

Neuronal Adaptation Reveals a Suboptimal Decoding of Orientation Tuned Populations in the Mouse Visual Cortex

Miaomiao Jin, Jeffrey M. Beck, and Lindsey L. Glickfeld

Department of Neurobiology, Duke University Medical Center, Durham, North Carolina 27710

Sensory information is encoded by populations of cortical neurons. Yet, it is unknown how this information is used for even simple perceptual choices such as discriminating orientation. To determine the computation underlying this perceptual choice, we took advantage of the robust visual adaptation in mouse primary visual cortex (V1). We first designed a stimulus paradigm in which we could vary the degree of neuronal adaptation measured in V1 during an orientation discrimination task. We then determined how adaptation affects task performance for mice of both sexes and tested which neuronal computations are most consistent with the behavioral results given the adapted population responses in V1. Despite increasing the reliability of the population representation of orientation among neurons, and improving the ability of a variety of optimal decoders to discriminate target from distractor orientations, adaptation increases animals' behavioral thresholds. Decoding the animals' choice from neuronal activity revealed that this unexpected effect on behavior could be explained by an overreliance of the perceptual choice circuit on target preferring neurons and a failure to appropriately discount the activity of neurons that prefer the distractor. Consistent with this all-positive computation, we find that animals' task performance is susceptible to subtle perturbations of distractor orientation and optogenetic suppression of neuronal activity in V1. This suggests that to solve this task the circuit has adopted a suboptimal and task-specific computation that discards important task-related information.

Key words: behavior; calcium imaging; circuits; decision-making; decoding; perception

Significance Statement

A major goal in systems neuroscience is to understand how sensory signals are used to guide behavior. This requires determining what information in sensory cortical areas is used, and how it is combined, by downstream perceptual choice circuits. Here we demonstrate that when performing a go/no-go orientation discrimination task, mice suboptimally integrate signals from orientation tuned visual cortical neurons. While they appropriately positively weight target-preferring neurons, they fail to negatively weight distractor-preferring neurons. We propose that this all-positive computation may be adopted because of its simple learning rules and faster processing, and may be a common approach to perceptual decision-making when task conditions allow.

Introduction

Sensory processing supports the transformation of signals from the outside world into a neural code represented by the spiking activity of cortical neurons (Hubel and Wiesel, 1959; Dubner and Zeki, 1971; Desimone et al., 1984). Decades of causal and correlative studies suggest that these representations are the basis for

perceptual choice (Salzman et al., 1990; Schiller, 1993; Britten et al., 1996; Hung et al., 2005). Many theories have been put forth for how these representations might be used by downstream circuits to make perceptual choices (Georgopoulos et al., 1986; Pouget et al., 2003; Jazayeri and Movshon, 2006; Ma et al., 2006; Graf et al., 2011; Zaidi et al., 2014). However, there is little evidence demonstrating how sensory representations are actually combined, and what information is used, to compute a perceptual choice.

One way to determine the transformation between representations and perceptual choice is to manipulate sensory encoding and determine the effects on behavior. Thus, stimulus-specific adaptation, with its strong and predictable effects on sensory encoding (Müller et al., 1999; Dragoi et al., 2000), can be a useful tool for evaluating how sensory information is used to guide perceptual choice. For instance, by preferentially decreasing the contribution of neurons that prefer the adapter orientation, adaptation will increase the discriminability of the neuronal repre-

Received Dec. 18, 2018; revised Feb. 15, 2019; accepted Feb. 21, 2019.

Author contributions: M.J. and L.L.G. designed research; M.J. and L.L.G. performed research; M.J., J.M.B., and L.L.G. analyzed data; M.J., J.M.B., and L.L.G. edited the paper; M.J. and L.L.G. wrote the paper; L.L.G. wrote the first draft of the paper.

This work was supported by an NIH Director's New Innovator Award (DP2-EY025439), the Pew Biomedical Trusts, and the Alfred P. Sloan Foundation (L.L.G.). We thank B. Gindley and J. Sims for assistance with behavioral training; K. Leonard, M. Fowler, and J. Isaac for surgical assistance; Z. Xu for assistance with software development; C. Hass, G. Field, G. Horwitz, C. Hull, J. Isaacson, S. Lisberger, M. Scanziani, A. Wilson, and X. Yao, for helpful discussions and comments on the paper.

The authors declare no competing financial interests.

Correspondence should be addressed to Lindsey L. Glickfeld at glickfeld@neuro.duke.edu.

<https://doi.org/10.1523/JNEUROSCI.3172-18.2019>

Copyright © 2019 the authors

sentations of the adapter stimulus and other nearby stimulus values while decreasing the discriminability of more distant stimulus values (Müller et al., 1999; Dragoi et al., 2000; Kohn and Movshon, 2004; Stocker and Simoncelli, 2006). Indeed, common perceptual illusions, such as the tilt after-effect and the waterfall illusion, are consistent with the resulting repulsive effects of adaptation on tuned populations (Levinson and Sekuler, 1976; Clifford, 2002; Krekelberg et al., 2006; Zavitz et al., 2016). In addition, by sparsifying and increasing the signal-to-noise of neuronal population responses, stimulus-specific adaptation has been shown to increase information about the stimulus and decrease discrimination thresholds (Ulanovsky et al., 2003; Krekelberg et al., 2006; Wark et al., 2007; Ollerenshaw et al., 2014). However, because adaptation also decreases firing rates (Müller et al., 1999; Dragoi et al., 2000), the impacts of adaptation on behavior will depend on the specific computation used to read out that sensory information. Indeed, despite the evidence that adaptation increases stimulus information, there are examples of adaptation impairing discrimination (Regan and Beverley, 1985; Zhang et al., 2009; Ollerenshaw et al., 2014) or having no effect at all (Barlow et al., 1976; Zhang et al., 2009).

Here we focus on understanding how decision-making circuits compute perceptual choices in a go/no-go orientation discrimination task. This is a quintessential computation that relies on the representations encoded in the primary visual cortex (V1; Glickfeld et al., 2013; Poort et al., 2015; Resulaj et al., 2018). We find that brief, oriented gratings drive robust adaptation in populations of neurons in mouse V1 that increases the reliability of decoding the presented orientation. However, adaptation increases orientation discrimination thresholds in behaving mice, suggesting that the perceptual choice circuit fails to make use of all of the information present in the adapted population response. Instead our data demonstrates that, in this task, the perceptual choice circuit has adopted a suboptimal computation that ignores rather than discounts the activity of distractor-preferring neurons. The use of only the minimal necessary information, despite costs to performance, may be the result of a prioritization of rapid processing and simple learning rules.

Materials and Methods

Animals. All animal procedures conformed to standards set forth by the NIH, and were approved by the IACUC at Duke University. Thirty-three mice (both sexes; 3–24 months old; singly and group housed (1–4 in a cage) under a regular 12 h light/dark cycle; C57/B6J (Jackson Laboratories, 000664) was the primary background with up to 50% CBA/CaJ (Jackson Laboratories, 000654) were used in this study. Ai93 (*tm93.1(tetO-GCaMP6f)Hze*; Jackson Laboratories, 024103; $n = 4$) and Ai94 (*tm94.1(tetO-GCaMP6s)Hze*; Jackson Laboratories, 024104; $n = 8$) were crossed to *EMX1-IRES-Cre* (Jackson Laboratories, 005628) and *CaMK2a-tTA* (Jackson Laboratories, 003010) to enable constitutive GCaMP6 expression for *in vivo* imaging experiments. *Pvalb-cre (tm1(cre)Arbr*; Jackson Laboratories, 008069; $n = 13$), *VGAT-ChR2-EYFP (Slc32a1-COP4*H134R/EYFP*; Jackson Laboratories, 014548; $n = 3$); and *Emx1-IRES-Cre (tm1(cre)Krj*; Jackson Laboratories, 005628; $n = 2$) were crossed to C57/B6J mice for *in vivo* extracellular electrophysiology ($n = 4$) and behavior ($n = 14$) experiments. *Gad2-IRES-cre (Gad2^{tm2(cre)Zjh}*; Jackson Laboratories, 010802; $n = 2$) and C57/B6J ($n = 1$) mice were crossed to CBA/CaJ for eye-tracking experiments.

Cranial window implant. Dexamethasone (3.2 mg/kg, s.c.) and meloxicam (2.5 mg/kg, s.c.) were administered at least 2 h before surgery. Animals were anesthetized with ketamine (200 mg/kg, i.p.), xylazine (30 mg/kg, i.p.), and isoflurane (1.2–2% in 100% O₂). Using aseptic technique, a headpost was secured using cyanoacrylate glue and C&B Metabond (Parkell), and a 5 mm craniotomy was made over the left hemisphere (center: 2.8 mm lateral, 0.5 mm anterior to lambda) allowing

implantation of a glass window [an 8 mm coverslip bonded to two 5 mm coverslips (Warner, no. 1) with refractive index-matched adhesive (Norland, no. 71)] using Metabond.

The mice were allowed to recover for 1 week before habituation to head restraint. Habituation to head restraint increased in duration from 15 min to >2 h over 1–2 weeks. During habituation, imaging, and electrophysiology sessions, mice were head restrained while allowed to freely run on a circular disc (InnoWheel, VWR). Wheel revolutions were monitored with a digital encoder.

Visual stimulation. Visual stimuli were presented on a 144 Hz LCD monitor (Asus) calibrated with an i1 Display Pro (X-Rite). The monitor was positioned 21 cm from the contralateral eye. Circular 30° Gabor patches containing static sine-wave gratings (0.1 cycles per degree) alternated with periods of uniform mean luminance (60 cd/m²). Visual stimuli for imaging, electrophysiology and behavior experiments were controlled with MWorks (<http://mworks-project.org>).

Two visual stimulus protocols were used for imaging experiments: (1) paired-pulse, same orientation (Fig. 1); and (2) paired pulse, different orientation (Fig. 2). In Protocol 1 ($n = 5$ mice), two static, 100 ms sine-wave gratings of the same orientation (0, 30, 60, 90, 120, or 150°) were successively presented with a variable interstimulus interval (ISI; 0.25, 0.5, 1, 2, or 4 s) and an intertrial interval (ITI) of 4 s. Measurement of adaptation was averaged across all orientations, except in Figure 1D. In Protocol 2 ($n = 12$ mice), a static, 100 ms sine-wave vertical grating (0°; “adapter”) was followed by a 100 ms grating (“test”) of varying orientation (0, 22.5, 45, 67.5, 90, 112.5, 135, or 157.5°) after a variable ISI (250 or 750 ms), with an ITI of 8 s. On 30% of trials, the first stimulus was omitted to measure the non-adapted (control) tuning curve.

For electrophysiology experiments only Protocol 1 ($n = 4$ mice; Fig. 1F, G) was used. Stimuli were the same as in the imaging protocols except only one orientation (0°) was used and the ITI was 10 s. For all stimulus protocols, all orientations and interval conditions were randomly interleaved.

Retinotopic mapping. Retinotopic maps generated from intrinsic autofluorescence, cortical reflectance (for VGAT-ChR2-EYFP mice) or GCaMP signals. For intrinsic autofluorescence imaging, the brain was illuminated with blue light [473 nm LED (Thorlabs) or 462 ± 15 nm band filter (Edmund Optics)], and emitted light was measured through a green and red filter (500 nm long-pass). Images were collected using a CCD camera (Rolera EMC-2, QImaging) at 2 Hz through a 5× air-immersion objective [0.14 numerical aperture (NA), Mitutoyo], using Micromanager acquisition software (NIH). Stimuli were presented at 4–6 positions (drifting, sinusoidal gratings at 2 Hz) for 10 s, with 10 s of mean luminance preceding each trial. Images were analyzed in ImageJ (NIH) to measure changes in fluorescence (dF/F ; with F being the average of all frames) to identify V1 and the higher visual areas. For cortical reflectance, the brain was illuminated with orange light (530 nm LED; Thorlabs), and all of the reflected light was collected; for GCaMP imaging, light was collected with a bandpass filter (520 ± 18 nm) and total trial duration was reduced to 5 s. Vascular landmarks were used to identify targeted sites for imaging, electrophysiology, and optogenetics experiments.

Viral injection. We targeted V1 in *Pvalb-cre* mice ($n = 2$) for expression of channelrhodopsin2 (ChR2). Dexamethasone (3.2 mg/kg, s.c.) was administered at least 2 h before surgery and animals were anesthetized with isoflurane (1.2–2% in 100% O₂). The coverslip was sterilized with 70% ethanol and the cranial window removed. A glass micropipette was filled with virus (AAV5.EF1.dFloxex.hChR2.YFP; UPenn CS0384), mounted on a Hamilton syringe, and lowered into the brain. Fifty nanoliters of virus were injected at 250 and 500 μm below the pia (30 nl/min); the pipette was left in the brain for an additional 10 min to allow the virus to infuse into the tissue. Following injection, a new coverslip was sealed in place, and an optical cannula (400 μm diameter; Doric Lenses) was attached to the cranial window above the injection site. Optogenetic behavioral experiments were conducted at least 2 weeks following injection to allow for sufficient expression.

Two-photon calcium imaging. Images were collected with a two-photon microscope controlled by Scanbox acquisition software (NeuroLabware). Excitation light (920 nm) from a Mai Tai eHP DeepSee laser

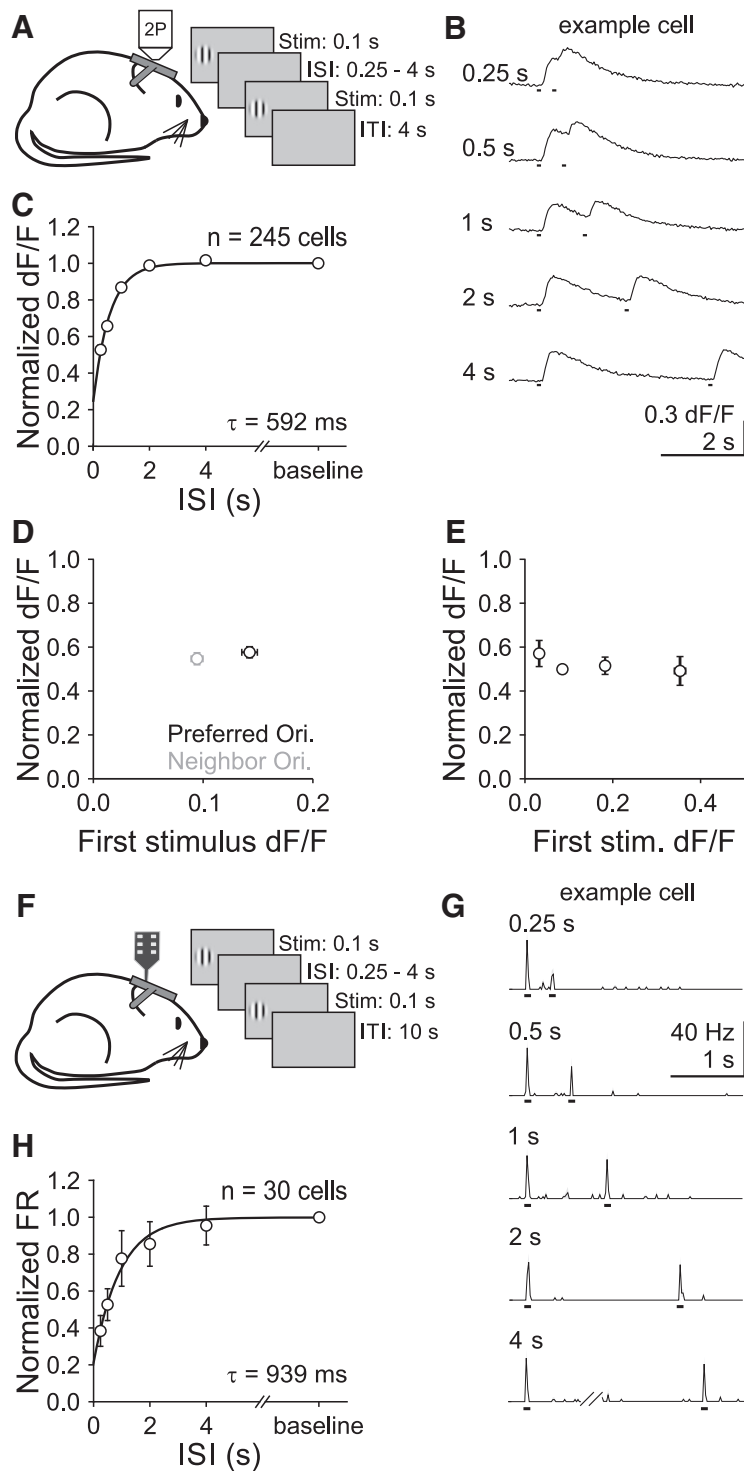


Figure 1. Layer 2/3 neurons in mouse V1 undergo strong adaptation. **A**, Schematic of *in vivo* two-photon calcium imaging (GCaMP6f or GCaMP6s) and visual stimulus protocol. Head-restrained mice passively view presentation of pairs of iso-oriented stimuli. **B**, Average fluorescence traces (dF/F) from an example neuron to pairs of iso-oriented gratings separated by increasing ISIs (top to bottom). **C**, Summary of the average amplitude of the second stimulus normalized to the amplitude of the first stimulus for each ISI for all cells ($n = 245$ cells, 5 mice). Data were fit with a single exponential decay with $\tau = 592$ ms, 95% CI: 499–698 ms. **D**, Average normalized dF/F (250 ms ISI) and average dF/F in response to the first stimulus for preferred (black) and neighboring (gray) orientations. **E**, Average normalized dF/F (250 ms ISI) for cells binned by their response to the first stimulus. **F–H**, Same as **A–C** for *in vivo* extracellular recording ($\tau = 939$ ms, 95% CI: 425–2111 ms, $n = 30$ cells, 4 mice). FR, Firing rate. Error bars are SEM across cells.

(Newport) was directed into a modulator (Conoptics) and raster scanned onto the brain with a resonant galvanometer (8 kHz; Cambridge Technology) through a $16\times$ (0.8 NA, Nikon) or $25\times$ (1.05 NA, Nikon) water-immersion lens. Average power at the surface of the brain was 30–50 mW. Frames were collected at 30 Hz (256 lines) for a field-of-view of $\sim 700 \times 400 \mu\text{m}$ on a side. Emitted photons were directed through a green filter (510 ± 42 nm band filter; Semrock) onto GaAsP photomultipliers (H10770B-40, Hamamatsu). Images were captured at a plane $207 \pm 4 \mu\text{m}$ below the pia (range 180–250 μm). Frame signals from the scan mirrors were used to trigger visual stimulus presentation for reliable alignment with collection.

Eye tracking. Images were collected at 30 Hz with a Genie Nano CMOS camera (Teledyne Dalsa) using a 695 nm LP filter (Midopt) controlled by Scanbox acquisition software. IR illumination (920 nm) was provided from the two-photon laser through the cranial window.

Extracellular electrophysiology. Electrophysiological signals were acquired with a 32-site polytrode acute probe [either A4x8-5mm-100-400-177-A32 (4 shanks, 8 site/shank at 100 μm spacing) or A1x32-Poly2-5mm-50s-177-A32 (1 shank, 32 sites, 25 μm spacing), NeuroNexus] through an A32-OM32 adaptor connected to a Cereplex digital headstage (Blackrock Microsystems). Unfiltered signals were digitized at 30 kHz at the headstage and recorded by a Cerebus multichannel data acquisition system (Blackrock Microsystems). Visual stimulation synchronization signals were also acquired through the same system via a photodiode directly monitoring LCD output.

On the day of recording, the cranial window was removed, and a small durotomy performed to allow insertion of the electrode into V1. A ground wire was connected via a gold pin cemented in a burr hole in the anterior portion of the brain. The probe was slowly lowered into the brain (over the course of 15 min with travel length of $\sim 800 \mu\text{m}$) until the most superficial recording site was in the brain and allowed to stabilize for 45–60 min before beginning recordings. A fluorescent dye (diI, Life Technologies) was painted on the back of the probe before recording and the probe position was thus confirmed *post hoc* in histological sections.

Behavioral task. Animals were water scheduled and trained to discriminate orientation by manipulating a lever. The behavior training and testing occurred during the light cycle. We first trained mice to detect full-field, 90° orientation changes from a static grating. Most mice ($n = 12$) were trained with a 0° distractor; however, two mice were trained with a 45° distractor. On the initial days of training, mice were rewarded for holding the lever for at least 400 ms (required hold time) but not >20 s (maximum hold time). At the end of the required hold time, the target grating appeared and remained until the mouse released the lever (or the maximum hold time expired). Typically, within 2 weeks of training, the mice began re-

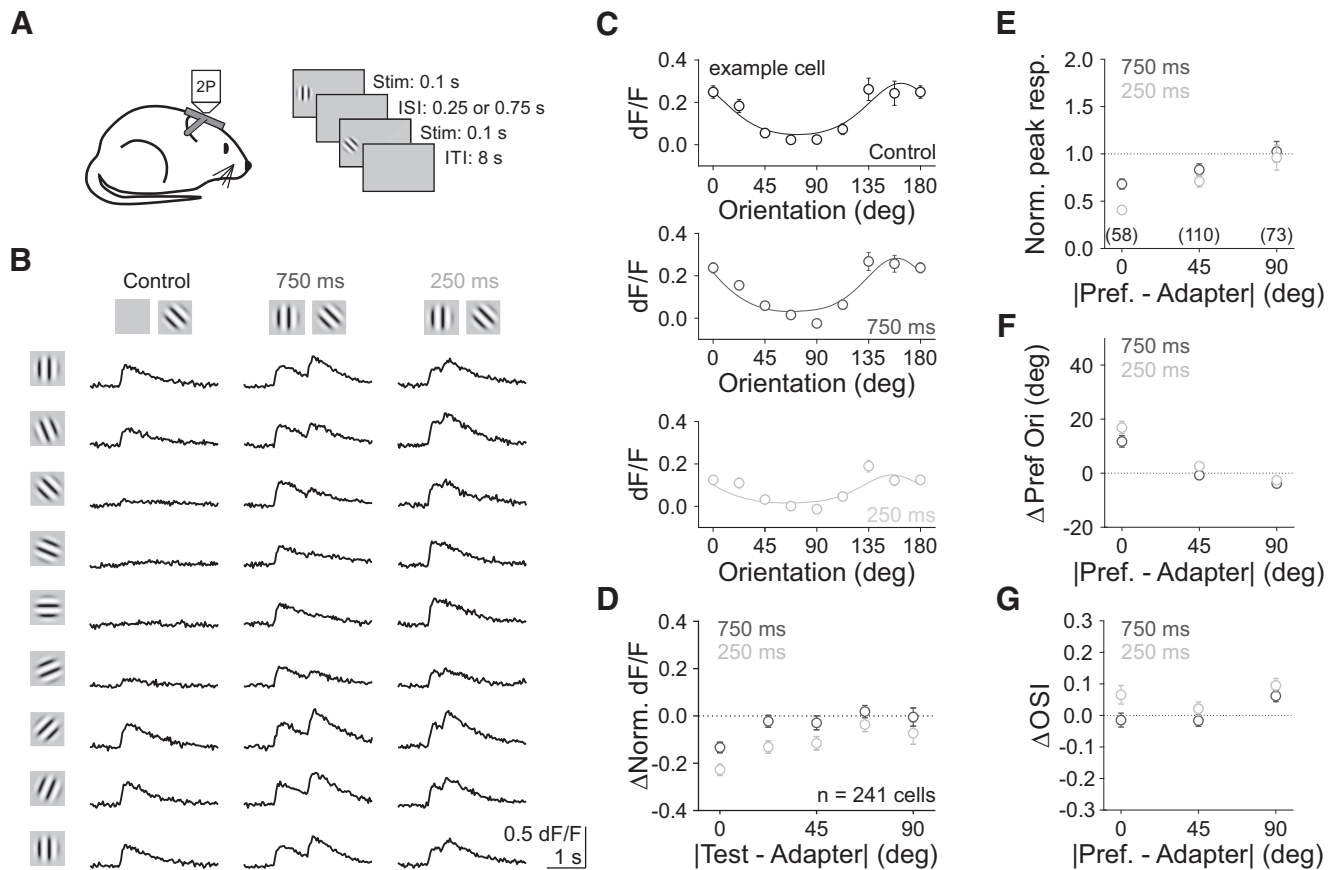


Figure 2. Adaptation changes the orientation tuning and preference of layer 2/3 neurons in mouse V1. **A**, Schematic of *in vivo* two-photon calcium imaging (GCaMP6f or GCaMP6s) and visual stimulus protocol. Head-restrained mice passively view presentation of pairs of stimuli with varying orientations and intervals. **B**, Average dF/F from an example cell to eight different orientations (rows) without adaptation (left column: Control) or after 750 (middle column) or 250 ms (right column) recovery from adaptation to a vertical (0°) grating. **C**, Average orientation tuning curves for the neuron in **B** measured in control (top, black) and after 750 (middle, dark gray) or 250 ms (bottom, light gray) recovery from adaptation. Average responses (error bars are SEM across trials) in each condition were fit with a von Mises function. **D**, Summary of the average difference in dF/F (each cell normalized to its own peak response without adaptation) as a function of stimulus distance from the adapter ($n = 241$ cells, 12 mice) after 750 (dark gray) or 250 ms (light gray) recovery from adaptation. Error bars are SEM across cells. **E**, Summary of the average normalized peak dF/F as a function of the distance of the cells' preferred orientation from the adapter. Cells are binned by their preferred orientation (determined from the peak of the fit in control conditions) into three groups: $<20^\circ$, between 20° and 70° , and $>70^\circ$ from the adapter ($n = 58$, 110 and 73 cells). **F**, Summary of the average change in preferred orientation as a function of the distance of the cells' preferred orientation from the adapter. Positive shift indicates repulsion and negative shift indicates attraction relative to the adapter. **G**, Summary of the average change in OSI as a function of the distance of the cells' preferred orientation from adapter. Positive change indicates increased selectivity and negative change indicates decreased selectivity relative to control.

leasing the lever as soon as the target appeared. Once the animals began reliably responding to the target stimulus, we added a random delay between lever press and the target presentation to discourage adoption of a timing strategy. Over the course of the next few weeks, the task was made harder by (in roughly chronological order): (1) increasing the random delay, (2) decreasing the target stimulus duration and reaction time window, (3) removing the stimulus during the ITI, (4) shrinking and moving the stimuli to more eccentric positions, (5) adding a mean-luminance ISI to mask the motion signal in the transition from distractor to target, and finally (6) reducing the difference between the distractor and the target. Delays after errors were also added to discourage lapses and early releases.

In the final form of the task, each trial was initiated when the ITI had elapsed and the mouse had pressed the lever. Trial start triggered the presentation of a 100 ms static sinusoidal, Gabor patch (30° in diameter, positioned at an eccentricity of $30\text{--}40^\circ$ azimuth and $0\text{--}10^\circ$ elevation) followed by an ISI randomly selected on a presentation-by-presentation basis (250, 500, or 750 ms). For a subset of mice ($n = 3$), the ISI was fixed for a given trial but randomly interleaved on trial-by-trial basis (250 or 500 ms). In this subset of experiments, the maximum number of distractors presented before the target depended on ISI (12 for 250 ms ISI trials; 7 for 500 ms ISI trials) to match the average trial length between these two conditions. In all cases, on each trial the target appeared on a variable presentation (flat distribution) after the first two distractor presentations

and the target orientation was randomly selected from a fixed set of values around each animal's threshold. Mice received water reward if they released the lever within 100–650 ms (sometimes extended to 1000 ms) after a target occurred. However, for calculating hit and false alarm (FA) rate (see Figs. 3, 4, 7, 8), we use a narrower reaction window (200–550 ms) to ensure that the majority of the releases in this window are because of stimulus driven responses and have independent reaction windows for adjacent stimuli with short ISIs. Mice were initially trained with a constant distractor stimulus (see Figs. 3, 4) and were later tested with randomly interleaved distractor orientations (the trained orientation and stimuli -15° and $+15^\circ$ from the trained orientation) selected on a trial-by-trial basis (see Fig. 7).

Of the 11 mice presented in Figures 3 and 4, some ($n = 5$) were initially trained with a single (250 ms) ISI, whereas others ($n = 6$) were immediately introduced to having the fully interleaved condition (presentation-by-presentation selection of 1 of 3 ISIs). Training history had no significant effect on the relationship between ISI and threshold (two-way ANOVA; main effect of ISI: $p < 0.005$, $DF = 2$; main effect of training: $p = 0.41$, $DF = 1$). Notably, the incorporation of targets close to the distractor occurs during the final stages of training; thus, the mice learn relatively late in training that the targets lie within only one quadrant of orientation space. Nonetheless, the FA rates from the variable distractor task (see Fig. 7) reveal that the mice understand this contingency: if the

mice continued to use the strategy learned when there was only a 90° target, then the 15° and –15° distractors should have similar FA rates.

For optogenetic stimulation (see Fig. 8), we delivered blue light to the brain through the cannula from a 473 nm LED (Thorlabs) or a 450 nm laser (Optoengine) and calibrated the total light intensity at the entrance to the cannula (0.1–0.3 mW). On randomly interleaved trials, the light was turned on at the time of lever press and remained on until lever release. Behavioral control was done with MWorks (<http://mworks-project.org>), and custom software in MATLAB (MathWorks) and Python (<http://enthought.com>).

Data processing

Image processing and analysis. All image processing and analysis was performed in MATLAB. The image stack was registered to a stable, average field-of-view using subpixel registration methods to correct for motion along the imaged plane (x – y motion). For segmentation of visually driven neurons, we used semiautomated segmentation algorithms to select regions-of-interest (cell masks) from the average change in fluorescence [dF/F , where F is the average fluorescence in the 20 frames (~660 ms) preceding the first stimulus in each trial] evoked in response to each stimulus type. Fluorescence time courses were generated by averaging all pixels in a cell mask. Neuropil signals were removed by first selecting a shell around each neuron (excluding neighboring neurons), estimating the neuropil scaling factor (by maximizing the skew of the resulting subtraction), and removing this component from each cell's time course.

Visually-evoked responses were measured as the difference in the dF/F before [baseline window: average of three frames (~100 ms) around stimulus onset] and during the response (response window: average of three frames around the peak response; window was selected separately for each experiment to account for variability in response latencies and indicator kinetics). “Responsive cells” were chosen as having statistically significant visually-evoked responses to at least one of the stimulus types (Bonferroni-corrected paired t test) or all stimuli (paired t test), and the maximum derivative in the dF/F occurred before the end of the response window (to eliminate cells strongly driven by the removal of the stimulus). Using these criteria, 245/279 and 473/587 cells were included for visual stimulation Protocols 1 and 2, respectively. All measurements are the average of at least seven trials of the same type.

Cells imaged in Protocol 2 (paired-pulse, different orientation) were further selected based on the reliability of their orientation tuning. Responses in the control (no adapter) condition were fit with a von Mises function:

$$B + Re^{\kappa(\cos(2(\theta - \mu)) - 1)},$$

where B is the baseline firing rate, R is the modulation rate, κ is the concentration, and μ is the preferred orientation. To measure the reliability of the fit, the fit was repeated 1000 times resampling trials with replacement. Only cells for which 90% of the bootstrapped fits were within 22.5° of the original fit were included in further analysis (241/473 cells).

Analysis of the effects of adaptation on tuning [e.g., the preferred orientation and the orientation selectivity index (OSI); Fig. 2E–G] were derived from the von Mises fit to the data. OSI was measured as follows:

$$OSI = \frac{R_{pref} - R_{orth}}{R_{pref} + R_{orth}},$$

where R_{pref} is the cell's response at the preferred orientation (pref: maximum of the fit) and R_{orth} is the response to the orthogonal orientation. In the case that R_{orth} was negative, it was set to 0. Cells were grouped according to the distance of their preferred orientation from the adapter (cells that prefer 22.5° and 157.5° were in the same group; Fig. 2E–G).

For Protocol 1, each cell's responses were normalized to the average response to the first stimulus. For Protocol 2, each cell's responses were normalized to R_{pref} .

Electrophysiology processing and analysis. Individual single units were isolated using the SpyKing CIRCUS package (<http://spyking-circus.readthedocs.io/en/latest/>). Raw data were first high-pass filtered (>500

Hz) and spikes were detected when a filtered voltage trace crossed threshold (9–13 median absolute deviations computed on each channel). A combination of density-based clustering and template matching algorithms were used to automatically cluster the spikes. The resulting clusters were then inspected and adjusted manually using a MATLAB GUI. Clusters with refractory period violations (<2 ms, >1% violation) in the autocorrelogram and that were not stable across the whole recording session were discarded from the dataset. Clusters were combined if they met each of three criteria by inspection: (1) similar waveforms, (2) coordinated refractory periods in the cross-correlogram, and (3) similar interspike interval distribution shape. Unit position with respect to the recording sites was calculated as the average of all site positions weighted by the waveform amplitude of each site. All the subsequent analysis was performed in MATLAB.

Local field potential and current source density analysis. For local field potential (LFP) recording, the extracellular raw signal was bandpass filtered from 1 to 200 Hz and downsampled to 10 kHz. The current source density (CSD) was computed from the average LFP by taking the discrete second derivative across the electrode sites that were linearly spaced across cortical depth, and interpolated to produce a smoothed visually driven CSD profile. This analysis transformed the LFP signal into the locations of current sources and sinks, revealing a patterned laminar distribution of sinks in V1 after the visual onset: an initial sink in layer 4 (latency: ~50 ms), followed by a sink in layer 2/3 and finally a weak and sustained sink in layer 5. Therefore, guided by the visually-evoked CSD map, retrospective histology, and relative depth of recordings relative to the pia surface, layer 2/3 units were identified and chosen for comparison with the two-photon imaging dataset.

Visually-evoked responses of each unit in layer 2/3 of V1 were measured based on average peristimulus time histograms (PSTHs; bin size: 20 ms) over repeated presentations (>25 trials) of the same stimulus. Response amplitudes were measured on a trial-by-trial basis: by subtracting the firing rate at the time of the visual stimulus onset from the value at the peak of the average PSTH within a window of 0–100 ms after the visual onset. Responsive cells were chosen as having statistically significant visually-evoked responses (first baseline response, averaged over 0–100 ms before the visual onset, vs visually-evoked responses, averaged over 0–100 ms after the visual onset; paired t test; this analysis window excluded off-responsive units from analysis). For Protocol 1, we only included the responsive units that had no significant difference in response to the first stimulus (the adapter) across five ISIs (one-way ANOVA). Using these criteria, 30/39 layer 2/3 cells were included.

Behavior processing and analysis. All behavioral processing and analysis were performed in MATLAB. All trials were categorized as either an early release (releases before target appears, a subset of which are FAs), hit, or miss based on the time of release relative to target onset: responses occurring earlier than 100 ms after the target were considered early releases; responses occurring between 200 and 550 ms after a target were considered hits; failures to respond before 550 ms after the target were considered misses. Behavioral sessions were manually cropped to include only stable periods of performance (removing consecutive trials at the beginning or end of a session) and were further selected based on the following criteria: (1) at least 50% of trials were hits, and (2) <35% of trials were early releases. Based on these criteria, the data in Figures 3 and 4 included 17 ± 3 sessions (range: 5–46) for each mouse with an average of 6348 ± 815 trials per mouse (range: 2593–11857); the data in Figure 7 included 20 ± 5 sessions (range: 6–40) for each mouse with an average of 4904 ± 941 trials per mouse (range: 2284–8236); and the data in Figure 8 included 24 ± 5 (range: 11–35) sessions for each mouse with 7524 ± 1582 trials (range: 5021–11710), respectively.

Hit rate was computed from the number of hits and misses for each stimulus type:

$$Hit\ rate = \frac{hit}{hit + miss}.$$

Lapse rates were measured as (1–Hit rate) for 90° targets. Most mice had low lapse rates (11/11 mice were <10%) for the task in Figures 3 and 4. However, as mice age their reaction times become slower, thereby

inflating the lapse rate; we think that this effect explains the increased lapse rate during optogenetic suppression of V1 (only 1/4 mice were <10% in Fig. 8).

Hit rates across stimulus types were fit with a Weibull function to determine discrimination thresholds (50% of the upper asymptote to account for lapse rate). No correction was made for FA rate. Threshold confidence intervals (CIs) were estimated via nonparametric bootstrap resampling trials with replacement.

All distractor stimulus presentations were categorized as either a correct rejection (CR) or a FA: responses occurring between 200 and 550 ms after a distractor stimulus were considered FAs; presentations where the mouse held the lever for at least 550 ms after the distractor stimulus were considered CRs. FA rate was computed from the total number of FAs and CRs in the session:

$$FA \text{ rate} = \frac{FA}{FA + CR}$$

Signal detection theory (Green and Swets, 1966) was applied to calculate bias (c) given the measured hit and FA rate as follows:

$$c = -\frac{Z(\text{hit rate}) + Z(\text{FA rate})}{2},$$

where Z is the inverse of the cumulative distribution function of the normal Gaussian distribution.

For matching the trial length across ISIs (Fig. 3G), trial length was first binned every 0.5 s within the range of 1.2–6.2 s (for hits and misses) and 1.2–4.2 s (for FAs and CRs) respectively. Within each bin, the same number of trials was chosen for all ISIs to ensure that the average trial length was not significantly different. The selected trial number for each bin was determined by the minimum number of trials across ISIs in that bin for each mouse.

For calculating fraction of rewarded trials (P_{rew}) given the 22.5°-Hit and FA rates of each ISI (250 and 750 ms), we first simulated 1000 trials where the 22.5° target is presented after six distractor 0° presentations. For a known FA rate, we calculated the fraction of early trials (P_{early}) that were unrewarded assuming every 1/FA number of distractor presentations would generate an early release. Then the P_{rew} was calculated as follows:

$$P_{\text{rew}} = (1 - P_{\text{early}}) \text{Hit}.$$

Eye-tracking processing and analysis. All eye-tracking data were analyzed in MATLAB. The size of the pupil on each frame was extracted using the built-in function `imfindcircles`. Pupil size was normalized to the maximum radius measured during the session and quantified as the average measured in a 250 ms window preceding the target stimulus.

Modeling. For all modeled decoders, neurons imaged in different mice were analyzed separately and only datasets with at least 10 well-fit neurons (using the criteria above) were included in these analyses (10/12 experiments). As a result, error bars in Figures 5 and 6 represent SEM across datasets.

To determine the properties of a neural decoder capable of fitting animals' behavioral choice, we summarized the behavioral data in the two adapted conditions for the 11 animals used in Figure 3. Because different orientations were sampled for each of the mice, we performed a spline fit of each psychometric function and averaged across all animals (see Fig. 5A). For each neuronal dataset we assumed that behavior was generated by sampling from a posterior distribution that had the form of a logistic regression, i.e.:

$$\log p(\text{detect}|n_i) = \sum_{i=0}^N a_i n_i - \log Z,$$

where n_0 is once again assumed to be 1 and Z is the normalizer. The a_i 's were then determined by minimizing the symmetrized Kullback–Leibler divergence between the average of this neurally predicted detection probability raised to some power and the behaviorally observed detection probability across stimulus and adaptation conditions. Once again, we

used gradient descent to perform this optimization. The resulting mean squared error between the neurally generated psychometric curves and the behaviorally observed ones was $2e-4$. These behaviorally generated weights strongly correlated (correlation coefficient = 0.54) with the weights discovered by using the neuronal data to decode whether the stimulus was a target or distractor (using MATLAB's `glmfit` routine; see Fig. 5D).

Log likelihood functions were generated using two methods: the first followed the equation (Jazayeri and Movshon, 2006):

$$\log L(\theta) = \sum_{i=1}^N n_i \log f_i(\theta) - \sum_{i=1}^N f_i(\theta),$$

where N is the number of neurons in the population, $f_i(\theta)$ is each neuron's normalized tuning function and n_i is each neuron's response to a given stimulus. The tuning functions were obtained from well-fit responsive neurons, in the non-adapted condition (under the assumption that the decoder is unaware of the effects of adaptation; Serriès et al., 2009), and were used to decode the presented stimulus given single-trial responses (n_i) that varied with adaptation.

Because our neuronal data does not adhere to the assumptions that spiking is Poisson and independent, we also used a second more empirical method based upon the assumption that our population of neurons represented von Mises over orientation using a linear PPC (Ma et al., 2006). This approach assumes that the likelihood function takes the form:

$$\log L(\theta) = \sum_{i=0}^N a_i n_i \cos(\theta) + b_i n_i \sin(\theta),$$

where n_0 is assumed to be 1 and the a_i 's and b_i 's are discovered by maximizing the likelihood of the empirically observed joint distribution of the presented stimulus and the neural response. This results in a convex optimization problem which we solved using gradient ascent. Because this method is prone to overfitting resulting in poor cross-validation performance when the number of units greatly exceeds the number of trials per condition, in datasets with >15 well-fit neurons, we used only the 15 best-fit neurons (as measured by their 90% CI). As a control, we also preprocessed our neural data using principle component analysis and eliminated all but the 15 dominant modes of variability and refit the empirically generated log likelihood. This had no effect on our results. The number 15 was chosen to optimize performance on the cross-validated dataset, however, we note that choosing values between 12 and 16 did not substantially change our results.

The single-trial log likelihood functions were then used to determine either the MAP estimate (Method 1: Fig. 6E; Method 2: Fig. 6A–C) or the posterior probability distribution ("optimal decoder": Fig. 6D). We also used a standard population vector decoder to estimate the orientation (see Fig. 6F). For this decoder a preferred stimulus value, θ_p , was extracted from a parameterized fit to the tuning curve of each well-fit unit. Estimated orientation for each trial was then obtained from the equation:

$$\theta = \text{angle} \left(\sum_i n_i \begin{bmatrix} \cos(\theta_i) \\ \sin(\theta_i) \end{bmatrix} \right).$$

Estimate biases were computed by taking the mean estimate as a function of stimulus and adaptation condition. Because there was no clear trend in estimate SD as a function of presented orientation, we computed estimate variance in each adaptation condition by removing the bias associated with each stimulus value and concatenated the resulting estimate residuals into a single vector. To account for differences in the information content of each dataset, estimate variances (and SD) were normalized by the variance of the control condition of each dataset ($35 \pm 14^\circ$). We then concatenated the resulting normalized residuals from each dataset into a single dataset to measure the mean and associated SE of the residual variances.

To compute the area under the receiver operating characteristic (auROC) we treated our estimates of orientation (or in the case of the de-

coders in Figs. 5E, 6D: the probability that the orientation was not the distractor orientation) as decision variables and computed false-positives and correct detections for 400 uniformly sampled values of the decision criterion chosen to span all observed values of the decision variable. For the sum decoder (see Fig. 5C) we simply treated the total population activity on each trial as the decision variable. The auROC was then computed numerically using the trapezoid rule.

Note that in all cases we report only cross-validated results (using leave one out cross-validation in the control condition). This is because parameters of our neural decoders were all fit using only the control condition and not our two adaptation conditions. The only exception to this is for the calculation of neuronal weights using the logistic regression to optimize discriminating between target and distractor stimuli; in this case, to make it comparable to the estimation of neuronal weights from the fit to animals' behavioral choice, we trained the decoder on the 750 and 250 ms conditions. The results for all decoders were largely unchanged if we obtained parameters from either the 750 ms condition or all conditions simultaneously and cross-validate using a leave one out procedure. However, we note that logistic regression weights are not sensitive to this choice of training datasets.

To assess the degradation in performance that results from use of a suboptimal decoder we computed percentage correct in the following way. First the weights obtained from the fit to decode the animals' behavioral choice and the fit to decode stimulus identity were used to generate a set of decision variables for each dataset and stimulus condition. A task relevant measure of percentage correct was computed from these values by mimicking the statistics of the behavioral task (i.e., distractors are 8 times more common than targets and target stimulus values are uniformly distributed) for a range of potential decision criteria. The optimal decision criterion was selected by determining which provided the maximum value of percentage correct. This resulted in two values for optimal percentage correct, one for the behavioral weights and one for stimulus weights for each dataset. Average and SD for these percentage correct values were then computed across datasets.

Experimental design and statistical analysis

Data were tested for normality using a Lilliefors test. Whereas all behavioral measures and auROC estimates were normally distributed, distributions of neuronal responses were not. Thus, in the case that two distributions were compared we used a *t* test in behavioral measures and auROCs, and a Wilcoxon rank sum test for neuronal responses; however, because ANOVA and *post hoc* Tukey HSD tests have been shown to be robust to non-normality (Driscoll, 1996), these tests were used for all data. Sample sizes were not predetermined by statistical methods, but our sample sizes of the neurons and behavior animals are similar to other studies. The numbers of cells, animals, or experiments were provided in the corresponding text, figures, and figure legends. All error values in the text are SEM unless otherwise specified. Data collection and analysis were not performed blind to experimental conditions, but all visual presentation conditions in either calcium imaging, extracellular recording, or behavior testing are randomized. Moreover, the strength and time course of adaptation on neuronal responses was measured using two methods (Fig. 1) with data collected by two experimenters.

Data and code availability

All relevant data and code will be made available upon reasonable request.

Results

Adaptation has prolonged effects on the amplitude and selectivity of visual responses

To understand how adaptation impacts sensory encoding, we first sought to characterize the time course of adaptation in the mouse V1. Using video-rate two-photon imaging, we measured visually-evoked responses in layer 2/3 of V1 in alert mice transgenically expressing the calcium indicator GCaMP6f (GCaMP6f; $n = 4$; GCaMP6s; $n = 8$). Mice passively viewed pairs of brief, identical static gratings (an adapter followed by a test; each 100

ms duration) presented at a range of ISIs (0.25–4 s; Fig. 1A). At short intervals, neurons in V1 had significantly reduced responses to the test stimulus relative to the adapter and gradually recovered ($\tau = 592$ ms, 95% CIs: 499–698 ms) with increasing ISI [$n = 245$ cells, 5 mice; one-way ANOVA ($p < 10^{-17}$; $DF = 4$) with *post hoc* Tukey HSD compared with non-adapted responses (250 ms: $p < 10^{-7}$; 500 ms: $p < 10^{-7}$; 1 s: $p < 0.0001$; 2 s: $p = 1.0$; 4 s: $p = 0.96$); Fig. 1B,C]. We do not think that this strong adaptation is an artifact of either indicator or spike rate saturation because there was no relationship between response amplitude and degree of adaptation either within (normalized dF/F after 250 ms ISI for preferred vs neighboring orientation: $p = 0.37$; Wilcoxon rank sum test; Fig. 1D) or across cells (linear regression: $r^2 = 0.003$, $p = 0.42$; Fig. 1E). Further, the degree of adaptation measured with extracellular single-unit recording was similar to, though significantly stronger than, data collected with calcium imaging (two-way ANOVA: main effect of recording method: $p < 0.001$; $DF = 1$; Fig. 1F–H). Thus, the effects of adaptation were strong and relatively long-lasting compared with the duration of the stimulus.

A large component of cortical adaptation is stimulus-specific, and the effects of adaptation can thus depend on the difference between a neuron's preferred orientation and the orientation of the adapter stimulus (Müller et al., 1999; Dragoi et al., 2000; Stroud et al., 2012; Patterson et al., 2013). To determine how adaptation alters orientation tuning in mouse V1, we measured the orientation tuning of a population of layer 2/3 neurons with and without a vertical adapter grating (Fig. 2A–C). Adaptation significantly reduced responses to stimuli near the adapter orientation (difference in normalized dF/F : two-way ANOVA, main effect of orientation: $p < 10^{-6}$, $DF = 4$; $n = 241$ cells, 12 mice; Fig. 2D), and this effect was larger when the ISI was short (two-way ANOVA, main effect of interval: $p < 0.0001$; $DF = 1$; Fig. 2D). Moreover, the peak responses of neurons with preferred orientations near the adapter orientation were significantly reduced (normalized peak amplitude: two-way ANOVA, main effect of orientation: $p < 10^{-5}$, $DF = 2$; main effect of interval: $p < 0.05$, $DF = 1$; Fig. 2E) and their tuning curves repelled away from the adapter (change in preferred orientation: two-way ANOVA, main effect of orientation: $p < 10^{-24}$; main effect of interval: $p < 0.01$; Fig. 2F). In addition, neurons with preferred orientations orthogonal to the adapter had a significant increase in OSI (difference from control OSI: two-way ANOVA, main effect of orientation: $p < 0.01$; main effect of interval: $p < 0.01$; Fig. 2G), likely because of selective adaptation of responses on the flanks of their tuning curves (Dragoi et al., 2002). Thus, adaptation altered the amplitude, preference and selectivity of neuronal responses in V1 in a manner very similar to what has been previously observed in carnivores and primates (Müller et al., 1999; Dragoi et al., 2000; Patterson et al., 2013).

Adaptation increases orientation discrimination thresholds

To determine how adaptation affects orientation discrimination, we designed a multi-interval, go/no-go task in which the mouse needs to use information about the orientation of visual stimuli to earn reward. In this task, head-restrained mice were trained to press a lever to initiate a trial and release it to report any change in the stimulus orientation (Fig. 3A; Movie 1). On each trial, the lever press triggered the serial presentation of multiple gratings (2–9 presentations per trial) of the same orientation [“distractors”; 100 ms duration; mice were trained with either a 0° ($n = 9$) or 45° ($n = 2$) distractor] followed by a grating that is different

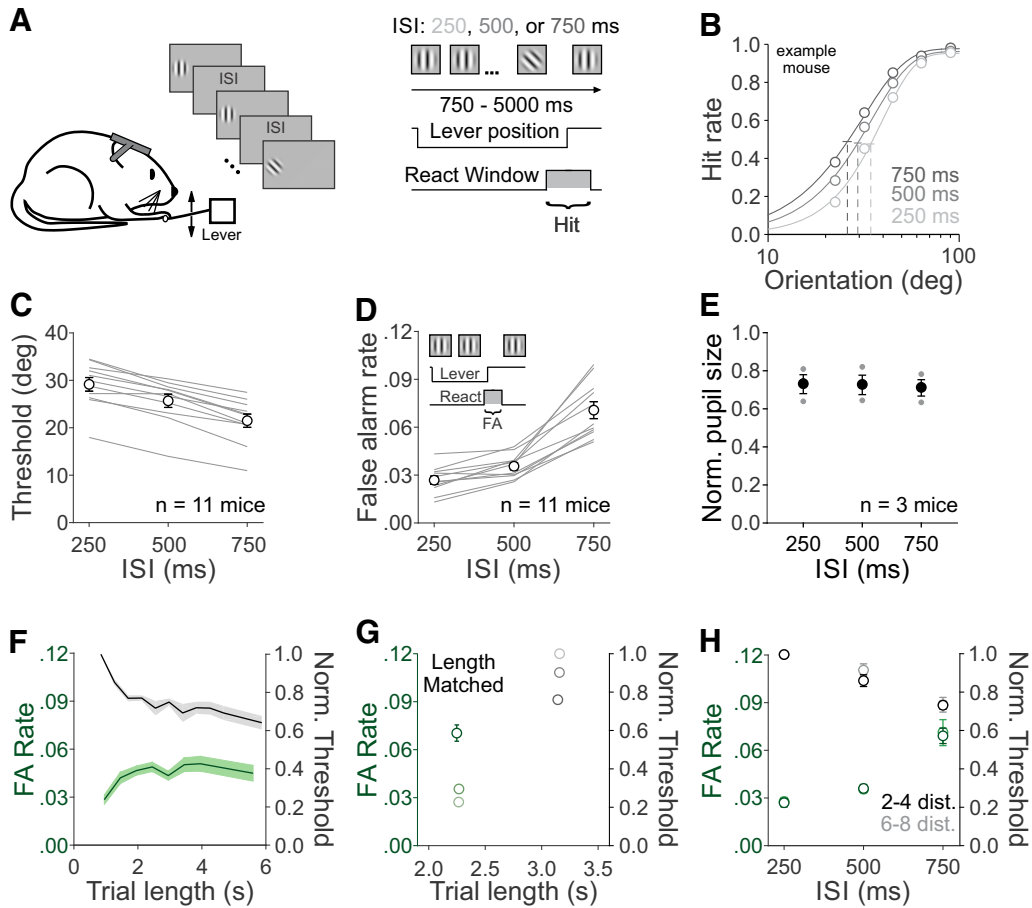
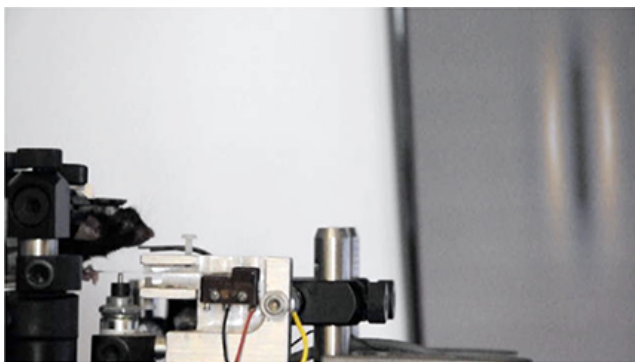


Figure 3. Increase in threshold and decrease in FA rate following shorter ISIs in an orientation discrimination task. **A**, Schematic of behavioral setup and trial progression. Head restrained mice press a lever to initiate a trial, triggering the repeated presentation of a vertically oriented grating (100 ms) with a randomly interleaved ISI [250 (light gray), 500 (gray), or 750 (dark gray) ms]; the mouse must release the lever within a short window following the presentation of a non-vertical grating to receive a reward. **B**, Hit rate for an example mouse in which target orientations were sorted according to the preceding ISI. Data are fit with a Weibull function; vertical lines denote threshold, error bars are 95% CIs. **C**, Summary of the threshold across ISIs. Open circles are the average of all mice ($n = 11$ mice), connected gray lines are individual mice; error bars are SEM across mice. **D**, Summary of the average FA rate across mice. Inset, Schematic illustrating the window for a FA. **E**, Summary of average pupil radius (normalized to the maximum radius measured during the experiment) for each mouse (gray circles) and across mice (black circles) by ISI. **F**, Discrimination threshold (black; right y-axis) and FA rate (green; left y-axis) as a function of trial length (bin size: 0.5 s). For threshold, values are normalized by the value at the first bin (center; 0.95 s). Shaded area is the SEM across 11 mice. **G**, Summary of the change in threshold (normalized to 250 ms ISI; gray shades, right y-axis) and FA rate (non-normalized; green shades, left y-axis) as a function of average trial length, only including trials with matched length across ISIs. **H**, Summary of the change in threshold (each normalized to own 250 ms condition; gray shades, right y-axis) and FA rate (green shades, left y-axis) as a function of ISI, for short (2–4 distractor presentations; dark) or long (6–8 distractor presentations; light) trials.



Movie 1. Movie of a mouse performing the orientation discrimination task. Two consecutive trials are shown in the orientation discrimination task. The outcome of the first trial is a hit and the outcome of the second trial is a FA.



from the distractor orientation (“targets”; drawn randomly from a 9° to 90° range on trial-by-trial basis) to probe the animal’s discrimination threshold. The number of distractor stimuli on each trial was variable to prevent the mouse from anticipating the target presentation. If the mouse released the lever within a window 200–550 ms following the onset of a target stimulus, it was considered a hit; if the mouse released the lever within the same window following a distractor stimulus, it was considered a FA.

To test how adaptation affects orientation discrimination, each stimulus presentation was separated by a randomly selected ISI (250, 500, or 750 ms; the randomization prevents the mouse from anticipating the upcoming interval). These ISIs were selected because they are on the steep portion of the recovery from adaptation (Fig. 1). Thus, by conditioning hit and FA rates on the interval that preceded each target or distractor stimulus, this task allowed us to compare the discrimination threshold and FA rate in different adaptation states.

We found that the animals’ discrimination thresholds were significantly increased when the ISI was short (one-way ANOVA

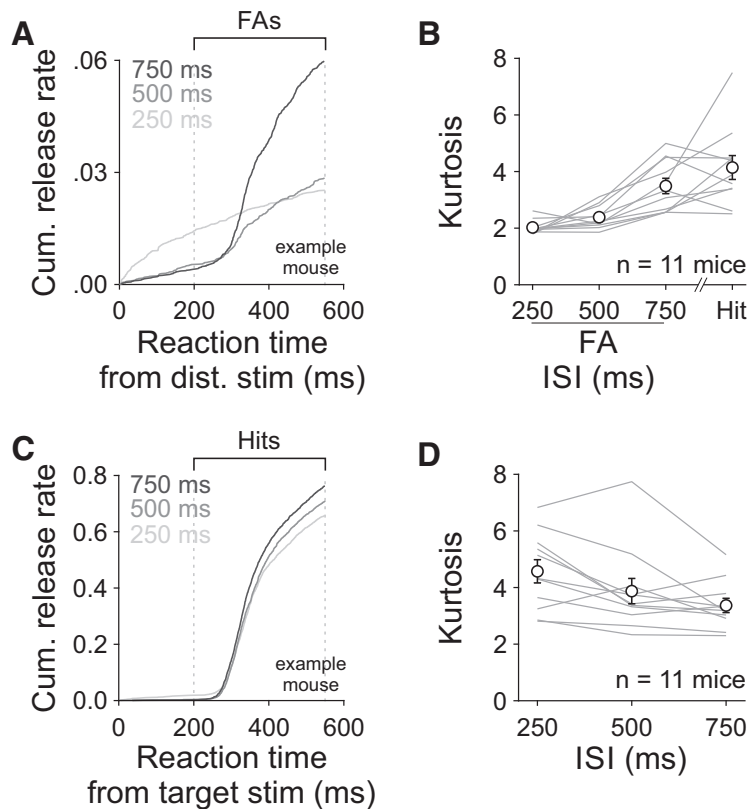


Figure 4. Decrease in FA rate following shorter ISIs is because of a decrease in stimulus-reactive responses. **A**, Cumulative release rate (as fraction of distractor presentations) as a function of reaction times relative to the time of distractor stimulus onset sorted according to the preceding ISI; same mouse as in Figure 3B. Vertical lines represent react window used to calculate FA rate. **B**, Summary of the average kurtosis of the reaction time distributions of FAs across mice for the three intervals. Kurtosis of the response distribution to all targets collapsed across ISIs is also included for a direct comparison. Connected gray lines are individual mice ($n = 11$). **C, D**, Same as **A** and **B** for hits at all targets.

with *post hoc* Tukey HSD compared with 250 ms: 500 ms: $p = 0.21$; 750 ms: $p = 0.002$; $DF = 2$; $n = 11$ mice; Fig. 3B,C). Although we propose that this change in threshold is because of the effects of adaptation on sensory processing, there are a number of possible cognitive confounds that could contribute to this effect. For instance, effect of ISI on threshold could also be because of effects on task engagement or arousal. However, our data suggest that these parameters were stable across ISIs because there is (1) no significant difference in lapse rate (measured as 1-Hit rate for 90° targets) across ISIs (250 ms: 0.05 ± 0.01 ; 500 ms: 0.05 ± 0.01 ; 750 ms: 0.05 ± 0.01 ; $p = 0.97$; one-way ANOVA; $n = 11$ mice), and (2) no difference in pupil size preceding targets of each ISI ($p = 0.96$, one-way ANOVA; $n = 3$ mice; Fig. 3E). We also considered the possibility that the uncertainty in the timing of stimulus appearance might be influencing the animals' behavior, for instance by generating surprise at the appearance of a stimulus earlier or later than expected. However, we found animals also had a lower threshold for longer intervals when ISIs were interleaved on a trial-by-trial rather than presentation-by-presentation basis (500 vs 250 ms: $p = 0.01$, paired t test; $n = 3$ mice).

The effect of ISI on threshold may also be because of the temporal properties of the task. Trials with a short pre-target ISI are, by definition, shorter than those with long ISIs ($p < 10^{-13}$, $DF = 2$, one-way ANOVA) and the animals' discrimination thresholds decreased with trial length ($p < 10^{-3}$, $DF = 10$; one-way ANOVA; Fig. 3F). However, when trials were selected such that all ISIs were matched for length, the ISI-dependent changes

in threshold remained intact (two-way ANOVA, main effect of method: all versus matched trials: $p = 0.18$, $DF = 1$; Fig. 3G). Similarly, the effect of ISI on discrimination threshold was also not significantly different for targets that appear after 2–4 distractor presentations compared with those that follow 6–8 (two-way ANOVA: main effect of distractor number: $p = 0.36$, $DF = 1$; Fig. 3H). Together, these data suggest that the effects of ISI on discrimination threshold were not because of differences in trial length or number of stimuli.

We also found a decrease in FA rate with shorter ISIs (one-way ANOVA with *post hoc* Tukey HSD compared with 250 ms: 500 ms: $p = 0.22$; 750 ms: $p < 10^{-8}$; $DF = 2$; Fig. 3D,F,G). One interpretation of the observed increase in FA rate following longer ISIs is that there is an increase in impatient responses during the extended interval between stimuli. If true, this would lead to an increase in release probability shortly after stimuli following a long ISI. However, inspection of the distribution of reaction times revealed the opposite effect: the distribution of releases to distractor stimuli following short ISIs had shorter latencies than those following longer ISIs [one-way ANOVA ($p < 10^{-8}$, $DF = 2$) with *post hoc* Tukey HSD compared with 250 ms: mean reaction time: 500 ms: $p < 10^{-5}$; 750 ms: $p < 10^{-8}$; $n = 11$ mice; example mouse in Fig. 4A). The distributions of reaction times were con-

sistent with there being two classes of FAs: (1) releases following distractor stimuli in which the mouse guessed that it was a target, and (2) "spontaneous" releases because of non-sensory factors (i.e., impatience). The relatively flat distribution of reaction times in the 250 ms ISI condition is a hallmark of this latter, nonsensory behavior (Tiefenau et al., 2006; though note the initial increase in release rate immediately after the stimulus presentation, likely because of intentional releases from the previous presentation). In contrast, the comparatively skewed distribution following the longer intervals suggests that the majority of these were stimulus-driven releases [one-way ANOVA ($p < 10^{-5}$, $DF = 3$) with *post hoc* Tukey HSD compared with 250 ms: 500 ms: $p = 0.76$; 750 ms: $p = 0.002$; all targets: $p < 10^{-5}$; Fig. 4B]. In fact, the distribution of responses following the 750 ms ISI closely resembled the reaction time distribution when a target stimulus was presented (one-way ANOVA of 750 ms ISI distractor compared with all targets: $p = 0.3$). Notably, ISI had no detectable effect on the distribution of responses to targets (one-way ANOVA; $p = 0.1$, $DF = 2$; Fig. 4C,D), suggesting that the majority of these responses were stimulus driven in all conditions. Thus, short ISIs reduced the FA rate by decreasing the likelihood of a stimulus-driven response to a distractor.

The effects of adaptation on behavior are consistent with a perceptual choice circuit that ignores distractor preferring neurons

To determine how the perceptual choice circuit might be combining the adapted cortical population activity in such a way that

increases the animals' discrimination threshold, we used a linear combination of neuronal activity (data from Fig. 2) to decode the animals' average behavioral choices (data from Fig. 3; $n = 11$ mice; Fig. 5A; see Materials and Methods). This fit involved assigning a weight to each neuron that defined its contribution to the behavior: positive weights reveal a positive relationship between the spiking of a neuron and the mouse releasing the lever (i.e., hits and FAs); negative weights reveal a positive relationship between the spiking of a neuron and the mouse not releasing the lever (i.e., misses and CRs); and no weight reveals no relationship between spiking and behavior. As expected, neurons preferring target orientations tended to have positive weights [orientation preference $> 11.25^\circ$ from the distractor ($n = 123$ cells): $p < 0.001$; Student's t test; Fig. 5B]. However, although the weights assigned to neurons preferring stimuli closer to the distractor orientation (orientation preference within 11.25° of the distractor; $n = 22$ cells) were significantly different from target preferring neurons ($p < 0.01$; unpaired t test), they were not significantly different from zero ($p = 0.3$, Student's t test). This was surprising because an optimal classifier should both positively weight the target orientations (to increase the probability of identifying a target stimulus) and negatively weight the distractor orientation (to decrease the probability of mistaking it for a target).

This suggests that the animals' behavior may be generated by positively weighting the target neurons while largely ignoring the neurons that are tuned to the distractor. To test whether this all-positive weighting of the neuronal population activity is sufficient to explain the direction of the effects of adaptation on behavior, we tested the effect of different linear combinations of neuronal activity on the discriminability of targets and distractors. To simplify the potential set of neuron weights, we defined the weights of each neuron as 1, 0, or -1 according to its orientation preference. We then scaled each neuron's response on each trial by its weight, and summed the activity of the simultaneously recorded population, to generate a population response for each trial. We quantified how well an ideal observer is able to differentiate targets and distractors from the single trial population responses by measuring the auROC. This is a criterion-free approach for comparing the degree of overlap in the distributions of neuronal responses, and thus is a good proxy for how well the stimuli can be discriminated. Indeed, across a range of different neuronal weights, the auROC was significantly larger, and there-

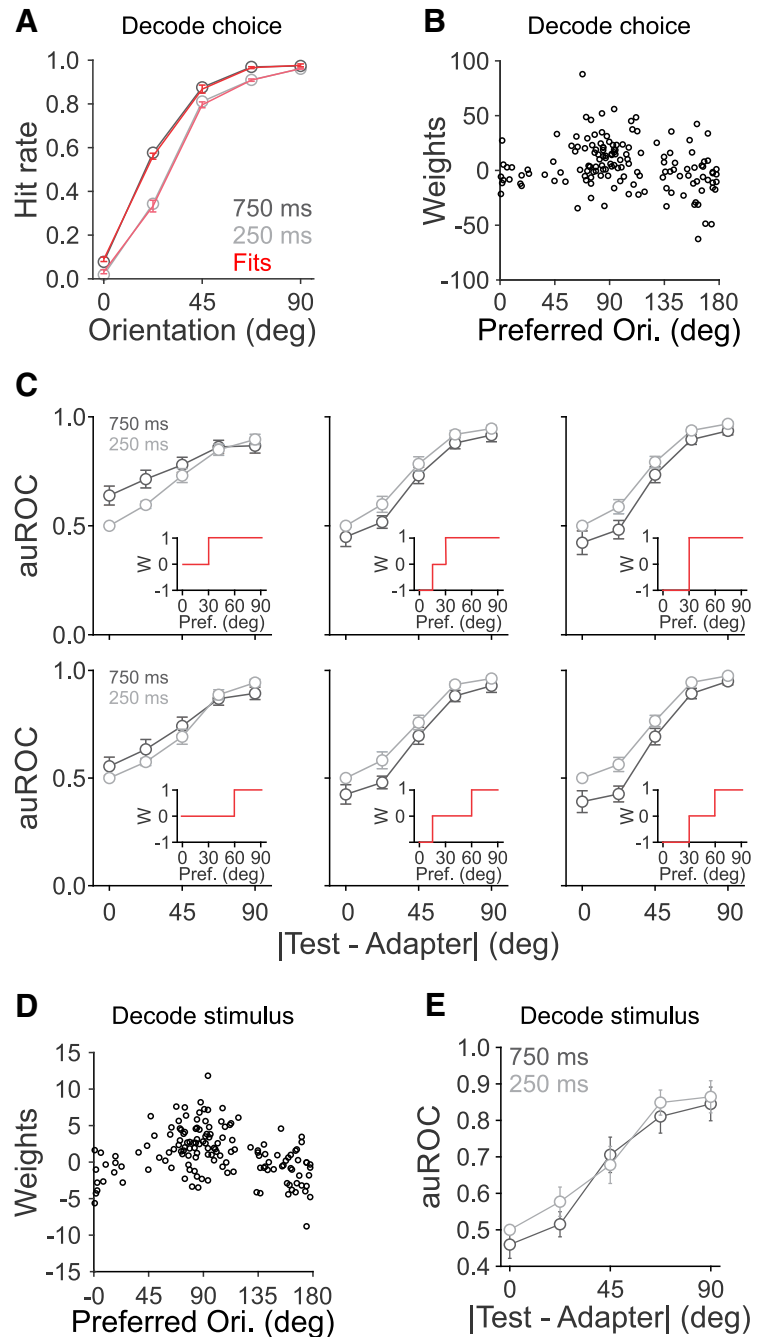


Figure 5. Predicted weights are biased and positive. **A**, Fit (red) of the neuronal data optimized to decode the hit rate from the average behavioral data (gray) from the experiments in Figure 3; hit rates are responses to targets (orientations $> 0^\circ$) and distractors (0°) after 750 (dark gray) or 250 (light gray) ms recovery from adaptation ($n = 10$ mice). **B**, Summary of weights found by the fit to decode the behavioral choice in **A** as a function of neuron orientation preference. **C**, Summary of the average auROC as a function of stimulus distance from the adapter after 750 (dark gray) or 250 (light gray) ms recovery from adaptation, found by a weighted sum of neuronal activity ($n = 10$ mice). Rows: positive weight for neurons with orientation preference $> 30^\circ$ (top) or 60° (bottom); Columns: negative weight for neurons with orientation preferences $< 15^\circ$ (middle) or 30° (right), or no negative weights (left). Insets, The weighting scheme for each panel. Error bars are SEM across experiments. **D**, Same as **B** for weights found using the neuronal data to decode the presented stimulus (target or distractor). **E**, Summary of the average auROC, found when optimizing the neuronal weights to decode the stimulus, as a function of stimulus distance from the adapter.

fore the distributions were more discriminable, for targets farther from the distractor (main effect of target orientation: $p < 10^{-9}$, $DF = 3$; two-way ANOVA; $n = 10$ mice; Fig. 5C).

When we generated a suboptimal weighting of inputs, where those neurons with preferences $< 30^\circ$ were set to zero, whereas

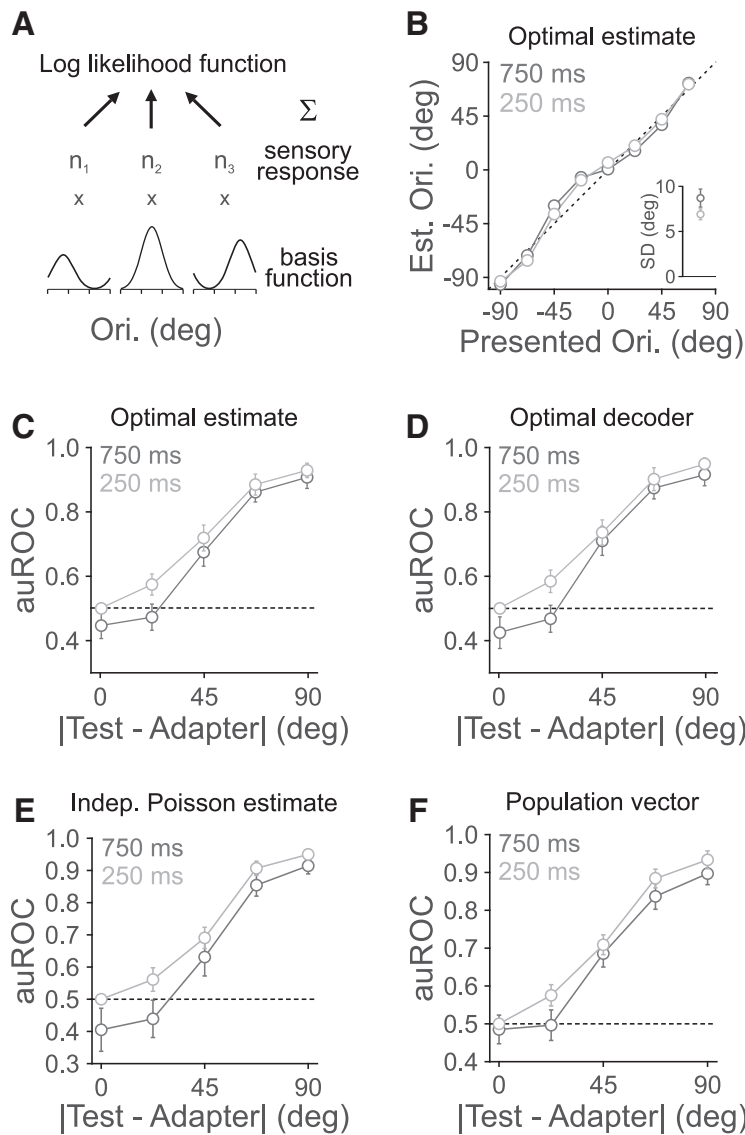


Figure 6. Feature identification models predict that adaptation improves discrimination. **A**, Schematic of an optimal estimator of orientation. The likelihood function is determined by scaling each neuron's basis function (see Materials and Methods) to the amplitude of its response, n_i , to a given test stimulus and then summing across the population. The peak of this likelihood function is the estimated orientation. **B**, Average estimate of orientation as a function of the presented orientation after 750 (dark gray) or 250 (light gray) ms recovery from adaptation ($n = 10$ mice). Error bars are SEM across experiments. Inset, Average SD in estimate of orientation as a function of adaptation state. **C**, Summary of the average auROC for discriminating target and distractor orientations, when estimating the presented orientation using an optimal estimator of orientation, as a function of stimulus distance from the adapter after 750 or 250 ms recovery from adaptation. **D–F**, Same as **C** using the average posterior probability as the decision variable (**D**), or when estimating the presented orientation using estimator that assumes Poisson statistics (**E**), or a population vector (**F**).

those $>30^\circ$ were set to one, the decoder reliably predicted that shorter ISIs would decrease discriminability for targets near the distractor orientation (auROC for 250 vs 750 ms: $22.5^\circ: p < 0.01$; $n = 10$ mice; paired t test; Fig. 5C) and distractors (auROC for 250 vs 750 ms: $0^\circ: p < 0.05$). Thus, this weighting predicted an increase in discrimination threshold and a decrease in FA rate with adaptation. The effect was not significant, though trends in the same direction, if the threshold for positive weights was set at 60° ($22.5^\circ: p = 0.15$; $0^\circ: p = 0.25$; paired t test). However, any addition of negative weights to distractor preferring neurons inverted the relationship between adaptation states such that shorter ISIs predicted a decrease in threshold and no change in

false alarm rate ($<15^\circ$: auROC- $22.5^\circ: p < 0.01$; $0^\circ: p = 0.25$; $<30^\circ$: auROC- $22.5^\circ: p < 0.05$; $0^\circ: p = 0.19$; paired t test). Thus, the observed effects of adaptation on orientation discrimination are consistent with a suboptimal computation in which the downstream perceptual choice circuit performs an all-positive integration of the population activity.

To confirm that an optimal classifier would negatively weight distractor preferring neurons, we trained a logistic regression to correctly discriminate distractor and target stimuli. Indeed, neurons preferring target orientations tended to have positive weights ($p < 10^{-5}$; Student's t test; Fig. 5D), and neurons preferring stimuli closer to the distractor orientation tended to have negative weights ($p < 0.05$). Moreover, these optimal weights did not generate the behaviorally observed effects of adaptation on hit and FA rate (auROC for 250 vs 750 ms: $22.5^\circ: p = 0.06$; $0^\circ: p = 0.23$; Fig. 5E).

Behavioral evidence for a task-specific circuit that preferentially weights target-preferring neurons

Notably, these computations for identifying target stimuli are highly task-specific, and do not take advantage of population-level computations of the presented orientation. However, optimal approaches for estimating the presented orientation from the neuronal population response (Fig. 6A,B; Ma et al., 2006) predicted that adaptation would decrease thresholds on a discrimination task (250 vs 750 ms: $22.5^\circ: p < 0.001$; $0^\circ: p = 0.14$; Fig. 6C). This improvement in the performance of the neuronal population at discriminating orientation after adaptation was likely because of an adaptation-dependent decrease the variability of the orientation estimate (SD: 250 ms: $6.9^\circ \pm 0.6^\circ$; 750 ms: $8.7^\circ \pm 1.0^\circ$; $p < 10^{-5}$; F test; Fig. 6B) as we did not observe a significant effect of adaptation on the accuracy of the estimate (bias: 250 ms: $4.1^\circ \pm 2.5^\circ$; 750 ms: $6.8^\circ \pm 2.5^\circ$; $p = 0.14$; paired t test; Fig. 6B).

Similar results were found when using the area under the posterior probability distribution as the decision variable (auROC- $22.5^\circ: p < 0.0001$; $0^\circ: p = 0.08$ paired t test; Fig. 6D), or orientation estimates that assume independent Poisson statistics (auROC- $22.5^\circ: p < 0.05$; $0^\circ: p = 0.19$; paired t test; Fig. 6E; Jazayeri and Movshon, 2006), or that were found using a population vector (auROC- $22.5^\circ: p < 0.01$; $0^\circ: p = 0.71$; paired t test; Fig. 6F). This suggests that adaptation actually increases the stimulus information available in the mouse visual cortex. Moreover, of the many computations tested, the only neuronal computation consistent with the observed effects of adaptation on behavior relies on the summed output of target-preferring neurons.

To test whether the perceptual choice is performed by a task-specific computation that relies on the activity of target-preferring neurons, we perturbed the task parameters. We tested animals on a variant of the task where the distractor orientation could either be the trained orientation (0°) or rotated 15° or -15° from the trained orientation (randomly interleaved on a trial-by-trial basis), with the set of target orientations rotated $9-90^\circ$ counter-clockwise relative to the distractor orientation (Fig. 7A). If the mouse uses a more general strategy (for instance comparing the orientation of the distractor to the target), the psychometric curves for the three distractor conditions should be the same. This is because task difficulty depends only on the difference between the target and distractor orientations, not on the absolute target orientation (Fig. 7B). However, we found that all six mice had lower discrimination thresholds in the 15° condition (rotated toward learned targets) compared with the 0° condition, which in turn had lower discrimination thresholds than in the -15° condition (rotated away from learned targets; one-way ANOVA, $p < 10^{-4}$, $DF = 2$, $n = 6$ mice; Fig. 7D,E). The effects on FA rates were not symmetric: all six mice had higher FA rates in the 15° condition compared with the 0° condition (one-way ANOVA with *post hoc* Tukey HSD compared with 0° , $p < 0.05$, $n = 6$ mice; Fig. 7E), but no difference in FA rate between the 0 and -15° conditions (one-way ANOVA with *post hoc* Tukey HSD -15° compared with 0° , $p = 0.97$). This indicates that the circuit has adopted a task-specific computation that treats all stimuli with positive orientation as targets. This can be accomplished by a perceptual choice circuit that more strongly weights neurons with preferred orientation close to 90° .

Behavioral evidence for a perceptual choice circuit that relies on the summed output of visual cortical neurons

The effects of adaptation on behavior suggest that the perceptual choice circuit performs a weighted sum of visual cortical activity and is detecting increases above a learned threshold. To test whether the decision is indeed sensitive to the summed output of the visual cortex, we optogenetically activated inhibitory neurons to transiently suppress activity in V1 on randomly interleaved trials (Fig. 8A,B). The light power was titrated to decrease hit rates for small orientation differences (22.5° , $p < 0.05$, paired *t* test, $n = 4$ mice) without affecting performance on easy trials (90° , $p = 0.54$, paired *t* test). Similar to decreasing the ISI, suppression of V1 increased the discrimination threshold ($p < 0.005$; paired *t* test; Fig. 8B,C) and reduced the FA rate ($p < 0.01$). This effect of suppressing V1 activity is consistent with a perceptual choice circuit that

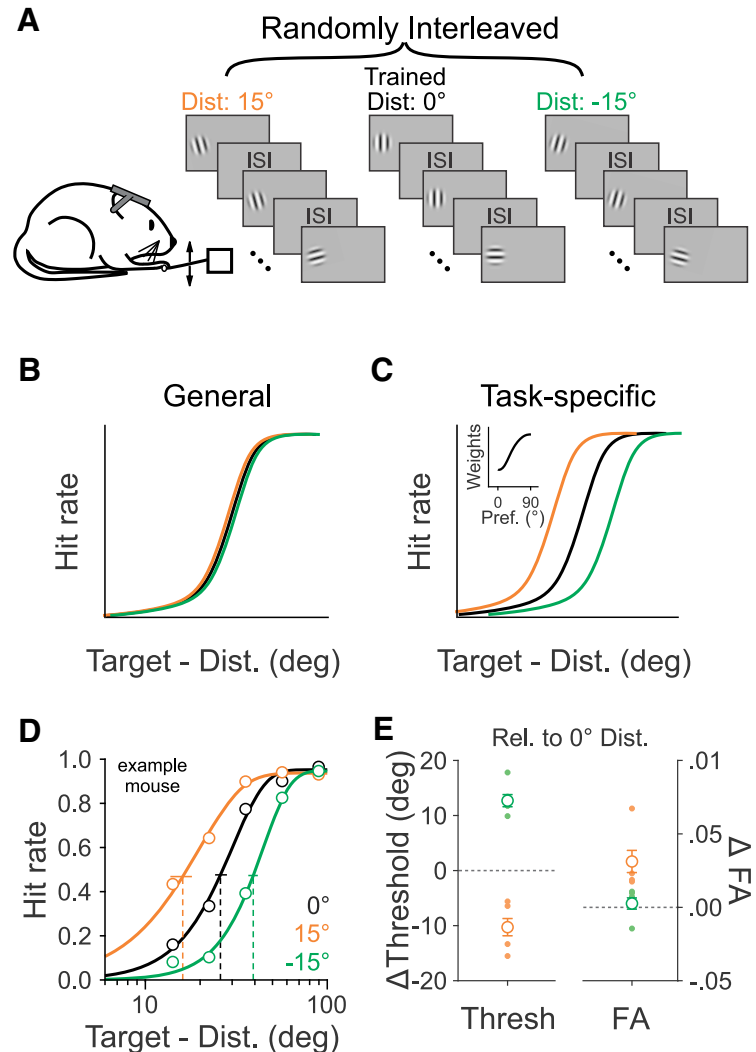


Figure 7. Behavior is inconsistent with a general change detection strategy, but can be explained by a task-specific circuit with biased weights. **A**, Schematic of the trial progression. For each trial, the distractor (Dist) orientation could be 15° (orange), 0° (black), or -15° (green). The target orientations are $9-90^\circ$ counter-clockwise change from the distractor orientation. **B**, If the mouse adopts a general change detection strategy, the discrimination threshold should be similar across different distractor orientations. **C**, If the mouse adopts a task-specific strategy that discriminates a learned quadrant of target orientation space (positive orientations), the discrimination threshold would be lower for 15° distractor and higher for -15° distractor compared with 0° distractor. Inset, Schematic of biased weights across neurons tuned for positive orientations. **D**, Hit rate for an example mouse in which target orientations were sorted according to the distractor orientations. **E**, Summary of the average change in the threshold (left) and FA rate (right) for each distractor orientation relative to the 0° condition across mice. Open circles are the average of all mice ($n = 6$ mice), small filled circles are individual mice; error bars are SEM across mice.

sums and thresholds the output of visual cortical circuits. Moreover, the observed decrease in FA rate with suppression of V1 is also consistent the hypothesis that the perceptual choice circuit ignores distractor preferring neurons: if distractor preferring neurons were to be negatively weighted, then we would expect to see an increase in FA rate when they are suppressed.

Suppression of V1 also significantly reduced the dependence of threshold (two-way ANOVA, main effect of V1 inhibition: $p < 0.01$, $DF = 1$; Fig. 8D,E) and FA rate (two-way ANOVA, main effect of V1 inhibition: $p < 0.01$; Fig. 8F,G) on ISI. This decrease in the effect of ISI on behavior is consistent with a role for adaptation in the visual cortex in the effects that ISI has on task performance. Together, these data provide evidence for a perceptual choice circuit that detects increases in the activity of target preferring visual cortical neurons.

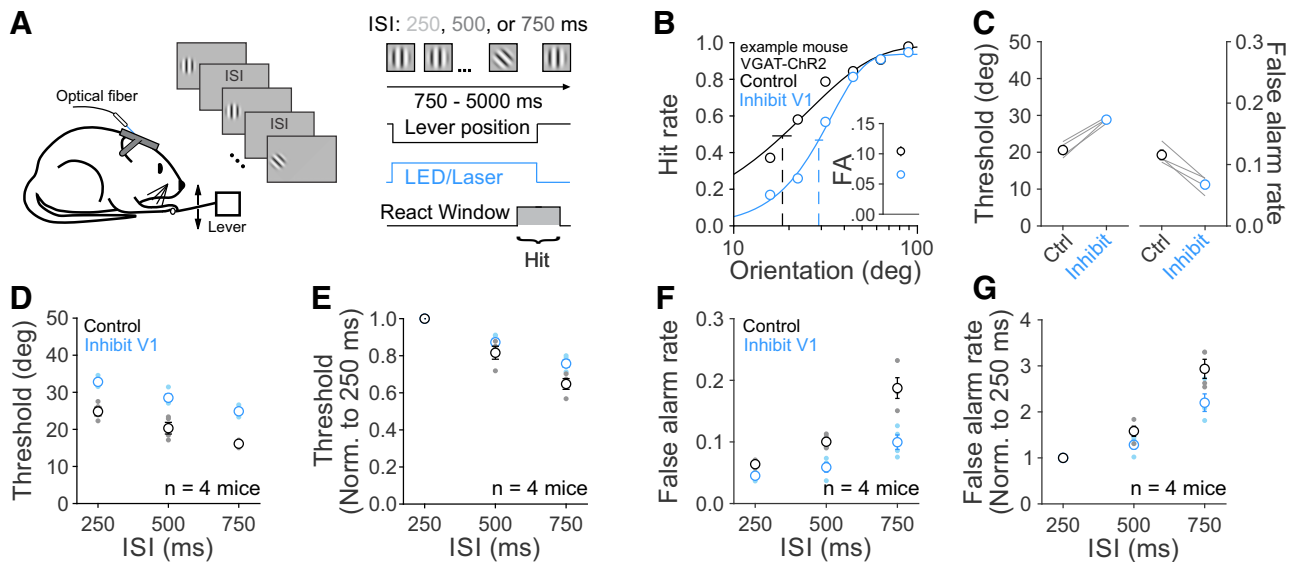


Figure 8. Suppression of V1 increases threshold, decreases FA rate and decreases the dependence of behavior on ISI. **A**, Schematic of behavioral setup and trial progression. V1 inhibition was achieved via optogenetic excitation of Chr2-expressing inhibitory interneurons using VGAT-ChR2 ($n = 2$) and PV::Cre-AAV.flex.ChR2 ($n = 2$). **B**, Hit rate on randomly interleaved control (black) and V1 inhibition (blue) trials, averaged across all ISIs, for an example VGAT-ChR2 mouse; same conventions as in Figure 3B. Inset, FA rate. **C**, Summary of V1 inhibition on threshold (left) and FA rate (right). Connected lines are individual mice ($n = 4$). Error bars are SEM across mice. **D**, Summary of the average threshold for control and V1 inhibition by ISI ($n = 4$ mice). **E**, Summary of change in threshold, normalized to the threshold in the 250 ms condition, for each ISI. **F**, **G**, Same as **D** and **E** for FA rate.

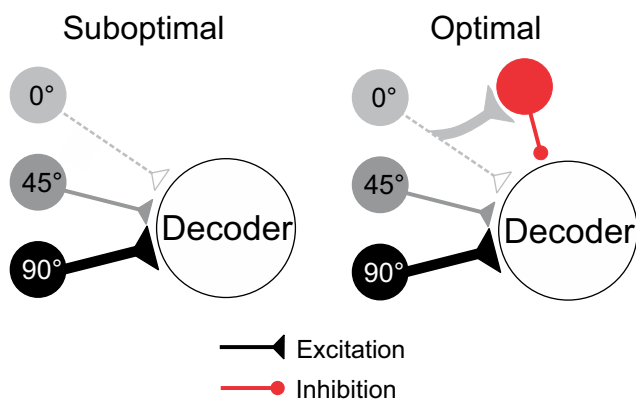


Figure 9. Model for perceptual choice circuit for our orientation discrimination task. Schematic of perceptual choice circuits for suboptimal (left) and optimal (right) computations. Orientation tuned excitatory neurons (shades of gray) converge onto the decoder with weights biased toward target-preferring neurons. We proposed that mice adopt a suboptimal computation implemented in a feedforward excitatory circuit that lacks lateral inhibition (red) from distractor-preferring neurons.

Discussion

To determine how decision-making circuits use information from orientation tuned neuronal populations, we trained the mice to perform a multi-interval, go/no-go orientation discrimination task. We find that adaptation impairs the animals’ ability to discriminate targets near the distractor orientation both in terms of threshold and fraction of rewarded trials (22.5° target, accounting for difference in false alarm rate: 250 ms, 24% trials rewarded; 750 ms: 34.7% trials rewarded). These behavioral data are consistent with a suboptimal computation in which the perceptual choice circuit relies on neurons that prefer target stimuli but fails to appropriately negatively weight neurons that prefer the distractor (Fig. 9). This suggests that sensory information may sometimes be integrated suboptimally when solving perceptual tasks.

In this study we use the term *optimal* to describe the computation purely from the perspective of the best possible use of the information present in the neuronal population activity that can be extracted linearly. Indeed, the optimal computation (using the neuronal weights from the fit optimized to decode the stimulus presented) does significantly better on our discrimination task than the suboptimal computation (using the neuronal weights fit optimized to decode the animals’ behavioral choices) that discards the activity of distractor-preferring neurons (percentage correct, optimal: 97%; suboptimal: 90%; $p < 10^{-6}$). Thus, the finding that the brain favors a suboptimal computation to discriminate orientation may tell us something about the constraints of the circuit. All of the optimal computations that we considered require the negative weighting of specific populations. Such a negative weighting may require a disynaptic circuit to recruit stimulus-specific inhibition and so may be more difficult to learn. Such a computation might also come at a cost. For instance, the optimal circuit imposes at least one synaptic delay and requires additional integration time in the recurrent network, which may slow decision-making. Indeed, there is evidence that the human cortex uses suboptimal computations that ignore distractors or weakly informative features, especially under time constraints (Ho et al., 2012; Oh et al., 2016).

In contrast, the suboptimal computation suggested by our neuronal fit to animals’ behavioral choice relies only on an excitatory, feedforward circuit. Our behavioral experiments, including the distractor orientation perturbations and V1 population inhibition, support this proposed circuit. However, direct evidence for the absence of negative weights can only be provided by applying new tools that allow for the specific activation of functional subsets of neurons in behaving animals (Mardinly et al., 2018). This excitatory computation may have been learned through experience, where the downstream circuit comes to rely on the activity of neurons that increase their activity in response to targets. This computation is amenable to a simple associative learning rule through the repeated pairing of targets and reward (Law and Gold, 2009; Znamenskiy and Zador, 2013; Xiong et al.,

2015). For instance, the positive weighting of target neurons might be achieved through Hebbian long-term synaptic potentiation, whereas decreasing the weight of distractor-preferring neurons might be achieved through long-term depression.

It is possible that the strategy used to train the mice to perform this task, or the specific task parameters used, supported the development of this suboptimal computation. In particular, the strong adaptation of the distractor preferring neurons may decrease the need for a negative weighting of their activity. There may also be something specific to the structure or incentives of go/no-go tasks that bias toward this computation: mice are more robust to increases in stimulus contrast (which increase firing rates) when performing a two-alternative forced choice orientation discrimination compared with a go/no-go version of the task (Long et al., 2015). Importantly, the suboptimal computation is only viable when the distractor orientations are the same across trials. Thus, it is possible that mice initially trained on a task that varies the distractor orientation on a trial-by-trial basis would adopt an optimal perceptual choice circuit.

Although the task-specific, suboptimal perceptual choice circuit is viable, it leaves the animal vulnerable to systematic error. By relying on the absolute firing rate of a subset of neurons, anything that increases firing rates can be mistaken for a target. We think that this is why the mice have a high FA rate following long intervals: the recovery from adaptation with long intervals results in larger than expected firing rates, making the animal respond as if a target had been presented (Fig. 4). Conversely, short ISIs reduce the firing rates of target preferring neurons, reducing the likelihood of a stimulus-evoked release. Indeed, reducing the firing rates in V1 by optogenetic activation of inhibitory neurons also results in a decrease in hit and FA rates, consistent with the hypothesis that the perceptual choice circuit is summing the total activity in this area.

Suppression of V1 and decreasing ISI both reduced hit and FA rate. Such concomitant changes in hit and FA rate are associated with changes in bias (c) as measured using signal detection theory (Green and Swets, 1966). Indeed, decreasing ISI did significantly increase c (22.5° target, 250 ms: 1.28 ± 0.06 ; 500 ms: 1.08 ± 0.05 ; 750 ms: 0.75 ± 0.06 ; $p < 10^{-5}$; one-way ANOVA; $n = 11$ mice). However, this is most likely because of an adaptation-dependent reduction in firing rates in response to both targets and distractors (i.e., signal and noise), which will decrease the optimal criterion (and increase bias because bias is measured relative to optimal). Thus, changes in sensory processing in V1, in the absence of cognitive mechanisms whereby the mouse increases its criterion, are sufficient to explain this change in bias (Witt et al., 2015; Jin and Glickfeld, 2018). The lack of effect of ISI on lapse rate also argues against a cognitive mechanism for the effects of ISI on behavior. Finally, the disproportional increase in threshold (and decrease in FA rates) with suppression of V1 is consistent with effects of ISI on behavior acting through effects on sensory coding, and not through more cognitive mechanisms like forward masking and attentional blink (Raymond et al., 1992; Macknik and Livingstone, 1998; Alwis et al., 2016). Notably, these phenomena also tend to act on much shorter time-scales (tens of milliseconds) than intervals used in this study, making them unlikely candidates to explain the effects of ISI on behavior. Together, these observations support the argument that varying ISI affects discrimination thresholds through altering adaptation state in the visual cortex.

The neuronal data that we used to generate the model predictions was collected from naive mice that were passively viewing the visual stimuli. This was done to generate full orientation tun-

ing curves in the adapted and unadapted conditions, as well as to avoid contamination of non-sensory signals. However, this means that the effects of training or active behavioral engagement are not included in our model predictions. Because the tuning of visual cortical neurons can be affected by visual experience (Schoups et al., 2001; Kreile et al., 2011; Goltstein et al., 2013), it is possible that our task training paradigm (abundance of 0° distractors and 90° targets, 6 d/week, >3 months) induced a change in the representation of orientation in V1 neurons. However, the orientation identification models are designed to account for skewed distributions, and thus we do not expect that experience-dependent changes in the representation of orientation would substantially affect orientation identification models' predictions. On the other hand, the degree of adaptation is dependent on both the number of distractors preceding the target as well as task engagement (Keller et al., 2017), and therefore the mapping of the neural data onto the behavioral data may not be straightforward. Nonetheless, we were able to reliably fit the neural data to the behavioral data, suggesting that there may in fact be a linear transform of firing rates across behavioral state.

Our behavioral, physiological, and computational approaches reveal that the circuit adopts an all-excitatory computation to solve an orientation discrimination task. Future work will determine the conditions that support the use of this specific computation, and in which tasks and species it arises. Nonetheless, these results reveal that the perceptual choice circuit likely does not always take advantage of all available sensory information, perhaps in the interest of expediency. Specifically, ignoring important information leaves the mouse vulnerable to the effects of adaptation, but may enable both robust associative learning and fast decision making via a simple feedforward circuit.

References

- Alwis DS, Richards KL, Price NS (2016) Masking reduces orientation selectivity in rat visual cortex. *J Neurophysiol* 116:2331–2341.
- Barlow HB, Macleod DI, van Meeteren A (1976) Adaptation to gratings: no compensatory advantages found. *Vision Res* 16:1043–1045.
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon J (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.
- Clifford CW (2002) Perceptual adaptation: motion parallels orientation. *Trends Cogn Sci* 6:136–143.
- Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051–2062.
- Dragoi V, Sharma J, Sur M (2000) Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron* 28:287–298.
- Dragoi V, Sharma J, Miller EK, Sur M (2002) Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nat Neurosci* 5:883–891.
- Driscoll WC (1996) Robustness of the ANOVA and Tukey–Kramer statistical tests. *Comput Ind Eng* 31:265–268.
- Dubner R, Zeki SM (1971) Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res* 35:528–532.
- Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. *Science* 233:1416–1419.
- Glickfeld LL, Histed MH, Maunsell JH (2013) Mouse primary visual cortex is used to detect both orientation and contrast changes. *J Neurosci* 33:19416–19422.
- Goltstein PM, Coffey EB, Roelfsema PR, Pennartz CM (2013) *In vivo* two-photon Ca^{2+} imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. *J Neurosci* 33:11540–11555.
- Graf AB, Kohn A, Jazayeri M, Movshon JA (2011) Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci* 14:239–245.
- Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

- Ho T, Brown S, van Maanen L, Forstmann BU, Wagenmakers EJ, Serences JT (2012) The optimality of sensory processing during the speed-accuracy tradeoff. *J Neurosci* 32:7992–8003.
- Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 148:574–591.
- Hung CP, Kreiman G, Poggio T, DiCarlo JJ (2005) Fast readout of object identity from macaque inferior temporal cortex. *Science* 310:863–866.
- Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. *Nat Neurosci* 9:690–696.
- Jin M, Glickfeld LL (2018) Contribution of sensory encoding to measured bias. *bioRxiv*. Advance online publication. Retrieved October 16, 2018. doi.org/10.1101/444430.
- Keller AJ, Houlton R, Kampa BM, Lesica NA, Mrcic-Flogel TD, Keller GB, Helmchen F (2017) Stimulus relevance modulates contrast adaptation in visual cortex. *eLife* 6:e21589.
- Kohn A, Movshon JA (2004) Adaptation changes the direction tuning of macaque MT neurons. *Nat Neurosci* 7:764–772.
- Kreile AK, Bonhoeffer T, Hübener M (2011) Altered visual experience induces instructive changes of orientation preference in mouse visual cortex. *J Neurosci* 31:13911–13920.
- Krekelberg B, van Wezel RJ, Albright TD (2006) Adaptation in macaque MT reduces perceived speed and improves speed discrimination. *J Neurophysiol* 95:255–270.
- Law CT, Gold JJ (2009) Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat Neurosci* 12:655–663.
- Levinson E, Sekuler R (1976) Adaptation alters perceived direction of motion. *Vision Res* 16:779–781.
- Long M, Jiang W, Liu D, Yao H (2015) Contrast-dependent orientation discrimination in the mouse. *Sci Rep* 5:15830.
- Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. *Nat Neurosci* 9:1432–1438.
- Macknik SL, Livingstone MS (1998) Neuronal correlates of visibility and invisibility in the primate visual system. *Nat Neurosci* 1:144–149.
- Mardinly AR, Oldenburg IA, Pégard NC, Sridharan S, Lyall EH, Chesnov K, Brohawn SG, Waller L, Adesnik H (2018) Precise multimodal optical control of neural ensemble activity. *Nat Neurosci* 21:881–893.
- Müller JR, Metha AB, Krauskopf J, Lennie P (1999) Rapid adaptation in visual cortex to the structure of images. *Science* 285:1405–1408.
- Oh H, Beck JM, Zhu P, Sommer MA, Ferrari S, Egnér T (2016) Satisficing in split-second decision making is characterized by strategic cue discounting. *J Exp Psychol Learn Mem Cogn* 42:1937–1956.
- Ollerenshaw DR, Zheng HJV, Millard DC, Wang Q, Stanley GB (2014) The adaptive trade-off between detection and discrimination in cortical representations and behavior. *Neuron* 81:1152–1164.
- Patterson CA, Wissig SC, Kohn A (2013) Distinct effects of brief and prolonged adaptation on orientation tuning in primary visual cortex. *J Neurosci* 33:532–543.
- Poort J, Khan AG, Pachitariu M, Nemri A, Orsolich I, Krupic J, Bauza M, Sahani M, Keller GB, Mrcic-Flogel TD, Hofer SB (2015) Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* 86:1478–1490.
- Pouget A, Dayan P, Zemel RS (2003) Inference and computation with population codes. *Annu Rev Neurosci* 26:381–410.
- Raymond JE, Shapiro KL, Arnell KM (1992) Temporary suppression of visual processing in an RSVP task: an attentional blink? *J Exp Psychol Hum Percept Perform* 18:849–860.
- Regan D, Beverley KI (1985) Postadaptation orientation discrimination. *J Opt Soc Am A* 2:147–155.
- Resulaj A, Ruediger S, Olsen SR, Scanziani M (2018) First spikes in visual cortex enable perceptual discrimination. *eLife* 7:e34044.
- Salzman CD, Britten KH, Newsome WT (1990) Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346:174–177.
- Schiller PH (1993) The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis Neurosci* 10:717–746.
- Schoups A, Vogels R, Qian N, Orban G (2001) Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412:549–553.
- Seriès P, Stocker AA, Simoncelli EP (2009) Is the homunculus “aware” of sensory adaptation? *Neural Comput* 21:3271–3304.
- Stocker AA, Simoncelli EP (2006) Sensory adaptation within a Bayesian framework for perception. *Adv Neural Inf Process Syst* 18:1291–1298.
- Stroud AC, Ledue EE, Crowder NA (2012) Orientation specificity of contrast adaptation in mouse primary visual cortex. *J Neurophysiol* 108:1381–1391.
- Tiefenau A, Neubauer H, von Specht H, Heil P (2006) Correcting for false alarms in a simple reaction time task. *Brain Res* 1122:99–115.
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398.
- Wark B, Lundstrom BN, Fairhall A (2007) Sensory adaptation. *Curr Opin Neurobiol* 17:423–429.
- Witt JK, Taylor JE, Sugovic M, Wixted JT (2015) Signal detection measures cannot distinguish perceptual biases from response biases. *Perception* 44:289–300.
- Xiong Q, Znamenskiy P, Zador AM (2015) Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* 521:348–351.
- Zaidi Q, Marshall J, Thoen H, Conway BR (2014) Evolution of neural computations: mantis shrimp and human color decoding. *Iperception* 5:492–496.
- Zavitz E, Yu HH, Rowe EG, Rosa MG, Price NS (2016) Rapid adaptation induces persistent biases in population codes for visual motion. *J Neurosci* 36:4579–4590.
- Zhang P, Bao M, Kwon M, He S, Engel SA (2009) Effects of orientation-specific visual deprivation induced with altered reality. *Curr Biol* 19:1956–1960.
- Znamenskiy P, Zador AM (2013) Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497:482–485.