, the same expert mouse is performing an olfactory discrimination task while ignoring the grating patterns on the corridor wall during the same recording session (see also Figure 4 and Experimental Procedures). The tube visible in the movie that is directed to the nose of the mouse was used to deliver odors. Moving patterns on the monitors displaying the virtual reality appear distorted only due to low video recording quality.
SUPPLEMENTAL EXPERIMENTAL PROCEDURES

1. Animals and surgical procedures
2. Two-photon calcium imaging
3. Behavioral training and visual discrimination task
4. Olfactory discrimination and switching task
5. Bilateral optogenetic silencing of V1 activity
6. Imaging of intrinsic signals
7. Decoder analysis
8. Selectivity index for approach corridor features
9. Locomotion control analyses and GLM
10. Analysis of eye position and movements and pupil size during visual discrimination task
11. Visual discrimination task with fixed visual flow

1. Animals and surgical procedures

All experimental procedures were carried out in accordance with institutional animal welfare guidelines and were licensed by the UK Home Office and the Swiss cantonal veterinary office. Experiments were performed on male C57Bl/6J mice. Mice aged between postnatal days 49-57 were anaesthetized with a mixture of Fentanyl (0.05 mg kg⁻¹), Midazolam (5 mg kg⁻¹), and Medetomidin (0.5 mg kg⁻¹). Additional drugs provided analgesia (Rimadyl 6 mg/kg), anti-inflammatory effects (dexamethasone, 1.6 mg/kg), and reduced mucus secretions (atropine, 0.04 mg kg⁻¹). Eye cream (Maxitrol) was applied to the eyes, and anesthesia was maintained by re-injecting one-sixth of the initial dose approximately every hour. A circular piece of scalp was removed and the underlying skull was cleaned and dried. A circular craniotomy (4 mm diameter) was made over the right primary visual cortex (V1) located using stereotaxic coordinates. A virus expressing either GCaMP6f or GCaMP6m (AAV2/1-hsyn-GCaMP6-WPRE (Chen et al., 2013), vector core, University of Pennsylvania Gene Therapy Program) was injected in 1-3 sites using glass pipettes and a pressure micro-injection system (Picospritzer III, Parker). Each recorded site was later confirmed to be in monocular V1 by intrinsic signal imaging and by identifying the approximate receptive field position of imaged neurons (see below). The craniotomy was then sealed with a glass cover slip and cyano-acrylic glue (Loctite) and a custom machined aluminum head-plate was cemented onto the skull using
dental cement (C&B Superbond). Antibiotic (Baytril 0.4 mg kg\(^{-1}\)) and analgesia (buprenorphine 0.08 mg kg\(^{-1}\)) was given at the end of surgery and repeatedly during recovery of the animal. Imaging and behavioral training started approximately three weeks after surgery.

2. Two-photon calcium imaging

Imaging was performed using a custom-built resonant scanning two-photon microscope and a Spectra Physics MaiTäi eHP DeepSee laser (< 70 fs pulse width, 80 MHz repetition rate) at 920 nm using a Nikon 16x 0.8 NA objective. Images of 750 by 750 pixels and a field of view of 375 × 375 µm were acquired using a 12 kHz resonant scanner (Cambridge Technology) and an FPGA module (PXIe-7965R FlexRIO, National Instruments), yielding an effective imaging frame rate of 32 Hz. The microscope was controlled using custom software written in Labview (National Instruments). Imaging and behavioral training started approximately three weeks after surgery, and two-photon calcium imaging of GCaMP6-labelled neurons in layer 2/3 (typically 150-200 um below the cortical surface) was performed during 104 training sessions in 15 mice (including controls).

In 5 mice, imaging was carried out during all training sessions (on average 12 sessions per mouse), in 3 other mice imaging was performed before and after learning (on average 8 imaging sessions: typically 4 imaging sessions before learning, 6 behavior sessions with no imaging, 4 imaging sessions after learning), and 3 mice were imaged only after learning (6 training sessions before imaging, on average 4 imaging sessions after learning). Before each recording session the same imaging site was found by matching anatomical landmarks. Mice with bone re-growth under the window, poor viral expression or a large number of brightly-labelled cells with nuclear GCaMP6 expression were excluded from the study.

We determined the approximate receptive field positions of the recorded neurons at the end of the first and the last training session, while mice ran freely on the cylinder. The monitor in front of the contralateral eye (covering ~ 100x60 degrees of visual space) was divided into a 4x3 grid and rectangles alternating between black and white at 2 Hz were presented at each grid position on a grey background in randomized order (10 repetitions). Stimuli were generated using Psychtoolbox-3 in Matlab (Mathworks). Only cortical regions showing clearly defined visual responses in the monocular visual field in at least a subset of neurons before and after learning were included in the analysis.

To compare neuronal selectivity and pre-stimulus activity during the task and during anesthesia in the same neurons, one day after the last behavioral imaging session mice were
initially sedated (Chlorprothixene, 0.7 mg/kg) then lightly anesthetized with isoflurane (0.5-1% in O₂) delivered via a nose cone. Eyes were covered with eye ointment (Maxitrol), reduced to a thin layer during imaging. Atropine (0.1 mg/kg) was given to slightly dilate the pupil and reduce mucus secretions. Mice were kept at 37°C on a heating pad (DC Temperature Controller, FHC). We used the visual stimulus produced by the mouse’s interaction with the virtual environment in the preceding training session and played it back to the animal during anesthesia. Data during anesthesia were collected for 11 animals, one of which contributed two recording sessions.

Image stacks were first corrected for motion by maximizing the cross-correlation of all frames with a reference image (Guizar-Sicairos et al., 2008). The reference image was an average of 30 successive frames (~1s) and was selected by computing for 30 candidate reference images (at linearly spaced positions in the recording stack) the x- and y-shift to 100 single frames (also linearly spaced) to obtain the most typical reference image with the smallest average shift. Regions of interest (ROIs) were selected by manually assigning pixels to individual cells by inspecting individual frames, as well as the average and maximum projections of the imaging stacks. Frames with motion shifts exceeding 5 microns (Euclidean distance) relative to neighboring frames were rejected (on average 0.06% of all frames, range 0 to 0.9%).

3. Behavioral training and visual discrimination task

Mice were trained in a visual discrimination task in a virtual corridor over the course of one to two weeks. Mice had free access to water, but were food deprived to maintain at least 85% of their free-feeding body weight (typically 85-90%, 2-3 g of standard food pellets per animal per day). Animals became accustomed to handling and gentle restraint over two to three days, before they were head-fixed and trained to run on a styrofoam cylinder (20 cm diameter, on a ball-bearing mounted axis) in the dark for two to four days. This period was also used to find suitable imaging sites. A reward delivery spout was positioned near the snout of the mouse, and licks were detected using a piezo disc sensor and custom electronics. The reward was a 10% solution of soya infant milk powder (SMA Wysoy) delivered by opening a pinch valve (NResearch) controlled through custom electronics. We recorded images of both mouse eyes with CMOS cameras (Imaging Source, 30 Hz frame rate) and tracked the pupil position and diameter offline (Sakatani and Isa, 2004). The mice’s running speed on the cylinder was detected with an optical mouse (Logitech G700, USB polling rate of 125 to 1000 Hz). This signal was used to control the speed at which mice moved through a virtual environment that
was presented on two screens in front of them. The optical mouse was calibrated with an incremental rotary encoder (Kübler) which gave 100 digital pulses for every full rotation of the cylinder. The virtual environment consisted of linear corridors created in a game engine (Unity), and the position in the environment was controlled by custom software written in Labview (National Instruments). The luminance of visual stimuli was corrected using a luminance meter (Konica Minolta, LS-100).

A trial started when the mouse was positioned at a random starting point (chosen from a uniform distribution) in an approach corridor with walls showing black (0.1cd m⁻²) and white (10.9cd m⁻²) circles on a grey background (luminance 5.5cd m⁻²). Over the course of learning the length of the approach corridor was usually slightly increased (length typically varied from 185 to 370 cm). When the mouse had reached a specific point in the corridor, it was randomly teleported to one of two grating corridors (length 111 cm) with either a vertical grating pattern (square wave gratings, 100% contrast) or an angled grating pattern (rotated 40° relative to vertical) on the walls. The spatial frequency of the gratings in the center of the screen was ~ 0.05 cycles per degree, and temporal frequency at the average running speed ~ 4 Hz. In the vertical grating corridor, the mouse could trigger the delivery of a reward, a drop of soy milk, by licking the spout after it had entered a ‘reward zone’ a short distance (55.5 cm) into the grating corridor (mice typically started licking in anticipation of the reward zone). This was considered a 'hit' trial. If an animal did not lick by the end of the reward zone, shortly before the end of the vertical grating corridor (92.5 cm from the start), it was given a drop of reward by default and this was considered a 'miss' trial (this rarely happened after the first few sessions). In the angled grating corridor, the mouse did not receive a reward, and a single lick or more in this corridor was considered a 'false alarm' (FA) trial. No punishment was given. Animals typically learned this task in 3 to 6 days eventually discriminating with > 90% accuracy by selectively licking in the vertical corridor, and running without licking through the angled corridor ('correct rejection' (CR) trial). During initial stages of training the probability of non-rewarded trials was occasionally increased to discourage unrewarded licking (probability of angled grating trials 50-70%). We quantified the performance of the mouse using a behavioral d-prime: \( bd' = \Phi^{-1}(H) - \Phi^{-1}(F) \), where \( \Phi^{-1} \) is the normal inverse cumulative distribution function, H is the rate of hit trials and F is the rate of false alarm trials.

4. Olfactory discrimination and switching task

A subset of mice was trained in an olfactory discrimination task after reaching high levels of performance in the visual discrimination task. While running through the same virtual reality
environment as in the visual discrimination task, mice were presented with one of two odor stimuli at a fixed location in the approach corridor (after 185 to 370 cm): 10% soya milk odor (go stimulus, licking was rewarded with a drop of soya milk) or 10% soya milk with 0.1% limonene mixture odor (no-go stimulus, no reward for licking). Odors were delivered through a flow dilution olfactometer calibrated with a mini PID (Aurora) at 10-20% saturated vapor concentration of the above solutions, and at 1 L/min flow rate. Once animals were performing accurately in the olfactory discrimination task (typically after 20 to 50 trials), they were trained to switch between blocks of the olfactory and visual discrimination task. Mice typically learnt to switch successfully within 1-2 days. In the olfactory blocks, mice were presented with the vertical and angled grating corridor at different positions in the approach corridor. In each olfactory session, 3 to 4 fixed positions in the approach corridor were assigned at which to present the irrelevant grating corridors, one of which was randomly chosen on each trial (either before the odor onset or partially overlapping with the odor). Gratings were presented on 70% of olfactory trials. Mice learnt to ignore these irrelevant grating stimuli while accurately discriminating the odors. On switching to the visual block, mice started licking selectively to the rewarded grating as before (see below). Mice typically alternated between two olfactory and visual blocks in each training session. Each block typically contained 70-150 trials.

Mice took varying numbers of trials to switch their behavior and perform accurately after each block transition. To exclude this initial period in which the behavior of the mice was ambivalent, we determined in each block the trial boundary after which the mouse performed accurately. In the odor block, we found the center of an 11-trial sliding window in which the accuracy of ignoring vertical stimuli first became > 50%. If for the remaining trials of the same trial type accuracy exceeded 75%, this was considered the trial boundary for stable switching and only subsequent trials in that block were included in the analysis. The same procedure was used to find the trial boundary for successfully discriminating odors. If no such trial boundary was found for either case, the block was considered invalid. In visual blocks we determined the first trial in which the mouse stopped ignoring the vertical grating. Visual discrimination performance accuracy in the subsequent trials always exceeded 80%. We included only sessions that had at least one odor block with valid withholding of licking to irrelevant grating stimuli, and at least one valid visual block (N = 4 mice, 11 sessions). Neuronal responses to gratings in the olfactory block appearing at different positions in the approach corridor did not significantly differ and were pooled together. Grating responses when odor was present overlapping with the grating corridors were removed from the analysis.
5. Bilateral optogenetic silencing of V1 activity

For bilateral silencing of visual cortex 4 transgenic mice (3 males, 1 female) expressing channelrhodopsin-2 in parvalbumin-expressing interneurons were used, obtained by crossing PV-Cre and ChR2(H134R)-EYFP mice (Jackson Labs, stock number 012569 (Madisen et al., 2012) and 008069 (Hippenmeyer et al., 2005), respectively). Three male wild type C57Bl/6J mice were also used to control for non-specific light effects. Mice were implanted with two cranial windows over both visual cortices, and with a custom machined stainless steel head-plate as described above. After recovery from surgery for at least 5 days mice underwent intrinsic imaging through both windows under light isoflurane anesthesia (see below). After determining the spatial extent of V1, all regions excluding V1 were covered with black paint. Mice were then trained on the discrimination task described above until they reached > 90% performance levels. Optical fibers ending in a cone of black fabric were placed over both cranial windows to create uniform illumination in a 5 mm circle, and sealed around the edges using petroleum jelly mixed with black pigment. Using a computer-controlled 470 nm LED (Thorlabs), light was delivered at one of four intensities (randomly interleaved), starting at a fixed distance (74 cm) in the approach corridor before appearance of the grating corridor, until the end of the grating corridor. Typically 30% of the trials were randomly chosen as trials without light stimulation. Light intensity was calibrated with a thermal light power sensor (Thorlabs) and the power per square millimeter was calculated using the average area of exposed V1 across mice. The same mice were trained on an olfactory discrimination task as described above. Once mice performed above 90% accuracy (usually 50-100 trials), V1 was silenced shortly before (74 cm distance in the approach corridor) and during presentation of the odors. Mice also performed the odor discrimination task while running in the dark, with similar results. We recorded a total of 10 sessions in which the 4 transgenic mice performed the visual task, 4 sessions in which the same mice performed the odor discrimination, and 7 sessions in which 4 WT mice performed the visual task. WT mice underwent the same surgical and experimental procedures as transgenic mice.

6. Imaging of intrinsic signals

To determine the location of primary visual cortex, mice underwent optical imaging of intrinsic signals. Mice were anaesthetized with isoflurane, supplemented with chlorprothixene, as described in Experimental Procedures. Visual cortex was illuminated with 700 nm light split from an LED source into 2 light guides. Imaging was performed with a tandem lens macroscope defocused 500 µm into the cortex and a bandpass filter centered at 700 nm with 10
nm bandwidth (67905, Edmund optics). Images were acquired with a rate of 6.25 Hz with a 12-bit CCD camera (1300QF, VDS Vosskühler), frame grabber (PCI-1422, National Instruments) and custom software written in Labview. The visual stimulus was a white vertical bar, 3–4 degrees in width, which drifted left, right, up or down at 0.09 Hz on a black background. Retinotopic maps were obtained by determining the phase of the Fourier component at the frequency of the drifting bar (Kalatsky and Stryker, 2003).

7. Decoder analysis

To quantify the accuracy by which two conditions (either trials with vertical and angled grating corridors (Figure 3B and 5C,D) or FA and CR trials (Figure 6D) could be classified at time \( t \) relative to grating onset, a cumulative decoder was employed which took into account all time points from grating onset to \( t \). Input to the decoder consisted of either calcium traces of all simultaneously recorded neurons, or the mouse running speed, aligned to grating onset. Due to the relatively long decay times of the calcium signal (Chen et al., 2013), the signal at the onset of the grating corridor will contain a decaying calcium component from action potentials at earlier time points, for instance from responses to features in the approach corridor. Because we wanted to determine the earliest time point at which stimulus information was available after grating onset, we used causal high-pass filtering to linearly deconvolve the calcium signal and therefore reduce the history dependence of the signal and increase the signal-to-noise ratio.

The filter was optimized for each neuron separately, based on a simple linear model of calcium

\[
e_{Ca}(t+1) = \exp(-1/\tau_{Ca}) e_{Ca}(t) + n(t),
\]

where \( e_{Ca} \) is calcium concentration, \( \tau_{Ca} \) is the decay constant of calcium, and \( n(t) \) is spiking activity. Fluorescence is a function of the calcium concentration with added zero mean Gaussian Noise:

\[
\Delta F / F(t) = e_{Ca}(t) + N(0, \sigma^2),
\]

and to approximate spiking activity we applied an optimal deconvolution based on the linear model, using a causal Wiener filter (Holekamp et al., 2008; Vogelstein et al., 2010). From training data (30 trials of both conditions), the decoder constructed for each neuron \( n \) a model of the response using as parameters the mean response to the vertical (\( \mu_n^{V}(t) \)) and angled grating corridor (\( \mu_n^{A}(t) \)) and the variance of the noise \( \sigma_n \) to maximize the observed log-likelihood of the data under a Gaussian noise model. On test trials (the remaining trials that were not used as training data), the log-likelihood at time \( t \) that trial \( k \) belongs to condition \( C \) (where \( C \) was for instance the vertical (V) or angled grating corridor (A) condition) is proportional to

\[
L_C(t) = -\sum_n^{N_{cell}} \sum_0^{T_{start}} (D_{n,k}(t - T_{start}) - \mu_n^{V}(t - T_{start}))^2 / (2\sigma_n^2),
\]

where \( D \) indicates deconvolved \( \Delta F/F \). If \( L_V > L_A \), the trial was assigned to the vertical condition, otherwise to the angled
grating condition. To obtain at each time point $t$ the cumulative likelihood $L_C$, the summation only included time points starting from $T_{\text{start}}$, which was the time of the grating onset, up until time $t$. Note that without the temporal accumulation of log-likelihood, the decoder would be equivalent to a linear discriminant analysis. To determine the time point at which there was a detectable divergence of running speed between vertical and angled grating trials, we performed a Wilcoxon rank-sum test on the average speed in individual trials in non-overlapping, consecutive 50 ms windows. The time of divergence was defined as the center of the first window with $P<0.01$ followed by $P<0.01$ in at least 4 consecutive windows. For Figure 5D, we defined post-learning sessions with delayed divergence as sessions with behavioral d-prime > 2 and time of running speed divergence greater than 400 ms (N=8 sessions in N=7 mice, average d-prime 2.59). We paired each of these sessions with an unique session with the smallest difference in behavioral d-prime, but with time of divergence < 400 ms (N=8 sessions in N=6 mice, average d-prime 2.61).

8. Selectivity index for approach corridor features

To analyze responses to the visual features in the approach corridor, we first computed the neuronal response at each corridor position bin $i$ (length corridor 230cm, divided into 125 position bins) for individual trials (because mice ran in each trial with different speeds we used linear interpolation to determine $dF/F$ at each position), and the mean response $R_i$ and standard deviation $s_i$ across trials. For Figure S7E and S7F, we grouped the 125 position bins into 20 coarser bins, but results were similar when using finer (without additional binning, 125 bins) or coarser binning (10 bins). Average peak response for each cell was the maximum mean response out of the 20 bins. To obtain a selectivity index for this continuous stimulus space, approach corridor selectivity was calculated as the standard deviation of the mean response across positions $S_R$ divided by the pooled standard deviation across positions: $(2 \cdot S_R) / s_p^{\text{pos}}$, where $s_p^{\text{pos}} = \sum_{i=1}^{N_{\text{pos}}} (n_i - 1)s_i^2 / \sum_{i=1}^{N_{\text{pos}}} (n_i - 1)$ where $N_{\text{pos}}$ is the number of corridor positions. In the case of only two conditions, the measure is equivalent to the absolute of the grating selectivity index, because $\text{abs}(\bar{R}_V - \bar{R}_A) = 2 \cdot S_R$, which enabled us to compare the effect of learning on the approach corridor selectivity and the absolute grating selectivity (Figure S7E).

9. Locomotion control analyses and GLM

To determine the effect of running speed on neuronal activity, we computed for each cell the correlation between $\Delta F/F$ and the animal's running speed. The correlation was computed for
different time lags, ranging from -2s (speed leading ΔF/F) to 2s (speed lagging ΔF/F). We observed that the correlation typically peaked around -0.5s (15 frames) and we therefore used the correlation at this lag in Figure S8 and S12, but results were similar for lags from -1 to 0s. We used the square of the correlation coefficient ($R^2$, coefficient of determination) to quantify the strength of the modulation of neural responses by running speed. We used a similar approach to quantify the effect of licking on neural responses. Licking was represented by a vector consisting of 0's and 1's at the time of each lick. This vector was smoothed with an exponential function ($f(t') = \exp(-t'/\tau)$ where $\tau$ is 30 frames) and the modulation of neural responses by licking was computed as described above.

To capture more complex dependencies between running behaviour and licking and neural activity, we also formulated a generalized linear statistical model (GLM) of the neural responses $R_n(t,x)$ of each neuron $n$ as a function of time $t$ and position in the virtual environment $x$. The model took into account several factors: the history of running speed including the instantaneous running speed $v(t)$, the tuning to visual features in the run-up corridor as well as the tuning to the onset and offset of the gratings:

$$R_n(t,x) = I_{\text{speed}}(t) + I_{\text{licks}}(t) + s(x) + s_{\text{on}}(t) + s_{\text{off}}(t) + \text{noise},$$

where $I_{\text{speed}}(t) = \sum_{t'=0}^{T} v(t-t')f(t')$ is the convolution of the history of the running speed with a flexible filter $f(t')$ fit as a sum of exponential basis functions $f(t') = \sum_{k=1}^{4} a_k \exp(-t'/\tau_k)$ with 4 basis functions, each with a decay timescale specified in frames as $\tau = \{4,8,20,50\}$. Different sets of timescales were tried and found to give similar results, but the set used for the final analysis was chosen because it had the highest validation set accuracy. We analogously modelled the influence of licking on the neural signals. Lick predictors were represented as 0-1 vectors containing a 1 at the time of each lick. The basis functions were used to smooth these predictors, thus generating time-varying lick-dependent predictors. Because cells had strong onset (and offset) responses to the two gratings, we additionally modelled these as inputs $s_{\text{on}}$ and $s_{\text{off}}$ that depend on the time since the onset (or offset) of a grating. The temporal profile of these inputs $s_{\text{on}}$ and $s_{\text{off}}$ was modelled as a sum of decaying exponentials (timescale $\tau = \{5,10,20,60,200\}$). This model component takes into account not only the latencies of each cell but also their dynamical response patterns and the decay profile of calcium signals. The coefficients of the linear model were estimated with ridge regression, where the penalty coefficient of the ridge regression was chosen based on cross-validation set prediction accuracy. We formed 100 different cross-validation splits of the data, fit the model on the training set and computed statistical measure of predictor quality on the
test set. First we computed the fraction of model-explained variance in neuron \( n \) as

\[
1 - \frac{\text{Var}(R_n - \tilde{R}_n)}{\text{Var}(R_n)}
\]

where \( \tilde{R}_n \) is the model prediction. We defined running-explained variance as the difference in model-explained variances between the full model, and a model where running predictors are excluded. Lick-related variance explained was defined similarly.

10. Analysis of eye position and movements and pupil size during visual discrimination task

We tracked the pupil position and diameter offline, and computed the angular eye position based on the method described by Sakatani and Isa (2004). To rule out that variations in eye position, eye movements, and pupil size could have influenced the neuronal learning effects we observed, we performed several control analyses (Figure S13, see also Keller et al., 2012; Ayaz et al., 2013; Erisken et al., 2014). For these analyses we used the left eye (contralateral to the right hemisphere that was imaged). We computed the average horizontal and vertical angular eye position in the first second after grating onset. We used a stratification approach where we selected a subset of trials with similar distributions of eye position in all conditions (see e.g. Roelfsema et al., 1998). In each session, we binned responses into \( 8 \times 8 \) bins of eye position (bin width = \( 4 \times 4 \) deg), and we made the number of trials in each bin equal to the minimum number of trials across vertical and angled grating trials by randomly removing surplus trials (28% of trials rejected). As a result, no significant differences were present between the distribution of horizontal and the distribution of vertical eye position in the two conditions within each session (Wilcoxon rank-sum tests in each of the 93 sessions, all \( P > 0.05 \)). We then selected a subset of trials before learning and after learning for which the horizontal and vertical eye position was similar by making the relative proportions (number of trials normalized by the maximum number of trials) in each of the bins the same by randomly removing trials in bins where the proportion was higher than the minimum across conditions (37% of total number of trials rejected, as a result the relative proportions are identical for all conditions after stratification). After stratification of horizontal and vertical eye position the neuronal learning effect was similar to the complete data set (Figure S13D, compare with Figure 2D).

We performed the same stratification procedure as described above for pupil size to select for each session a subset of trials with similar pupil size in the two grating corridors (21 bins, bin width = 0.1 deg, 16% of trials rejected), such that no significant differences were present after stratification (all Wilcoxon rank-sum tests, \( P > 0.05 \)). Next, we selected a subset of trials
before and after learning for which the relative proportion was the same (35% of total trials rejected). With this selection we replicated the neuronal learning effects of the whole dataset (Figure S13E). Eye movements were detected by computing the eye velocity and computing for each session a minimum velocity threshold for saccades by multiplying the standard deviation of the velocity distribution by a constant factor of 12 (median threshold across sessions of 57 deg/s, Engbert and Kliegl, 2003; Sakatani and Isa, 2007). Excluding all trials in which eye movements were present (9% of trials rejected) did not change the learning-related increase in neuronal selectivity (Figure S13F). We performed the same analyses on eye position, eye movements and pupil size when mice switched between the visual and olfactory discrimination task (Figure S13G-J). In the stratification procedure for eye position and pupil size 60% and 46% of trials were rejected, respectively. Rejection of trials with eye movements resulted in removing 4% of trials during the visual and olfactory task. We also analyzed the pre-grating increase in activity (Figure 7C) after removing trials with eye movements 0-1 s before stimulus onset (Figure S13K, 4% of trials rejected).

11. Visual discrimination task with fixed visual flow

Six additional mice were trained in a modified virtual reality environment, in which the approach corridor walls were uniformly gray, and the length of the approach corridor on each trial was chosen from an exponential distribution (see Figure S11). The mice thus had a uniform expectation of the stimulus appearing at any point in time (a flat hazard rate, Zariwala et al., 2013). To present identical task-relevant visual stimuli before and after learning for each animal, the pre-learning grating speed profiles 0 - 1 s after grating onset (taken from an imaging session before learning, approximately 140 trials) were played back during the task after learning in 3 mice (Figure S11A). Thus, the grating stimuli were uncoupled from the animals’ running speed for the first second of grating stimulus presentation, and replaced with a grating stimulus with the same optic flow dynamics as in the pre-learning trials (Figure S11A,B).

We applied a similar approach to the task in which trained animals switched between the visual and olfactory discrimination task. We determined the average grating speed profile from a preceding session with visual and olfactory blocks, and presented the gratings with this speed profile, uncoupled from the animals’ running speed, for the first 1 s of grating presentation in the visual and the olfactory discrimination blocks (Figure S11C). The approach corridor preceding the gratings had grey walls and was of random length as described above. To ensure the animals performed well in these tasks, we used similar criteria as before to select trials and
blocks: in the odor block we found the center of an 11-trial sliding window in which the accuracy of ignoring vertical stimuli first became > 50%. If for the remaining trials accuracy of ignoring the gratings exceeded 75%, this was considered the trial boundary for stable switching and only subsequent trials in that block were included in the analysis. The same procedure was used to find the trial boundary for successfully discriminating odors in the olfactory blocks and gratings in the visual blocks. If no trial boundaries were found, the block was considered invalid. We included only sessions that had at least one valid olfactory block and one valid visual block (N = 4 mice, 4 sessions). In 2 of the visual blocks, task performance declined towards the end. In these blocks we therefore excluded the trials at the end during which performance had dropped below the criterion (8 and 10 trials excluded, respectively).

In these control experiments with fixed visual flow we recorded a smaller number of sessions than in the original task (N=3 sessions pre and post learning and N=4 sessions with visual and olfactory blocks). Statistical tests were therefore carried out across cells instead of sessions, using the same approach as described in the main Experimental Procedures. We used bootstrapping where we randomly picked the same number of cells (the minimum across conditions) from both conditions, and repeated this 10000 times. We then computed in both conditions the average proportion of cells of each category, and we also computed the proportion after randomly assigning cells to one of the two conditions. The p-value was given by the number of bootstraps in which the proportion change in the actual data was greater than the proportion change with randomly assigned condition labels.

**Supplemental references**


