Supplementary methods

Animal preparation

Ferrets were anesthetized with ketamine (50mg/kg), shaved and scrubbed. The femoral vein was cannulated for delivery of 5% dextrose in lactated ringers and paralytic, a tracheotomy performed, and the ferret placed in a stereotaxic head frame. A mixture of nitrous oxide and oxygen (2:1) with isoflurane (2-2.5%) was administered and adjusted if necessary based on EKG and expired CO₂. Body temperature was maintained at 37°C, and silicone oil was used to protect the corneas. A craniotomy was performed above primary visual cortex and the dura and arachnoid mater removed.

In intrinsic imaging experiments, a stainless steel chamber 2cm in diameter was mounted to the skull around the craniotomy using tissue adhesive and dental acrylic. The chamber was filled with saline and sealed with a glass coverslip. In 2-photon experiments, a wire mesh was attached to the skull with epoxy to provide an anchor point for dental caulk (Densply), and a stainless steel head plate with a 1 cm opening was cemented above the craniotomy site.

After completing surgical procedures, incisions and pressure points were infiltrated with bupivacaine and isoflurane concentration was lowered to 0.75-1%. The animal was paralyzed with rocuronium bromide to prevent eye movements.

Visual stimulation

Visual stimulation was provided by an Apple Macintosh PowerMac G4 and a GDM-C520 CRT (Sony), gamma-corrected using ColorVision Spyder (Pantone, Carlstadt, NJ), mean. Custom stimulation software was developed using Matlab (MathWorks, Natick, MA) and the Psychophysics Toolbox¹,². All stimuli were grayscale sinusoidal gratings at 100% contrast with spatial frequency 0.04-0.08 cycles per degree and temporally modulated at 4Hz. Spatial and temporal frequency parameters were chosen to match optimal values observed in pilot intrinsic imaging experiments in naïve ferrets. Stimuli were 5 s in duration with a 5 s interstimulus interval. For direction tuning curves, gratings drifted in a direction orthogonal to the orientation of the grating. For flash stimulation, stimuli were modulated sinusoidally between the grating and a gray background (luminance 45 cd/m²).

For motion training, drifting gratings that moved in one of two opposite directions were shown for 5 s, followed by a 10 s interstimulus interval. For each 5 s stimulus, one of the two opposite directions was pseudorandomly selected for presentation. Stimulation continued for 20 minutes followed by 10 minutes of no stimulation. This protocol was maintained for 12-24 hours for intrinsic signal imaging and 3-6 hours for 2-photon imaging.
The flash training protocol was identical to motion training except that stimuli did not drift and instead were sinusoidally modulated between the grating and a gray screen. Spatial phase was chosen randomly for each flash stimulus.

**Intrinsic-signal optical imaging**

The cortical surface was visualized through a tandem lens macroscope attached to a low-noise CCD camera, and the cortex illuminated with 705 nm light. The timing of stimulus presentation and collection of images was controlled by software from Optical Imaging Inc. Data for 20 stimulus presentations were averaged to yield the final activity map for a given stimulus condition.

Images were analyzed with custom software written in MATLAB (Mathworks, Natick, MA). Difference images of direction or orientation selectivity were generated by subtracting images obtained from a pair of opposite directions of motion or a pair of orthogonal orientations, respectively. To assess selectivity, an orientation selectivity index (OSI) or direction selectivity index (DSI) was computed from the difference images by clipping the floating point distribution (at ± 3 s.d. from the median), calculating the s.d. of the distribution of gray levels within V1, and subtracting from this value the average selectivity index obtained in an identical manner from optical images acquired during presentations of a blank stimulus. Then, the value was normalized by the maximal OSI or DSI obtained from a cohort of adult ferrets (Li et al 2006) in order to restrict index values between 0 and 1 (see Supplementary Figure S1). All quantitative analyses of cortical response are performed on raw, 32-bit images. One-way analysis of variance was used to assess the significance of selectivity indices.

**Two-photon calcium imaging**

Calcium dye was prepared by dissolving 50 µg Oregon Green 488 BAPTA-1AM (Invitrogen) in 4 µl 20% pluronic acid (Invitrogen), vortexing for 2-3 minutes, diluting with 45 µl of a pipette solution (150 mM NaCl, 2.5 mM KCl, 10 mM Hepes, pH 7.4) and 1 µl of an AlexaFluor 594 solution (for visualization), vortexing for 2-3 additional minutes, and sonicating in a bath of ice water for 10-30 minutes. Finally, the solution was passed through a 0.22 µm-pore centrifuge filter (Millipore, UFC30GV25) and stored on ice for up to 2 hours before use. Alexa-Fluor 594 solution was made by dissolving 1 mg AlexaFluor 594 hydrazide (Invitrogen) in 660 µl of water. Both stock solutions were stored at room temperature for months.

Filamented glass pipettes were pulled on a vertical puller and tips broken to 3-5 µm. Pipettes were filled with 2 µl of calcium dye and were tested in the saline above the brain to ensure patency. Pipettes were introduced to the cortex at a 31° angle (MP-285 manipulator, Sutter Instruments) and initially inserted to a depth of 50-100 µm. Next, warm 2% agarose (Sigma, A9793-100G) in pipette solution was added to the craniotomy to reduce cortical pulsations. After the agarose solidified, pipettes were driven to a depth of 250-350 µm for injection. To avoid damage and to ensure clear imaging, care was...
taken to avoid being near surface blood vessels at both the insertion location and the injection site.

Injections were visualized by examining the AlexaFluor 594 signal with epi-fluorescent optics on our 2-photon microscope (light source X-cite 120, EXFO; filter cube exHQ575/50x, emHQ640/50, bsQ610lp, Chroma Technology; 4x objective UPLFLN4X, Olympus). Some 40-80 pressure pulses (5-12 psi, 0.5-2 s) were administered with an MPPI-2 pressure injector (Applied Scientific Instrumentation, Inc.), and the injection site was periodically checked to ensure that a 200-400 µm sphere of AlexaFluor 594 signal was evident. If no dye was apparent or if the initial flow was much faster than expected due to a broken pipette, the pipette was discarded and a new injection was made. Care was taken to minimize epi-fluorescent illumination of the injection site. Stosieck and colleagues³ have estimated that the concentration of OGB-1 after bulk-loading is about 20 µM, well below the 1-10 mM BAPTA concentration that is used in many experiments that seek to block synaptic plasticity (e.g., Sjöström et al., 2003⁶).

After successful dye delivery, the pipette was withdrawn and the brain was left undisturbed for 45-90 minutes to allow dye loading. Following this loading period, the agarose used for dye injection was removed and replaced with fresh warm agarose. A 5 mm circular coverslip (Warner Instruments) was gently pressed into the solidifying agarose such that the coverslip rim touched the bone around the craniotomy and so that the agarose slightly oozed out on top of the coverslip rim to create a stable imaging chamber.

Cells were imaged with a 2-photon microscope (Ultima IV, Prairie Technologies, Madison, WI) using 810 nm laser light (Chameleon, Coherent) and a 40 X saline-immersed objective (LUMPLFL40XW/IR, Olympus) with total output power < 50 mW. To shield stimulus monitor light from the microscope, the objective and head plate were covered with a custom-sewn sleeve made from light-tight fabric (Thor Labs) and elastic (Michaels Stores Inc.), and augmented as needed with light-tight tape (Thor Labs). Light block quality was assessed by setting photomultiplier gain to maximum, shuttering the laser, and scanning while turning the monitor on and off to verify that no modulation could be observed. During visual stimulation, 512 by 512 pixel image frames were acquired continuously every 1.5-1.8 s, and 10-12 repetitions of each stimulus were recorded.

Cells were recorded in layer 2/3 at depths ranging from 120 µm to 300 µm. In all cases, we examined the cytoarchitecture at the surface and focused downward to be sure we were lower than cortical layer 1. Often, images were acquired at more than one depth in each animal, and these recordings were displaced in Z by at least 20-40 µm to avoid recording the same cells twice. All recordings were carefully examined to ensure stability in the Z axis, and any recordings showing significant Z drift were discarded.

After training, we manually matched each recorded field with images that were acquired before training. Many cells were clearly visible in both recordings, but due to slight changes in imaging position, cells could disappear or appear. The total amount of
imaging time was limited by how long the calcium dye fluorescence remained intense; this varied from as little in 5 hours in some experiments to as long as 10 hours.

**Two-photon imaging analysis**

Images were analyzed with custom software written in Matlab and C. Small horizontal drift over time was corrected by measuring correlation at different pixel offsets and realigning the images according to the best match. Cells were identified by the experimenter. For each frame, the fluorescence value of each cell was computed by averaging all pixels in a circle of radius 12 pixels (5.5 µm) that was centered on the soma. These circular regions of interest were smaller than the soma itself to reduce possible intrusion of signals from the surrounding neuropil.

Image series taken at different times were aligned with the correlation procedure above. Each cell was examined to see if it was still clearly present in the second recording; that is, it remained distinct from background and retained its spatial position relative to its neighbors (see Figure 3a). The correlation procedure was generally successful in identifying the cell center location in the second recording, although small manual corrections were made as necessary. The entire process of cell identification was done blind to subsequent analyses of physiological responses.

The response to each stimulus was calculated as $\frac{\Delta F}{F} = \frac{(F_{\text{STIM}} - F_0)}{F_0}$, where $F_{\text{STIM}}$ is the average response during each frame when the stimulus was on, and $F_0$ is the average response during the final 3 seconds of the interstimulus period before stimulus onset.

Cells with significant responses were identified by plotting each trial's response as a point in orientation space: $\sum_{i} S(\theta_i)e^{\frac{2\pi i \theta_i}{180^\circ}}$, where $S(\theta_i)$ is raw $\frac{\Delta F}{F}$ at direction $\theta_i$. Cells were considered significant responders if the mean of the points was different from (0, 0) by Hotelling's $t^2$-test with $P < 0.05$.

Responses to drifting gratings were fit with a 2-peak Gaussian function:

$$R(\theta) = R_{\text{OFFSET}} + R_{\text{PREF}}e^{-\frac{\text{ang}(\theta - \theta_{\text{PREF}})^2}{2\sigma^2}} + R_{\text{OPP}}e^{-\frac{\text{ang}(\theta + 180^\circ - \theta_{\text{PREF}})^2}{2\sigma^2}},$$

where $R_{\text{OFFSET}}$ is a constant offset, $\theta_{\text{PREF}}$ is the preferred direction angle, $R_{\text{PREF}}$ is the above-offset response to the preferred direction, $R_{\text{OPP}}$ is the above-offset response to the opposite direction, $\sqrt{\log 4\sigma}$ is the tuning width (half-width at half height) and $\text{ang}(x) = \min(x, x - 180, x + 180)$ wraps angular difference values onto the interval 0° to 180°.
Direction index values were calculated as

\[ DI = \min(\frac{\text{PREFERRED} - \text{OPPOSITE}}{\text{PREFERRED} - \text{BLANK}}, 1) \]

where PREFERRED = \( R(\theta_{\text{PREF}}) \) and OPPOSITE = \( R(\theta_{\text{PREF}} + 180) \). This direction index varies from 0 (no direction selectivity) to 1 (no response to opposite direction). The minimum function restricts DI to be between 0 and 1 in some cells where the OPPOSITE response was below BLANK.

Orientation index values were calculated as

\[ OI = \min(\frac{\text{PREFERRED\_ORI} - \text{ORTHOGONAL}}{\text{PREFERRED\_ORI} - \text{BLANK}}, 1) \]

where PREFERRED\_ORI = mean(\( R(\theta_{\text{PREF}}), R(\theta_{\text{PREF}} + 180) \)) and ORTHOGONAL = mean(\( R(\theta_{\text{PREF}} + 90), R(\theta_{\text{PREF}} - 90) \)). Once again the minimum function restricts OI to be between 0 and 1 in cells when ORTHOGONAL is below BLANK.

**Timecourse differences between intrinsic imaging and 2-photon**

Increases in direction selectivity were observed after 3-6 hours of motion training in 2-photon recordings, but were not evident until 8 hours of training in intrinsic imaging experiments. We assume these differences reflect the higher signal to noise ratio and higher spatial resolution in 2-photon recordings as compared to intrinsic imaging.

**Statistical abbreviations and analysis**

“KW test” is Kruskal Wallis non-parametric test, “t-test” is student’s t-test, “ANOVA” is analysis of variance type I, “bootstrap test” is described below. Null hypotheses were rejected if \( \alpha < 0.05 \). Actual P values are reported except when P < 0.001, in which case we report “P < 0.001”. The Kruskal-Wallis test was used when making multiple comparisons of DI, OI, and LCI; if the multiple comparison showed evidence of significant differences, the Kruskal-Wallis test was again used to evaluate pairwise comparisons. The Bonferroni correction was applied when examining points in the naïve and flash training curves in Figure 4b.

The direction index above is based on mean responses and does not consider variation across trials. To statistically test direction selectivity for each tuning curve, we used the bootstrap algorithm\(^8\) and simulated 100 recordings of 12 trials by drawing from the actual recorded trials randomly with replacement. We then calculated direction angle preferences and direction index values for these 100 simulated recordings. If the direction angle preference for each simulated tuning curve was greater than 90° from the median simulated direction preference, the DI was multiplied by –1. These 100 values served as an estimated "likelihood distribution" for the true value of the latent variable.
DI, which can be inferred but not directly observed. We took DI to be significantly greater than 0 if 95% of the simulated values were greater than 0. We considered DI after training to be significantly greater than before training if 95% of DI values before training were less than the 5th percentile value of DI after training. For this test, the sign of the DI values for both tuning curves was determined by the "after" tuning curve.

**Local coherency index and likelihood of direction angle reversal**

Local coherency index for a cell was defined to be the percentage of nearby neurons that had the same direction angle preference (within 45°) minus the percentage of nearby neurons that had the opposite direction angle preference (within 45°). Distance was measured in the horizontal plane: if multiple depths were recorded at the same point on the cortical surface, only horizontal distance between cells was considered. For each cell, we computed 100 values of the local coherency index by randomly drawing the direction angle preference of every neighboring cell from the likelihood distribution estimated by the bootstrap procedure above. We took the median of these 100 values to be the local coherency index for each cell. We used this Monte Carlo approach because we felt it provided a better estimate of the true local coherency index rather than simply using the raw data. These simulations take into account the biological and measurement uncertainty of each cell's direction angle preference.

To estimate the likelihood that a direction angle reversal occurred after training, we calculated the percentage of the 100 bootstrap simulations of direction tuning curves from before and after training that exhibited a change in direction of 90° or greater.

**Equation relating changes in local coherence to direction preference reversal, changes in uncertainty, and changes in orientation preferences**

Due to intrinsic biological variability and measurement noise, we do not know the “true” latent direction preferences of the cells in our sample, so we have employed the bootstrap algorithm to produce 100 likely values of the true direction preference before and after training. We computed the local coherence index (LCI) by calculating the percentage of neighboring cells with the same direction preference (within 45°) minus the percentage of neighboring cells with the opposite direction preference (within 45°), averaged across all simulations. Here we describe how to provide an account of what fraction of the change in local coherence following training is due to changes in orientation preference and what fraction is due to changes in direction preference. We then estimated the likelihood that the changes due to direction are due to reversals of direction preference and how much is due to changes in uncertainty (not due to reversals):
The equation for local coherence is as follows:

\[ LCI_i = \frac{100}{N_i} \sum_{s=1}^{N_s} \frac{1}{N_s} \sum_{j=1}^{N_i} \text{same}(\theta_{i,s}^a, \theta_{j,s}^a) - \text{opposite}(\theta_{i,s}^a, \theta_{j,s}^b), \]

where \( N_i \) is the number of neighbors of neuron \( i \), \( j \) loops over the neighbors of neuron \( i \), \( N_s \) is the number of simulations, and \( \theta_{i,s} \) indicates direction preference for neuron \( i \) for simulation \( s \). The function \( \text{same}(x,y) \) returns 1 if \( x \) and \( y \) are within 45° of one another and 0 otherwise, and \( \text{opposite}(x,y) \) returns 1 if \( x \) and \( y \) differ by more than 135° and 0 otherwise. The factor of 100 establishes percentage as the unit of measure.

If we let the superscripts \( b \) and \( a \) indicate values before and after training, respectively, then the change in LCI is as follows:

\[ \Delta LCI_i = \frac{100}{N_i} \sum_{s=1}^{N_s} \frac{1}{N_s} \sum_{j=1}^{N_i} \text{same}(\theta_{i,s}^a, \theta_{j,s}^a) - \text{opposite}(\theta_{i,s}^a, \theta_{j,s}^b) - (\text{same}(\theta_{i,s}^b, \theta_{j,s}^a) - \text{opposite}(\theta_{i,s}^b, \theta_{j,s}^b)). \]

For short, we can let the contribution of neuron \( j \) to \( \Delta LCI \) of neuron \( i \) be \( \Delta LCI_{i,j}^{\text{ori}} \). This contribution can be expressed as a sum of changes due to orientation and direction preferences:

\[ \Delta LCI_i = \frac{100}{N_i} \sum_{j=1}^{N_i} \Delta LCI_{i,j}^{\text{ori}} = \frac{100}{N_i} \sum_{j=1}^{N_i} \Delta LCI_{i,j}^{\text{ori}} + \Delta LCI_{i,j}^{\text{dir}}. \]

We want the term \( \Delta LCI_{i,j}^{\text{dir}} \) to represent changes due to direction only, so it should only include simulations where orientation preferences do not change; that is, where \( \text{same}(\theta_{i,s}^a, \theta_{j,s}^b) = 1 \) or \( \text{opposite}(\theta_{i,s}^a, \theta_{j,s}^b) = 1 \) and \( \text{same}(\theta_{j,s}^a, \theta_{j,s}^b) = 1 \) or \( \text{opposite}(\theta_{j,s}^a, \theta_{j,s}^b) = 1 \). The equation for this term is the following:

\[ \Delta LCI_{i,j}^{\text{dir}} = \frac{1}{N_i} \sum_{s=1}^{N_s} \text{same}(\theta_{i,s}^a, \theta_{j,s}^b) + \text{opposite}(\theta_{i,s}^a, \theta_{j,s}^b) \text{same}(\theta_{j,s}^a, \theta_{j,s}^b) + \text{opposite}(\theta_{j,s}^a, \theta_{j,s}^b) \times \text{same}(\theta_{i,s}^a, \theta_{j,s}^b) - \text{opposite}(\theta_{i,s}^a, \theta_{j,s}^b) - (\text{same}(\theta_{i,s}^b, \theta_{j,s}^a) - \text{opposite}(\theta_{i,s}^b, \theta_{j,s}^a)), \]

and the number of simulations that do not exhibit a change in orientation preferences is...
\[ \Delta LCI, = \Delta LCI^{\text{ori}} + \Delta LCI^{\text{reversal}} + \Delta LCI^{\text{unc}}, \]

where

\[ \Delta LCI^{\text{ori}} = \frac{100}{N_i} \sum_{j=1}^{N_i} \Delta LCI_{i,j}^{\text{ori}} = \frac{100}{N_i} \sum_{j=1}^{N_i} \Delta LCI_{i,j}^{\text{ori}} - \Delta LCI_{i,j}^{\text{dir}}, \]

\[ \Delta LCI^{\text{reversal}} = \frac{100}{N_i} \sum_{j=1}^{N_i} L_{i,j}^{\text{reversal}} \Delta LCI_{i,j}^{\text{dir}}, \] and

\[ \Delta LCI^{\text{unc}} = \frac{100}{N_i} \sum_{j=1}^{N_i} L_{i,j}^{\text{no reversal}} \Delta LCI_{i,j}^{\text{dir}}. \]
Supplementary Figures and Observations

Orientation responses measured with intrinsic signal imaging are unchanged after motion training

Supplementary Figure S1. Orientation selectivity as assessed by intrinsic imaging is not altered by training with motion stimuli (a) Orientation selectivity index (OSI) assessed at the start and end of motion training; each color codes the same animal as in Figure 1, arrows indicate median values. Unlike DSI, median OSI shows no significant change after training (t-test, p = 0.12). (b) After training, cortical responses in the trained and untrained orientation domains were measured; each color codes the same animal as in Figure 1. These data rule out the possibility that the cortical regions that were preferentially activated by the training stimulus are simply more responsive to visual stimuli and thus contribute to a stronger direction difference response.
Direction tuning curves before and after training

Supplementary Figure S2. Responses of 9 of 59 single cells that showed significant increases in direction index after motion training (bootstrap analysis). Error bars are SEM, fits are 2-peak Gaussians, dashed lines are mean responses to a gray screen. Blue curves are before motion training, red curves are after training. DI before and after is indicated above each pair of tuning curves.
Response time courses before and after motion training

Before training

- DI = 0.03
- DI = 0.032
- DI = 0.016
- DI = 0.061
- DI = 0.19

After motion training

- DI = 0.49
- DI = 0.52
- DI = 0.26
- DI = 0.66
- DI = 0.78

Supplementary Figure S3. Tuning curves and average response time courses for 5 of 59 cells before and after motion training. Average time courses are computed over all available frame data within a 1 s sliding window. Before training, responses at the preferred and opposite directions are relatively similar; that is, cells are primarily orientation-selective and are not strongly direction-selective. After training, the response to the preferred direction is much larger than the response to the opposite direction.
Single cell orientation angle preference changes slightly after motion training

Supplementary Figure S4. After motion training, orientation angle preferences are shifted slightly towards the training stimulus. (a) Orientation preference before and after motion training. Data have been rotated so training orientation is 90°. Orientation preference is a wrapped variable, so points near 0° and 180° are similar in preference. Solid line is unity line. (b) Change in orientation preference plotted against preference before training, rotated so training angle is 90°. Cells exhibit a significant tilt towards the training orientation, indicated by the negative slope of the regression line (F test, P < 0.001). Median tilt is 1.5° towards training orientation. (c, d) Orientation preference changes following flash training. There is no significant change in orientation preference following flash training (F test, P = 0.725).
Both motion stimuli and flash stimuli drive cortical activity

Supplementary Figure S5. (a) Mean responses to drifting gratings (motion) and static gratings that are modulated sinusoidally (flash). For drifting gratings, the direction with the largest response was used for calculating the mean for each cell. For flashing gratings, the spatial phase with the largest response was used. These values were then averaged for each animal, and the error bars indicate standard error over animals. The differences in mean responses were not significant (t-test, P=0.7). (b) Change in DI before and after training plotted against change in response magnitude to a grating in the cell’s preferred direction. No correlation was found, indicating that training induced increases in DI were not a function of changes in response magnitude.

Reliability of calcium imaging after training

Two lines of evidence suggest that signals acquired from 2-photon calcium imaging after training are reliable. First, the bootstrap analysis includes the effects of noise, so one can be confident that cells which show significant increases in direction selectivity in the bootstrap analysis do so despite noise present before and after training. Second, Figure 3C and Supplemental Figure S4 indicate that orientation indexes and orientation preferences are well matched before and after motion or flash training. So, in the orientation domain, responses before and after training are similar. Finally, one might imagine that increases in DI after motion training could be artifactual if responses were saturated before training or if systematic reductions in responsiveness occurred during motion training. However, Supplementary Figure 5b indicates there is no correlation between change in DI and change in response magnitude to a grating in the cell’s preferred direction, suggesting that these artifactual explanations are not applicable.
Choice of direction index does not alter conclusions

To ensure that our conclusions were robust and did not depend on a particular direction index or a particular definition of baseline fluorescence, we reanalyzed our single cell data using several common direction indices:

\[
DI_2 = \frac{\text{PREFERRED} - \text{OPPOSITE}}{\text{PREFERRED} - \text{BLANK}}, \text{ same as our } DI \text{ but without restricting to } [0,1],
\]

\[
DI_3 = \frac{\text{PREFERRED} - \text{OPPOSITE}}{\text{PREFERRED} + \text{OPPOSITE}}, \text{ a direction modulation index, and}
\]

\[
DI_4 = \text{abs} \left( \sum_{\theta_i} S(\theta_i) e^{i \frac{\pi \theta_i}{180}} \right), \text{ a vector index, where } S(\theta) \text{ is raw } \frac{\Delta F}{F} \text{ at direction } \theta_i.
\]

Our qualitative conclusions were identical regardless of which direction index was used. There was a significant difference between before and after motion training (Kruskal-Wallis test, \(P < 0.001\) for all), and there was no difference before and after flash training (\(P = 0.58\), \(P = 0.58\), \(P = 0.25\), respectively). Median values before training were 0.14, 0.33, and 0.07 for \(DI_2\), \(DI_3\), and \(DI_4\), respectively, and these increased to 0.39, 0.77, and 0.19 after motion training; after flash training, these values were 0.19, 0.37, and 0.09. Of 262 cells with significant responses before and after motion training, the number of single cells exhibiting significant direction selectivity increases by bootstrap statistics were 56 cells distributed among 4 of 5 animals for \(DI_2\), 58 cells distributed among 5 of 5 animals for \(DI_3\), and 28 cells distributed among 4 of 5 animals for \(DI_4\). The number of cells out of 135 showing significant increases after flash training was 2 with our standard DI, 1 with \(DI_2\), 2 with \(DI_3\), and 0 with \(DI_4\).
Training effect occurs in cells that are identified to be neurons

The fluorescent calcium dye Oregon Green BAPTA-1 AM is known to label both neurons and glia\(^9\). In one experiment, we incorporated the glial cell marker Sulforhodamine 101. In that experiment, 6 of 6 single cells that exhibited significant increases in direction index after training as assessed by the bootstrap test were confirmed to be neurons. In addition, glial cells are selective for stimulus orientation in ferret visual cortex, but the onset of their responses are delayed 2-4 seconds relative to those of neurons\(^10\). Many cells in our study exhibited short response latencies more characteristic of neurons rather than glia. Supplemental Figure S6 shows response time courses for the same cells that are plotted in Supplemental Figure S2.

Supplementary Figure S6. Response time courses suggest that many cells imaged with 2-photon microscopy were neurons rather than glia. Average response time courses to stimulation with a drifting grating in the preferred direction for 12 cells. Stimulus presentation time is shown as a solid black bar 5 s in duration. The responses have been smoothed with a sliding window 1.5 s wide that moves in steps of 0.5 s. It is most likely that these short-latency responses arose from neurons rather than glia\(^10\).
Responses below baseline are not artifacts

We sometimes observed responses that fell below the baseline response to a blank stimulus, particularly at orientations orthogonal to the preferred. Responses below the spontaneous firing rate at high contrasts have been commonly reported in extracellular recordings: e.g., in 14 of 45 complex cells in Sclar and Freeman’s study\textsuperscript{11} in cat (Figure 2), in 60\% of cells in Ringach’s et al.’s study\textsuperscript{12} in macaque monkey (Figure 9), and in 26 of 40 cells in Alitto and Usrey’s study\textsuperscript{13} in ferret (Figure 7).

Supplementary Figure S7. Responses below baseline. (a) Direction tuning curves of 3 example cells. (b) Response time courses to preferred stimulus, orthogonal orientation, and a blank screen. The drop below baseline in response to the orthogonal stimulus began at stimulus onset. (c) Unprocessed fluorescence measurements during single stimulus presentations for cells 1 and 2. Responses were consistently above baseline levels for the preferred stimulus but below baseline for orthogonal stimulus. Here, baseline has been drawn manually. Inset: individual frames for cell 1 for stimuli numbered 1-12 above following drift correction. Three frames were recorded at a sampling rate of 0.56Hz during each 5s stimulus presentation. The region of interest (ROI) was centered on the cell for all stimuli, indicating that measurements were not corrupted by uncorrected drift.
References