Systems/Circuits

Temporal Contingencies Determine Whether Adaptation Strengthens or Weakens Normalization

Amir Aschner,[1] Samuel G. Solomon,[2] Michael S. Landy,[3] David J. Heeger,[3] and Adam Kohn[1,4,5]

[1]Dominik Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461, [2]Department of Experimental Psychology, University College London, London, United Kingdom WC1H 0AP, [3]Department of Psychology and Center for Neural Science, New York University, New York, New York 10003, [4]Department of Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, Bronx, New York 10461, and [5]Department of Systems and Computational Biology, Albert Einstein College of Medicine, Bronx, New York 10461

A fundamental and nearly ubiquitous feature of sensory encoding is that neuronal responses are strongly influenced by recent experience, or adaptation. Theoretical and computational studies have proposed that many adaptation effects may result in part from changes in the strength of normalization signals. Normalization is a "canonical" computation in which a neuron's response is modulated (normalized) by the pooled activity of other neurons. Here, we test whether adaptation can alter the strength of cross-orientation suppression, or masking, a paradigmatic form of normalization evident in primary visual cortex (V1). We made extracellular recordings of V1 neurons in anesthetized male macaques and measured responses to plaid stimuli composed of two overlapping, orthogonal gratings before and after prolonged exposure to two distinct adapters. The first adapter was a plaid consisting of orthogonal gratings and led to stronger masking. The second adapter presented the same orthogonal gratings in an interleaved manner and led to weaker masking. The strength of adaptation's effects on masking depended on the orientation of the test stimuli relative to the orientation of the adapters, but was independent of neuronal orientation preference. Changes in masking could not be explained by altered neuronal responsivity. Our results suggest that normalization signals can be strengthened or weakened by adaptation depending on the temporal contingencies of the adapting stimuli. Our findings reveal an interplay between two widespread computations in cortical circuits, adaptation and normalization, that enables flexible adjustments to the structure of the environment, including the temporal relationships among sensory stimuli.

Key words: adaptation; cross-orientation suppression; macaque; V1

Significance Statement

Two fundamental features of sensory responses are that they are influenced by adaptation and that they are modulated by the activity of other nearby neurons via normalization. Our findings reveal a strong interaction between these two aspects of cortical computation. Specifically, we show that cross-orientation masking, a form of normalization, can be strengthened or weakened by adaptation depending on the temporal contingencies between sensory inputs. Our findings support theoretical proposals that some adaptation effects may involve altered normalization and offer a network-based explanation for how cortex adjusts to current sensory demands.

Introduction

Sensory statistics vary over time, imposing changing demands on the brain. Sensory neuronal responses adjust to the recent pattern of inputs, or adaptation. This adjustment is a fundamental and nearly ubiquitous feature of sensory encoding (Clifford et al., 2007; Kohn, 2007; Schwartz et al., 2007; Wark et al., 2007; Rieke and Rudd, 2009; Solomon and Kohn, 2014; Webster, 2015).

Adaptation has diverse effects on sensory neurons, including changes in responsivity and altered tuning or selectivity (for review, see Solomon and Kohn, 2014). Theoretical and computational work has proposed that much adaptation phenomenology could be explained through adjustments of normalization signals (Heeger, 1992; Wainwright et al., 2002; Lochmann et al., 2012;

Solomon and Kohn, 2014; Snow et al., 2016; Westrick et al., 2016). Normalization is a "canonical" computation in which a neuron's response is divisively modulated by the activity of other neurons, a normalization pool (Heeger, 1992; Carandini and Heeger, 2011). Normalization can explain nonlinear response properties of cortical visual neurons such as cross-orientation suppression (hereafter, masking; Morrone et al., 1982; Carandini et al., 1997a; Priebe and Ferster, 2006) and spatial contextual effects (surround suppression; Cavanaugh et al., 2002; Coen-Cagli et al., 2012, 2015). Normalization has also been invoked to explain phenomena as varied as olfactory encoding in *Drosophila* and value encoding in primates (Carandini and Heeger, 2011).

Previous work has provided some evidence that adaptation alters the suppression attributed to normalization. For instance, adaptation of the receptive field (RF) surround weakens the suppression it provides in the lateral geniculate nucleus (LGN) and primary visual cortex (V1), leading to response facilitation (Webb et al., 2005; Camp et al., 2009; Wissig and Kohn, 2012; Patterson et al., 2013). How adaptation affects masking within the RF is unclear. One study found that masking was unaltered in V1 after adaptation (Freeman et al., 2002), but others reported it was weakened (Li et al., 2005; Dhruv et al., 2011; see also Kaliuk-hovich and Vogels, 2016). Critically, no systematic study has reported that the suppression attributed to normalization can be strengthened by adaptation. The ability of recent sensory experience to strengthen normalization is a critical component of most normalization-based models of adaptation effects.

The Hebbian normalization model (Westrick et al., 2016) was proposed to explain adaptation effects in visual cortex and provides clear predictions about when normalization should be strengthened or weakened by adaptation. Specifically, it predicts that changes in normalization signals between neurons depend on their recent history of coactivation (also see Barlow and Földiák, 1989; Wainwright et al., 2002 and Hosoya et al., 2005 for related suggestions). When the normalization pool and a target neuron are consistently coactive, normalization should be strengthened. Conversely, when the neuron and normalization pool are driven asynchronously, normalization should be weakened.

We tested these predictions using extracellular recordings of macaque V1 neurons. We compared the strength of masking before and after consistent pairing of a target grating and mask (contingent adaptation) or the asynchronous presentation of these stimuli. We show that masking can be strengthened or weakened depending on whether mask and target are consistently paired or separated in time.

Materials and Methods

Surgery. We made recordings from five adult male macaque monkeys (*Macaca fascicularis*). Animals were administered glycopyrrolate (0.01 mg/kg) and diazepam (1.5 mg/kg) shortly before the induction of anesthesia with ketamine (10 mg/kg). Animals were then intubated and provided isoflurane (1–2%). Intravenous catheters were placed in the saphenous veins of each leg. Animals were positioned in a stereotaxic device and a craniotomy and durotomy were performed over V1 (~5 mm posterior to the lunate sulcus and 10 mm lateral to the midline). A 10×10 microelectrode array (400 μm spacing, 1 mm length; Blackrock Microsystems) was implanted and the craniotomy was covered with agar to prevent desiccation. Postsurgical anesthesia was maintained with a venous infusion of sufentanil citrate (6–24 μg/kg/h, adjusted as needed) in Normosol solution with dextrose. Vecuronium bromide (150 μg/kg/h) was provided intravenously to minimize eye movements. Vital signs, including heart rate, SpO_2, ECG, blood pressure, EEG, end-tidal CO_2 partial pressure, core temperature, urinary output, and airway pres-

sure, were constantly monitored to ensure adequate anesthesia and physiological state. Heating pads maintained rectal temperature near 37°C. Topical atropine was used to dilate the pupils. Corneas were protected with gas-permeable contact lenses. Supplementary lenses were used to bring the retinal image into focus. Antibiotics (Baytril, 2.5 mg/kg) and a corticosteroid (dexamethasone, 1 mg/kg) were administered daily.

All procedures were approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine and were in compliance with the guidelines set forth in the National Institutes of Health's *Guide for the Care and Use of Laboratory Animals*.

Recording and visual stimuli. Extracellular voltage signals were filtered from 0.5–7.5 kHz. Waveforms that exceeded a user-defined voltage threshold (usually 5 times the root-mean-square signal on each channel) were digitized at 30 kHz. Waveforms were classified using the Plexon Offline Sorter into single-unit and multiunit clusters. We computed signal-to-noise ratios (SNRs) for each unit as the ratio of the amplitude of the average waveform to the SD of the individual waveforms (Kelly et al., 2007). Units with an SNR $\geq$ 3.5 were classified as single units. Results were similar for these single units (13% of units) and for multiunit clusters and are therefore presented together (see also Wissig and Kohn, 2012). Due to the duration of the adaptation experiments, ~2 h, each recording was sorted separately.

Visual stimuli were generated with custom software based on OpenGL (EXPO; P. Lennie) and displayed on a calibrated cathode ray tube monitor (Hewlett Packard p1230; 1024 $\times$ 768 pixels; 100 Hz frame rate, ~40 cd/m^2 mean luminance) viewed at a distance of 110 cm and subtending ~20° of visual angle. Spatial RFs of each unit were estimated by occluding one eye and presenting small patches of drifting gratings (0.5° diameter; 4 orientations, 1 cycle/°, 3 Hz drift rate, 250 ms presentation) at 225 distinct positions spanning a 3° $\times$ 3° region of visual space. Subsequent stimuli were centered in the aggregate RF of the recorded units.

Stimulus orientations for subsequent experiments were determined by presenting a continuous, pseudorandom sequence of 16 full-contrast sinusoidal gratings (1.5° patch diameter, 1 cycle/°, 3 Hz drift rate) at equally spaced orientations (22.5° steps; 1 s presentation). Gratings were presented monocularly, in a hard-edged circular window.

To characterize normalization, we measured masking (or cross-orientation suppression; Morrone et al., 1982; Bonds, 1989; DeAngelis et al., 1992; Carandini et al., 1997a) by presenting a drifting, sinusoidal grating (the target) with an overlapping, orthogonal grating (the mask). Both gratings were presented monocularly in a hard-edged circular window 1.5° in diameter with a spatial frequency of 1.5 cycle/° and drift rate of 3 Hz. On each trial, the contrast of the target and mask gratings varied independently over 5 values (0%, 6.25%, 12.5%, 25%, and 50%), generating a matrix of 25 test stimuli (Fig. 1A). For each neuron, we defined the target as the grating that evoked the stronger response at 50% contrast when presented alone.

To measure the strength of masking in control conditions (i.e., before adaptation), we used 1 s presentations of each of the 25 test stimuli interleaved with 5 s presentations of a uniform gray screen (Fig. 1C). The order of presentation was randomized within each block of trials and each stimulus was presented 20 times. We refer to these as preadaptation responses even though they are more accurately termed responses measured during adaptation to a gray, uniform screen. We then measured responses after adaptation, using an "adapt-test-top-up" paradigm (Fig. 1C; Kohn and Movshon, 2004). After 40 s of initial exposure to the adapter, we presented each of the 25 test stimuli for 1 s (block randomized, as above) interleaved with 5 s presentations of the adapter. Each stimulus was presented 20 times.

We used two types of adapters (Fig. 1B). For asynchronous adaptation, the target and mask grating (each 50% contrast) alternated in time, with each drifting grating displayed for 250 ms and no interstimulus interval. For contingent adaptation, the same two gratings were presented simultaneously (forming a plaid) for 250 ms, alternated with 250 ms of a uniform gray screen. The presentation of a blank screen between plaids ensured that the time-averaged contrast of each adapter type was equal between adaptation paradigms.

To assess the stimulus (orientation) specificity of adaptation effects, we measured responses to mask and target stimuli that were rotated
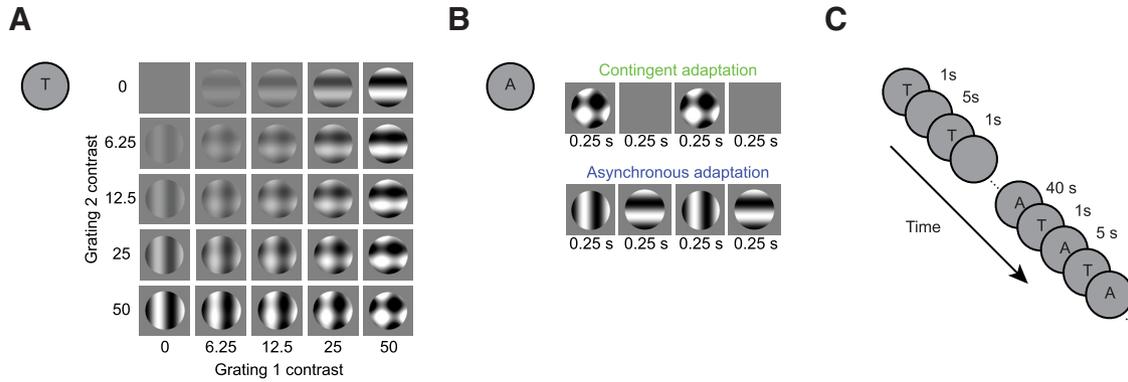
Figure 1. Stimulus protocol. *A*, Ensemble of test stimuli (T). *B*, Temporal structure and form of the contingent and asynchronous adapters (A). *C*, Temporal structure of the experiment, which involved measuring responses before adaptation, adapting, and then measuring responses using a top-up/test paradigm.

relative to the adapter by 45° (e.g., the effects of adaptation to 0–90° plaids were evaluated with 45–135° plaids).

Finally, to determine how adaptation effects depended on the orientation preference of the recorded unit, we conducted additional experiments in which masking was measured with a reduced test stimulus ensemble combined with interleaved presentations of gratings of different orientations. Specifically, we presented the target and mask at contrasts of 0%, 6.25%, 12.5%, 25%, and 50% either in isolation or paired with the 50% contrast orthogonal grating. Tuning was measured with 50% contrast gratings spanning 180° of orientation in 22.5° steps. The temporal structure of these experiments was identical to that described above.

Analyses. Data analysis was performed in MATLAB (The Math-Works). We measured responses for two cycles of the grating (666 ms) beginning 333 ms after stimulus onset. For each neuron, we measured both the mean response (F0) and its modulation by the grating drift frequency (F1) for the high-contrast target. We used the response measure (F0 or F1) that gave the higher value for all subsequent analyses.

We quantified masking strength with an index based on the area-under-the-curve (AUC) of the contrast response functions for the target grating, in the presence and absence of the mask (similar to Carandini et al., 1997b, 1998; Wissig and Kohn, 2012; using log contrast values as the *x*-axis except for the placement of zero contrast, as in Fig. 2). We used this approach, rather than fitting descriptive functions to the data (as in Freeman et al., 2002; Dhruv et al., 2011), because postadaptation responses often showed little evidence of contrast saturation, leaving model parameters poorly constrained by the data.

We measured the AUC for each mask contrast separately (AUC_{TM}) after subtracting the response to the mask when presented alone. The AUC for the target alone (AUC_T) was that obtained for masks of zero contrast. Negative AUC_{TM} values were set to zero, so that our masking index (MI) was constrained to lie between −1 and 1. The MI was defined as follows:

$$MI = \frac{(AUC_T - AUC_{TM})}{(AUC_T + AUC_{TM})} \qquad (1)$$

An MI near 1 indicates strong masking ($AUC_T \gg AUC_{TM}$), whereas a value near 0 indicates little masking ($AUC_T \approx AUC_{TM}$). Negative values indicate that the response to the target alone (AUC_T) was smaller than the difference between the response to the plaid and the response to the mask alone (AUC_{TM}).

To ensure that we could measure masking, if it occurred, we only analyzed data from cells for which the response to the 50% contrast target was greater than the mean +3 SEM of the spontaneous firing rate both before and after adaptation (38% of units excluded for contingent adaptation experiments; 57% for asynchronous adaptation). The relatively high proportion of excluded units arose because we presented gratings at only two orientations and one spatial and temporal frequency. Therefore, these stimuli failed to drive robust responses in many of the recorded units.

Note that more units were excluded for asynchronous adaptation experiments than for contingent adaptation because responses were more strongly reduced after asynchronous adaptation (see Figs. 4 and 5). Further, we found that units with particularly weak responses after adaptation tended to have higher preadaptation MIs. As a result, the preadaptation MI was lower for asynchronous than contingent adapters among the units analyzed. To ensure that this mismatch in preadaptation MI did not contribute to the observed differences in the effects of contingent and asynchronous adaptation, we repeated our analyses after matching the preadaptation MIs for the two datasets. Specifically, we binned the preadaptation MIs (bin width of 0.2) in each dataset and then randomly selected for each bin the same number of units in the two datasets. This matching led to preadaptation masking values that were statistically indistinguishable. All of the differences between the effects of contingent and asynchronous adaptation that are reported herein were equally evident in this subset of data (data not shown).

To control for any confounds due to adaptation-induced changes in responsivity, we performed additional analyses. We first calculated a suppression index (SI) from the postadaptation responses defined as follows:

$$SI = 1 - \frac{R_{TM}}{R_T + R_M} \qquad (2)$$

where R_T and R_M are the responses to the 50% contrast target and mask, respectively, and R_{TM} is the response to the two presented together (i.e., the plaid).

We then identified target and mask contrasts in the preadaptation condition that produced responses equivalent to those observed after adaptation. To do so, we needed to interpolate the preadaptation measurements, so we fitted those responses (R) for each cell with a descriptive function as follows:

$$R(T_c, M_c) = B + \frac{(A_T T_c + A_M M_c)^n}{1 + (d_T T_c)^n + (d_M M_c)^n} \qquad (3)$$

where T_c and M_c are the contrasts of the target and mask, respectively, B is the spontaneous firing rate, A_T and A_M determine the drive provided by each grating, n approximates an expansive nonlinearity, and d_T and d_M capture the weight of each grating in the normalization signal. These parameters (A_T, A_M, d_T, d_M, and n) were estimated by maximizing the log likelihood of the data given the model predictions, assuming Poisson spiking statistics (El-Shamayleh and Movshon, 2011). Fit quality was characterized by the normalized log likelihood, where the lower bound (a value of 0) was the likelihood of a model with predicted responses equal to the average response across all conditions, and the upper bound (a value of 1) was the likelihood calculated by using the data as the model (Stocker and Simoncelli, 2006). The mean fit quality of the analyzed units (see below) was 0.86.

We used the model fit to identify contrasts of the target and mask that generated predicted responses equal to those observed after adaptation

(R_T and R_M of Eq. 2) and to estimate responses to plaid stimuli composed of the target and mask at these "rate-matched" contrasts. We then calculated a preadaptation rate-matched SI, as in Eq. 2, from these responses. This analysis allowed us to compare SI values before and after adaptation for which, by definition, responsivity to the component gratings was equal. Therefore, we could compare how summation was affected by adaptation when the efficacy of the individual gratings was identical before and after adaptation.

For this analysis, we included units only if the following requirements were met. First, the preadaptation and postadaptation responses to both the 50% contrast mask and the 50% target had to be at least 3 SEMs above the mean spontaneous rate because we could only measure summation for mean-matched responses when both stimuli generated measurable responses. This criterion excluded 57% of units in the contingent-adaptation experiment and 77% in the asynchronous-adaptation experiments. Second, the model fit quality to the preadaptation data had to be at least 0.7 so that the matching to preadaptation responses was meaningful (18% of remaining units excluded). Third, we had to be able to identify a contrast that evoked a matched response in the preadaptation epoch (30% of the remaining units excluded). Because of these requirements, we were able to perform this analysis on 12% of the recorded units.

To estimate neuronal orientation preference, we fitted the responses to gratings spanning a 180° range of orientations with a von Mises function using the maximum-likelihood fitting procedure described above. To ensure that our measurements of preference were meaningful, we only analyzed responses for which fit quality exceeded 0.7 (mean fit quality of selected neurons was 0.86). We also used the fitted functions to estimate tuning-curve gain before and after adaptation defined as the maximum predicted evoked response.

To quantify statistical significance, we used t tests (two-tailed) unless otherwise indicated. All error estimates indicate 1 SE from the mean unless indicated otherwise.

Response-product homeostasis model. We performed simulations to compare our neurophysiological results with those predicted by a recently proposed rule for updating normalization weights based on stimulus history (the "Hebbian normalization model" of Westrick et al., 2016).

Our variant of the Hebbian normalization model consisted of 120 neurons with preferred orientations equally spaced from 0° to 178.5°. Responses were simulated by computing a feedforward drive for each neuron and then normalizing by the feedforward drive to the other neurons. Specifically, the feedforward drive for cell i was defined as follows:

$$F_i(\theta) = C * \exp[k * \cos(\theta - \theta_i) - 1] + B \tag{4}$$

where θ is the grating orientation and C is its contrast, θ_i is the preferred orientation of cell i, and B is an offset. The response of the cell was then computed as follows:

$$R_i(\theta) = \frac{F_i(\theta)^2}{\sigma^2 + \sum_{j=1}^{N} W_{i,j} F_j(\theta)^2} \tag{5}$$

where W_{ij} is the weight that defines unit j's contribution to i's normalization pool and σ is the contrast saturation constant.

The normalization weight between each pair of neurons was updated based on their responses to the current stimulus as follows:

$$W_{i,j}^{t+1} = W_{i,j}^t + \alpha(R_i^t R_j^t - H_{i,j}) \tag{6}$$

where α is a factor that determines the update rate and $H_{i,j}$ is a homeostatic target defined as the mean response product of the pair.

We also considered a variant of the model that included a "fatigue" factor that was defined as follows:

$$G_i^{t+1} = G_i^t + \beta(R_i^t / Rmax_i) \tag{7}$$

where $Rmax_i$ is the response of the neuron to its preferred grating at full contrast. The response of each neuron, R_i, as defined in Equation 5, was then scaled by $1 - G_i$ at each time step. The initial value of G_i was 0 for all units.

For the model simulations without fatigue (see Figs. 6 and 7), we used the following parameters: k was 3, producing a bandwidth (full width at half-height) of $\sim$40°; B was 0.1, σ was 0.35, α was 0.005, and $H_{i,j}$ was defined as the mean of the preadaptation response product of the pair to gratings of all orientation presented at 36% contrast. The initial normalization weights were 0.027 and the model was run for 200 time steps. For the simulations with fatigue (see Fig. 7), the parameters were as above except that B was 0.3, α was 0.01, $H_{i,j}$ was defined with gratings at 50% contrast, and the model was run for 1000 time steps. For fatigue, the parameter β was set to 0.015 and values of $G > 0.55$ were set to 0.55.

Code availability. All MATLAB code and analysis will be made available upon reasonable request to the corresponding author.

Results

We recorded neurons in the superficial layers of V1 in anesthetized macaque monkeys using microelectrode arrays. Spatial RFs were located in the lower visual field, at an eccentricity of 2–4°. Recordings consisted of both well isolated single units (13% of cases) and small multiunit clusters that passed inclusion criteria (see Materials and Methods). Results for single-unit and multiunit clusters were not distinguishable and are reported together.

We first measured responses to a drifting sinusoidal grating (target) alone and when paired with an overlapping, orthogonal grating (mask). These stimuli have been used extensively in previous work to measure masking in visual cortex (Morrone et al., 1982; Bonds, 1989; DeAngelis et al., 1992; Carandini et al., 1997b, 1998; Freeman et al., 2002). As expected, the mask reduced responses to target stimuli, particularly those of low contrast, as shown for an example unit in Figure 2A. To quantify the strength of masking, we computed a masking index (MI) for which a value of 0 indicates no masking and a value of 1 indicates complete suppression of responses to the target (see Materials and Methods). For the example unit of Figure 2A, the MI before adaptation was 0.15 for a 6% contrast mask and 0.49 for a 50% mask. Across units, the average MIs before adaptation were 0.21 ± 0.01 (6% mask), 0.30 ± 0.01 (12% mask), 0.38 ± 0.02 (25% mask), and 0.58 ± 0.01 (50% mask).

We then tested the effect of contingent adaptation by consistently presenting the mask and target grating together (Fig. 1B,C). For the example unit, contingent adaptation increased the suppressive influence of the mask; for the 25% contrast mask, the MI increased from 0.33 to 0.55 after adaptation (cf. light and dark shaded areas in Fig. 2A, top vs bottom). Across the population, contingent adaptation caused a consistent increase in the MI. For the 25% contrast mask, the index increased by 0.21 ± 0.02, or $>50\%$ (Fig. 3A, green; $p < 0.001$). With the exception of the lowest contrast mask, for which masking is weakest, contingent adaptation consistently strengthened the suppressive effect of the mask (Fig. 3B, green; $p < 0.001$ for 12%, 25% and 50% masks).

In stark contrast to the effects of contingent adaptation, asynchronous adaptation—interleaving the presentation of the target and mask—reduced masking. In the example unit of Figure 2B, the MI for the 25% contrast mask was reduced from 0.39 before adaptation to 0 after adaptation. Across the population, the MI for the 25% mask was reduced by 0.32 ± 0.04, or nearly 85% (Fig. 3A, blue). Similar effects were evident for masks of other contrasts (Fig. 3B, blue; $p < 0.001$ for 6%, 25%, and 50% masks, $p = 0.04$ for the 12% mask).

As evident for the example cells of Figure 2, both contingent and asynchronous adaptation also reduced responsivity to all
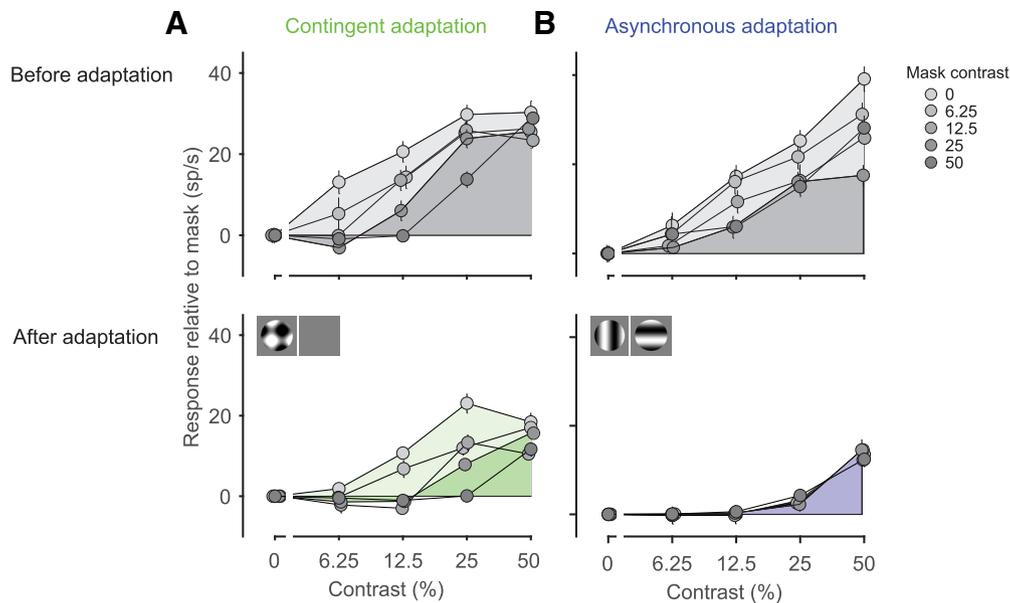
Figure 2. Example neurons. *A*, Example unit for contingent adaptation. Top, Contrast–response functions for target stimuli presented in isolation or with masks of different contrasts (indicated by symbols with different shades of gray). Responses are measured relative to the response evoked by each mask. Fill indicates the AUC used to calculate the MI. Positions of the symbols along the abscissa have been jittered slightly to improve visibility. Bottom, Responses of the same cell after contingent adaptation. *B*, Example unit for asynchronous adaptation following the conventions of *A*. Error bars indicate 1 SEM.
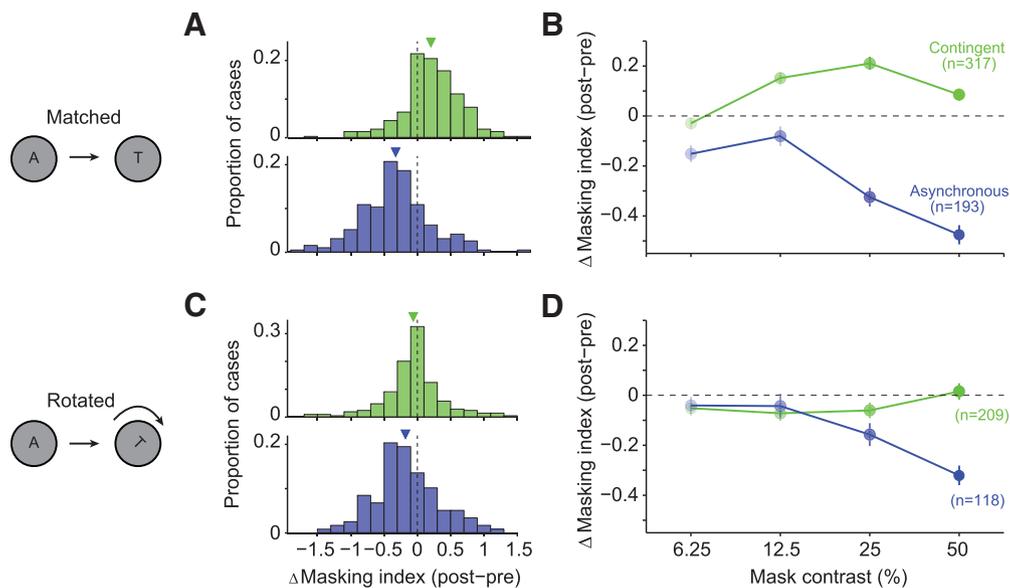
Figure 3. Population summary. *A*, Change in MI for the 25% contrast mask (postadaptation values − preadaptation values) when test stimuli are matched in orientation to the adapter. Data for contingent adaptation are shown in green; those for asynchronous adaptation in blue. Values larger than zero indicate stronger masking; those less than zero indicate weaker masking. Arrowheads indicate mean of the distributions. *B*, Mean change in MI after contingent (green) or asynchronous (blue) adaptation as a function of mask contrast. *C*, *D*, Effects of contingent and asynchronous adaptation on test stimuli with orientation rotated by 45° from the adapter following the conventions of *A* and *B*. Error bars indicate 1 SEM.

stimuli. We consider the influence of altered responsivity on our measurements of masking below.

We tested the orientation specificity of adaptation-induced changes in masking by making separate measurements of masking with test stimuli that were rotated 45° relative to the adapters. Contingent adaptation weakened masking for rotated test stimuli on average except for the highest contrast masks (Fig. 3C,D, green; $p = 0.09$ for 6%, $p = 0.02$ for 12%, $p = 0.05$ for 25%, and $p = 0.64$ for 50% masks). Contingent adaptation's effect on masking strength was significantly affected by test stimulus orientation ($F_{(1,2096)} = 58.62$, $p <$

0.001, ANOVA), with a significant interaction between mask contrast and stimulus orientation ($F_{(3,2096)} = 9.68$, $p < 0.001$, ANOVA). Asynchronous adaptation weakened masking for rotated test stimuli, particularly at higher mask contrasts (Fig. 3C,D, blue, $p = 0.23$ for 6%, $p = 0.38$ for 12%, and $p < 0.001$ for 25% and 50% masks). The degree to which masking was weakened depended on whether test stimuli were matched in orientation to the adapters ($F_{(1,1236)} = 16.48$, $p < 0.001$, ANOVA). There was no significant interaction between contrast and test stimulus orientation on the change in MI ($F_{(3,1236)} = 1.03$, $p = 0.38$, ANOVA).

In summary, the suppression recruited by a mask can be strengthened by consistently pairing the mask with a target (contingent adaptation) or weakened by interleaving its presentation with the target (asynchronous adaptation). These effects were orientation specific: when test and adapting stimuli had different orientations, we saw less adaptation-induced change in masking. The orientation specificity of adaptation-induced changes in masking is consistent with similar orientation specificity for adaptation effects on contrast sensitivity (Sengpiel and Banhoeffer, 2002; Crowder et al., 2006; Dhruv et al., 2011) and orientation tuning (Müller et al., 1999; Dragoi et al., 2000; Patterson et al., 2013) and suggests a cortical contribution.

Dependence of adaptation effects on neuronal response properties

In the preceding analyses, we focused on average changes in masking across our entire sample. We then investigated whether adaptation's effect on masking can be predicted by the functional properties of V1 neurons: their phase sensitivity (i.e., whether they are simple vs complex), their orientation preference, and the properties of masking before adaptation.

Previous work has shown that the effects of adaptation may depend on whether neurons are simple or complex (Giaschi et al., 1993; but see Ohzawa et al., 1985 and Crowder et al., 2006) and that masking may be stronger in simple cells (Bonds, 1989). More critically, sensitivity to grating phase may influence measurements of masking, when cells are sensitive to both grating components. We thus tested for a relationship between a cell's phase sensitivity (the logarithm of the F1/F0 ratio; Skottun et al., 1991) and masking both before and after adaptation. We found little relationship between the cell's phase sensitivity and the MI before contingent ($r = 0.11$, Spearman correlation; $p = 0.05$) or asynchronous ($r = 0.04$, $p = 0.53$) adaptation, in contrast to Bonds (1989). We also found no relationship between phase sensitivity and the change in masking after either contingent ($r = -0.03$, $p = 0.48$) or asynchronous adaptation ($r = 0.06$, $p = 0.39$; Fig. 4A). In cells with weak phase sensitivity (F0 > F1), masking was stronger after contingent adaptation (increased by 0.24 ± 0.03, $p < 0.001$, $n = 217$) and weaker after asynchronous adaptation (decreased by 0.38 ± 0.03, $p < 0.001$, $n = 140$). In cells with strong phase sensitivity (F1 > F0), masking was also stronger after contingent adaptation (increased by 0.13 ± 0.05, $p = 0.02$, $n = 100$) and weaker after asynchronous adaptation (decreased by 0.16 ± 0.09, $p = 0.07$, $n = 53$), although the latter effect was not statistically significant due likely to a smaller number of cells.

Previous studies have found that the effects of adaptation on V1 tuning depend on the relationship between the adapter and a neuron's stimulus preference (Dragoi et al., 2000; Crowder et al., 2006; Dhruv et al., 2011; Wissig and Kohn, 2012; Benucci et al.,
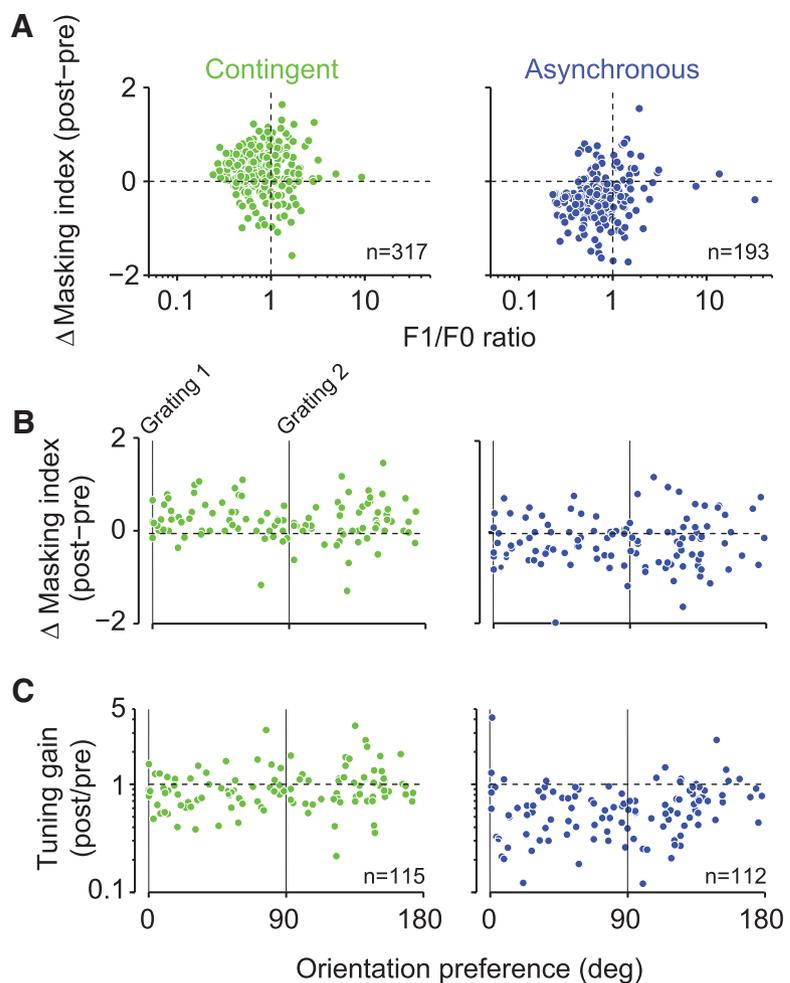
Figure 4. Dependence of changes in masking on neuronal properties. **A**, Relationship between adaptation-induced changes in MI and phase sensitivity, as measured by the F1/F0 response ratio. Each dot represents effects for 25% contrast masks for one unit. **B**, Relationship between adaptation-induced changes in MI and each neuron's orientation preference, where 0° and 90° indicate preferences aligned with the component gratings (indicated by vertical thin black lines). **C**, Relationship between adaptation-induced changes in tuning gain and each neuron's orientation preference following the conventions of **B**.

2013; Patterson et al., 2013). To assess whether adaptation-induced changes in masking depended on neuronal orientation preference, we recorded additional data in which we presented a reduced ensemble of test and mask gratings interleaved with gratings of different orientations (see Materials and Methods).

We found that changes in masking were largely independent of neuronal orientation preference (Fig. 4B). After contingent adaptation, masking was strengthened in cells with a preference within 22.5° of either component of the plaid adapter (defined as 0° and 90°, thin black lines in Fig. 4B), as well as in cells preferring orientations away from the plaid components (preferences offset by >22.5°). For the 50% mask, masking increased by 0.12 ± 0.05 ($p = 0.007$, $n = 53$) and 0.31 ± -0.06 ($p < 0.001$, $n = 62$), respectively ($p = 0.02$ for the comparison). Similarly, asynchronous adaptation led to weaker masking in cells with aligned preferences (-0.18 ± 0.07, $p = 0.03$, $n = 47$) and offset preferences (-0.22 ± 0.07, $p < 0.001$, $n = 65$; $p = 0.69$ for the comparison). We conclude that adaptation-induced changes in masking depend little, if at all, on neuronal orientation preference.

Our finding that changes in masking are evident in neurons with widely different preferences stands in contrast to the strong dependence of adaptation-induced changes in V1 tuning on preference reported in previous studies. This divergence in outcome

could reflect different specificity of adaptation-induced changes in masking and tuning. Alternatively, the different outcomes could arise because our masking experiments involved adapting with two orthogonal gratings instead of single gratings as in previous work (Dragoi et al., 2000; Crowder et al., 2006; Dhruv et al., 2011; Wissig and Kohn, 2012; Patterson et al., 2013). To distinguish between these possibilities, we measured how changes in tuning gain depend on neuronal preference (see Materials and Methods) for contingent and asynchronous adapters. Gain was slightly weakened after contingent adaptation (Fig. 4C, green), but this effect was indistinguishable in cells with well aligned preferences (geometric mean ratio of 0.92, $p = 0.003$) and offset preferences (0.95, $p = 0.009$; $p = 0.97$ for comparison between groups). Similarly, asynchronous adaptation caused a decrease in tuning gain that was similar in cells with aligned and offset preferences (Fig. 4C, blue; 0.61, $p < 0.001$ vs 0.64, $p < 0.001$; $p = 0.21$ for comparison between groups). Therefore, changes in tuning curve gain are also broadly shared across neurons with different preferences.

Finally, we assessed whether the change in masking with adaptation depended on the strength of masking observed before adaptation. Specifically, we sought to determine whether the strengthening of masking after contingent adaptation was driven by cases in which the cross-oriented mask was facilitatory rather than suppressive (i.e., the preadaptation MI < 0). Carandini et al. (1998) reported that adaptation with plaids caused a specific reduction in V1 responsivity to plaids, which could contribute to an apparent increase in masking. The reduction in plaid responsivity in that study was particularly prevalent in neurons in which responses to targets were enhanced by the co-presentation of a mask (i.e., those that showed cross-orientation facilitation). In our data, masks were rarely facilitatory (5–15% of cases depending on mask contrast). Further, we found that contingent adaptation strengthened masking even after excluding units showing cross-orientation facilitation in control conditions (increase of 0.10 ± 0.02, 0.17 ± 0.02, and 0.05 ± 0.02 for masks of 12%, 25%, and 50% contrast, respectively; $p < 0.02$ for all cases; and decrease in masking of 0.08 ± 0.02, $p < 0.001$ for 6% masks). We conclude that the strengthening of masking after contingent adaptation cannot be attributed to a loss of cross-orientation facilitation, as in Carandini et al. (1998).

Controlling for adaptation-induced changes in responsivity
A notable consequence of both contingent and asynchronous adaptation is reduced neuronal responsivity. Because both adapters reduced responsivity, their opposite effects on masking cannot be trivially explained by altered responsivity. Additionally, adaptation's effect on responsivity (defined as the change in response to the high-contrast target after adaptation) was not related to the change in masking after contingent adaptation (Fig. 5A; green, Spearman correlation of -0.06, $p = 0.26$) and only weakly related to the change in masking after asynchronous adaptation (blue, $r = 0.18$, $p = 0.01$).

The reduced responsivity after adaptation may nevertheless complicate inferences about how masking is altered by adaptation. For instance, adaptation might weaken masking because it fatigues the normalization pool rather than because it alters the interaction between the neuron and its normalization pool. We thus sought to determine how adaptation altered masking for stimuli of equal potency before and after adaptation.

To do so, we measured suppression through response summation. Specifically, we applied a "rate-matched" SI, which was defined as follows:

$$SI = 1 - \frac{R_{TM}}{R_T + R_M} \qquad (8)$$

where R_{TM} is the response to the full contrast plaid and R_T and R_M are the responses to the component gratings presented in isolation (Carandini et al., 1997b; Wissig and Kohn, 2012; Ruff et al., 2016). SI values near 0 indicate nearly linear summation of responses to the component gratings (consistent with weak normalization); values near 1 indicate that the response to the plaid is much weaker than expected from the responses to the component gratings (consistent with strong normalization).

We calculated the SI after adaptation using responses to the 50% contrast mask, the 50% contrast target, and the plaid made by their combination. We compared this SI with one measured before adaptation obtained from target and mask gratings that generated responses equal to those observed for the 50% contrast gratings after adaptation (i.e., matched for response strength rather than for contrast; Fig. 5B). Because the set of measured preadaptation responses rarely contained a perfect match to the postadaptation responses, we fitted the preadaptation responses to a descriptive model and used the model to identify the necessary target and mask contrasts (see Materials and Methods). The model also allowed us to estimate responses to plaid stimuli composed of these rate-matched target and mask gratings. Therefore, by comparing the SI before and after adaptation, we could test whether adaptation alters response summation (i.e., normalization) even for component gratings of matched potency.

Contingent adaptation caused the rate-matched SI to increase by 0.15 ± 0.03 ($p < 0.001$) from an initial value of 0.06 ± 0.03, a >3-fold increase (Fig. 5C, green; $n = 58$ units that passed selection criteria, see Materials and Methods). This increase is consistent with a substantial strengthening of normalization. Conversely, asynchronous adaptation caused the SI to decrease by 0.36 ± 0.08 ($p < 0.001$) from an initial value of -0.13 ± 0.03, consistent with substantial weakening of normalization (Fig. 5C, blue, $n = 51$ units). Note that preadaptation SI values were different for contingent and asynchronous adapters (0.06 vs -0.13; $p < 0.001$). This is because asynchronous adaptation reduced responsivity more than contingent adaptation, so the rate-matched preadaptation responses were from lower contrasts for asynchronous adaptation. Because masking is weaker for low-contrast stimuli, the preadaptation SI was lower in the asynchronous adaptation condition.

For responses to rotated test stimuli, contingent adaptation caused no significant change in SI (Fig. 5D, green; -0.04 ± 0.03 vs 0.01 ± 0.04; $p = 0.4$). Asynchronous adaptation caused an ~2-fold reduction in SI (Fig. 5D, blue; from -0.21 ± 0.07 to -0.43 ± 0.08; $p = 0.003$), consistent with a substantial weakening of normalization.

In summary, contingent adaptation caused responses to plaids to be even weaker than expected from the linear sum of the component gratings, consistent with stronger normalization; asynchronous adaptation, instead, caused summation to become more linear, or even more supralinear, consistent with weaker normalization. Therefore, adaptation-induced changes in response summation are evident even for stimuli that are equally potent before and after adaptation.

Modeling adaptation's effect on normalization strength
Westrick et al. (2016) proposed a specific learning rule for updating normalization signals based on stimulus history. Their model posits that the normalization signal received by a target neuron
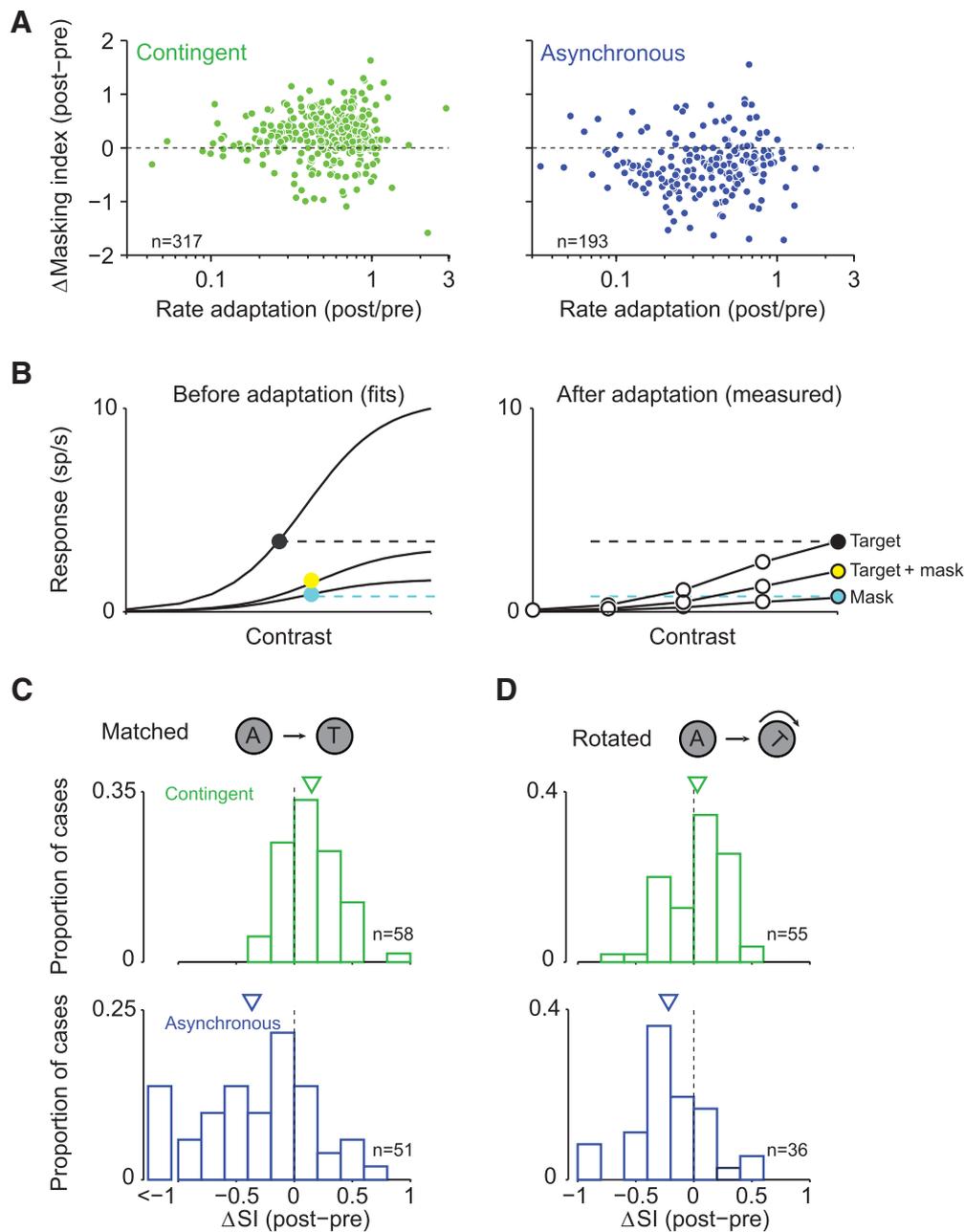
Figure 5. Controlling for rate adaptation. ***A***, Relationship between the change in MI and responsivity change measured as the ratio of response to the high-contrast target after versus before adaptation. Masking was measured using 25% contrast masks. Each dot indicates one unit. ***B***, Method for calculating rate-matched SI. Filled symbols in the right panel indicate the measured responses to the 50% contrast target (black), the 50% contrast mask (cyan), and the plaid formed by their combination (yellow). Filled symbols in the left panel indicate the target and mask contrasts that evoked matched responses (indicated by dashed horizontal lines) and the plaid formed by their combination. ***C***, Histogram of the change in rate-matched SI after contingent (green) and asynchronous (blue) adaptation. Arrowheads indicate distribution mean. ***D***, Histogram of the change in rate-matched SI for test stimuli with orientation rotated by 45° from the adapters. Conventions are as in ***C***.

arises from the response of a neuronal population, with the response of each neuron in the pool receiving a distinct weight, and that the weights are strengthened between the target neuron and the pool neurons that are consistently coactivated and weakened between those that are driven asynchronously (Fig. 6A). More precisely, the pairwise weight increases when the product of two neurons' responses is greater than their homeostatic target, defined as their expected average pairwise response (see Materials and Methods). Westrick et al. (2016) showed that the resultant changes in normalization could capture a range of adaptation effects on tuning and responsivity (Benucci et al., 2013). We

therefore sought to determine whether this model could also predict the adaptation-induced changes in masking that we observed.

The behavior of the model for contingent and asynchronous adaptation is illustrated in Figure 6. During contingent adaptation, neurons that prefer orientations near 0° and 90°, the orientations of the target and mask, respectively, will be strongly coactivated (Fig. 6C, arrow, lighter colors indicate stronger response products). Because the response product of these neurons is larger than typical—that is, larger than their homeostatic target (Fig. 6D, arrow)—the normalization weights for these neurons
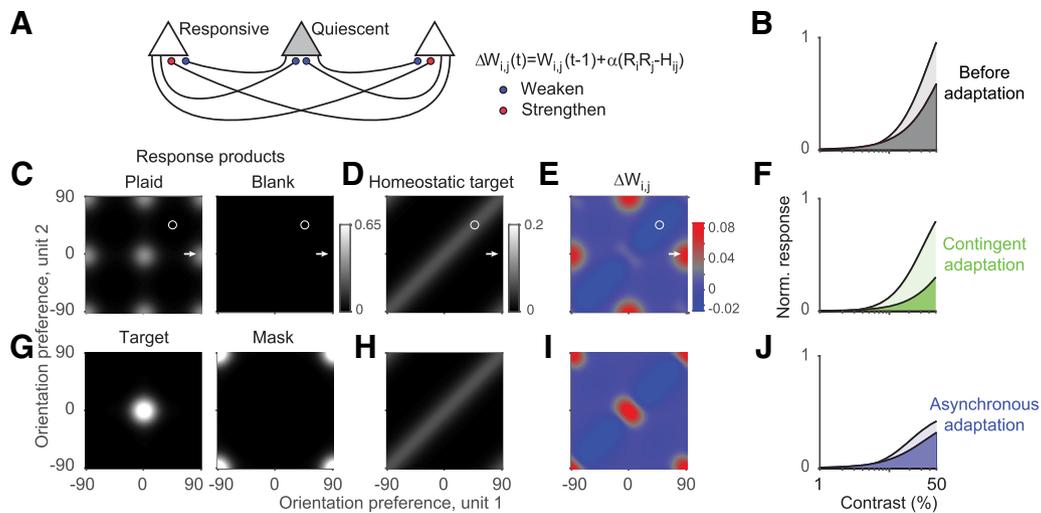
Figure 6. Hebbian normalization model. ***A***, Schematic of the normalization model and the learning rule (Westrick et al., 2016). The normalization signal received by each neuron arises from the weighted responses of other neurons in the population. The weights between neurons that are consistently coactivated (white triangles) are strengthened (red dots), whereas the weights are weakened between neurons that are driven asynchronously (blue dots). ***B***, Simulated contrast–response function before adaptation to the target alone (light gray) and the target presented with a 50% mask (dark gray) for a neuron preferring the orientation of the target grating. ***C***, Response products to the plaid contingent adapter. Lighter color indicates stronger response products. Arrow indicates neuron pair preferring 0° and 90°. Circle indicates neuron pair preferring 45°. ***D***, Homeostatic target defined as the average response products to a uniform distribution of oriented gratings. Markers indicate the same neuron pairs as in ***C***. ***E***, Change in normalization weights after contingent adaptation. Markers indicate the same pairs as in ***C***. ***F***, Contrast–response function after contingent adaptation using the same convention as ***B***. ***G–J***, The same conventions as ***C–F***, but after asynchronous adaptation.

will be strengthened (Fig. 6*E*, arrow, red indicates strengthening). Now consider neurons that prefer orientations near 45°. The contingent adapter provides little drive to these neurons (Fig. 6*C*, left, white circle), so their response products are smaller than their homeostatic target. Therefore, the normalization weights for these neurons weaken after contingent adaptation (Fig. 6*E*, blue). Note that the manner in which the normalization weights change after adaptation is not equivalent to that predicted by the initial response product. The change depends also on the homeostatic target; further, as the normalization weights adjust, the neurons' responses evolve as well.

The effect of asynchronous adaptation can be understood similarly. These adapters provide strong drive to neurons preferring orientations near 0° and 90°, but the two sets of neurons are active at different phases of the adapter (Fig. 6*G*). Therefore, normalization weights are strengthened for pairs responding to either 0° or 90° orientations, but weakened between 0–90° pairs (Fig. 6*I*).

Both contingent and asynchronous adaptation strengthen some normalization weights and weaken others, but the pattern of weight changes across the population of units differs between adapters, resulting in different effects on masking. For a neuron with a preference that is matched to the target (0°), masking is strengthened after contingent adaptation (Fig. 6*F*) compared with the masking evident before adaptation (Fig. 6*B*). There is more masking because the weights between neurons preferring 0° (the target) and 90° are strengthened, resulting in greater suppression. After asynchronous adaptation, these same weights are weakened, resulting in less masking (Fig. 6*J*). In addition, asynchronous adaptation strengthens weights between neurons with preferences near 0°, resulting in weaker responses to the target itself (Fig. 6*J*). Note that this change in responsivity, which was also evident in our neurophysiological data, arises solely from altered normalization because the model contains no other mechanism for adjusting to prolonged sensory input.

Across the population of units, the model predicts that contingent adaptation should strengthen masking (Fig. 7*A*, green,

dashed line), particularly for higher mask contrasts, as in the neuronal data. Asynchronous adaptation leads to weakened masking (Fig. 7*A*, blue, dashed line). However, this instantiation of the model underestimates the magnitude of altered masking in the data. Further, although changes in masking in the model depend on whether the test stimuli are matched to the adapter (cf. dashed lines in Fig. 7, *A* and *B*, which illustrates changes in masking for test stimuli offset by 45° from the adapter), the predicted effects are different from those that we observed. Finally, the model predicts that the degree to which masking is altered by adaptation depends strongly on neuronal preference (Fig. 7*C*,*D*, dashed lines), which was not the case in our data.

The behavior of the model depends critically on several parameters, including the tuning widths of the units (which may or may not be equivalent to V1 tuning, depending on the source of normalization), the time scale of the model, the contrast saturation of the model neurons, and the training contrast, which determines the homeostatic response target for each pair. For many parameter settings that produced stronger changes in normalization, the model tends to predict unrealistic response facilitation in units in which normalization signals are weakened. We therefore considered a simple extension to the model to mitigate this behavior (see Materials and Methods): a "fatigue" mechanism that reduces responsivity in proportion to recent activity levels, as described in previous neurophysiological studies (Schwindt et al., 1988; Sanchez-Vives, 2000a,b; see Carandini and Ferster, 1997 for related work). With this mechanism, the model was able to produce changes in masking similar to those in our data (Fig. 7*A*, solid lines), including the dependence on test stimulus orientation (Fig. 7*B*, solid lines), with limited response facilitation for moderate contrast stimuli. However, this model still incorrectly predicts that changes in masking depend on neuronal preference (Fig. 7*C*,*D*, solid lines).

In summary, a simple learning rule for updating normalization weights developed to account for an entirely distinct set of adaptation phenomena qualitatively matched many of our key physiological observations: opposite changes in masking after
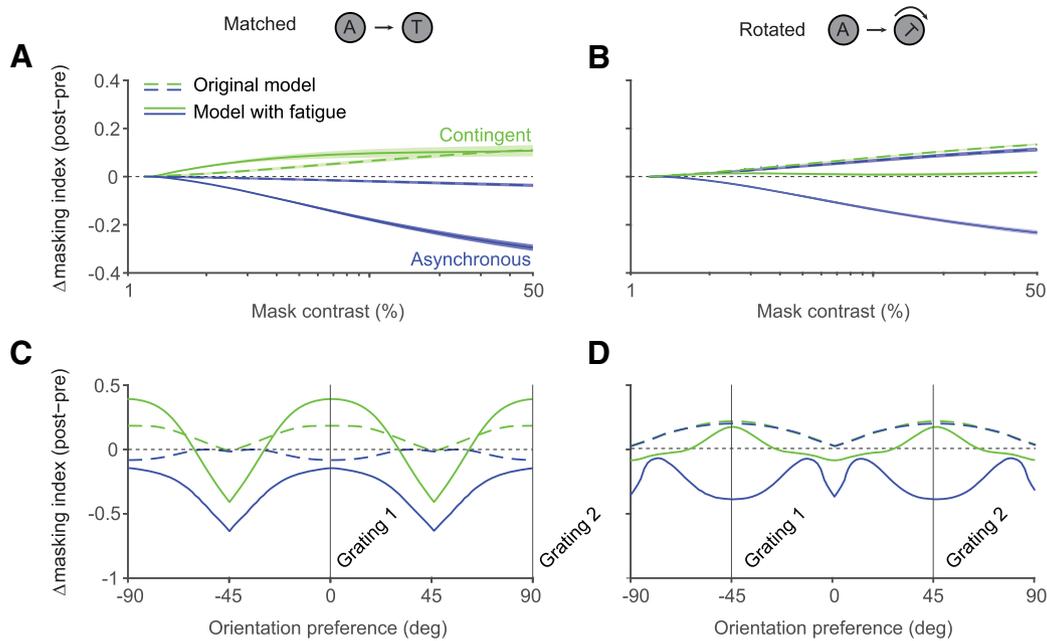
Figure 7. Model predictions for changes in masking after contingent or asynchronous adaptation. ***A***, Left, Changes in MI in simulated neurons as a function of mask contrast after contingent (green) and asynchronous (blue) adaptation for test stimuli matched in orientation to the adapters. Dotted lines indicate mean of simulated population of model neurons averaged across all orientation preferences. Solid lines indicate mean of simulated population of neurons from the extended model (i.e., with a fatigue mechanism). Shading indicates SD across model neurons with different orientation preferences. ***B***, Same as ***A*** for test stimuli offset in orientation from the adapters. Dashed curve for asynchronous adaptation has been scaled slightly for visualization. ***C***, Change in MI for the 50% contrast mask as a function of model unit orientation preference for test stimuli matched in orientation to the adapter. ***D***, Same as ***C*** for test stimuli offset in orientation from the adapters. Dashed curve for asynchronous adaptation has been scaled slightly for visualization.

contingent and asynchronous adaptation, a dependence of the changes in masking on stimulus contrast and the relative orientation of the adapters and test stimuli, and stronger loss of responsivity after asynchronous than contingent adaptation. An unresolved mismatch with the data is the model's prediction that changes in masking depend strongly on neuronal preference. One explanation for this discrepancy is that the V1 neurons with offset preferences for which we measured masking all had appreciable responses to the target grating, a requirement for measuring masking. As a result, these cells were well driven by the adapter and might be predicted to show stronger masking after adaptation. In the model, masking can be measured for arbitrarily small responses; therefore, the change in masking for units with offset preferences includes poorly driven units. We note that the model might be made to better fit our data (e.g., through a more exhaustive search of parameters), but such a data-fitting exercise would only be meaningful if it were constrained by a broader set of adaptation phenomena, including those used to develop the original model.

Discussion

We found that masking is enhanced when a mask and target grating are consistently paired, but weakened when those stimuli are presented asynchronously. Changes in masking depend on the orientation of the stimulus relative to the orientation of the adapter, but are shared broadly across the population independently of neuronal preference. Altered masking cannot be attributed to weaker stimulus potency after adaptation or to changes in neuronal responsivity. Our results thus show that a paradigmatic form of normalization in visual cortex can be either strengthened or weakened depending on the temporal contingencies between different visual inputs.

Relation to previous work
Previous reports have provided conflicting evidence for adaptation's effect on masking. Freeman et al. (2002) found no change in masking in cat V1 after adaptation with the mask grating. Li et al. (2005) found that dichoptic, but not monoptic, masking was reduced by adaptation. The authors concluded that there are two mechanisms of masking: one subcortical, monocular, and unadaptable and the other cortical, binocular, and adaptable. It is unclear why the cortical mechanism was not evident in their monocular adaptation experiments, but the adaptable mechanism of masking that they identified might underlie the weakened masking that we observed after asynchronous adaptation. In related work, Dhruv et al. (2011) measured the effects of adaptation in macaque V1. After adaptation with orthogonally oriented gratings, responses to high-contrast preferred stimuli were often elevated. By fitting a descriptive model to their data, the authors inferred that this facilitation was due to weakened normalization after adaptation, consistent with our measurements with asynchronous adapters. Previous evidence that adaptation can strengthen masking is scant. In a brief report, Carandini et al. (1998) measured the effect of adapting to plaids or gratings in eight cat V1 cells. They showed that adaptation altered cross-orientation interactions in five of these cells, including weakening of cross-orientation facilitation and strengthening of suppression. These observation are consistent with our more extensive and systematic observations, although we found little role for reduced cross-orientation facilitation in our results.

Our finding that masking can be strongly influenced by adaptation is consistent with recent reports that surround suppression, another form of normalization (Heeger, 1992; Cavanaugh et al., 2002; Carandini and Heeger, 2011; Coen-Cagli et al., 2012, 2015), can be altered by adaptation (Webb et al., 2005; Camp et

al., 2009; Patterson et al., 2014). Specifically, these studies showed that adaptation with an annular grating, a form of asynchronous adaptation in which the surround but not the center receives visual input, leads to weaker suppression. However, it has also been shown that adapting with large gratings, which should provide "contingent" coactivation of the RF and its surround, can lead to response facilitation (Wissig and Kohn, 2012; Patterson et al., 2013, 2014), implying weaker suppression. This may indicate that there are distinct rules by which suppressive signals within the RF and from the surround are updated by adaptation, consistent with their having distinct underlying mechanisms (Sengpiel et al., 1998). Alternatively, adaptation with large gratings may potentiate the suppression of normalization signals within the RF by the surround (i.e., a stronger suppression of suppression, leading to facilitation; Trott and Born, 2015).

Stronger masking was evident after the contingent display of the mask and target. Numerous perceptual studies have also reported contingent adaptation effects. Perhaps the most well known is the "McCollough effect" in which adaptation to colored, oriented gratings induces an orientation-dependent color aftereffect (McCollough, 1965; see also Hepler, 1968; Held and Shattuck, 1971; Favreau et al., 1972; Lovegrove and Over, 1972). The large number of possible stimulus contingencies makes it unlikely that these perceptual effects arise from the fatigue of cells selective for each pairing. Instead, they could be explained by altered interactions between neurons selective for different stimulus features (Barlow and Földiák, 1989; Barlow, 1990).

The attribution of contingent perceptual aftereffects to altered neuronal interactions is similar in spirit to the altered normalization suggested by our masking experiments because normalization is likely a network phenomenon. However, it is unlikely that our observations underlie the types of perceptual contingent aftereffects cited above. First, most of the perceptual effects persist for many hours, even after relatively brief (tens of seconds) adaptation (McCollough, 1965; Vul et al., 2008). The changes in masking that we observed are more transient, as shown in Figure 8: the effects of contingent (green) and asynchronous (blue) adaptation entirely dissipated after a 10–15 min recovery period. Second, the perceptual experiments involve a single adapter (consisting of paired stimuli) that induces distinct aftereffects depending on the test stimulus. Instead, we show distinct changes in masking with different adapters using a single ensemble of test stimuli.

Although our results are unlikely to underlie classic contingent aftereffects, our findings do have a direct perceptual correlate in human observers: contingent adaptation leads to greater perceptual masking, whereas asynchronous adaptation leads to weaker masking (Yiltiz et al., 2018; see also Foley and Chen, 1997).

Mechanisms

The biophysical and circuit mechanisms underlying masking within the RF (i.e., cross-orientation suppression) are not fully understood. Some have suggested that masking involves depression of thalamocortical synapses (Carandini et al., 2002; Freeman et al., 2002; but see Li et al., 2006). Others have suggested that masking arises from rectification and weak contrast saturation in LGN responses (Li et al., 2006), perhaps amplified by a nonlinear input–output transformation in cortex (Priebe and Ferster, 2006).

The changes in masking that we observed during adaptation are difficult to reconcile with any of these proposed mechanisms. Both contingent and asynchronous adapters should recruit robust responses in the LGN (Priebe and Ferster, 2006) and thus
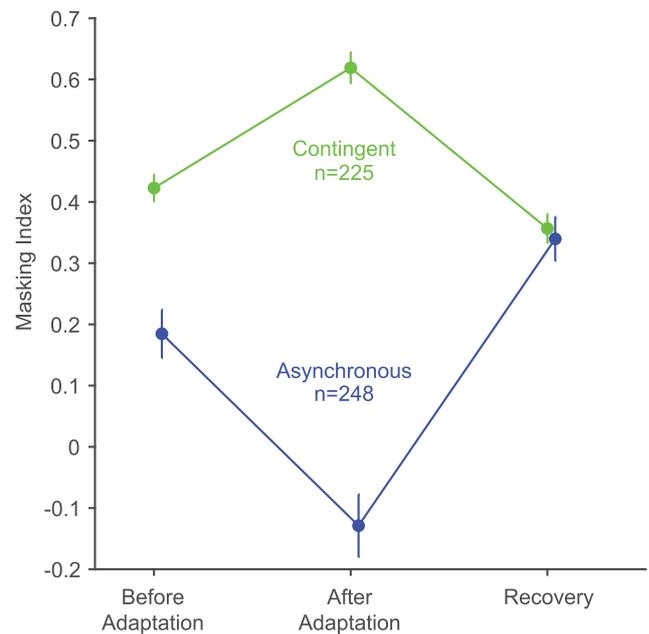
Figure 8. Recovery from adaptation. The average MI is shown for the 25% contrast mask before adaptation, after contingent (green) or asynchronous (blue) adaptation, and 10 –15 min later, after the continuous presentation of a gray screen. Adaptation-induced changes in masking dissipated entirely during the recovery period; in fact, they often showed a slight rebound effect, with masking in the recovery period slightly weaker (stronger) than the preadaptation measurements for contingent (asynchronous) adaptation. Error bars indicate 1 SEM. Contingent and asynchronous lines were separated by which adaptation paradigm was recorded first (both were always run back to back). Units shown are a subset of the full dataset, representing neurons with isolation that was stable throughout the recovery period.

depress thalamocortical synapses. In this adapted state, presenting a target and mask together would likely produce little additional depression (Boudreau and Ferster, 2005; Reig et al., 2006). Therefore, synaptic depression models may predict weaker masking after adaptation but cannot account for stronger masking after contingent adaptation. If, instead, cortical masking is largely inherited from the LGN, then our results would require that geniculate neurons adapt differently to contingent and asynchronous adapters. However, neurons in cat and monkey LGN adapt weakly (Movshon and Lennie, 1979; Ohzawa et al., 1985; Nelson, 1991; Sanchez-Vives et al., 2000a; but see Shou et al., 1996; Duong and Freeman, 2007) except when driven by stimuli of much higher temporal frequency than those we used (Solomon et al., 2004). In addition, adaptation effects in the LGN show no evidence of the orientation specificity that we observed for altered masking (Solomon et al., 2004).

Our results are qualitatively consistent with a recently proposed rule for updating normalization weights based on stimulus history (Westrick et al., 2016; see also related work by Hosoya et al., 2005). The model correctly predicts that masking is strengthened by contingent adaptation and weakened by asynchronous adaptation. It captures these effects solely through modulating the normalization weights between units with different tuning based on their degree of coactivation during adaptation. Because the model's behavior only depends on the pattern of response coactivation, it could also be used to predict changes in normalization for more complex stimuli (e.g., natural scenes) if based on RF models that accurately capture responses to those stimuli.

Although the mechanisms responsible for masking remain unclear, our finding that changes in masking depend on the temporal relationship between adapters suggests that a Hebbian-like

mechanism is responsible for their updating (Westrick et al., 2016). There is, of course, extensive evidence for Hebbian plasticity of synaptic strength on a range of time scales (Abbott and Nelson, 2000; Abbott and Regehr, 2004). Masking may thus be modified by the synaptic plasticity between target neurons and those neurons providing suppressive input. This synaptic plasticity could involve changes in inhibitory synapses or a more complex rebalancing of excitatory and inhibitory input (Nassi et al., 2015; Rubin et al., 2015; Sato et al., 2016).

We note that the masking that we measured was driven primarily by suppressive signals within the RF because we used small stimuli (1.5° diameter) centered on the aggregate RF of the recorded units. However, the size of our stimuli was slightly larger than the average spatial RF of V1 neurons at the targeted eccentricity (Cavanaugh et al., 2002), so our stimuli may have recruited some surround suppression. However, as noted above, surround signals are modified by adaptation in a manner seemingly distinct from the effects that we report here. Therefore, it is unlikely that the adaptation-induced changes in masking involved altered suppression from the RF surround.

Implications
Our results have several important implications for our understanding of normalization and of adaptation. Although initially developed to account for nonlinear response properties of V1 neurons, normalization has now been shown to be useful for explaining a broad set of phenomena (Carandini and Heeger, 2011). Across these contexts, normalization is often portrayed as a largely static computation, although recent work has shown that attention may modulate normalization signals dynamically (Lee and Maunsell, 2009; Reynolds and Heeger, 2009). Our results, and those on adaptation effects in the surround, indicate that normalization can also be modulated dynamically based on recent sensory input (Solomon and Kohn, 2014). Insofar as our results are consistent with the updating rule of Westrick et al. (2016), our results also provide an indication as to how the set-point of normalization signals may be calibrated to match the dominant statistics of the sensory environment.

Normalization is thought to be critical for a number of interrelated cortical functions, including improving representational efficiency (Schwartz and Simoncelli, 2001; Coen-Cagli et al., 2012); performing marginalization, a basic computation in probabilistic inference (Beck et al., 2011); implementing predictive coding (Spratling, 2010; Lochmann et al., 2012); and determining stimulus salience (Itti and Koch, 2000). Because normalization signals are strongly shaped by adaptation, a primary purpose of adaptation effects may be to modulate these computations. For instance, changes in neuronal tuning and responsivity may maintain or improve representational efficiency (Barlow and Földiák, 1989; Barlow, 1990; Wainwright et al., 2002) or highlight novel features of the environment (Hosoya et al., 2005; Solomon and Kohn, 2014). Alternative hypotheses of the functional benefits of adaptation would need to account for the strong modulation of normalization that occurs with adaptation.

Our finding that the updating of normalization-based suppressive signals is sensitive to temporal contingencies between visual inputs offers cortical networks the ability to adjust, not only to the persistence or frequency of occurrence of individual stimuli, but also to the relationships among stimuli. Neurons in higher visual cortex are known to become sensitive to temporal (sequential) pairings of stimuli through learning that occurs over weeks (Meyer and Olson, 2011). Our results show that sensitivity to temporal relationships among stimuli can emerge after much

briefer exposures, even in the earliest stages of cortical visual processing.

Finally, our findings lend credence to frameworks in which adaptation effects arise in part from altered normalization (Heeger, 1992; Wainwright et al., 2002; Lochmann et al., 2012; Solomon and Kohn, 2014; Snow et al., 2016; Westrick et al., 2016). Although such frameworks have been shown to account for a broad set of adaptation phenomena, their key assumption, that adaptation alters normalization, has received limited experimental support except for demonstrations that surround suppression can be weakened by adaptation. Our finding that normalization can be robustly strengthened or weakened opens the door to the development of a normalization-based framework for understanding adaptation. Such a framework might be able to predict effects for a much broader set of adaptation paradigms than usually considered (e.g., natural scenes, natural viewing), as well as offering new mechanistic, network-based explanations for how cortex adjusts to current sensory demands.

References
Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. Nat Neurosci 3:1178–1183. CrossRef Medline

Abbott LF, Regehr WG (2004) Synaptic computation. Nature 431:796–803. CrossRef Medline

Barlow HB (1990) A theory about the functional role and synaptic mechanism of visual aftereffects. In: Vision: coding and efficiency (Blakemore CB, ed), pp 363–375. New York: Cambridge University.

Barlow HB, Földiák P (1989) Adaptation and decorrelation in the cortex. In: The computing neuron (Durbin R, Miall C, Mitchison G, eds), pp 54–72. New York: Addison-Wesley.

Beck JM, Latham PE, Pouget A (2011) Marginalization in neural circuits with divisive normalization. J Neurosci 31:15310–15319. CrossRef Medline

Benucci A, Saleem AB, Carandini M (2013) Adaptation maintains population homeostasis in primary visual cortex. Nat Neurosci 16:724–729. CrossRef Medline

Bonds AB (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Vis Neurosci 2:41–55. CrossRef Medline

Boudreau CE, Ferster D (2005) Short-term depression in thalamocortical synapses of cat primary visual cortex. J Neurosci 25:7179–7190. CrossRef Medline

Camp AJ, Tailby C, Solomon SG (2009) Adaptable mechanisms that regulate the contrast response of neurons in the primate lateral geniculate nucleus. J Neurosci 29:5009–5021. CrossRef Medline

Carandini M, Ferster D (1997) A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. Science 276:949–952. CrossRef Medline

Carandini M, Heeger DJ (2011) Normalization as a canonical neural computation. Nat Rev Neurosci 13:51–62. CrossRef Medline

Carandini M, Heeger DJ, Movshon JA (1997a) Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17:8621–8644. CrossRef Medline

Carandini M, Barlow HB, O'Keefe LP, Poirson AB, Movshon JA (1997b) Adaptation to contingencies in macaque primary visual cortex. Philosophical Transactions of the Royal Society of London B: Biological Sciences 352:1149–1154. CrossRef Medline

Carandini M, Movshon JA, Ferster D (1998) Pattern adaptation and cross-orientation interactions in the primary visual cortex. Neuropharmacology 37:501–511. CrossRef Medline

Carandini M, Heeger DJ, Senn W (2002) A synaptic explanation of suppression in visual cortex. J Neurosci 22:10053–10065. CrossRef Medline

Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. J Neurophysiol 88:2530–2546. CrossRef Medline

Clifford CW, Webster MA, Stanley GB, Stocker AA, Kohn A, Sharpee TO, Schwartz O (2007) Visual adaptation: neural, psychological and computational aspects. Vision Res 47:3125–3131. CrossRef Medline

Coen-Cagli R, Dayan P, Schwartz O (2012) Cortical surround interactions

and perceptual salience via natural scene statistics. PLoS Comput Biol 8:e1002405. CrossRef Medline

Coen-Cagli R, Kohn A, Schwartz O (2015) Flexible gating of contextual influences in natural vision. Nat Neurosci 18:1648–1655. CrossRef Medline

Crowder NA, Price NS, Hietanen MA, Dreher B, Clifford CW, Ibbotson MR (2006) Relationship between contrast adaptation and orientation tuning in V1 and V2 of cat visual cortex. J Neurophysiol 95:271–283. CrossRef Medline

DeAngelis GC, Robson JG, Ohzawa I, Freeman RD (1992) Organization of suppression in receptive fields of neurons in cat visual cortex. J Neurophysiol 68:144–163. CrossRef Medline

Dhruv NT, Tailby C, Sokol SH, Lennie P (2011) Multiple adaptable mechanisms early in the primate visual pathway. J Neurosci 31:15016–15025. CrossRef Medline

Dragoi V, Sharma J, Sur M (2000) Adaptation-induced plasticity of orientation tuning in adult visual cortex. Neuron 28:287–298. CrossRef Medline

Duong T, Freeman RD (2007) Spatial frequency-specific contrast adaptation originates in the primary visual cortex. J Neurophysiol 98:187–195. CrossRef Medline

El-Shamayleh Y, Movshon JA (2011) Neuronal responses to texture-defined form in macaque visual area V2. J Neurosci 31:8543–8555. CrossRef Medline

Favreau OE, Emerson VF, Corballis MC (1972) Motion perception: a color-contingent aftereffect. Science 176:78–79. CrossRef Medline

Foley JM, Chen CC (1997) Analysis of the effect of pattern adaptation on pattern pedestal effects: a two-process model. Vision Res 37:2779–2788. CrossRef Medline

Freeman TC, Durand S, Kiper DC, Carandini M (2002) Suppression without inhibition in visual cortex. Neuron 35:759–771. CrossRef Medline

Giaschi D, Douglas R, Marlin S, Cynader M (1993) The time course of direction-selective adaptation in simple and complex cells in cat striate cortex. J Neurophysiol 70:2024–2034. CrossRef Medline

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181–197. CrossRef Medline

Held R, Shattuck SR (1971) Color-and edge-sensitive channels in the human visual system: tuning for orientation. Science 174:314–316. CrossRef Medline

Hepler N (1968) Color: a motion-contingent aftereffect. Science 162:376–377. CrossRef Medline

Hosoya T, Baccus SA, Meister M (2005) Dynamic predictive coding by the retina. Nature 436:71–77. CrossRef Medline

Itti L, Koch C (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Res 40:1489–1506. CrossRef Medline

Kaliukhovich DA, Vogels R (2016) Divisive normalization predicts adaptation-induced response changes in macaque inferior temporal cortex. J Neurosci 36:6116–6128. CrossRef Medline

Kelly RC, Smith MA, Samonds JM, Kohn A, Bonds AB, Movshon JA, Lee TS (2007) Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. J Neurosci 27:261–264. CrossRef Medline

Kohn A (2007) Visual adaptation: physiology, mechanisms, and functional benefits. J Neurophysiol 97:3155–3164. CrossRef Medline

Kohn A, Movshon JA (2004) Adaptation changes the direction tuning of macaque MT neurons. Nat Neurosci 7:764–772. CrossRef Medline

Lee J, Maunsell JH (2009) A normalization model of attentional modulation of single unit responses. PLoS One 4:e4651. CrossRef Medline

Li B, Peterson MR, Thompson JK, Duong T, Freeman RD (2005) Cross-orientation suppression: monoptic and dichoptic mechanisms are different. J Neurophysiol 94:1645–1650. CrossRef Medline

Li B, Thompson JK, Duong T, Peterson MR, Freeman RD (2006) Origins of cross-orientation suppression in the visual cortex. J Neurophysiol 96:1755–1764. CrossRef Medline

Lochmann T, Ernst UA, Denève S (2012) Perceptual inference predicts contextual modulations of sensory responses. J Neurosci 32:4179–4195. CrossRef Medline

Lovegrove WJ, Over R (1972) Color adaptation of spatial frequency detectors in the human visual system. Science 176:541–543. CrossRef Medline

McCollough C (1965) Color adaptation of edge-detectors in the human visual system. Science 149:1115–1116. CrossRef Medline

Meyer T, Olson CR (2011) Statistical learning of visual transitions in monkey inferotemporal cortex. Proc Natl Acad Sci U S A 108:19401–19406. CrossRef Medline

Morrone MC, Burr DC, Maffei L (1982) Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. Proceedings of the Royal Society of London B: Biological Sciences 216:335–354. CrossRef Medline

Movshon JA, Lennie P (1979) Pattern-selective adaptation in visual cortical neurones. Nature 278:850–852. CrossRef Medline

Müller JR, Metha AB, Krauskopf J, Lennie P (1999) Rapid adaptation in visual cortex to the structure of images. Science 285:1405–1408. CrossRef Medline

Nassi JJ, Avery MC, Cetin AH, Roe AW, Reynolds JH (2015) Optogenetic activation of normalization in alert macaque visual cortex. Neuron 86:1504–1517. CrossRef Medline

Nelson SB (1991) Temporal interactions in the cat visual system. I. Orientation-selective suppression in the visual cortex. J Neurosci 11:344–356. CrossRef Medline

Ohzawa I, Sclar G, Freeman RD (1985) Contrast gain control in the cat's visual system. J Neurophysiol 54:651–667. CrossRef Medline

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. CrossRef Medline

Patterson CA, Duijnhouwer J, Wissig SC, Krekelberg B, Kohn A (2014) Similar adaptation effects in primary visual cortex and area MT of the macaque monkey under matched stimulus conditions. J Neurophysiol 111:1203–1213. CrossRef Medline

Priebe NJ, Ferster D (2006) Mechanisms underlying cross-orientation suppression in cat visual cortex. Nat Neurosci 9:552–561. CrossRef Medline

Reig R, Gallego R, Nowak LG, Sanchez-Vives MV (2006) Impact of cortical network activity on short-term synaptic depression. Cereb Cortex 16:688–695. CrossRef Medline

Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185. CrossRef Medline

Rieke F, Rudd ME (2009) The challenges natural images pose for visual adaptation. Neuron 64:605–616. CrossRef Medline

Rubin DB, Van Hooser SD, Miller KD (2015) The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. Neuron 85:402–417. CrossRef Medline

Ruff DA, Alberts JJ, Cohen MR (2016) Relating normalization to neuronal populations across cortical areas. J Neurophysiol 116:1375–1386. CrossRef Medline

Sanchez-Vives MV, Nowak LG, McCormick DA (2000a) Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. J Neurosci 20:4267–4285. CrossRef Medline

Sanchez-Vives MV, Nowak LG, McCormick DA (2000b) Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. J Neurosci 20:4286–4299. CrossRef Medline

Sato TK, Haider B, Häusser M, Carandini M (2016) An excitatory basis for divisive normalization in visual cortex. Nat Neurosci 19:568–570. CrossRef Medline

Schwartz O, Simoncelli EP (2001) Natural signal statistics and sensory gain control. Nat Neurosci 4:819–825. CrossRef Medline

Schwartz O, Hsu A, Dayan P (2007) Space and time in visual context. Nat Rev Neurosci 8:522–535. CrossRef Medline

Schwindt PC, Spain WJ, Foehring RC, Chubb MC, Crill WE (1988) Slow conductances in neurons from cat sensorimotor cortex in vitro and their role in slow excitability changes. J Neurophysiol 59:450–467. CrossRef Medline

Sengpiel F, Bonhoeffer T (2002) Orientation specificity of contrast adaptation in visual cortical pinwheel centres and iso-orientation domains. Eur J Neurosci 15:876–886. CrossRef Medline

Sengpiel F, Baddeley RJ, Freeman TC, Harrad R, Blakemore C (1998) Different mechanisms underlie three inhibitory phenomena in cat area 17. Vision Res 38:2067–2080. CrossRef Medline

Shou T, Li X, Zhou Y, Hu B (1996) Adaptation of visually evoked responses of relay cells in the dorsal lateral geniculate nucleus of the cat following prolonged exposure to drifting gratings. Vis Neurosci 13:605–613. CrossRef Medline

Skottun BC, De Valois RL, Grosof DH, Movshon JA, Albrecht DG, Bonds AB (1991) Classifying simple and complex cells on the basis of response modulation. Vision Res 31:1079–1086. Medline

Snow M, Coen-Cagli R, Schwartz O (2016) Specificity and timescales of cortical adaptation as inferences about natural movie statistics. J Vis 16:1. CrossRef Medline

Solomon SG, Kohn A (2014) Moving sensory adaptation beyond suppressive effects in single neurons. Curr Biol 24:R1012–R1022. CrossRef Medline

Solomon SG, Peirce JW, Dhruv NT, Lennie P (2004) Profound contrast adaptation early in the visual pathway. Neuron 42:155–162. CrossRef Medline

Spratling MW (2010) Predictive coding as a model of response properties in cortical area V1. J Neurosci 30:3531–3543. CrossRef Medline

Stocker AA, Simoncelli EP (2006) Noise characteristics and prior expectations in human visual speed perception. Nat Neurosci 9:578–585. CrossRef Medline

Trott AR, Born RT (2015) Input-gain control produces feature-specific surround suppression. J Neurosci 35:4973–4982. CrossRef Medline

Vul E, Krizay E, MacLeod DI (2008) The McCollough effect reflects permanent and transient adaptation in early visual cortex. J Vis 8:4.1–12. CrossRef Medline

Wainwright MJ, Schwartz O, Simoncelli EP (2002) Natural image statistics and divisive normalization: Modeling nonlinearity and adaptation in cortical neurons. In: Probabilistic models of the brain: perception and neural function (Rao RPN, Olshausen BA, Lewicki MS, eds), pp 203–222. Cambridge, MA: MIT.

Wark B, Lundstrom BN, Fairhall A (2007) Sensory adaptation. Curr Opin Neurobiol 17:423–429. CrossRef Medline

Webb BS, Dhruv NT, Solomon SG, Tailby C, Lennie P (2005) Early and late mechanisms of surround suppression in striate cortex of macaque. J Neurosci 25:11666–11675. CrossRef Medline

Webster MA (2015) Visual adaptation. Annu Rev Vis Sci 1:547–567. CrossRef Medline

Westrick ZM, Heeger DJ, Landy MS (2016) Pattern adaptation and normalization reweighting. J Neurosci 36:9805–9816. CrossRef Medline

Wissig SC, Kohn A (2012) The influence of surround suppression on adaptation effects in primary visual cortex. J Neurophysiol 107:3370–3384. CrossRef Medline

Yiltiz H, Heeger DJ, Landy MS (2018) Contingent adaptation in masking and surround suppression. J Vis 18:259.