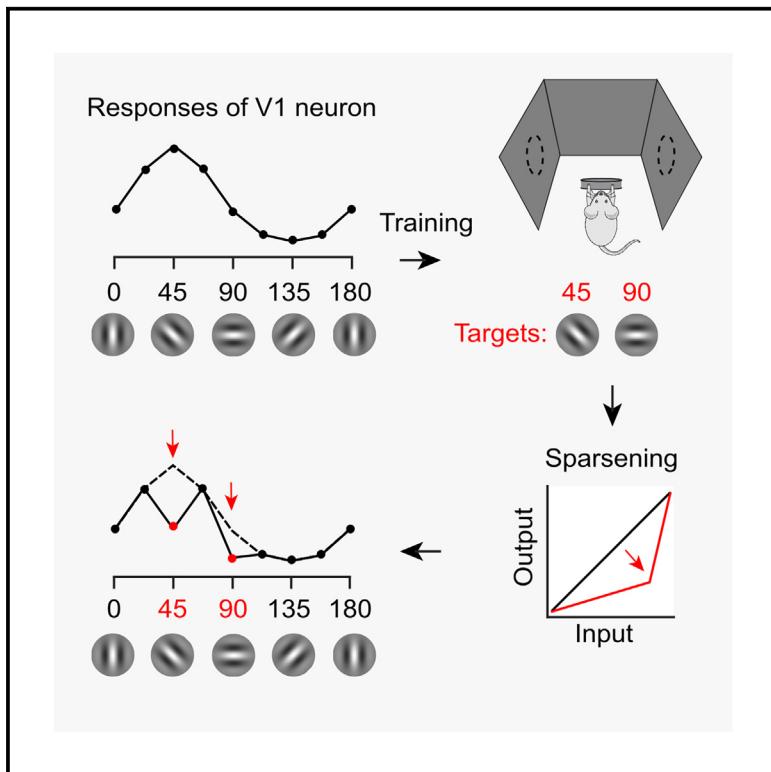


Visual experience orthogonalizes visual cortical stimulus responses via population code transformation

Graphical abstract



Authors

Samuel W. Failor, Matteo Carandini,
Kenneth D. Harris

Correspondence

s.failor@ucl.ac.uk (S.W.F.),
kenneth.harris@ucl.ac.uk (K.D.H.)

In brief

Failor et al. show that apparently complex plasticity of orientation tuning after visuomotor learning can be explained by a simple mathematical transformation of population codes. This sparsens and orthogonalizes population activity, improving the performance of a suboptimal decoder model with inductive bias as might be found in downstream readout circuits.

Highlights

- V1 neurons can develop bimodal or asymmetric tuning curves after task training
- This can be explained by a stimulus-dependent convex transformation of population codes
- Plasticity cannot improve performance of an optimal decoder, which is already perfect
- It does improve a suboptimal decoder with inductive bias as might be found downstream



Article

Visual experience orthogonalizes visual cortical stimulus responses via population code transformation

Samuel W. Failor,^{1,*} Matteo Carandini,² and Kenneth D. Harris^{1,3,*}

¹UCL Queen Square Institute of Neurology, University College London, London WC1N 3BG, UK

²UCL Institute of Ophthalmology, University College London, London EC1V 9EL, UK

³Lead contact

*Correspondence: s.failor@ucl.ac.uk (S.W.F.), kenneth.harris@ucl.ac.uk (K.D.H.)

<https://doi.org/10.1016/j.celrep.2025.115235>

SUMMARY

Sensory and behavioral experience can alter visual cortical stimulus coding, but the precise form of this plasticity is unclear. We measured orientation tuning in 4,000-neuron populations of mouse V1 before and after training on a visuomotor task. Changes to single-cell tuning curves appeared complex, including development of asymmetries and of multiple peaks. Nevertheless, these complex tuning curve transformations can be explained by a simple equation: a convex transformation suppressing responses to task stimuli specifically in cells responding at intermediate levels. The strength of the transformation varies across trials, suggesting a dynamic circuit mechanism rather than static synaptic plasticity. The transformation results in sparsening and orthogonalization of population codes for task stimuli. It cannot improve the performance of an optimal stimulus decoder, which is already perfect even for naive codes, but it improves the performance of a suboptimal decoder model with inductive bias as might be found in downstream readout circuits.

INTRODUCTION

Visual stimuli trigger patterns of neural activity across the visual cortex. These patterns can change following task training, with changes persisting even when the stimuli are presented outside the task context.^{1–17} This representational plasticity entails changes in single-cell tuning, but the nature of these changes appears complex and potentially contradictory. For example, some studies find increases^{1,8,11} and others decreases^{4,12,16} in the numbers of neurons representing task stimuli, some find broadening⁶ and others sharpening^{3,5,9,15} of orientation tuning curves, and others find diverse types of tuning curve changes, such as asymmetric shifts or slope increases.^{13,17} An alternative approach is to consider representational plasticity at the population level. The population response to a stimulus defines a representation in a high-dimensional vector space, similar to representations constructed by machine learning algorithms.^{18–20} If it were possible to mathematically summarize the effects of task training on these population responses, then this might help harmonize the diverse reported effects of training on single-cell tuning, providing clues to their biological mechanism and function.

A common hypothesis is that task training increases stimulus coding fidelity. Cortical responses vary between repeats of an identical stimulus, which could limit the ability of even an ideal observer to decode the stimulus from neuronal activity.^{21–23} It has been suggested that task training changes the size and cor-

relation structure of trial-to-trial variability, improving population code fidelity.^{1,7–9,11,12,24} This hypothesis, of course, presupposes that the population code in the naive cortex suffers from low fidelity. Although this may be true for some stimuli, representations of oriented gratings—a stimulus often used in learning experiments—are extremely reliable.²⁵ Furthermore, in some studies of the somatosensory cortex and olfactory bulb, training changes stimulus responses by sparsening without improving their fidelity.^{26,27} Thus, plasticity of cortical representations may serve a function other than noise reduction.

A second hypothesis pertains not to the abstract optimal decoder but to the kind of suboptimal decoder that might be found, e.g., in downstream brain structures. Even without noise, learning systems exhibit “inductive biases”: they learn some types of stimulus-response associations more readily than others, which may reflect “priors” on the type of associations likely to be encountered.^{28–31} Animals are likely to respond similarly to sensory stimuli evoking similar neural representations and to distinguish stimuli evoking different representations.^{28,29,32,33} For an animal to learn different associations to two stimuli, the cortical representations of the stimuli should become differentiated,³⁴ making firing vectors more orthogonal. This would make inductive bias in downstream brain structures more suited to the task, allowing more readily association of different representations with different actions. Among the types of single-cell plasticity reported in the visual cortex, two effects—reduction in the number of responsive cells and tuning



curve sharpening—are consistent with reduced overlap between populations responding to different stimuli¹⁷ and, thus, with orthogonalization of stimulus representations.

A quantitative understanding of how experience transforms stimulus representations may shed light on the underlying mechanisms. One can consider two types of transformations that would sparsen and orthogonalize stimulus representations. The first, which we term “tuning curve transformation,” sharpens the tuning of individual cells independently between cells. This kind of transformation might result from Hebbian plasticity of input synapses. The second, which we term “population code transformation,” transforms the response of the entire population together, independently between stimuli or even repeats of one stimulus, but consistently across neurons. This kind of transformation might result from network mechanisms such as increased inhibition or neuromodulatory activity.

We used two-photon calcium imaging to study how the tuning of V1 populations changes after mice learn to associate opposing actions with two oriented gratings. Stimulus coding was already perfectly reliable in naive animals so could not be improved further by task training. When analyzed at a single-cell level, plasticity appeared complex, with tuning curves transformed in multiple ways, including previously reported sharpening and asymmetric slope change, and novel phenomena, such as development of multimodal tuning curves. A tuning curve transformation equation failed to capture the changes we observed. However, these diverse single-cell phenomena could be accurately explained by population code transformation: a simple nonlinear transformation of population rate vectors by a function whose convexity varied between stimuli, which was largest for stimuli experienced during the task, especially stimuli associated with motor actions. The strength of transformation varied across repeats of a single stimulus but was consistent across the population, suggesting that it might emerge from circuit dynamics rather than static changes expected from cell-autonomous synaptic plasticity. This transformation caused the population responses to task stimuli to orthogonalize and improved the performance of a model decoder exhibiting inductive biases, such as what might be found in downstream brain circuits.

RESULTS

We trained mice in a visuomotor association task requiring orientation discrimination (Figures 1A, 1B, and S1). Mice were shown pairs of grating stimuli. Gratings of two orientations (45° and 90°) represented opposite motor contingencies (turn the wheel toward vs. away), while a third orientation (68°) was a distractor presented as frequently as the motor-associated stimuli. No other orientations were presented during task performance.

We assessed the orientation tuning of excitatory V1 neurons using two-photon calcium imaging (Figures 1C and 1D), first before task training began (naive condition), and again after training was complete (proficient condition). On average, we recorded $5,041 \pm 2,347$ and $4,197 \pm 1,486$ cells under naive and proficient conditions (mean \pm SD, $n = 5$ mice). We tracked 277 ± 232 of these cells between conditions using semi-supervised region of interest (ROI) matching (mean \pm SD, $n = 4$

mice). During recordings, we presented drifting gratings in a passive session where no rewards were given and the wheel was not coupled to visual stimuli, and we did not detect wheel or body movements from either the wheel rotary encoder or the video-graphic data. Grating presentation caused pupil constriction, more prominently following training but not specific to any orientation (Figures S2A and S2B). It evoked minimal whisking not significantly affected by training or orientation (Figures S2C and S2D). Thus, even though body movements modulate visual cortical activity,^{35–38} analyzing passive stimulus responses avoided this potential confound.

An ideal decoder perfectly distinguishes task stimuli even before training

Individual cells showed a range of tuning characteristics and formed a population code with extremely high fidelity in both naive and proficient mice. Some neurons in both naive and proficient mice showed sharp orientation tuning (Figures 1E–1G). Other neurons showed broader tuning, with multi-peaked tuning curves noticeable in proficient mice (Figures 1H–1J). Dimensionality reduction analysis (STAR Methods) indicated that population responses to different gratings showed essentially no overlap (Figure S3A). An optimal decoder (linear discriminant analysis) distinguished the two motor-associated orientations (45° vs. 90°) with essentially 100% cross-validated accuracy for all orientations in both naive and proficient mice (Figures 2A, 2B, and S3B).

This result suggests that correlated neural noise does not fundamentally limit the fidelity of stimulus coding in naive animals, at least for the stimuli used here. Because this hypothesis has been influential, we examined our contradictory evidence in detail to ensure its validity (Methods S1). This revealed that training does not improve the fidelity of stimulus encoding because a small population of neurons encoded the stimuli with extremely high reliability under both naive and proficient conditions. Indeed, training a sparse decoder by stepwise selection showed that one can decode the stimulus with 100% accuracy from just one to two neurons under both naive and proficient conditions (Figures 2C and 2D). We did not observe a difference in ideal decoder performance following training, even after handicapping the decoder, such as by limiting the number of neurons it could access (Methods S1).

Training sparsens and orthogonalizes population responses to task-associated orientations

Consistent with the inductive bias hypothesis, training differentiates population responses to task-associated stimuli by sparsening and orthogonalizing them. To visualize changes in the population code, we developed a “bullseye plot” (Figure 3A), which displays the responses of all recorded neurons to two stimuli. The responses of 16,000 cells (selected randomly to equalize numbers between conditions) are plotted as points in polar coordinates determined by the cell’s preferred orientation (angle) and orientation selectivity (distance). The cell’s response to the two task stimuli is represented using a two-dimensional colormap, with cells responding exclusively to 45° shown in green, cells responding exclusively to 90° shown in magenta, and cells responding to both shown in gray/black. Although both plots

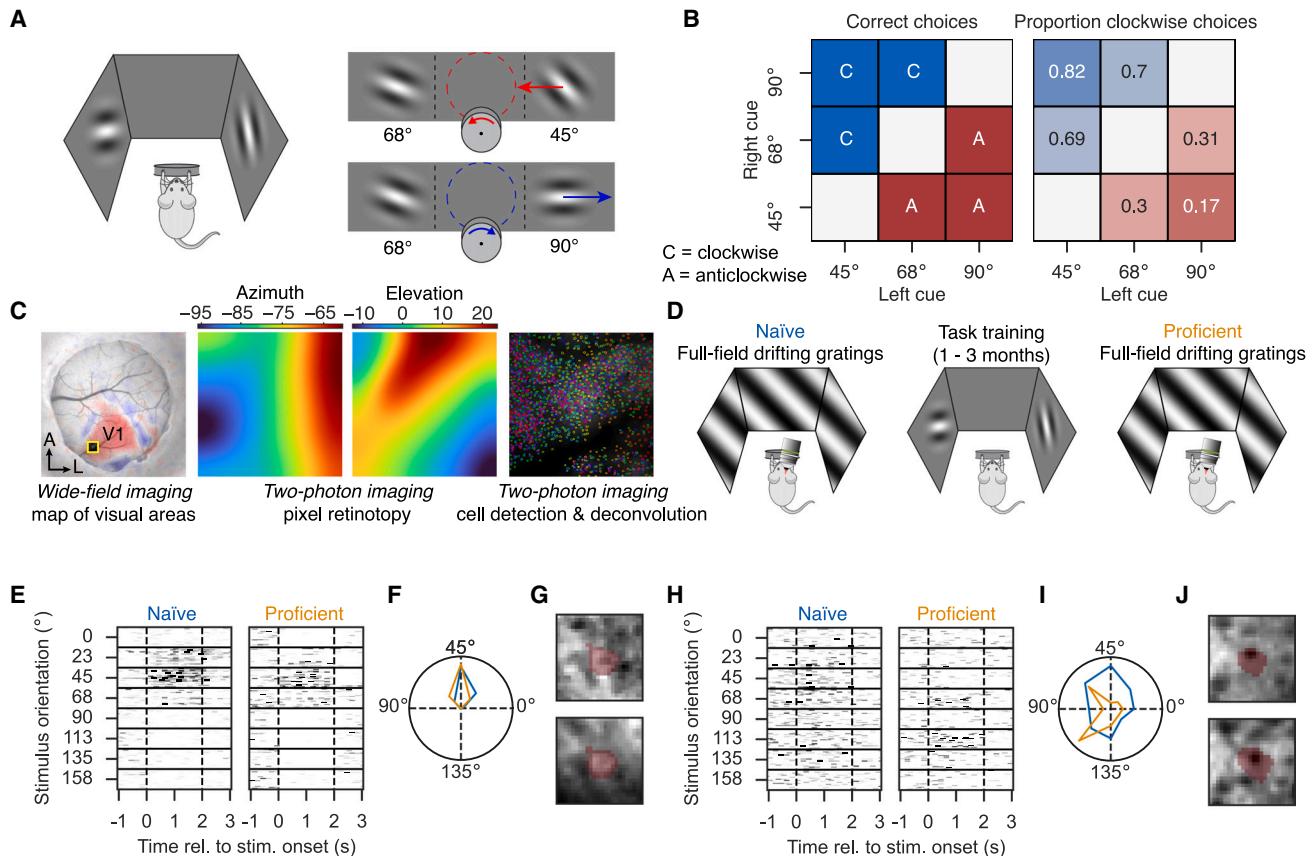


Figure 1. A visuomotor association task and two-photon calcium imaging methods

- (A) On each trial, mice are presented with two stimuli and then turn a wheel to move them on the screens. Turning toward the 45° stimulus or turning away from the 90° stimulus yields a reward, but 68° stimuli are distractors.
- (B) Correct choices for all stimulus pairings (left) and the average proportion of clockwise choices across mice taken from their 10 highest-performing sessions (right).
- (C) Pipeline for imaging neural activity. Left: V1 was located using wide-field imaging with sparse noise stimuli (red/blue, sign map; yellow outlined square, region selected for two-photon imaging). Center: retinotopy map for the two-photon field of view. Right: colored outlines of detected cells.
- (D) Timeline of experiments. Responses to drifting grating stimuli were recorded in naïve mice and in the same mice when proficient at the task.
- (E) Raster representation of responses to repeated grating stimuli for an example cell in a naïve mouse and the same cell when the mouse was proficient at the task.
- (F) Orientation tuning curves of the same cell before and after training, superimposed in polar coordinates (radius represents mean response of the cell to each orientation).
- (G) Mean intensity images of the region of V1 containing the ROI of the recorded cell. Top: naïve. Bottom: proficient. The cell's ROI is shown as a red overlay.
- (H–J) Same as (E)–(G) for a cell with weaker orientation selectivity.

contain the same numbers of neurons, fewer points are visible for proficient mice (right) because fewer neurons responded to the task stimuli (nonresponsive neurons are colored white and thus invisible). This sparsening was predominately due to weakly selective neurons (gray/black neurons in the center) becoming unresponsive, while highly selective neurons were generally unaffected. Moreover, the number of cells responding to both stimuli decreased (visible as a reduction in the number of gray/black points), reflecting orthogonalization of the codes for the two stimuli.

Population sparseness increased more for task stimuli (45°, 68°, and 90°) than for stimuli the mice did not experience during training and for motor-associated orientations (45° and 90°) more than for the distractor stimulus 68° (Figure 3B; 45° and

90° vs. 68°: $p = 0.04$; 45° and 90° vs. non-task: $p = 0.008$; 68° vs. non-task: $p = 0.003$; paired-samples t test, $n = 5$ mice; assessed with the Treves-Rolls measure,^{39,40} which quantifies the fraction of cells with near-zero activity to a stimulus). To visualize orthogonalization, we mapped population response vectors for the motor-associated stimuli into two dimensions using principal-component analysis (PCA), which indicated increased angles after training (Figure 3C). The cosine similarity of responses to the two motor-associated stimuli decreased significantly more than all other pairings (Figure 3D; similarity of 45° and 90° vs.: motor-associated and non-task, $p = 0.036$; motor-associated and distractor (68°), $p = 0.0215$; distractor and non-task, $p = 0.0256$; non-task and non-task, $p = 0.0136$; paired-samples t test, $n = 5$ mice; see Figure S3D for stimulus

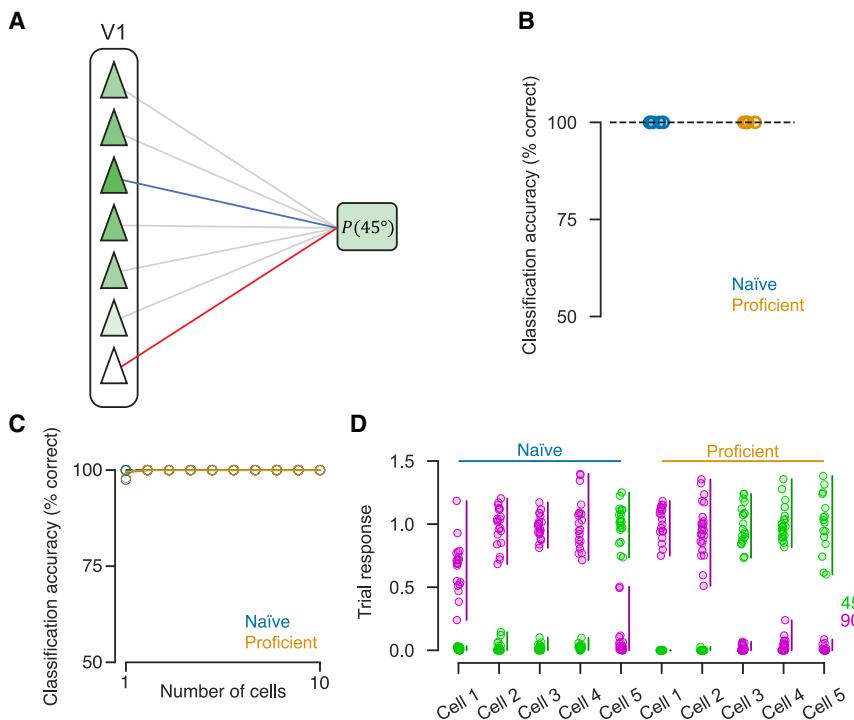


Figure 2. An ideal decoder performs perfectly even before training

(A) Schematic of a linear decoder trained to distinguish between 45° or 90° stimuli by discriminant analysis, which takes a weighted sum of cortical activity and applies a soft threshold to obtain a posterior stimulus probability.

(B) Cross-validated classification accuracy from naïve and proficient mice. The dashed line indicates perfect performance ($n = 5$ mice); shading for standard error is present but too small to see.

(C) Accuracy of decoding from an optimal subset of neurons, selected greedily from the population, as a function of subset size. No significant difference between naïve and proficient conditions was found.

(D) Single-trial responses of example cells, showing how the stimulus can be decoded perfectly from just one cell's activity. Each column shows one cell's activity on all trials (left 5 columns, naïve conditions; right 5 columns, proficient). Magenta circles, activity on individual trials with 90° stimuli; green circles, 45° stimuli. Bars to the right of each column show ranges of responses to the two stimuli, which are completely non-overlapping.

type definitions and all pairwise similarity comparisons). Cosine similarity was correlated with population sparseness across mice (Figure 3E; $p < 0.0001$, mixed effects model; $n = 5$ mice), but showed no further dependence on training conditions once sparsening was accounted for ($p = 0.12$, mixed effects model, $n = 5$ mice), suggesting that sparsening and orthogonalization are two reflections of a single process. Indeed, a correlation between sparsening and orthogonalization is expected for mathematical reasons; sparsening increases the number of near-zero components in the population response vectors and, thus, decreases the cosine similarity between stimulus pairs^{41,42} (Methods S2).

A computational model for the benefits of sparsening and orthogonalization

We next asked how population sparsening and orthogonalization might benefit decision-making. We defined a computational model focused not on the performance of an ideal decoder, which would perform perfectly under both naïve and proficient conditions (Figure 2B), but on a neural decoder with inductive bias related to the similarity of cortical representations. We hypothesized that orthogonalization might differentiate responses to the two task stimuli by reducing a prior bias to respond similarly to them.

The modeled decision circuit receives input from a population of cortical neurons with activity clamped to that measured in individual trials before or after training (Figure 3F). The location of this decision circuit is not relevant, but for concreteness one might imagine the striatum, whose circuitry our model is based on. The decision circuit contains three units: a “decision unit”

for both choices (turn toward or away) and a unit providing feed-forward divisive inhibition. We use single units for simplicity; the brain, of course, would contain populations for each. The decision units receive excitation from each cortical neuron with weights proportional to the cortical neuron's response to the appropriate stimulus, as could be learned by a simple Hebbian rule. The inhibitory unit receives excitation from all cortical neurons with constant equal weights. The probability of turning toward the task stimulus is a logistic function of the difference between the two decision units.⁴³ This model thus captures how decisions might be made by a striatum-like circuit with feedforward excitation and inhibition, although, of course, it is only one of many possible readout circuits.

The model performed significantly better when cortical activity was clamped to activity recorded under proficient compared with naïve conditions (Figure 3G). To understand why, we examined how activity in different model components depends on cortical population sparseness. Training reduced total excitatory input to the decision units (Figure S3E) but not the difference between the excitatory inputs received by the two decision units (Figure S3F). This is because training primarily reduces the activity of weakly tuned neurons (Figure 3A), which feed into both decision units. Instead, training boosted the decision units' activity by reducing feedforward inhibition (Figure S3G). Indeed, the decision network's output reflects the difference between the decision units' excitatory inputs divided by the amount of feedforward inhibition. Training-induced sparsening decreased the latter while leaving the former essentially unchanged, increasing the network's output and, therefore, the probability of a correct choice (Figure S3H).

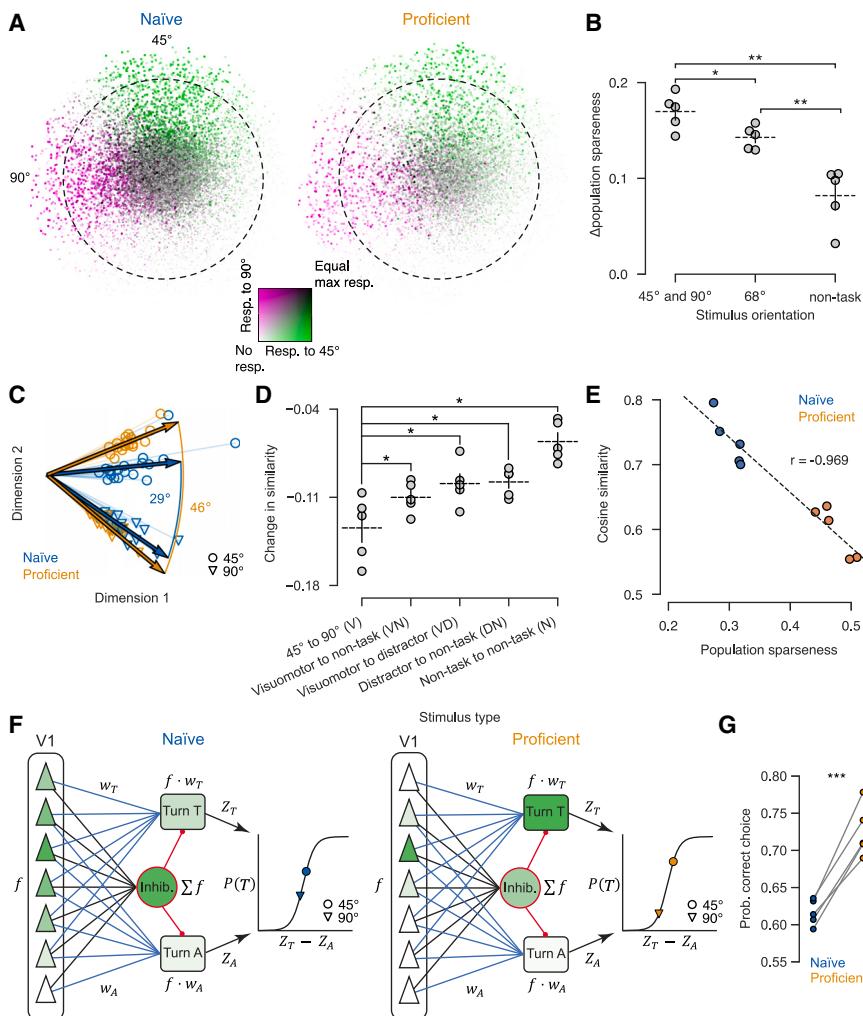


Figure 3. Task training sparsens and orthogonalizes cortical population codes

(A) Bullseye plots showing mean population responses to motor-associated orientations 45° and 90° for naive and proficient conditions. Each point represents a cell (16,000 randomly selected cells per plot) at a polar location determined by the cell's mean orientation preference (angle) and orientation selectivity (distance). Color represents the cell's response to the 45° (green) and 90° (magenta) stimulus orientations on an additive scale, so points responding to both stimuli appear gray; the point's size and brightness (light to dark) represents the cell's maximal response to these two stimuli. The dashed circle is the threshold for considering a neuron highly selective (distance of 0.64).

(B) Change in population sparseness between naive and proficient conditions for each stimulus type. Error bars: mean and SEM ($n = 5$ mice). Asterisks, paired-samples t test.

(C) PCA projection of 45° (circle) and 90° (triangle) trials from one mouse under naive (blue) and proficient (orange) conditions. Arrows: normalized mean response vectors for each stimulus and condition. The angle between mean population responses to the two stimuli goes from 29° (naive) to 46° (proficient).

(D) Change following training in cosine similarity between population responses to the two motor-associated stimuli and to other stimulus-type pairs. Error bars: mean and SEM ($n = 5$ mice); see Figure S7 for all stimulus pairs. Asterisks, paired-samples t test.

(E) Population sparseness vs. cosine similarity for naive and proficient experiments ($n = 5$ mice, naive and proficient conditions).

(F) Schematic of the downstream decoder model with inductive bias. Cortical cells (triangles), whose activity is clamped to experimental measurements, project to two decision units (one

for clockwise and one for anticlockwise turns) via excitatory weights proportional to the mean response of each presynaptic cell to the corresponding stimulus. The decision units also receive divisive feedforward inhibition proportional to the unweighted sum of presynaptic excitatory activity. The difference between the firing rates of the two decision units is passed into a sigmoid function to obtain the probability of a turn toward vs. away from the stimulus. Green shading, cartoon illustration of activity levels when cortical activity is dense in naive mice (left) or sparse in proficient mice (right). Circle and triangle on the logistic curve, model outputs when driven by recorded cortical responses to motor-associated stimuli.

(G) Model performance as assessed by probability of correct choice for models driven by activity recorded of naive and proficient mice ($n = 5$ mice). * $p < 0.05$, ** $p < 0.01$, *** $p < .001$.

Training suppresses responses to task stimuli in weakly tuned cells

Sparsening and orthogonalization are consequences of a change in the sensory population code but do not fully characterize this change. To understand the precise changes in code structure underlying trained-evoked orthogonalization and sparsening, we next examined single-cell orientation tuning curves (Figure 4A). In naive animals, tuning curves typically had standard single-peaked profiles. In proficient animals, however, tuning curves were often irregular and multipeaked. Multimodal tuning can be detected by comparing a cell's modal orientation (the orientation that drives it strongest; circles in Figure 4A) to its circular mean orientation. The circular mean vector is a sum of unit orientation vectors weighted by the cell's response to that orientation of grating (arrows in Figure 4A); this vector's direction

is the cell's mean orientation preference and its magnitude the cell's orientation selectivity. The modal and mean orientations are equal when tuning curves are unimodal and symmetrical but can differ when tuning curves are multimodal (Figures S4A and S4B). Comparing modal and mean orientations indicated that multimodal tuning curves were more common in proficient mice, typically for cells of weaker orientation tuning, with mean but not modal orientation close to the motor-associated orientations 45° and 90° (Figure S4C). This occurred because responses to the motor-associated orientations 45° and 90° were suppressed following training (Figures 4A–4C), more strongly than responses to the distractor stimulus 68°, which was, in turn, more suppressed than responses to orientations not presented during training (tracked cells, 45° and 90° vs. 68°: $p = 0.0007$; 45° and 90° vs. non-task: $p < 0.0001$; 68° vs.

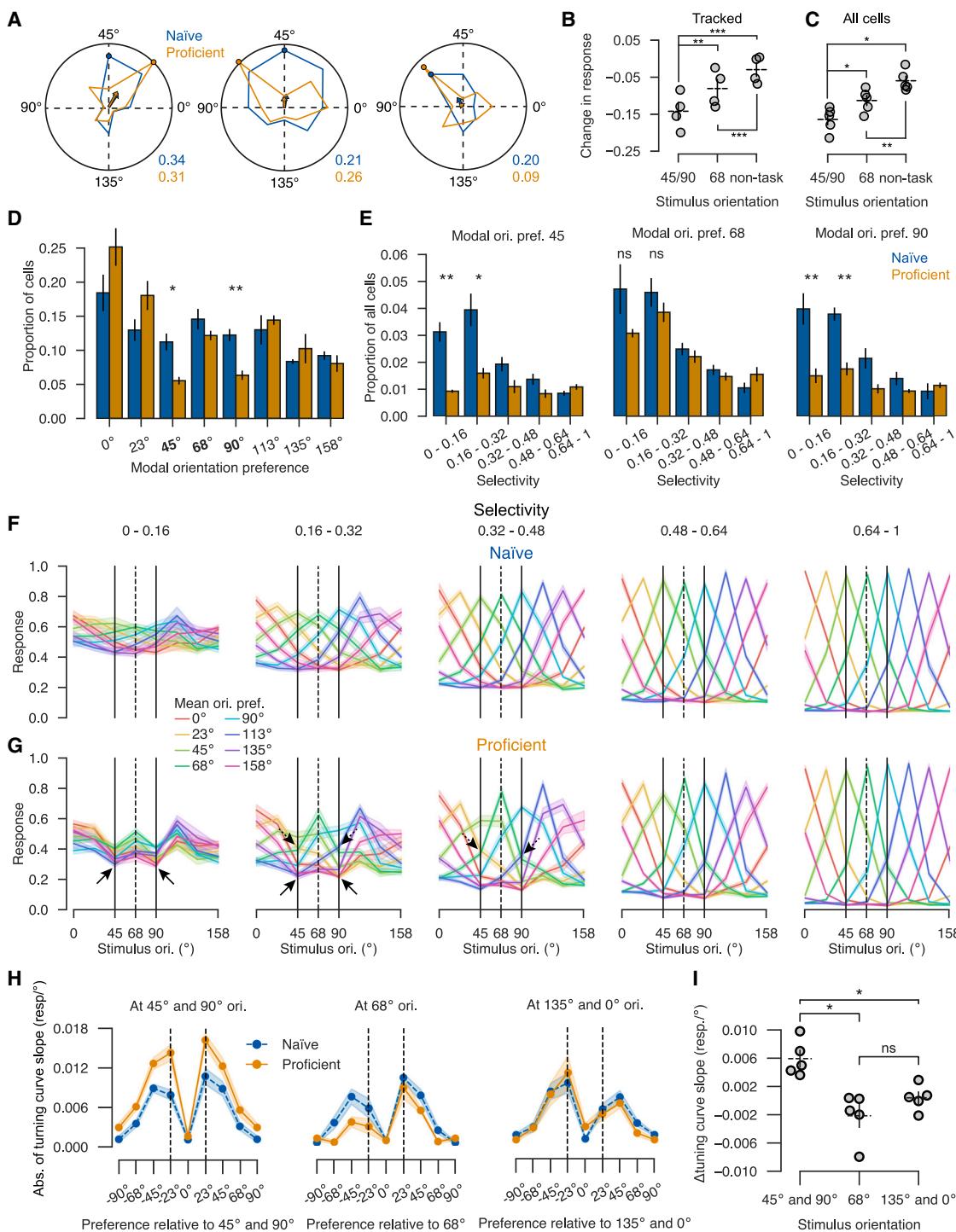


Figure 4. Training suppresses responses to task stimuli in weakly tuned cells

(A) Examples of naive and proficient orientation tuning curves for three cells. Colored polar curves, mean response to each orientation; dots, response to modal orientation; arrows, circular mean vectors representing mean orientation preference (angle) and orientation selectivity (length). Numbers at bottom right, naive and proficient selectivity indexes.

(B) Average training-induced change in responses to the motor-associated, distractor, and non-task orientations for tracked neurons. Error bars: mean and SEM ($n = 5$ mice). Asterisks, hierarchical linear mixed-effects model ($n = 1,107$ cells, $n = 5$ mice).

(C) The same for all recorded neurons. Error bars: mean and SEM ($n = 5$ mice). Asterisks, paired-samples t test ($n = 5$ mice).

(D) Proportion of cells with each modal orientation preference in naive and proficient mice. Error bars: SEM ($n = 5$ mice). Asterisks, paired-samples t test.

(legend continued on next page)

non-task: $p < 0.0001$; hierarchical linear mixed-effects model, $n = 1107$ cells, $n = 5$ mice; all cells, 45° and 90° vs. 68° : $p = 0.04$; 45° and 90° vs. non-task: $p = 0.01$; 68° vs. non-task: $p = 0.008$, paired samples t test, $n = 5$ mice).

Suppression of responses to task orientations changed the distribution of modal orientation preferences primarily in weakly tuned cells. Task training significantly decreased the fraction of cells modally preferring motor-associated (45° and 90°) but not distractor orientation (68°) (Figure 4D; 45° : $p = 0.014$; 68° : $p = 0.228$; 90° : $p = 0.006$, paired-sample t test, $n = 5$ mice). This decrease came specifically from cells of low orientation selectivity with no decrease in cells strongly tuned for motor-associated orientations (Figure 4E; 45° : $p = 0.005$ and 0.037 for selectivity 0–0.2 and 0.2–0.4; 68° : $p = 0.130$ and 0.390 ; 90° : $p = 0.001$ and 0.013 , paired samples t test, $n = 5$ mice). Analysis of cells tracked between naive and proficient recordings confirmed that cells with weak preference for the motor-associated orientations in naive mice were most likely to change their modal orientation (Figures S5A and S5B).

Training also changed tuning curve shapes, dependent on a cell's preferred orientation and selectivity (Figures 4F and 4G). We divided cells into groups according to their orientation selectivity and preference (using mean rather than modal preference as it is more stable; Figure S4) and plotted the mean tuning curves of cells in each group before and after training using held-out repeats. Naive tuning curves were unimodal and symmetrical for all mean orientations (Figure 4F). For proficient mice, however, a different structure appeared (Figure 4G). Weakly tuned neurons were suppressed by the motor-associated orientations, including cells whose mean orientation was motor associated. Analysis of neurons tracked between naive and proficient conditions gave similar results (Figures S5B and S5C). Suppression of motor-associated orientations also asymmetrically increased tuning curve slopes at these orientations, and as reported in primates,¹³ this increase in slope was specific to neurons whose mean orientation preference flanked one of the task stimuli (Figures 4H and 4I).

A simple equation for population code transformation

Although these training-related changes to single-cell tuning curves appeared complex, they could be accurately summarized by a simple population code transformation equation:

$$f'_{c,\theta} = g_\theta(f_{c,\theta})$$

This equation captures plasticity of V1 population codes by applying a stimulus-dependent convex transformation g_θ to

each cell's response (Figure 5A): if cell c's response to orientation θ was $f_{c,\theta}$ before training, then it becomes $g_\theta(f_{c,\theta})$ after training. Importantly, the function g_θ depends on the stimulus θ but not on the cell c: to predict a cell's proficient response to stimulus θ , it suffices to know the cell's naive response to this stimulus θ only; other properties of the cell, such as its responses to other stimuli, tuning sharpness, or preferred orientation, provide no additional information. We estimated g_θ as piecewise linear functions relating naive and proficient responses separately for each experiment and orientation, which accurately summarized the effects of task training (Figure 5B). The functions g_θ were most convex for motor-associated stimuli, indicating that cells that responded modestly to these stimuli have their responses suppressed further after training, but cells that responded either strongly or not at all are unaffected. For distractor or non-task orientations, the functions were closer to linear, indicating less plasticity of responses to these stimuli (Figure 5C; convexity of motor-associated vs. 68° : $p = 0.014$; motor-associated vs. non-task: $p = 0.009$; 68° vs. non-task: $p = 0.827$, paired samples t test, $n = 5$ mice). Applying this transformation to the naive tuning curves predicted proficient responses with remarkable accuracy (compare Figure 5D to Figure 4G). Similar results were observed for individual neurons tracked between naive and proficient conditions (Figure S5D). Thus, one simple equation summarizes the representational plasticity we observed: responses to motor-associated stimuli are suppressed by task learning but only in cells that responded to them at intermediate levels.

Population code transformation provides a simple, quantitative explanation for the apparently complex effects of training on single-cell tuning curves changes seen earlier (e.g., Figures 1I, 4A, and 4G). It explains why training affects mostly the cells that are broadly tuned and gives them multipeaked tuning curves; these cells' intermediate responses are most affected by the transformation, while strongly tuned cells always fire close to the minimum or maximum possible and so are unaffected. It explains why population sparseness increases most strongly to task stimuli (Figure 3B): we proved mathematically that sparsening inevitably follows a convex transformation of firing rates, accompanied by orthogonalization (Methods S2).

The plasticity we observed was better modeled by this stimulus-dependent transformation of population codes than by cell-autonomous sharpening of tuning curves. To show this, we evaluated a tuning curve transformation equation $f'_{c,\theta} = g_c(f_{c,\theta})$, which transforms each cell's tuning curve rather than transforming each stimulus' population response. Specifically, for each cell class c (i.e., each orientation preference and

(E) Proportion of the cell population that had modal orientation preference 45° (left), 68° (center), and 90° (right) and specified orientation selectivity. Asterisks, paired samples t test.

(F) Average orientation tuning curves for cell groups defined by mean orientation preference (color) and selectivity (column) in naive mice. Solid vertical lines, motor-associated orientations; dashed lines, distractor. Y-scale measures mean deconvolved response of all cells in the group, scaled by maximal response (STAR METHODS). Shading: SEM ($n = 5$ mice).

(G) The same plot for proficient mice. Solid arrows highlight suppression of cell responses to motor-associated orientations 45° and 90° . Dashed arrows highlight tuning curve steepening.

(H) Tuning curve slope analysis as in Schoups et al.¹³ The three plots show tuning curve slopes at motor-associated, distractor, and non-task orientations. Each point shows an average over cells grouped by mean preferred orientation relative to the orientation at which the slope is computed (x axis). Color, training condition; shading, SEM ($n = 5$ mice).

(I) Change in tuning curve slope at motor-associated, distractor, and non-task orientations, averaged over cells with adjacent orientation preferences for each experiment (at dashed lines in H). Error bars: mean and SEM ($n = 5$ mice). Asterisks, paired-samples t test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

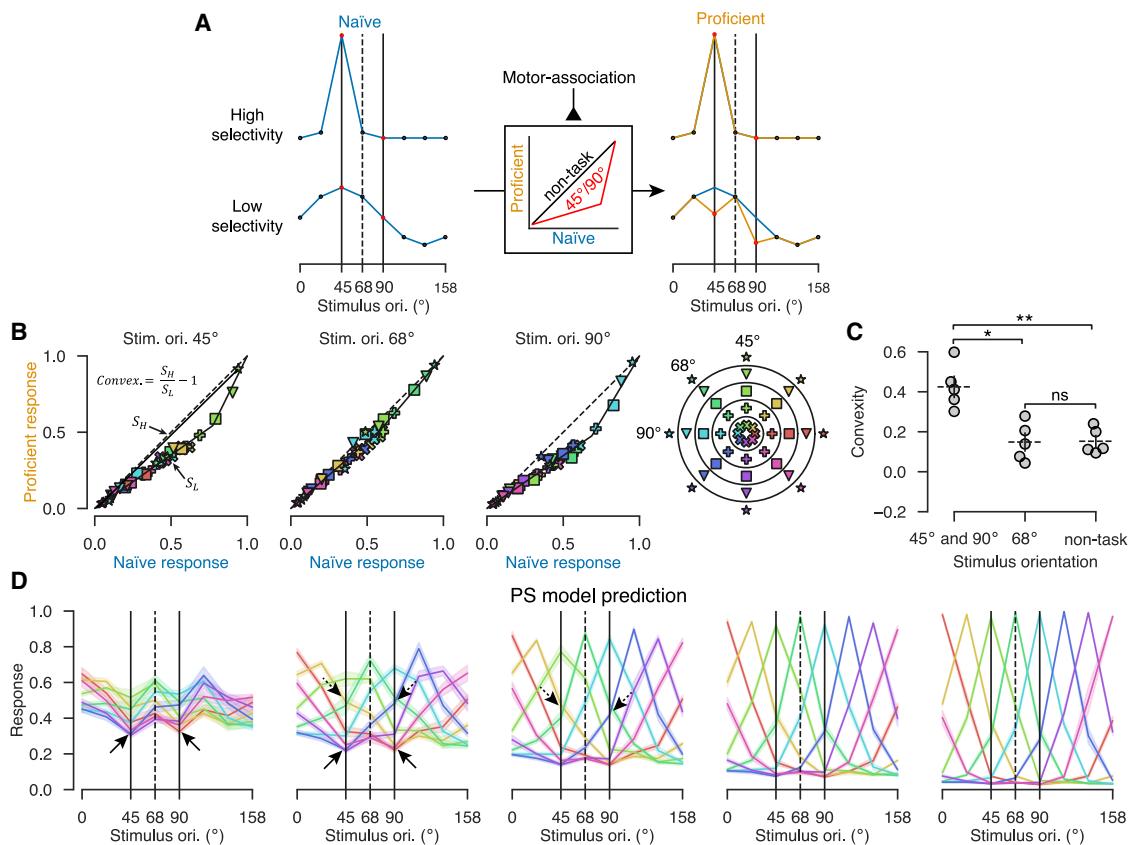


Figure 5. A simple equation summarizes how training changes population activity

(A) Schematic for population code transformation. Following task training, the naïve response $f_{c,\theta}$ of cell c to stimulus θ is transformed by nonlinear function g_θ , which depends on the stimulus θ but not the cell c . Blue curves (left) illustrate naïve tuning curves $f_{c,\theta}$ of two hypothetical cells. The center box illustrates the function g_θ , which is more convex for motor-associated stimuli (red) than for non-task stimuli (black). Orange curves (right) show proficient responses $g_\theta(f_{c,\theta})$ superimposed on original naïve responses (blue). This transformation suppresses moderate but not strong or zero responses to the motor-associated stimuli or any responses to non-task stimuli. Thus, a cell that was highly selective to 45° is unaffected (top), while a cell that was weakly selective to 45° develops a multi-peaked tuning curve (bottom).

(B) Empirical fits of the function g_θ for $\theta = 45^{\circ}, 68^{\circ}$, and 90° . Each symbol shows the mean response of the same cell groups as analyzed in Figure 2F to the orientation θ under naïve vs. proficient conditions. Each point shows the average response of cells from all experiments. Black solid lines are stimulus-specific fits of piecewise linear functions g_θ relating naïve responses to proficient responses. Symbol color indicates orientation preference, and glyphs indicate selectivity following the code illustrated in polar coordinates on the right.

(C) Convexity of g_θ (defined as the ratio of slopes of the two solid lines in (B) minus 1; STAR methods) for motor-associated orientations 45° and 90° , distractor orientation 68° , and all other non-task orientations. Points indicate individual mice. Error bars: mean and SEM ($n = 5$ mice). Asterisks, paired-samples t test.

(D) Proficient orientation tuning curves obtained by applying the functions fit in (B) to naïve tuning curves. Solid and dashed arrows highlight the same features seen in the actual proficient responses (Figure 4G). Shading: SEM ($n = 5$ mice). * $p < 0.05$, ** $p < 0.01$, paired-samples t test.

selectivity class, corresponding to one curve in Figure 4F), we fit a piecewise-linear transformation g_c relating activity under naïve and proficient conditions (Figure S6A). The tuning curve transformation equation did not predict the multi-peaked tuning curves and asymmetric slope shifts observed for weakly tuned neurons following training (Figure S6B) and had lower cross-validated accuracy (Figure S6C), indicating that cell-autonomous tuning curve sharpening insufficiently describes our results.

Population code transformation varies dynamically from trial to trial depending on behavioral state

Plasticity of cortical representations is often assumed to arise from plasticity of local excitatory synapses that changes the sen-

sory drive received by cortical neurons.^{44,45} However, the fact that plasticity is better fit by transformation of single-stimulus population responses than of single-cell tuning curves suggests an alternative hypothesis: V1 neurons' sensory drive is unaffected by training, but motor-associated stimuli engage a circuit process suppressing cells receiving weak sensory drive while sparing strongly driven cells. Multiple physiological mechanisms could underlie this, such as increased activation of a particular inhibitory cell class, input pathway, or neuromodulatory system. For example, a mathematical argument shows that convex transformation of firing rates should be expected from increased subtractive inhibition, even inhibition that increases nonspecifically in all cells (Methods S2).

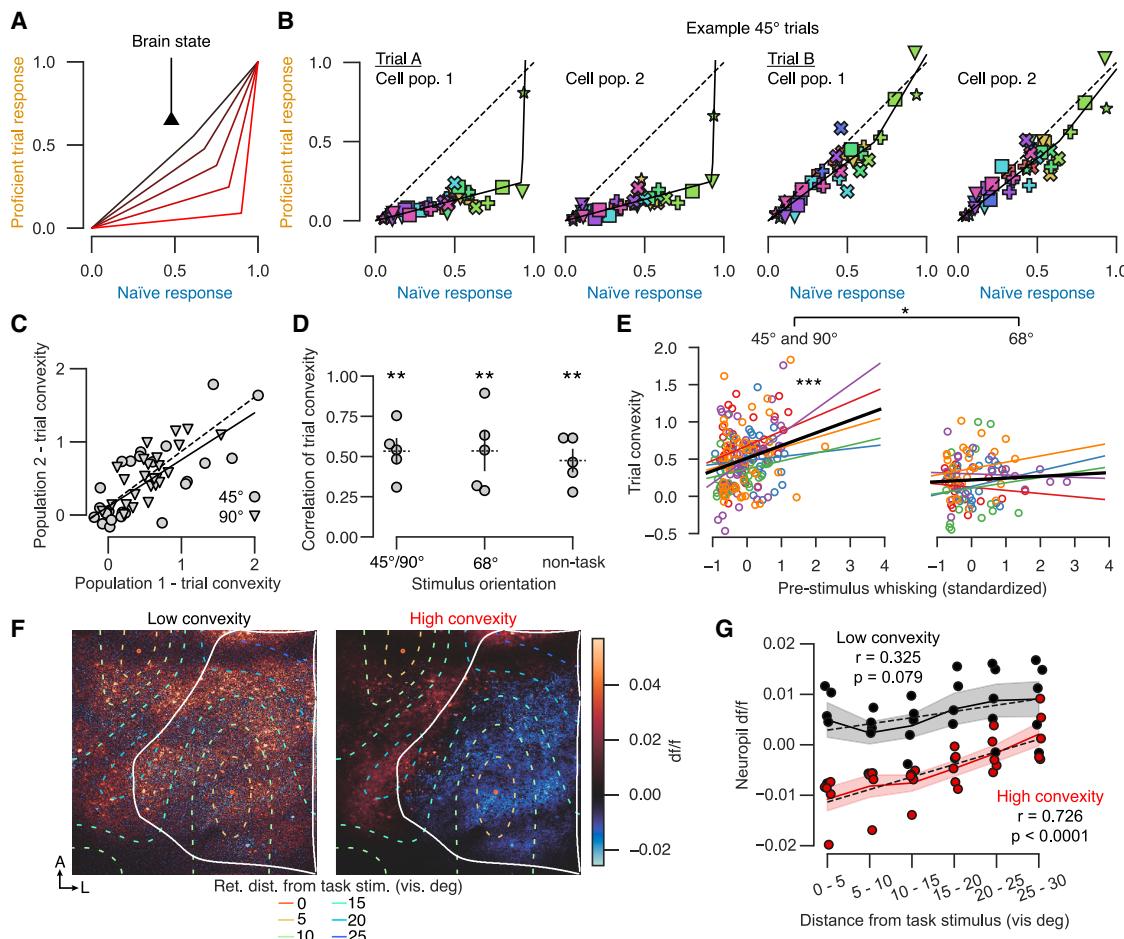


Figure 6. Population code transformation varies dynamically from trial to trial depending on behavioral state

- (A) Dynamic transformation model. Activity undergoes varying levels of transformation on different trials depending on instantaneous brain state.
- (B) Single-trial transformation functions for two example presentations of 45° gratings in the same recording session, plotted as in Figure 5B. For each trial, responses of separate halves of the cell population are shown.
- (C) Similarity of single-trial convexities in two different halves of the cell population for the recording in (B). Each point represents a single presentation of the 45° (circle) or 90° stimulus (triangle). Solid and dashed lines are fits to the 45° and 90° trials, respectively.
- (D) Correlation of single-trrial convexities between two halves of cells (average of 2,000 random splits; STAR Methods), with each point representing average over motor-associated, distractor, or non-task stimuli in one experiment. Error bars: mean and SEM ($n = 5$ mice). Asterisks, one sample t test for difference to 0.
- (E) Correlation of single-trial transformation convexity with pre-stimulus whisking. Each point represents a stimulus presentation color coded by mouse identity. Colored lines are linear regression fits for individual mice, and the black line is the mean over mice. Left: motor-associated stimuli. Right: distractor stimuli. Asterisks, linear mixed-effects model ($n = 303$ trials, $n = 5$ mice).
- (F) Trial-to-trial variability of neuropil responses. Left and right plots show mean change in fluorescence from baseline ($\Delta f/f$) (2 s post stimulus vs. 1 s pre stimulus) of two-photon imaging frames to motor-associated orientations. Left and right images show averages over trials for which the population convexity (defined by the ratio of slopes as in B) was low (<0) or high (>0.3), respectively. Colored contours: retinotopic distances from the task stimulus location (see legend).
- (G) V1 neuropil responses to motor-associated orientations, as a function of retinotopic distance from the task stimulus, for trials with low and high convexity. Dashed lines: linear regression fits. Shading: SEM ($n = 5$ mice). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

This hypothesis makes an experimental prediction: the convexity of population code transformation should vary from trial to trial (Figure 6A). While static synaptic plasticity would lead to plasticity that is constant across trials, a dynamic mechanism might be engaged more on some trials than others due to fluctuations in brain state. Thus, if the degree of population code transformation varies between stimulus repeats, it suggests a dynamic mechanism. Furthermore, since the circuit process would affect all neurons similarly, trial-to-trial fluctuations in

response transformation should be consistent across the population. Finally, the degree of transformation on each trial might correlate with behavioral state. Trial-to-trial variability in neuronal sensory responses has been reported to take additive and multiplicative forms.^{46–48} The current hypothesis predicts a further type of trial-to-trial variability in the responses to task-relevant stimuli.

Our data confirmed this prediction (Figures 6B–6E). We randomly divided cells into two groups, balanced for orientation

preference and selectivity, and within each group examined the transformation from trial-averaged naive population responses to single-trial proficient population responses. Transformation convexity varied substantially between trials, even within repeats of a single stimulus orientation, but was consistent across the cell groups (Figures 6B–6D; $p > 0$ at $p < 0.05$ for each stimulus orientation, one sample t test, $n = 5$ mice). Consistency between cell groups shows that trial-to-trial fluctuations in convexity result not from random noise (which would be independent across groups) but from consistent transformation of population responses.

The strength of the transformation on a given trial depended on behavioral state. Convexity was strongest when mice were whisking prior to stimulus onset, and this was specific to responses to motor-associated stimuli (Figures 6E and S7; linear mixed-effects model: $p < 0.0001$ for effect of whisking on convexity for motor-associated stimuli, $p = 0.014$ for motor-associated vs. distractor stimuli, $p = 0.001$ for motor-associated vs. non-task stimuli; $n = 303$ trials, $n = 5$ mice). This dependence on behavioral state did not result from stimulus-evoked body movements, which did not vary between different grating orientations (Figure S2).

Suppression of cortical activity by motor-associated stimuli was strongest in the part of V1 topographically representing the task stimulus (Figures 6F, 6G, and S8). This could be seen even in neuropil fluorescence, which was suppressed most strongly on trials where the transformation of population activity was most convex, specifically in parts of V1 aligned retinotopically with the stimulus (Figures 6F and 6G) and specifically for motor-associated orientations (Figure S8D). Analysis of individual cells gave consistent results, with cells in cortical areas within 10° of retinotopic visual angle from the stimulus significantly suppressed to task orientations but not cells farther than 20° (Figure S8E). These results are consistent with a circuit mechanism that suppresses cortical activity, specifically in the V1 region aligned with the stimulus, with the amount of local suppression varying between trials and strongest for motor-associated stimuli.

DISCUSSION

We have shown that training in a visuomotor task sparsens and orthogonalizes V1 population responses via population code transformation. After training, population responses are transformed by a nonlinear convex function that is the same for all neurons but whose convexity varies from trial to trial and is largest, on average, for task-associated stimuli. This transformation sparsens population responses by suppressing neurons responding at intermediate levels, making the resulting population vectors more orthogonal. This does not increase the fidelity of stimulus coding: an ideal observer could have perfectly decoded stimulus orientation from the population responses even before training. Rather, it might help downstream circuits produce different motor responses to the two task-associated orientations by biasing even suboptimal decoders to distinguish those task orientations.

The way task training transformed stimulus coding appeared complex when analyzing each cell's tuning individually but

became simple when described at the population level. Single-cell tuning curves showed diverse changes, including multimodality and asymmetry. These changes were accurately modeled by the population code transformation equation, which transforms each cell's activity by a nonlinear function that depended on the stimulus but not the cell. This transformation is simpler than one might have expected; to predict a cell's response to stimulus following learning, one only needs to know the cell's response to that stimulus before learning. No other details of the cell, such as its responses to other stimuli, preferred orientation, or orientation selectivity, are required. An alternative tuning curve transformation equation, in which the function depends on the cell but not the stimulus, gave a worse fit. Thus, the simplicity of the transformation from naive to trained responses is only apparent when analyzing all cells together.

The population code transformation equation can explain several apparently diverse plasticity phenomena observed previously at a single-cell level. It predicts a reduction in cells responding modally to trained orientations,^{4,16} an asymmetrical increase in tuning curve slope at trained orientations,¹³ suppression of neuronal activity for cells with a preferred orientation close to task stimuli,^{3,10} as well development of multimodal tuning, which we believe to be previously unreported. Some apparent discrepancies with previous studies may result from differing definitions. For example, while Poort et al.¹¹ reported recruitment of new neurons selective for task stimuli, this primarily reflected a reduction in neurons responding to multiple stimuli,¹⁰ consistent with our observations (we re-analyzed data from that study, confirming that population responses sparsen and orthogonalize after training; Figure S9). The fact that a single equation can explain diverse effects and even apparently contradictory phenomena demonstrates the importance of summarizing data with equations and suggests that these phenomena might arise from a single mechanism, for which we suggest possibilities below.

Despite concordance with previous results in the visual cortex, our findings do not appear to be fully congruent with results from the auditory and somatosensory cortex. In these regions, some studies^{49–51} (but not others^{26,52}) report that task training or neuromodulatory stimulation increases the number of neurons responding modally to task stimuli. We suggest three non-exclusive reasons for this apparent discrepancy. First, it would be surprising if there were only one mechanism for cortical representational plasticity, and different mechanisms may be employed in different cortical regions and tasks. In fact, other studies of associative learning in the somatosensory or auditory cortex did observe sparsening,^{26,52} and similar phenomena have been observed in the olfactory bulb.²⁷ Second, our study (like Gdalyahuet al.²⁶ and Chu et al.⁵²) used two-photon imaging to record excitatory cells in superficial layers. Other auditory and somatosensory studies used electrophysiological multi-unit recordings, which are biased toward fast-spiking interneurons, and increased activity of these cells is one possible mechanism for sparsening of pyramidal cell activity. Third, expansion of sites encoding task stimuli is transient: after continued training or stimulus exposure, expanded maps “renormalize” without compromising behavioral performance.⁵³ Furthermore, inducing map expansion can actually worsen task performance⁵⁴ by

increasing the rate of false responses to non-target stimuli.⁵⁵ Our task required a long training period, potentially allowing map expansion to reverse, and requires differentially responding to the two stimuli while not responding to the similar distractor stimulus, for which map expansion might impair performance.

We did not observe increased coding fidelity following training; stimuli were encoded with 100% accuracy even in naive mice. This contrasts with some previous studies,^{1,8,9,11,12} for which we offer three non-exclusive explanations. First, the visual stimuli we used, high-contrast, full-screen drifting gratings with orientations separated by 45° and no noise, were very distinct. The idea that cortical representations of these stimuli would have such low fidelity that decoding them is difficult is controversial. Indeed, one study could decode grating angles to 1° accuracy from naive mice.²⁵ The fact that other studies have failed to accurately decode such distinct stimuli from V1 activity does not prove that it cannot be done: two-photon microscopy is subject to artifacts such as brain movement and neuropil contamination, which, unless corrected with appropriate software, will introduce noise with correlations of precisely the form that compromises decoding.^{56,57} Second, activity in mouse V1 encodes not only visual stimuli but also non-visual features, such as ongoing movements,^{37,38} which may compromise decoder performance, particularly for recordings during task performance. Third, the performance of any decoder represents a lower bound on the ideal observer performance and is sensitive to training set size and regularization, particularly when decoding from large populations.

Our results do not determine the mechanisms of population code transformation but suggest a hypothesis: after training, task-associated stimuli drive localized but nonspecific inhibition. Representational plasticity is often assumed to arise from plasticity of excitatory synapses onto the recorded cells. This does not seem to be a likely mechanism for our results, as synaptic plasticity is presumably static, while the strength of population code transformation varies across trials. Instead, our results are more consistent with increased activity of an inhibitory population in the retinotopic location of the task stimulus, whose response to task-associated stimuli is strengthened by task experience. Our data do not shed light on the inhibitory class responsible, but work in the visual and auditory cortex^{16,52} has implicated increased somatostatin interneuron activity in reducing cortical responses following passive stimulus exposure. Nonspecific inhibition would explain why plasticity of all neurons could be explained by the same transformation and why suppression was also visible in neuropil fluorescence, which correlates with the summed activity of local neurons.⁵⁹ The fact that transformation strength fluctuates from trial to trial and correlates with pre-stimulus whisking is consistent with an inhibitory class that integrates behavioral and sensory information. A simple mathematical argument shows that convex transformation of firing rates should be expected from an increase in subtractive inhibition (*Methods S2*). Furthermore, previous work shows that local feedback inhibition contributes to V1 orientation tuning,⁶⁰ that stimulating parvalbumin-positive interneurons narrows tuning curves in a manner consistent with convex transformation of firing rates and improves behavioral orientation discrimination,⁶¹ and that learning can cause changes in inhibitory activity.^{62,63} The mechanism for plasticity of these inhibitory neurons' responses to motor-

associated stimuli is unclear. One possibility is strengthening of their inputs from local pyramidal cells tuned to task-associated stimuli. Alternatively, the feedback could arise from neuromodulation of local inhibitory classes^{64–66} or from long-range excitatory or inhibitory input¹⁶; for example, from secondary visual or motor areas, which, the predictive processing hypothesis⁶⁷ suggests, might suppress V1 visual responses, particularly for motor-associated stimuli. If a long-range input is involved, however, it would have to specifically target the retinotopic location of the task stimuli to explain localized suppression of both cellular and neuropil activity. Future experiments may be able to identify the precise types of inhibitory neurons involved and the mechanisms that cause them to increase responses to task-associated stimuli and to test the hypotheses that the resulting sparsening of excitatory population activity correlates with, and contributes to, improved behavioral performance.

Regardless of the mechanism, the fact that training orthogonalizes population responses to the motor-associated stimuli suggests a function for this process. Why animals so frequently make “incorrect” choices—i.e., choices leading to suboptimal reward—is one of the biggest puzzles in neuroscience. The “noise hypothesis” provides a superficially attractive explanation: suboptimality results from limitations of neural circuits, and when animals do not obtain a high reward, they simply cannot do better. Although the noise hypothesis might appear to be supported by recordings of small populations,^{21–23} it is ruled out by our own and previous²⁵ large-scale population recordings, which show that at least high-contrast gratings are encoded in naive cortical activity with 100% accuracy. The question thus becomes even more puzzling: why would an animal that could, in principle, obtain a high reward still not do so? We hypothesize that the answer lies in inductive biases or “priors” that have been shaped by evolution. Gratings are not natural stimuli, and if a mouse ever encountered one in the wild, then its orientation would unlikely be behaviorally significant. Thus, one might expect mice to be biased toward generalizing behavioral responses learned to one grating orientation to another. In our laboratory task, this inductive bias of default generalization between gratings leads to suboptimal performance that can only be overcome through extensive training. We hypothesize that sparsening and orthogonalization of cortical representations may contribute to breaking this inductive bias.

Limitations of the study

This study does not characterize whether task training transforms population codes more than would occur from passive stimulus exposure alone. The strength of population code transformation was significantly stronger for motor-associated than for the equally frequent distractor stimuli, but this conclusion could be strengthened by increasing the number of mice used or by counterbalancing target orientations across mice, as cardinal directions are known to behave slightly differently.⁵⁸

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Kenneth D. Harris (kenneth.harris@ucl.ac.uk).

Materials availability

This study did not generate any new unique reagents.

Data and code availability

- Experimental data have been deposited at Figshare and are publicly available at <https://doi.org/10.6084/m9.figshare.27855423.v1> as of the date of publication.
- All original code has been deposited at GitHub and is publicly available <https://doi.org/10.5281/zenodo.14205652> as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

We thank Charu Reddy for assistance with mouse surgeries; Miles Wells, Hamish Forest, and Laura Funnell for assistance with task training; and Michael Krumin for technical assistance with two-photon calcium imaging. We thank Jasper Poort for sharing data. This work was supported by the Wellcome Trust (grants 205093 and 108726 to M.C. and K.D.H.) and the BBSRC (grant BB/W015293/1). M.C. holds the GlaxoSmithKline/Fight for Sight Chair in Visual Neuroscience.

AUTHOR CONTRIBUTIONS

Conceptualization, S.W.F. and K.D.H.; methodology, S.W.F., M.C., and K.D.H.; investigation, S.W.F.; data curation, S.W.F.; formal analysis, S.W.F. and K.D.H.; funding acquisition, S.W.F., M.C., and K.D.H.; project administration, S.W.F.; resources, M.C. and K.D.H.; supervision, K.D.H.; visualization, S.W.F. and K.D.H.; writing, S.W.F., M.C., and K.D.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR METHODS

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

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
- [METHOD DETAILS](#)
 - Surgical procedure
 - Visuomotor association task
 - Recording visual responses in V1
 - Location of visual areas
 - Two-photon calcium imaging
 - Cell tracking
 - Sparse noise
 - Drifting gratings
 - Face recording
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
 - Pixel map of retinotopy
 - Visual sign maps
 - Pixel map of orientation responses
 - Cell detection
 - Characterizing single-cell orientation tuning
 - Discriminability index
 - Population sparseness
 - Orthogonalization of population responses
 - Dimensionality reduction
 - Stimulus prediction
 - Population code transformation equation
 - Pupil area and whisking
 - Significance testing

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2025.115235>.

Received: July 15, 2024

Revised: September 26, 2024

Accepted: January 6, 2025

Published: January 30, 2025

REFERENCES

1. Adab, H.Z., and Vogels, R. (2011). Practicing Coarse Orientation Discrimination Improves Orientation Signals in Macaque Cortical Area V4. *Curr. Biol.* 21, 1661–1666. <https://doi.org/10.1016/j.cub.2011.08.037>.
2. Chowdhury, S.A., and DeAngelis, G.C. (2008). Fine Discrimination Training Alters the Causal Contribution of Macaque Area MT to Depth Perception. *Neuron* 60, 367–377. <https://doi.org/10.1016/j.neuron.2008.08.023>.
3. Corbo, J., McClure, J.P., Erkat, O.B., and Polack, P.-O. (2022). Dynamic Distortion of Orientation Representation after Learning in the Mouse Primary Visual Cortex. *J. Neurosci.* 42, 4311–4325. <https://doi.org/10.1523/JNEUROSCI.2272-21.2022>.
4. Ghose, G.M., Yang, T., and Maunsell, J.H.R. (2002). Physiological Correlates of Perceptual Learning in Monkey V1 and V2. *J. Neurophysiol.* 87, 1867–1888. <https://doi.org/10.1152/jn.00690.2001>.
5. Goltstein, P.M., Meijer, G.T., and Pennartz, C.M. (2018). Conditioning sharpens the spatial representation of rewarded stimuli in mouse primary visual cortex. *eLife* 7, e37683. <https://doi.org/10.7554/eLife.37683>.
6. Goltstein, P.M., Coffey, E.B.J., Roelfsema, P.R., and Pennartz, C.M.A. (2013). In vivo two-photon Ca²⁺ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. *J. Neurosci.* 33, 11540–11555. <https://doi.org/10.1523/JNEUROSCI.1341-12.2013>.
7. Gu, Y., Liu, S., Fetsch, C.R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G.C., and Angelaki, D.E. (2011). Perceptual Learning Reduces Interneuronal Correlations in Macaque Visual Cortex. *Neuron* 71, 750–761. <https://doi.org/10.1016/j.neuron.2011.06.015>.
8. Henschke, J.U., Dylda, E., Katsanevakis, D., Dupuy, N., Currie, S.P., Amvrossiadis, T., Pakan, J.M.P., and Rochefort, N.L. (2020). Reward Association Enhances Stimulus-Specific Representations in Primary Visual Cortex. *Curr. Biol.* 30, 1866–1880.e5. <https://doi.org/10.1016/j.cub.2020.03.018>.
9. Jurjut, O., Georgieva, P., Busse, L., and Katzner, S. (2017). Learning Enhances Sensory Processing in Mouse V1 before Improving Behavior. *J. Neurosci.* 37, 6460–6474. <https://doi.org/10.1523/JNEUROSCI.3485-16.2017>.
10. Poort, J., Wilmes, K.A., Blot, A., Chadwick, A., Sahani, M., Clopath, C., Mrsic-Flogel, T.D., Hofer, S.B., and Khan, A.G. (2022). Learning and attention increase visual response selectivity through distinct mechanisms. *Neuron* 110, 686–697.e6. <https://doi.org/10.1016/j.neuron.2021.11.016>.
11. Poort, J., Khan, A.G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G.B., Mrsic-Flogel, T.D., and Hofer, S.B. (2015). Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron* 86, 1478–1490. <https://doi.org/10.1016/j.neuron.2015.05.037>.
12. Raiguel, S., Vogels, R., Mysore, S.G., and Orban, G.A. (2006). Learning to See the Difference Specifically Alters the Most Informative V4 Neurons. *J. Neurosci.* 26, 6589–6602. <https://doi.org/10.1523/JNEUROSCI.0457-06.2006>.
13. Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412, 549–553. <https://doi.org/10.1038/35087601>.
14. Yan, Y., Rasch, M.J., Chen, M., Xiang, X., Huang, M., Wu, S., and Li, W. (2014). Perceptual training continuously refines neuronal population codes

- in primary visual cortex. *Nat. Neurosci.* 17, 1380–1387. <https://doi.org/10.1038/nn.3805>.
15. Yang, T., and Maunsell, J.H.R. (2004). The Effect of Perceptual Learning on Neuronal Responses in Monkey Visual Area V4. *J. Neurosci.* 24, 1617–1626. <https://doi.org/10.1523/JNEUROSCI.4442-03.2004>.
 16. Makino, H., and Komiyama, T. (2015). Learning enhances the relative impact of top-down processing in the visual cortex. *Nat. Neurosci.* 18, 1116–1122. <https://doi.org/10.1038/nn.4061>.
 17. Makino, H., Hwang, E.J., Hedrick, N.G., and Komiyama, T. (2016). Circuit Mechanisms of Sensorimotor Learning. *Neuron* 92, 705–721. <https://doi.org/10.1016/j.neuron.2016.10.029>.
 18. Goodfellow, I., Bengio, Y., and Courville, A. (2017). *Deep Learning* (The MIT Press).
 19. Schlkopf, B., Smola, A.J., and Bach, F. (2018). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and beyond* (The MIT Press).
 20. DiCarlo, J.J., Zoccolan, D., and Rust, N.C. (2012). How does the brain solve visual object recognition? *Neuron* 73, 415–434. <https://doi.org/10.1016/j.neuron.2012.01.010>.
 21. Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. *Nat. Neurosci.* 17, 1410–1417. <https://doi.org/10.1038/nn.3807>.
 22. Rumyantsev, O.I., Lecoq, J.A., Hernandez, O., Zhang, Y., Savall, J., Chrapkiewicz, R., Li, J., Zeng, H., Ganguli, S., and Schnitzer, M.J. (2020). Fundamental bounds on the fidelity of sensory cortical coding. *Nature* 580, 100–105. <https://doi.org/10.1038/s41586-020-2130-2>.
 23. Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370, 140–143.
 24. Jeanne, J.M., Sharpee, T.O., and Gentner, T.Q. (2013). Associative learning enhances population coding by inverting inter-neuronal correlation patterns. *Neuron* 78, 352–363. <https://doi.org/10.1016/j.neuron.2013.02.023>.
 25. Stringer, C., Michaelos, M., Tsyboulski, D., Lindo, S.E., and Pachitariu, M. (2021). High-precision coding in visual cortex. *Cell* 184, 2767–2778.e15. <https://doi.org/10.1016/j.cell.2021.03.042>.
 26. Gdalyahu, A., Tring, E., Polack, P.-O., Gruver, R., Golshani, P., Fanselow, M.S., Silva, A.J., and Trachtenberg, J.T. (2012). Associative Fear Learning Enhances Sparse Network Coding in Primary Sensory Cortex. *Neuron* 75, 121–132. <https://doi.org/10.1016/j.neuron.2012.04.035>.
 27. Chu, M.W., Li, W.L., and Komiyama, T. (2016). Balancing the Robustness and Efficiency of Odor Representations during Learning. *Neuron* 92, 174–186. <https://doi.org/10.1016/j.neuron.2016.09.004>.
 28. Bordelon, B., and Pehlevan, C. (2022). Population codes enable learning from few examples by shaping inductive bias. Preprint at bioRxiv. <https://doi.org/10.1101/2021.03.30.437743>.
 29. Sinz, F.H., Pitkow, X., Reimer, J., Bethge, M., and Tolias, A.S. (2019). Engineering a Less Artificial Intelligence. *Neuron* 103, 967–979. <https://doi.org/10.1016/j.neuron.2019.08.034>.
 30. Wolpert, D.H. (1996). The Lack of A Priori Distinctions Between Learning Algorithms. *Neural Comput.* 8, 1341–1390. <https://doi.org/10.1162/neco.1996.8.7.1341>.
 31. Wolpert, D.H. (1996). The Existence of A Priori Distinctions Between Learning Algorithms. *Neural Comput.* 8, 1391–1420. <https://doi.org/10.1162/heco.1996.8.7.1391>.
 32. Hong, H., Yamins, D.L.K., Majaj, N.J., and DiCarlo, J.J. (2016). Explicit information for category-orthogonal object properties increases along the ventral stream. *Nat. Neurosci.* 19, 613–622. <https://doi.org/10.1038/nn.4247>.
 33. Majaj, N.J., Hong, H., Solomon, E.A., and DiCarlo, J.J. (2015). Simple Learned Weighted Sums of Inferior Temporal Neuronal Firing Rates Accurately Predict Human Core Object Recognition Performance. *J. Neurosci.* 35, 13402–13418. <https://doi.org/10.1523/JNEUROSCI.5181-14.2015>.
 34. Barak, O., Rigotti, M., and Fusi, S. (2013). The Sparseness of Mixed Selectivity Neurons Controls the Generalization–Discrimination Trade-Off. *J. Neurosci.* 33, 3844–3856. <https://doi.org/10.1523/JNEUROSCI.2753-12.2013>.
 35. Bimbard, C., Sit, T.P.H., Lebedeva, A., Reddy, C.B., Harris, K.D., and Carandini, M. (2023). Behavioral origin of sound-evoked activity in mouse visual cortex. *Nat. Neurosci.* 26, 251–258. <https://doi.org/10.1038/s41593-022-01227-x>.
 36. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* 22, 1677–1686. <https://doi.org/10.1038/s41593-019-0502-4>.
 37. Niell, C.M., and Stryker, M.P. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron* 65, 472–479. <https://doi.org/10.1016/j.neuron.2010.01.033>.
 38. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brain-wide activity. *Science* 364, 255.
 39. Willmore, B., and Tolhurst, D.J. (2001). Characterizing the sparseness of neural codes. *Netw. Bristol Engl* 12, 255–270.
 40. Treves, A., and Rolls, E. (1991). What determines the capacity of autoassociative memories in the brain? *Netw. Comput. Neural Syst.* 2, 371–397. <https://doi.org/10.1088/0954-898X/2/4/004>.
 41. Marr, D. (1969). A theory of cerebellar cortex. *J. Physiol.* 202, 437–470.
 42. Babadi, B., and Sompolinsky, H. (2014). Sparseness and Expansion in Sensory Representations. *Neuron* 83, 1213–1226. <https://doi.org/10.1016/j.neuron.2014.07.035>.
 43. Carandini, M. (2024). Sensory choices as logistic classification. Preprint at bioRxiv. <https://doi.org/10.1101/2024.01.17.576029>.
 44. Cooke, S.F., and Bear, M.F. (2014). How the mechanisms of long-term synaptic potentiation and depression serve experience-dependent plasticity in primary visual cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130284. <https://doi.org/10.1098/rstb.2013.0284>.
 45. Cooper, L.N. (2004). Theory of cortical plasticity. World Scientific. <https://doi.org/10.1142/5462>.
 46. Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science* 273, 1868–1871. <https://doi.org/10.1126/science.273.5283.1868>.
 47. Goris, R.L.T., Movshon, J.A., and Simoncelli, E.P. (2014). Partitioning neuronal variability. *Nat. Neurosci.* 17, 858–865. <https://doi.org/10.1038/nn.3711>.
 48. Lin, I.-C., Okun, M., Carandini, M., and Harris, K.D. (2015). The Nature of Shared Cortical Variability. *Neuron* 87, 644–656. <https://doi.org/10.1016/j.neuron.2015.06.035>.
 49. Buonomano, D.V., and Merzenich, M.M. (1998). CORTICAL PLASTICITY: From Synapses to Maps. *Annu. Rev. Neurosci.* 21, 149–186. <https://doi.org/10.1146/annurev.neuro.21.1.149>.
 50. Weinberger, N.M. (2004). Specific long-term memory traces in primary auditory cortex. *Nat. Rev. Neurosci.* 5, 279–290. <https://doi.org/10.1038/nrn1366>.
 51. Feldman, D.E., and Brecht, M. (2005). Map Plasticity in Somatosensory Cortex. *Science* 310, 810–815. <https://doi.org/10.1126/science.1115807>.
 52. Kato, H.K., Gillet, S.N., and Isaacson, J.S. (2015). Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance. *Neuron* 88, 1027–1039. <https://doi.org/10.1016/j.neuron.2015.10.024>.
 53. Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V., and Kilgard, M.P. (2011). Cortical Map Plasticity Improves Learning but Is Not Necessary for Improved Performance. *Neuron* 70, 121–131. <https://doi.org/10.1016/j.neuron.2011.02.038>.

54. Han, Y.K., Köver, H., Insanally, M.N., Semerdjian, J.H., and Bao, S. (2007). Early experience impairs perceptual discrimination. *Nat. Neurosci.* 10, 1191–1197. <https://doi.org/10.1038/nn1941>.
55. Thomas, M.E., P, L.C., Chaudron, Y.J.M., Cisneros-Franco, J.M., and Vil-lers-Sidani, É. de (2020). Modifying the adult rat tonotopic map with sound exposure produces frequency discrimination deficits that are recovered with training. *J. Neurosci.* 40, 2259–2268.
56. Harris, K.D., Quiroga, R.Q., Freeman, J., and Smith, S.L. (2016). Improving data quality in neuronal population recordings. *Nat. Neurosci.* 19, 1165–1174. <https://doi.org/10.1038/nn.4365>.
57. Pachitariu, M., Stringer, C., and Harris, K.D. (2018). Robustness of Spike Deconvolution for Neuronal Calcium Imaging. *J. Neurosci.* 38, 7976–7985. <https://doi.org/10.1523/JNEUROSCI.3339-17.2018>.
58. Hagiwara, K.M., Murakami, T., Yoshida, T., Tagawa, Y., and Ohki, K. (2015). Neuronal activity is not required for the initial formation and maturation of visual selectivity. *Nat. Neurosci.* 18, 1780–1788. <https://doi.org/10.1038/nn.4155>.
59. Peters, A.J., Fabre, J.M.J., Steinmetz, N.A., Harris, K.D., and Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature* 591, 420–425. <https://doi.org/10.1038/s41586-020-03166-8>.
60. Ringach, D.L., Hawken, M.J., and Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature* 387, 281–284. <https://doi.org/10.1038/387281a0>.
61. Lee, S.H., Kwan, A.C., Zhang, S., Phoumthipphavong, V., Flannery, J.G., Masmanidis, S.C., Taniguchi, H., Huang, Z.J., Zhang, F., Boyden, E.S., et al. (2012). Activation of specific interneurons improves V1 feature selectivity and visual perception. *Nature* 488, 379–383. <https://doi.org/10.1038/nature11312>.
62. Khan, A.G., Poort, J., Chadwick, A., Blot, A., Sahani, M., Mrsic-Flogel, T.D., and Hofer, S.B. (2018). Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nat. Neurosci.* 21, 851–859. <https://doi.org/10.1038/s41593-018-0143-z>.
63. Bannon, N.M., Chistiakova, M., and Volgushev, M. (2020). Synaptic Plasticity in Cortical Inhibitory Neurons: What Mechanisms May Help to Balance Synaptic Weight Changes? *Front. Cell. Neurosci.* 14, 204.
64. Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R.A., Huang, Z.J., and Stryker, M.P. (2014). A cortical circuit for gain control by behavioral state. *Cell* 156, 1139–1152. <https://doi.org/10.1016/j.cell.2014.01.050>.
65. Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. *Science* 345, 660–665. <https://doi.org/10.1126/science.1254126>.
66. Kuchibhotla, K.V., Gill, J.V., Lindsay, G.W., Papadoyannis, E.S., Field, R.E., Sten, T.A.H., Miller, K.D., and Froemke, R.C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.* 20, 62–71. <https://doi.org/10.1038/nn.4436>.
67. Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron* 100, 424–435. <https://doi.org/10.1016/j.neuron.2018.10.003>.
68. Burgess, C.P., Lak, A., Steinmetz, N.A., Zatka-Haas, P., Bai Reddy, C., Jacobs, E.A.K., Linden, J.F., Paton, J.J., Ranson, A., Schröder, S., et al. (2017). High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice. *Cell Rep.* 20, 2513–2524. <https://doi.org/10.1016/j.celrep.2017.08.047>.
69. Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: Flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* 2, 13. <https://doi.org/10.1186/1475-925X-2-13>.
70. Sereno, M.I., McDonald, C.T., and Allman, J.M. (1994). Analysis of Retinotopic Maps in Extrastriate Cortex. *Cerebr. Cortex* 4, 601–620. <https://doi.org/10.1093/cercor/4.6.601>.
71. Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L.F., Dalgleish, H., Carandini, M., and Harris, K.D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at bioRxiv, 061507. <https://doi.org/10.1101/061507>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed neural data	This paper	https://doi.org/10.6084/m9.figshare.27855423.v1
Experimental models: Organisms/strains		
tetO-GCaMP6s mice	Jackson Labs	Tg(tetO-GCaMP6s)2Niell
CaMK2a-tTA mice	Jackson Labs	B6.Cg-Tg(Camk2a-tTA)1Mmay/Dbo
Software and algorithms		
Custom analysis code	This paper	https://doi.org/10.5281/zenodo.1420562 https://github.com/kdharris101/orthogonalization-visual-cortex

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All experimental procedures were conducted according to the UK Animals Scientific Procedures Act (1986). Experiments were performed at University College London under personal and project licenses released by the Home Office following appropriate ethics review. CaMK2a-tTA (Strain #:007004) and tetO-GCaMP6s (Strain #024742) mice were obtained from Jackson labs. Mice were housed in individually-ventilated cages at no more than four per cage, with free access to food, and with free water until the study began. Mice were under a 12-h light-dark cycle (light on from 8 a.m. to 8 p.m.). Five mice were used in the study (3 males and 2 females). Comparison of task performance or brain activity between sexes was not attempted due to lack of statistical power.

METHOD DETAILS

Surgical procedure

Five transgenic adult mice (60 days or older) expressing GCaMP6s in excitatory neurons (CaMK2a-tTA; tetO-GCaMP6s) underwent a procedure to implant cortical windows over right primary visual cortex (V1). Mice were anesthetized with isoflurane, an ophthalmic ointment was applied to the eyes, and injections of carprofen and dexamethasone were administered. The hair on the head at the planned incision site was shaved away, and the mouse was transferred to a stereotaxic apparatus where its skull was secured with ear bars. The scalp was cleaned with 70% ethanol to remove loose hairs and other detritus, after which a lidocaine ointment was applied. Following a final application of iodine and ethanol, the scalp over visual cortex was excised, and the edges of the incision were sealed to the skull with a cyanoacrylate adhesive. Using dental acrylic resin, a sterilized metal head plate with a circular well was cemented onto the skull. A 4 mm circular craniotomy was made over right V1 using a biopsy punch, and a glass window was sealed in place with a cyanoacrylate adhesive and dental acrylic resin. At the end of the procedure, mice were removed from anesthesia and placed on a heating pad to recover. Carprofen was added to the mice's drinking water for three days following surgery to mitigate post-operative pain, and mice were checked daily for any adverse outcomes.

Following recovery, mice were habituated for handling and head-fixation before carrying out recordings.

Visuomotor association task

The task is a modification of a two-alternative forced choice contrast discrimination task previously developed by our lab.⁶⁸ Mice were head-fixed with their body and hindlimbs resting on a stage, leaving their front forepaws free to turn a small wheel clockwise or anticlockwise. Three computer screens surrounded the mouse, spanning -135 to +135 visual degrees (deg) along the azimuth axis and -35 to +35 v° along the elevation axis. Trials began after 1–2 s of continuous quiescence (no wheel movement), after which two full contrast Gabors with sigmas of 18 deg and spatial frequencies of 0.04 cycles/deg were presented simultaneously and centered at -80 and +80 deg azimuth. These Gabors were randomly oriented at either 45°, 68°, or 90°, though the pair were never identical. After an additional quiescence period of approximately 1 s, an auditory cue (12 kHz, 100 ms) would sound, signaling to the mouse that the horizontal position of the Gabors could be manipulated via wheel movement. If the mouse moved the wheel before the auditory cue, the Gabors remained stationary while the quiescence requirement remained in force. When a Gabor was moved to the center screen, a choice was recorded for that trial, and a feedback period was initiated. Correct choices (driving a 45° stimulus to the center, or a 90° stimulus away) were rewarded with 1–5 µL of water and a short 0.25 s delay, while incorrect choices (driving a 90° stimulus to the center, or a 45° stimulus away) resulted in a 1–2 s burst of white noise. The Gabor was locked at the center position during the

feedback, following which it would disappear, and the next pre-trial period of enforced quiescence would begin. During task training, mice were water restricted in line with the approved project license. Mice were considered proficient at the task when they consistently made the correct choice on over 70% of trials.

Recording visual responses in V1

Two sessions of two-photon calcium imaging were performed: one before task training (naive) and one after mice had achieved high performance in the task (proficient). Imaging in the proficient condition was performed immediately after a behavioral session and in the same apparatus.

Location of visual areas

Prior to the first two-photon imaging session, we determined the location of V1 in each mouse's cortical window by recording cortical responses to sparse noise under mesoscopic wide-field calcium imaging and then generating a visual sign map, as previously described.⁵⁹ Mice were placed on a stage of the same type used in the task, and white squares of width 7.5° visual degrees were shown on a black background at a frame rate of 6 Hz for 10 min. Squares appeared randomly at fixed positions in a 12 by 36 grid, spanning the retinotopic range of the computer screens. 12% of the squares shown at any one time.

Two-photon calcium imaging

Layer 2/3 in V1 was imaged using a commercial two-photon microscope (Bergamo II, Thorlabs Inc) controlled by ScanImage.⁶⁹ A Ti:sapphire laser (Chameleon Vision, Coherent) was set to a wavelength between 940 and 980 nm, and the beam was focused with a 16X water-immersion objective (0.8 NA, Nikon). Images were acquired at a frequency of 30 Hz across six planes (5 Hz per plane), a resolution of 512 x 512 pixels, with a frame width between 730 and 810 μm. The fly-back plane was excluded from further analysis. During recordings, mice were head-fixed and placed on the same type of stage used for the task. Three computer screens surrounded the mouse, spanning -135 to +135 v° along the azimuth axis and -35 to +35 v° along the elevation axis.

Cell tracking

Neurons were tracked between naive and proficient recording sessions using a custom-built toolbox written in Python (<http://github.com/sfaior/srt4s2p>). Mean pixel intensity images, pixel correlation maps, and ROI maps were visualized, and recognizable cells and other landmarks were designated in a GUI. The coordinates of these matches were used to perform a perspective transformation to align the naive and proficient ROI maps. The aligned maps were compared and ROIs whose overlaps met a threshold were manually curated.

Sparse noise

To map the retinotopy of V1 under two-photon imaging (Figure 1C, middle), sparse noise stimuli were presented. Black or white squares of width 4.5° visual degrees were shown on a gray background at a frame rate of 5 Hz for 8–30 min. Squares appeared randomly at fixed positions in a 16 by 60 grid, spanning the retinotopic range of the computer screens. 1.5% of the squares were shown at any one time.

Drifting gratings

At least 16 blocks of drifting grating stimuli were presented in each recording. In each block, gratings spanning 16 directions (22.5° intervals) and a blank stimulus were each presented once in a randomized sequence. Each grating lasted 2 s, with an inter-trial interval sampled randomly from a uniform distribution with a range of 2–3 s. Drifting gratings were full contrast and sinusoidal, with a spatial frequency of 0.04 cycles/deg and a temporal frequency of 4 cycles/s, that either encompassed all three screens (full-field, three mice) or the entire left screen (two mice), contralateral to the recorded hemisphere. Data from the two directions for each of the eight orientations covering 180° were analyzed together.

Face recording

An infrared LED illuminated the mouse's face, and a camera with an infrared filter was used to capture any changes in pupil area or whisking behavior.

QUANTIFICATION AND STATISTICAL ANALYSIS

Pixel map of retinotopy

To obtain a retinotopic map of the two-photon imaging frame (Figure 1C middle; Figure S8A), we analyzed the two-photon recordings during sparse noise stimuli on a pixel-by-pixel basis, without cell detection. Analyses were performed after singular value decomposition (SVD) to accelerate the computation and denoise the data, producing valid results as these computations are linear. First, we z-scored each pixel's time course independently. Next, we applied single-value decomposition (SVD) on the z-scored image frames, $F = USV^T$, where F was the full movie encoded as a matrix of size $N_{pixels} \times T$, U was size $N_{pixels} \times N_{SVDs}$, S was a diagonal matrix of

singular values, and V was size $T \times N_{SVDs}$ with T being the number of two-photon imaging frames. A matrix Y was computed summarizing the mean response of each of the first 100 columns of V to each noise frame, as the time-averaged activity in a window 0.2 to 0.6 s after stimulus onset minus the time-averaged activity in a 1 s pre-stimulus window. This matrix was of size $F \times 100$, where F is the number of noise stimulus frames. The dependence of these responses on individual noise pixels was estimated using ridge regression: $\beta = (X^T X + \lambda I)^{-1} X^T Y$, where X was a $F \times N_{noise_squares}$ matrix containing 1 if a particular square was white or black on a particular frame (0 if it was gray), λ was a ridge parameter ($\lambda = 100$), and I was the identity matrix. The stimulus dependence of each pixel was then obtained by matrix multiplication $R = US\beta$, resulting in a matrix R of size $N_{pixels} \times N_{noise_squares}$, encoding the receptive field map of each 2p imaging pixel. To generate retinotopic maps of the imaging frame, each pixel's receptive field map was smoothed with a Gaussian (sigma 12 v°) and a peak was found, giving a retinotopic position along the elevation and azimuth axes for each pixel.

Pixel retinotopy maps were used to ensure that the two-photon imaging frames were retinotopically aligned with the position of the left task stimulus (0 v° elevation, -80 v° azimuth) during drifting grating recordings. When the optimal imaging location in V1 was identified in naive mice, an image of the cortical vasculature was saved for positioning subsequent imaging experiments.

Visual sign maps

Due to the retinotopic eccentricity of the imaging location in V1 and the large field of view used, it was occasionally the case that areas outside V1 were also recorded. To differentiate V1 from adjacent visual areas, visual sign maps were obtained using the above pixel retinotopy maps averaged across planes (Figure S8). First, elevation and azimuth maps were smoothed with a median (width 10 pixels) and a Gaussian (sigma 60 pixels) filter. Similar to the process described in ref. 70, the sine of the difference in angle between the gradients of the elevation and azimuth maps was calculated. This sign map was then thresholded to values above 0.31, and pixels that were members of the largest patch were considered to be in V1. This process was consistent in isolating V1, as verified by visual inspection of the elevation and azimuth retinotopic maps.

Pixel map of orientation responses

To obtain a pixel map of oriented grating responses (Figure S8B), the average df/f of each pixel was calculated in response to each stimulus orientation. For each trial, df was defined as the average fluorescence in a post-stimulus window spanning 0–2 s, minus the baseline defined as the average fluorescence in a pre-stimulus window spanning -1 to 0 s relative to stimulus onset. This value was divided by f_0 , the baseline measurement. To isolate neuropil responses (Figure S8D), only pixels that did not belong to a cell, as determined by Suite2P and subsequent manual curation, were included in the analysis.

Cell detection

Registration, cell detection, neuropil correction, and deconvolution of the two-photon imaging data were carried out using Suite2P.⁷¹ Imaged planes were aligned with non-rigid registration (four blocks, 128 × 128), and spiking activity was deconvolved from calcium fluorescence using a kernel with a timescale of 2 s.

Characterizing single-cell orientation tuning

All cells identified by Suite2P were analyzed for orientation responses. First, each cell's trial responses were computed by time-averaging its deconvolved activity on each trial over a window of width 0–2 s from drifting grating onset. Next, the mean response of each cell to each orientation and to the blank stimulus was computed by averaging over the respective stimulus trials. These mean responses were then normalized by dividing by its mean response to its preferred stimulus condition (either the modal orientation, or mean activity during the blank stimulus if this exceeded mean responses to all orientations).

A cell's orientation preference was defined in two ways: the orientation it responded maximally to (modal preferred orientation) or its mean preferred orientation, the argument of the complex number $z = \frac{\sum r_\theta e^{2i\theta}}{\sum r_\theta}$, where r_θ is the cell's mean response to orientation θ .

The orientation selectivity of a cell was defined as the modulus of z . To determine the tuning curve of each cell as a function of its orientation preference and selectivity (Figures 4F–4G), a cross-validated approach was used to avoid erroneously detecting tuning due to random fluctuations in responses. Each cell's preferred mean orientation and selectivity were calculated using odd-numbered trials, and tuning curves were generated using the mean response to each orientation on even-numbered trials.

Tuning curve slope (Figures 4H and 4I) was quantified as the difference between the cell's response at a stimulus orientation, and the orientation 22.5° closer to the cell's mean preferred orientation, divided by 22.5. The cell's tuning curve slope at its preferred mean orientation was defined as the difference between orientations -22.5° or +22.5° from preferred, divided by 45. Thus, if these responses were equal, the tuning curve slope at the preferred orientation would be zero. Tuning curve slopes were separately averaged for all stimulus orientations, and all preferred mean orientations, resulting in nine values per stimulus, per recording. These were displayed as a function of the difference between the stimulus and the mean orientation preference (i.e., -90 to 90 with -90 and 90 being identical). To compare slope magnitudes between stimuli and training conditions, we applied a t test to the absolute value of the slopes.

Discriminability index

The discriminability index (d') of a cell, its ability to discriminate between two orientations (θ_a and θ_b), was defined as $\sqrt{\frac{\mu_{\theta_a} - \mu_{\theta_b}}{\sigma_{\theta_a}^2 + \sigma_{\theta_b}^2}}$

where μ and σ^2 are the mean and variance of the respective orientation responses. The mean and variance for each stimulus orientation was the average of the mean and variance of the two corresponding stimulus directions.

Population sparseness

Population sparseness was calculated using the Treves-Rolls formula,³⁹

$$S = 1 - \frac{\left(\sum_{i=1}^N r_i / N \right)^2}{\sum_{i=1}^N (r_i^2 / N)},$$

where r_i is the response of neuron i , and N is the number of neurons. The Treves-Rolls measure can be thought of as counting the fraction of cells with near-zero activity. For example, if M of N cells respond with equal rate, and the remaining $N-M$ cells are silent, this measure takes the value $S = \frac{N-M}{N}$.

Orthogonalization of population responses

To calculate the orthogonalization of population responses between different stimulus orientations (Figures 3D and S3D), we split the trials into odd and even halves, and computed the N_{cells} -dimensional population response vectors $P_i(\theta)$ to orientation θ for the trial set i ($i = 1$: odd trials; $i = 2$: even trials). We computed the cosine similarity between orientations θ_1 and θ_2 as $\frac{P_1(\theta_1) \cdot P_2(\theta_2)}{\|P_1(\theta_1)\| \|P_2(\theta_2)\|}$. This process resulted in an eight-by-eight matrix of similarity values for each mouse and training condition. Computing this similarity between two separate halves ensured that the diagonal was not 1 by definition.

Dimensionality reduction

To illustrate the decodability of population responses in a 2-dimensional plot (Figure S3A), we trained a linear regression model to decode a 2-dimensional vector $(\cos 2\theta, \sin 2\theta)$ for each trial from the N_{cells} -dimensional population response vector on that trial, where θ is the stimulus orientation ranging from 0° to 180° . The model was trained on odd trials, and then applied to population responses on even trials to obtain a two-dimensional projection of population activity that separates points by stimulus orientation.

To visualize the orthogonalization of the 45° and 90° population responses (Figure 3C), we split the responses to 45° and 90° into odd- and even-numbered trials. Projections were determined by singular value decomposition of the odd-numbered trials, and the even-numbered trials were projected onto the resulting top two components.

Stimulus prediction

Orientation was also decoded from population activity using linear discriminant analysis (LDA; Figures 2B–2C and S3B). An LDA model was fit using the population responses in odd trials, and its performance was assessed on even trials. To build the model, we used the class *LinearDiscriminantAnalysis* from the Python library scikit-learn, with solver set to “eigen” and the shrinkage coefficient automatically calculated.

Population code transformation equation

For each mouse, cells in the naive and proficient recordings were divided into classes by binning mean orientation preference (eight bins, 0° : 168.75° – 11.25° , 23° : 11.25° – 33.75° , 45° : 33.75° – 56.25° , 68° : 56.25° – 78.75° , 90° : 78.75° – 101.25° , 113° : 101.25° – 123.75° , 135° : 123.75° – 146.25° , 158° : 146.25° – 168.75°) and selectivity (five bins, 0 – 0.16 , 0.16 – 0.32 , 0.32 – 0.48 , 0.48 – 0.64 , 0.64 – 1). The mean response of each cell class to each stimulus was determined by cross-validation, using odd trials to determine the cell’s tuning class, and using even trials to compute its tuning, as described above. The mean response of each cell class c to each stimulus θ in naive mice is denoted by $f_{c,\theta}$.

The population code transformation equation predicts the mean response in trained mice of each cell class to each stimulus as

$$f'_{c,\theta} = g_\theta(f_{c,\theta})$$

Here g_θ is a function which depends on the stimulus θ but is common across cell classes c . The functions were constrained to pass through $(0,0)$, $(a(\theta), b(\theta))$, and $(1,1)$, where $a(\theta)$ and $b(\theta)$ are parameters that vary between stimuli θ . This was achieved using the form:

$$g_{a,b}(x) = \begin{cases} xb/a, & r_n \leq a \\ (x-1)\frac{b-1}{a-1} + 1, & r_n > a \end{cases}$$

The parameters $a(\theta)$ and $b(\theta)$ were fit for each mouse and stimulus by nonlinear least squares (Python library SciPy, *optimize.curve_fit*), constrained to values between 0 and 1.

The convexity of the transformation from naive to proficient population responses to a stimulus θ was quantified as $C = \frac{m_{pref}}{m_{non-pref}} - 1$, where m_{pref} was the slope of a line from the origin to the point representing the cell class of preferred orientation θ and maximal selectivity, and $m_{non-pref}$ was the slope of a linear regression on the points corresponding to all cell classes of orientation preference other than θ . This approach was used to measure convexity on mean responses, relating the trial-averaged population response in the same mouse prior and after training (Figures 5B and 5C), and on single trials, where the population responses in single trial in a proficient mouse was compared to the trial-averaged population response in that mouse prior to training (Figure 6B).

To assess the consistency of trial-to-trial fluctuations in sparsening across the population (Figures 6C and 6D), we randomly divided the proficient cells into two populations balanced for orientation preference and selectivity. Trial-by-trial convexity was measured, as described above, for each cell population, and the correlation coefficient of these convexities was computed. This process was repeated 2000 times, and the average correlation in convexity over orientations was found for each mouse. This approach of independently computing the convexity from two non-overlapping cell populations ensured that the measured trial-to-trial variability did not arise from noise or overfitting, which would be independent between the two groups, but instead reflected genuine and consistent transformations of population activity.

Pupil area and whisking

Facial video recordings were processed with the toolkit FaceMap (www.github.com/MouseLand/FaceMap) to obtain traces of pupil area and whisking intensity. The pupil area was defined as the area of a Gaussian fit on thresholded pupil frames, where pixels outside the pupil were set to zero. Whisking intensity was defined as the average change in individual pixels between frames for a region of interest limited to the whisker pad. From these resulting traces, trial-evoked changes in pupil area and whisking were calculated. First, traces of pupil area and whisking intensity were z-scored for each recording session. Second, for each trial pupil area and whisking were averaged in a post-stimulus time window spanning 0 to 2 s. Lastly, stimulus-evoked changes in pupil area and whisking were calculated by subtracting a pre-stimulus baseline, defined as the average pupil area and whisking in a -1 to 0 s window.

Significance testing

Simultaneously recorded neurons are not statistically independent, so analyses of them must use one of two approaches.

The first approach is to treat the experiment, rather than the neuron, as the unit of independent variability: to summarize the activity of all cells in an experiment in one number, and to use a traditional method such as a t test, with N being the number of experiments rather than the number of neurons. We did this for the analyses in figures: 2C-E, 2I, 3C, 4D, 4G, 5B, 5D, 6B, S2B, S2D, S3C, S4E, S6E, S8C-D, A1H-K.

The second approach is to use hierarchical models, such as linear mixed effects models, which account for correlated differences in all cells of one recording. We did this for the analyses in figures: 2B, 4E, 5E, S4B, S5A-D, A1A-D, using the following models: 2B: change in response ~ orientation (1 + mouse); 4E: trial convexity ~ whisking * stimulus (1 + mouse); 5E: cosine similarity ~ population sparseness + condition (1 + mouse | condition); S4B: stable ~ stimulus + (1 + mouse) where stimulus was composed of the stimuli being compared within a selectivity bin; S5A-D: trial convexity ~ behavior * stimulus (1 + mouse); A1A-D: mean response/coefficient of variation/response index/d-prime ~ condition + (1 + mouse | condition).