

# Cross-orientation suppression in human visual cortex

Gijs Joost Brouwer and David J. Heeger

Department of Psychology and Center for Neural Science, New York University, New York, New York

Submitted 13 June 2011; accepted in final form 17 July 2011

**Brouwer GJ, Heeger DJ.** Cross-orientation suppression in human visual cortex. *J Neurophysiol* 106: 2108–2119, 2011. First published July 20, 2011; doi:10.1152/jn.00540.2011.—Cross-orientation suppression was measured in human primary visual cortex (V1) to test the normalization model. Subjects viewed vertical target gratings (of varying contrasts) with or without a superimposed horizontal mask grating (fixed contrast). We used functional magnetic resonance imaging (fMRI) to measure the activity in each of several hypothetical channels (corresponding to subpopulations of neurons) with different orientation tunings and fit these orientation-selective responses with the normalization model. For the V1 channel maximally tuned to the target orientation, responses increased with target contrast but were suppressed when the horizontal mask was added, evident as a shift in the contrast gain of this channel's responses. For the channel maximally tuned to the mask orientation, a constant baseline response was evoked for all target contrasts when the mask was absent; responses decreased with increasing target contrast when the mask was present. The normalization model provided a good fit to the contrast-response functions with and without the mask. In a control experiment, the target and mask presentations were temporally interleaved, and we found no shift in contrast gain, i.e., no evidence for suppression. We conclude that the normalization model can explain cross-orientation suppression in human visual cortex. The approach adopted here can be applied broadly to infer, simultaneously, the responses of several subpopulations of neurons in the human brain that span particular stimulus or feature spaces, and characterize their interactions. In addition, it allows us to investigate how stimuli are represented by the inferred activity of entire neural populations.

functional magnetic resonance imaging; vision; contrast suppression; forward model; primary visual cortex

NEURONS IN PRIMARY VISUAL CORTEX (V1) exhibit cross-orientation suppression: a grating orthogonal to a neuron's preferred orientation (mask) suppresses the response to a simultaneously presented grating at the neuron's preferred orientation (target) (Carandini et al. 1997; DeAngelis et al. 1992; Geisler and Albrecht 1992; Morrone et al. 1982). Psychophysically, this leads to impairment in the detection of the target grating, a phenomenon known as cross-orientation masking (Foley 1994). The suppression in neural activity is best characterized as a shift in the contrast gain of the neuron's response and can be accurately captured by a model based on contrast normalization. This normalization model (Heeger 1992) encompasses a linear receptive field, soft-thresholding, and divisive suppression. The divisive signal increases with the overall contrast in the stimulus across all orientations and suppresses (or normalizes) the activity produced by the grating of the neuron's preferred orientation. The normalization model has been proposed to explain stimulus-evoked responses in various cortical areas including V1 (Carandini and Heeger 1994; Carandini et

al. 1997; Heeger 1992), MT (Rust et al. 2006; Simoncelli and Heeger 1998), and inferotemporal cortex (Zoccolan et al. 2009), multisensory integration in MST (Ohshiro et al. 2011), the representation of value in LIP (Louie and Glimcher 2010), olfactory processing in *Drosophila* antennal lobe (Olsen et al. 2010), and modulatory effects of attention on visual cortical neurons (Reynolds and Chelazzi 2004; Reynolds and Heeger 2009). Similarly, the normalization model accurately predicts the activity of large populations of neurons in cat primary visual cortex (measured with electrode arrays) and visually evoked potentials in human subjects (Busse et al. 2009).

Measuring the responses of distinct subpopulations of neurons (e.g., different orientation-selective channels) in the human brain poses a serious challenge. Even though functional magnetic resonance imaging (fMRI) allows us to noninvasively measure human brain activity, each voxel in visual cortex contains a large number of orientation-selective neurons, most likely encompassing the full range of possible orientations, and the voxel's response amplitudes reflect the pooled activity of all these neurons. Consequently, it has been difficult to test the normalization model in humans with fMRI, although some progress has been made (Busse et al. 2009; Moradi and Heeger 2009).

In the present study, we employed a forward modeling technique (Brouwer and Heeger 2009; Kay et al. 2008) to transform voxel responses to orientation-selective channel responses. The analysis relied on the well-established finding that there are slight biases in orientation preferences across voxels (Freeman et al. 2011; Haynes and Rees 2005a; Kamitani and Tong 2005). This allowed us to measure cross-orientation suppression and demonstrate the validity of the normalization model in human visual cortex.

## MATERIALS AND METHODS

**Observers and scanning sessions.** Four healthy observers between the ages of 24 and 35 yr participated in this study. Observers provided written informed consent. Experimental procedures were in compliance with the safety guidelines for MRI research and were approved by the University Committee on Activities Involving Human Subjects at New York University. Observers had normal or corrected-to-normal vision.

Each observer participated in one or two experimental sessions of the weight estimation experiment (see below), consisting of 8–10 runs each, one or two experimental sessions of the main cross-orientation experiment, and one or two experimental sessions of the control experiment (see below), with sessions of each experiment consisting of 8 runs each. Observers also participated in a retinotopic mapping session and a session in which a high-resolution anatomical volume was acquired.

The weight estimation experiment, the main contrast suppression experiment, and the control experiment were performed in separate sessions. We would have preferred to acquire all the data in one scanning session, but the combined number of different conditions

Address for reprint requests and other correspondence: G. J. Brouwer, Dept. of Psychology and Center for Neural Science, New York Univ., New York, NY 10003 (e-mail: gbrouwer@nyu.edu).

between experiments (26) was too high to robustly estimate response amplitudes for all of these conditions at once. Therefore, we opted to dedicate each session to one experiment only, so that the response amplitudes could be estimated robustly. This does require coregistration of the data acquired in different sessions to align the same voxels from session to session. However, imperfections in the registration across scanning sessions were not problematic. The orientation biases in each voxel reflect a coarse-scale bias for radial orientation across the retinotopic map, not the fine-scale columnar architecture for orientation (Freeman et al. 2011). Consequently, neighboring voxels have a very similar orientation bias, i.e., very similar channel weightings.

**Visual stimulus presentation.** Visual stimuli were presented with an LCD projector (Eiki LC-XG100; Eiki, Rancho Santa Margarita, CA) with a pixel resolution of  $1,024 \times 768$  and a 60-Hz refresh rate. Subjects viewed the image from the LCD projector on a rear projection screen placed inside the bore of the magnet at a distance of 57 cm, yielding a field of view of  $32 \times 20^\circ$ . The monitor was calibrated by using a spectroradiometer (SpectraColorimeter PR650; Photo Research, Chatsworth, CA) to achieve a linear gamma.

**MRI acquisition.** MRI data were acquired with a 3-T, head-only MRI scanner (Allegra; Siemens, Erlangen, Germany) using a head coil (NM-011; NOVA Medical, Wakefield, MA) for transmitting and an eight-channel phased-array surface coil (NMSC-071; NOVA Medical) for receiving. Functional scans were acquired with gradient-recalled echo-planar imaging to measure blood oxygen level-dependent (BOLD) changes in image intensity (Ogawa et al. 1990). Functional imaging was conducted with 24 slices oriented perpendicular to the calcarine sulcus and positioned with the most posterior slice at the occipital pole (repetition time, 1.5 s; echo time, 30 ms; flip angle,  $75^\circ$ ;  $2 \times 2 \times 2.5$  mm;  $64 \times 64$  grid size). A T1-weighted magnetization-prepared rapid gradient echo (MPRAGE,  $1 \times 1 \times 2.5$  mm) anatomical volume was acquired in each scanning session with the same slice prescriptions as the functional images. This anatomical volume was aligned with a robust image registration algorithm (Nestares and Heeger 2000) to a high-resolution anatomical volume. The high-resolution anatomical volume, acquired in a separate session, was the average of several MPRAGE scans ( $1 \times 1 \times 1$  mm) that were aligned and averaged and was used not only for registration across scanning sessions but also for gray matter segmentation and cortical flattening (see below).

**Defining visual cortical areas.** Primary visual cortex was defined by standard retinotopic mapping methods (Engel et al. 1994, 1997; Larsson and Heeger 2006; Sereno et al. 1995). Visual area boundaries were drawn by hand on the flat maps, following published conventions (Larsson and Heeger 2006), and the corresponding gray matter coordinates were recorded.

**Stimuli and experimental protocols.** Stimuli were contrast-reversing sinusoidal gratings (spatial frequency: 1 cycle/ $^\circ$  of visual angle; temporal frequency: 1.33 cycles/s), within an annular aperture (inner radius:  $0.5^\circ$  of visual angle; outer radius:  $8^\circ$  of visual angle) at different orientations (weight estimation experiment) and/or different contrasts (cross-orientation suppression and control experiments). The mean luminance of the stimulus and background was  $526 \text{ cd/m}^2$ . At maximum contrast, the minimum luminance of the stimulus was  $31 \text{ cd/m}^2$  and the maximum luminance  $1,083 \text{ cd/m}^2$ . In the weight estimation experiment (Fig. 1A), we presented stimuli at maximum contrast, with six different possible orientations ( $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ ,  $120^\circ$ , and  $150^\circ$ ). In the cross-orientation suppression experiment (Fig. 1B), we presented a vertical ( $0^\circ$ ) target grating at five different contrasts (1.56%, 3.125%, 6.25%, 12.5%, and 50%) either in isolation (target-only condition) or superimposed with a horizontal ( $90^\circ$ ) mask grating of a constant contrast of 50% (target + mask condition). Size, extent, and spatial and temporal frequencies of the stimuli were identical to those in the weight estimation experiment. In the control experiment, we used identical gratings; however, the target and mask were temporally interleaved: the target grating modulated from zero

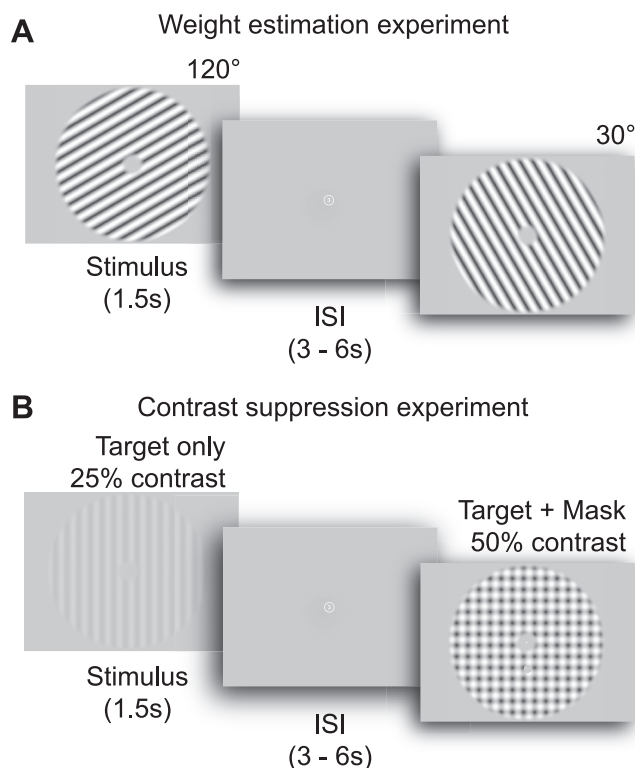


Fig. 1. Stimulus and experimental protocol. Stimuli were contrast-reversing sinusoidal gratings, within an annular aperture. *A*: in the weight estimation experiment, stimuli were full-contrast gratings, with 6 different orientations. ISI, interstimulus interval. *B*: in the cross-orientation suppression experiment, stimuli were vertical target gratings with different contrasts either in isolation (target-only condition) or superimposed with a high-contrast, horizontal, mask grating (target + mask condition). In the control experiment (not shown) we used identical gratings, but the target and mask were temporally interleaved and doubled in contrast.

contrast to the maximum contrast and back to zero contrast, followed by the mask grating, which modulated from zero contrast to the maximum contrast and back to zero contrast. Two of these cycles made up each stimulus presentation. In addition, we doubled the contrasts of the target stimuli (3.125%, 6.25%, 25%, and 100%) and the mask stimulus (100%) to compensate for the 50% reduction in duty cycle compared with the cross-orientation suppression experiment.

Stimuli were presented for 1.5 s in randomized order, interleaved with interstimulus intervals (ISIs) that ranged from 3 to 6 s, in steps of 1.5 s. In the weight estimation runs, all six possible orientations were presented eight times in each run, along with eight blank trials. This created a total of 56 trials per run (including blank trials), with one run lasting 5 min and 42 s. In the cross-orientation suppression and control experiments, all 10 different stimuli (5 different contrasts, target-only/target + mask condition) were presented 6 times in each run, along with 6 blank trials. This created a total of 66 trials per run (including blank trials), with one run lasting 6 min and 42 s.

Observers performed a two-back detection task continuously throughout each run to maintain a consistent behavioral state and to encourage stable fixation. A sequence of digits (0 to 9) was displayed at fixation (each appearing for 400 ms). The observer's task was to indicate, by means of a button press, whether the current digit matched that from two steps earlier. We have used this or a similar protocol in previous fMRI experiments and have found that the measured fMRI responses are more reliable under these conditions, even though the subjects' attention is diverted away from the target stimuli in the periphery. Without any attentional control, or if subjects are attending the peripheral/parafoveal target stimuli, we (and others)

have reported large and highly variable (trial to trial) attentional effects in visual cortex (Gandhi et al. 1999). Diverting attention away from the target stimuli yields a measure of the stimulus-evoked responses that is not confounded with such attentional modulation. Moreover, the single-unit and multiunit electrophysiological measurements of cross-orientation suppression and normalization have mostly been performed with anesthetized animals, i.e., without attending the target stimuli. Diverting attention to fixation allowed us to compare our results with these above-mentioned studies.

**Response time courses and response amplitudes.** fMRI data were preprocessed with standard procedures. The first four images of each run were discarded to allow the longitudinal magnetization to reach steady state. We compensated for head movements within and across runs with a robust motion estimation algorithm (Nestares and Heeger 2000), divided the time series of each voxel by its mean image intensity to convert to percentage signal change and compensate for distance from the RF coil, and linearly detrended and high-pass filtered the resulting time series with a cutoff frequency of 0.01 Hz to remove low-frequency drift. A V1 region of interest (ROI) was defined, separately for each observer, with retinotopic mapping procedures (see above). The aperture for the stimuli in the experiments reported in this study was identical to the aperture of the checkerboard stimuli used during retinotopic mapping, ensuring that the V1 ROI contained only voxels to which the stimulus was visible. The hemodynamic impulse response function (HIRF) for the V1 ROI was estimated with deconvolution (Dale 1999) using the same procedure that we have described in detail previously (Brouwer and Heeger 2009).

The response amplitudes for each trial type (contrast-orientation combination) were computed separately for each voxel in the V1 ROI and separately for each run by linear regression. A regression matrix was constructed for the ROI by convolving the ROI-specific HIRF and its numerical derivative with binary time courses corresponding to the onsets of each trial type (with 1s at each stimulus onset and 0s elsewhere). For the weight estimation experiment, this resulted in a regression matrix with 12 columns: 6 columns for the HIRF convolved with each of the 6 stimulus onsets and 6 columns for the HIRF-derivative convolved with each of the 6 stimulus onsets. For the cross-orientation suppression and control experiments, the regression matrix had 20 columns: 10 columns for the HIRF convolved with each of the 10 stimulus onsets (5 contrasts, with or without the mask) and 10 columns for the HIRF-derivative convolved with each of the 10 stimulus onsets. Each column of the regression matrix was linearly detrended and high-pass filtered, identically to the preprocessing of the fMRI measurements. Response amplitudes were estimated by multiplying the pseudoinverse of this regression matrix with the measured (and preprocessed) fMRI response time courses. The values (beta weights) obtained for the derivative regressors were discarded after response amplitudes were estimated (Brouwer and Heeger 2009). We included the HIRF-derivative in the regression, even though the associated beta weights were discarded, because the HIRF of an individual voxel may have differed from the mean HIRF of the V1 ROI. The HIRF and its derivative are not mutually orthogonal, so including the derivative in the regression accounted for some of the variance in the measured response time courses and affected the response amplitudes associated with the HIRF. The variance of the estimated response amplitudes across runs was indeed smaller with the derivative included than without it. We thus obtained, for each voxel and each run, one response amplitude measurement for each of the different trial types (6 orientations or 10 combinations of contrast  $\times$  condition).

**Voxel selection.** To maximize the signal-to-noise ratio in the cross-orientation suppression and control experiments, we selected voxels that showed the highest differential responses between orientations (Fig. 2A). Specifically, we computed the ANOVA  $F$ -statistic of response amplitudes in the weight estimation experiment across orientations for each voxel. Voxels were included whose  $F$ -statistic was

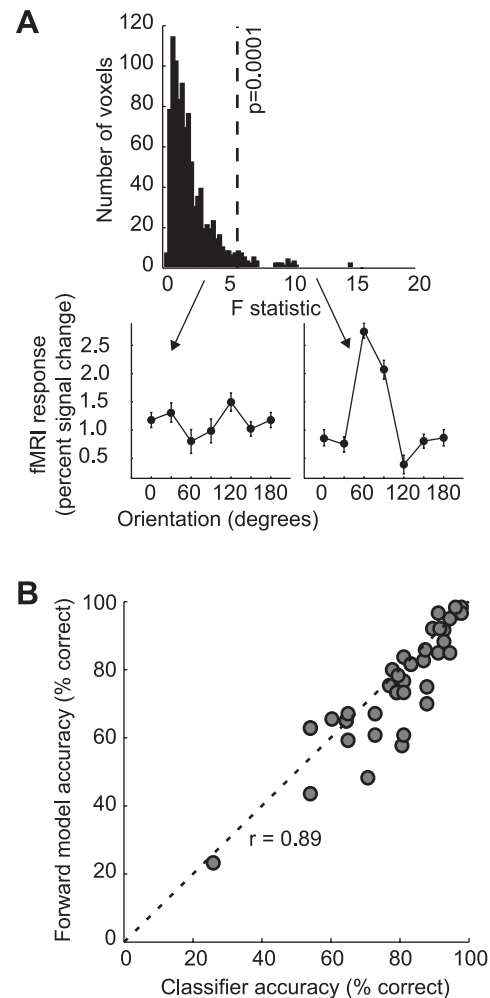


Fig. 2. *A*: voxel selection. Distribution of  $F$ -statistic values taken from V1 of 1 representative subject. The  $F$ -statistic quantifies how well a single voxel differentiates between stimulus orientations: a voxel with a low  $F$ -statistic (*left*) shows no significant bias for orientation, while a voxel with a high  $F$ -statistic (*right*) shows a clear (and significant) tuning, centered on  $60^\circ$ . fMRI, functional magnetic resonance imaging. *B*: orientation decoding with the forward model. The accuracy of orientation decoding using the forward model is plotted against the accuracy using a conventional classifier. Each data point represents a scanning session. The forward model reduced the high-dimensional (no. of voxels) voxel space to a low-dimensional (no. of channels = 6) channel space. This dimensionality reduction did not result in a considerable loss of information. The decoding accuracies were nearly the same, but the conventional classifier utilized all the information in the full, high-dimensional voxel space. Classification was performed with a 8-way maximum likelihood classifier, implemented by the Matlab (Mathworks) function 'classify' with the option 'diaglinear'.

above the median  $F$ -statistic of all voxels in the V1 ROI, selecting 50% of the original voxels for the subsequent analysis of the responses in the cross-orientation suppression and control experiments. The median split was arbitrary and was used solely to remove noisy voxels. A range of  $F$ -statistic thresholds (25th–75th percentile) yielded similar results and supported the same conclusions.

**Baseline removal.** Before the channel responses were computed, a baseline was removed from each voxel's response, separately for each run, in each scanning session. Specifically, let  $v$  be the number of voxels and  $c$  the number of conditions (e.g., orientations/contrasts), giving us, for each run, a matrix of estimated response amplitudes  $\mathbf{B}$  of size  $v \times c$ . For each  $\mathbf{B}$ , we computed the mean voxel responses across all stimulus conditions, yielding a vector  $\mathbf{m}$  of mean response amplitudes of length  $v$  (1 per voxel). This vector was normalized to a unit vector and removed by linear projection from the responses to each

stimulus condition:  $\mathbf{B} = \mathbf{B} - \mathbf{m}(\mathbf{m}^T\mathbf{B})$ . The baseline removal was done as an additional preprocessing step before transforming the voxel responses to the channel responses (described next); this analysis step was not performed when computing the mean responses across voxels.

*Forward model.* Following our previous work on color vision (Brouwer and Heeger 2009), we used a forward model of orientation tuning to separate the voxel responses into a smaller number of channel responses, each tuned to a different orientation. The forward model assumed that each voxel contained a large number of orientation-selective neurons, each tuned to a different orientation. We characterized the orientation selectivity of each neuron as a weighted sum of six hypothetical channels, each with an idealized orientation tuning curve (or basis function) such that the transformation from orientation to channel outputs was one-to-one and invertible. The shape of the tuning curves was selected (as described below) so that the response tuning of any one neuron could be expressed as a weighted sum of the six basis functions and these basis functions would closely resemble measured neural tuning curves. We further assumed that the response of a voxel was proportional to the summed responses of all the neurons in that voxel. This assumption is justified by the observation that the residual hemodynamic responses have been shown to be tightly linked with spiking activity, after removing the baseline response component (see DISCUSSION). The underlying orientation tuning curve of each neuron was presumed to be a property of that neuron, such that it did not depend on the stimulus (the neuron's response was presumed to depend on the stimulus, but the underlying tuning curve was presumed to be independent of the stimulus). Hence, the weights on each voxel likewise were presumed to be independent of the stimulus.

Each basis function was a sinusoid raised to the fifth power. Raising to the fifth power made the tuning curves narrower and thereby comparable to physiological findings. A tuning curve with any possible orientation preference (i.e., intermediate to the 6 basis functions) can be expressed exactly as a weighted sum of the six basis functions (Freeman and Adelson 1991). We evenly spaced the six basis functions along the orientation axis, so that one channel responded maximally to the vertical target stimulus and another channel responded maximally to the horizontal mask stimulus.

In the first stage of the analysis, we used the data from the weight estimation experiment to estimate the weights on the six hypothetical channels separately for each voxel. With these weights in hand, the second stage of analysis computed the channel outputs associated with the spatially distributed pattern of activity across voxels evoked by the stimuli in the cross-orientation suppression and control experiments. This allowed us to transform the voxel responses to the channel responses, each tuned to a different orientation. Let  $k$  be the number of channels,  $m$  the number of voxels, and  $n$  the number of repeated measurements (i.e., 6 orientations times the number of runs for the weight estimation experiment). The matrix of estimated response amplitudes in the weight estimation experiment ( $\mathbf{B}_w$ ,  $m \times n$ ) was related to the matrix of hypothetical channel outputs ( $\mathbf{C}_w$ ,  $k \times n$ ) by a weight matrix ( $\mathbf{W}$ ,  $m \times k$ ):

$$\mathbf{B}_w = \mathbf{W}\mathbf{C}_w \quad (1)$$

The least-squares estimate of the weights was computed with linear regression:

$$\widehat{\mathbf{W}} = \mathbf{B}_w\mathbf{C}_w^T(\mathbf{C}_w\mathbf{C}_w^T)^{-1} \quad (2)$$

The channel responses ( $\mathbf{C}_c$ ) associated with the cross-orientation suppression and control experiment responses ( $\mathbf{B}_c$ ) were estimated by using the weights ( $\widehat{\mathbf{W}}$ ):

$$\widehat{\mathbf{C}}_c = (\widehat{\mathbf{W}}^T\widehat{\mathbf{W}})^{-1}\widehat{\mathbf{W}}^T\mathbf{B}_c \quad (3)$$

For these matrices to be invertible, the number of voxels must be greater than the number of channels, and there must be an uneven

weighting of the orientation channels in at least a subset of the voxels so that the voxels exhibit sufficiently different responses to the different orientations (specifically, the space spanned by the voxel responses must be at least as large as the number of channels). These requirements posed no difficulty for our analysis. First, the average size of V1 across subjects, at our scanning resolution, was  $\sim 800$  voxels, much larger than the number of channels. Second, there are stable biases in the responses of voxels to different orientations, exhibiting a robust, coarse-scale, radial organization (Freeman et al. 2011). We combined the intermediate channels (not tuned to either horizontal or vertical) such that responses from the two channels closest to the horizontal channel were estimated separately but later averaged together to form one single intermediate channel and responses from the two channels closest to the vertical channel were estimated separately but later averaged into another intermediate channel. For each main experimental session, the estimated channel responses were then averaged across runs, separately for each stimulus condition, within that session. Finally, we computed the mean and SE of the channel responses across subjects, separately for each stimulus condition and separately for the cross-orientation suppression and the control experiment.

*Weight estimation stability and forward model fit.* The forward model provided a good fit to the orientation biases in the voxel responses. We evaluated this in several ways. First, the average  $r^2$  value (explained variance of the fit) for the weight estimation experiment was 0.76 across the subjects. This was much larger than expected by chance ( $P < 0.001$ ). Second, we used decoding accuracy to assess the model fit. Within-session accuracy for decoding orientation from the voxel responses was around 68%, well above the chance level of 16% ( $P < 0.001$ ). Decoding accuracies obtained with the forward model were highly correlated to those obtained with a conventional classifier (Fig. 2B). This demonstrates that there was little or no information lost by replacing the voxel responses with the channel responses, i.e., by reducing the high-dimensional voxel data with the lower-dimensional space of six channels. Third, we found that the best-fit weights were stable across scanning sessions on different days (Fig. 3A). We used data from one weight estimation session to calculate the weights and then applied those weights to measure how well they fit the data from a second weight estimation session. This revealed  $r^2$  values (mean 0.54) that were still significantly larger than expected by chance ( $P < 0.001$ ). In addition, orientation decoding (using the forward model and estimated weights) was also stable across sessions (Fig. 3B). We used data from one weight estimation session to calculate the weights and then applied these weights to decode orientation from a second weight estimation session. Session-to-session decoding accuracies were  $\sim 50\%$  on average, much larger than expected by chance ( $P < 0.001$ ). Finally, the "preferred orientations" (determined using the best-fit weights) of voxels were stable from session to session (Fig. 3C). We computed the preferred orientation for each voxel as follows. The six channels were tuned for six different orientations. This provided us with six unit vectors, of unit length, each pointing to the preferred orientation of one orientation channel in the forward model. We multiplied each of these vectors with each of the six corresponding estimated weights and computed the vector sum of the resulting six vectors. The angle of the resulting vector sum was taken to be the "preferred orientation" of that voxel. This calculation was repeated for each voxel in two separate sessions. Preferred orientations were highly correlated across sessions, for all subjects, significantly higher than what would be expected by chance ( $P < 0.01$ ).

*Main experiment forward model fit.* After estimating the weights, how well did the forward model fit the voxel responses from the main experiment and the control experiment? If there was no neuronal interaction (i.e., suppression) between the vertical and horizontal gratings, a plaid would evoke channel responses that are the sum of the channel responses evoked by the vertical and horizontal gratings in isolation. Such a scenario would have allowed us to determine the

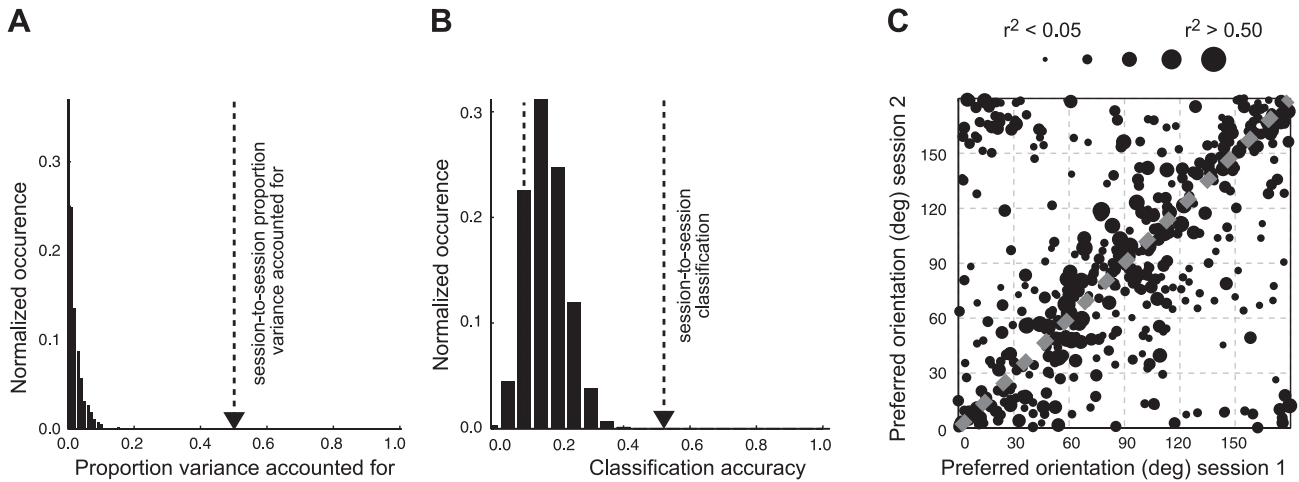


Fig. 3. Stability of weight estimation across sessions. *A*: estimating the weights in one session and then applying these weights to fit the data of a second session revealed  $r^2$  values significantly higher than chance: the mean  $r^2$  value was 0.54, many standard deviations away from the null distribution of  $r^2$  values. The null distribution was obtained by shuffling the weights between the voxels. *B*: accuracy of the forward model in decoding orientation in one session, using weights estimated from a different session. Decoding accuracy was significantly higher than expected by chance: the mean accuracy of 0.56 was in the 99.75th percentile of the null distribution. The null distribution was obtained by shuffling the weights between the voxels. *C*: session-to-session comparison of the preferred orientation. Each data point represents a voxel. The preferred orientation (a continuous measure) of each voxel was computed by using the response amplitudes to each stimulus orientation in the weight estimation experiment. Points cluster around the diagonal, indicating that the preferred orientations of most voxels were stable between sessions. The size of each point represents  $r^2$ , the proportion of the variance in the voxel's response time course that was accounted for by the regression model (i.e., the regression matrix and hemodynamic impulse response function that was used to estimate the response amplitudes). Voxels with a robust responses to the stimuli (higher  $r^2$  values) tended to have stable weights across sessions (closer to the diagonal).

fit between channel responses computed from the voxel responses and predicted channel responses. This was not possible, however, because of suppression. We therefore used a different approach to determine the quality of the forward model fit in the main experiment. The channel responses were computed by multiplying the voxel responses with the inverse of the weight matrix (from the weight estimation experiment). We then computed a set of predicted voxel responses by multiplying the channel responses with the weights and compared them with the measured voxel responses. Specifically, we determined the proportion of the variance in the voxel responses that was accounted for by the forward model (i.e.,  $r^2$ ). The resulting  $r^2$  values were then compared with a null distribution of  $r^2$  values computed by randomly shuffling the weights between the voxels. We found that the  $r^2$  values for both the main and control experiments (in all subjects) were modest (mean  $r^2 = 0.33$ ) but significantly above the 99th percentile of the null distribution.

*Normalization model and fitting procedures.* The normalization model (Heeger 1992) was used to fit the channel responses from the cross-orientation suppression experiment and the control experiment. The model for the cross-orientation suppression experiment included cross-orientation suppression terms. The response of each channel  $r_i$  as a function of target and mask contrast was modeled as:

$$r_i = r_{\max} \left( \frac{c_t^n v_i(\theta_t) + c_m^n v_i(\theta_m)}{\left( \sqrt{c_t^2 + c_m^2} \right)^n + \sigma^n} \right) + b \quad (4)$$

where  $c_t$  is the target contrast,  $c_m$  is the mask contrast,  $v_i$  is the tuning curve of the  $i$ th channel,  $\theta_t$  is the orientation of the target grating, and  $\theta_m$  is the orientation of the mask grating. For the channel tuned to the target orientation,  $v_i(\theta_t) = 1$  and  $v_i(\theta_m) = 0$  (the channel did not respond to gratings orthogonal to its preferred orientation). In the channel tuned to the orientation of the mask grating,  $v_i(\theta_t) = 0$  and  $v_i(\theta_m) = 1$ . The intermediate channels are associated with intermediate values for  $v_i(\theta_t)$  and  $v_i(\theta_m)$ . The model had four parameters,  $\sigma$ ,  $n$ ,  $r_{\max}$ , and  $b$ , determining, respectively, the gain, slope, saturation, and baseline of the resulting contrast-response functions. For the channel responses during the control experiment, we used a modified version of the model without cross-orientation suppression:

$$r_i = r_{\max} \left( \frac{c_t^n v_i(\theta_t)}{c_t^n + \sigma^n} + \frac{c_m^n v_i(\theta_m)}{c_m^n + \sigma^n} \right) + b \quad (5)$$

Thus in this modified version we weighted the normalizing signal originating from each grating by the contrast energy of that grating, while in the cross-orientation suppression experiment we normalized the channel activity by the summed contrast energy of all gratings, regardless of the channel's sensitivity to those gratings. We fitted the channel responses simultaneously from both the cross-orientation suppression and control experiments to all contrasts and both conditions (target only, target + mask). We used a single  $n$ ,  $r_{\max}$ , and  $b$ , but allowed  $\sigma$  to vary between the two experiments. Thus there were five free parameters ( $\sigma$  main,  $\sigma$  control,  $n$ ,  $r_{\max}$ , and  $b$ ). In addition, we fitted the models to each experimental data set separately, to determine which version (with or without cross-orientation suppression) was more appropriate for the data from each experiment.

Statistical significance of the model fits was determined with cross-validation. We divided the data from each experiment in half. Both models were fitted to the first half of the data. Using the resulting parameters, we created predicted channel responses and computed the amount of variance in the remaining half of the data explained by the predicted channel responses ( $r^2$ ). Repeating this procedure a large number of times with different (random) subdivisions of the data generated distributions of  $r^2$  values, one distribution for each version of the model (with and without cross-orientation suppression). Taking the ratio between these two distributions generated a new distribution for which we determined the median and 5th and 95th percentiles. If the  $r^2$  values were not statistically different, the distribution of their ratios would have been centered on 1, with the 5th percentile being smaller and the 95th percentile being larger than 1. If, for the distribution of ratios for *model A* over *model B*, we found that the 5th percentile was larger than 1, we concluded that *model A* provided a statistically better fit to the data than *model B*. If, on the other hand, the distribution of ratios for *model A* over *model B* yielded a 95th percentile smaller than 1, we concluded that *model B* provided a statistically better fit to the data than *model A*.

**Contrast gain fit.** To test and measure the shift in gain between the target-only and the target + mask conditions, we fitted a simplified version of normalization model to the responses in the target channel:

$$r_i = r_{\max} \left( \frac{c^n}{c^n + \sigma^n} \right) + b \quad (6)$$

For these fits, we fixed all parameters to the values obtained in the simultaneous fit (described above), with the exception of  $\sigma$  (the parameter determining the gain of the function). By fitting the target-only and target + mask conditions separately, we thus obtained a value of  $\sigma$  for each condition. If cross-orientation suppression indeed led to an increase in gain, the  $\sigma$  parameter for the target + mask condition should have been larger than the  $\sigma$  in the target-only condition. For the control experiment,  $\sigma$  values should have been similar, as there was no suppression of the mask on the target. The statistical significance of any difference in  $\sigma$  values was determined by means of bootstrapping. We randomly resampled the data, with replacement. The model was then fitted to the resampled data set. Repeating this procedure a large number of times provided a distribution of  $\sigma$  parameter values that was statistically compared between conditions and experiments, using the ratio method described above.

**Mean responses fit.** The normalization model was also fit to the responses averaged across all voxels in the V1 ROI. The goal of this analysis was to determine whether the mean responses reflected the underlying cross-orientation suppression observed in the channel responses. The data from both experiments (cross-orientation suppression and control experiment) were fit with two versions of the normalization model, one including cross-orientation suppression:

$$r = r_{\max} \left( \frac{c_t^n + c_m^n}{\sqrt{(c_t^2 + c_m^2)^n + \sigma^n}} \right) + b \quad (7)$$

and one without suppression:

$$r = r_{\max} \left( \frac{c_t^n}{c_t^n + \sigma^n} + \frac{c_m^n}{c_m^n + \sigma^n} \right) + b \quad (8)$$

These equations specifically modeled the average responses to the gratings, rather than the responses of separate channels. The fits to the data from both experiments were compared with the same cross-validation procedure described above (see *Normalization model and fitting procedures*).

**Psychophysics.** We measured the strength of cross-orientation masking in a separate experiment (outside the scanner). Mimicking the stimuli used in the fMRI experiment, we presented subjects with 1) a vertical target grating presented in isolation for 375 ms (target-only condition), 2) the vertical target superimposed with the horizontal mask stimulus (target + mask condition), and 3) the vertical target temporally interleaved with a horizontal mask stimulus (100% contrast) appearing for 375 ms before and after the target grating (control condition). In two additional conditions, we presented subjects the vertical target superimposed with a horizontal mask stimulus with a spatial frequency 4 or 0.25 times the spatial frequency of the vertical target. Orientation-discrimination thresholds were measured separately for each of these conditions, in separate blocks. Cross-orientation masking should be observed only in *condition 2*, in which the vertical target and horizontal mask are spatially and temporally superimposed and have the same spatial frequency.

The strength of cross-orientation masking was measured as the orientation-discrimination threshold at a fixed target (3.125%) and mask (50%) contrast. In a two-interval forced choice task (2IFC), subjects were presented with two stimuli, one in which the target was rotated slightly clockwise relative to vertical and the other in which it was rotated slightly counterclockwise of vertical. The 75% correct threshold was determined by running four interleaved staircases for each of the five conditions, one 2-up-1-down staircase and one 3-up-1-down staircase starting at zero orientation difference and one

2-up-1-down staircase and one 3-up-1-down staircase starting at an easily detected orientation difference (which varied slightly depending on the subject). The subjects' responses were converted to percent correct at each orientation difference, and the resulting data points were fitted to a cumulative normal distribution with two free parameters: mean and variance. From these fits, we computed the 75% correct orientation difference thresholds for each subject and condition.

## RESULTS

**Psychophysical results.** The effectiveness of our stimuli (Fig. 1) to induce cross-orientation suppression was determined psychophysically by measuring cross-orientation masking. We determined subjects' orientation discrimination thresholds at a fixed target (3.125%) and mask (50%) contrast, using a 2IFC orientation-discrimination task. Orientation-discrimination thresholds (75% correct responses) were low for the target-only condition: very small orientation differences were easily detected by all subjects (Fig. 4). Thresholds were considerably and statistically significantly higher when the target was presented with the mask superimposed, indicative of

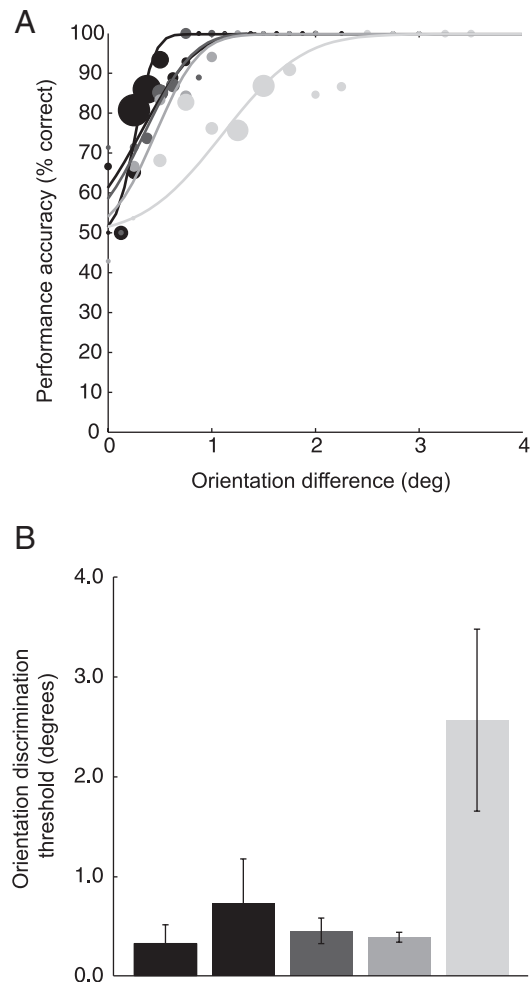


Fig. 4. Psychophysical results. *A*: psychometric functions (see MATERIALS AND METHODS) for 1 representative subject. Symbol size is proportional to number of trials, which differed because of the staircase procedure (see MATERIALS AND METHODS). SF, spatial frequency. *B*: orientation-discrimination thresholds (75% correct), averaged across subjects. Error bars, SE across subjects. Colors indicate the different conditions as described in *A*.

cross-orientation masking. When the mask stimulus was temporally interleaved with the target, thresholds were similar to those obtained for the target-only condition. Superimposing the vertical target with a horizontal mask of higher ( $4\times$ ) or lower ( $0.25\times$ ) spatial frequency also resulted in thresholds comparable to the target-only condition, demonstrating that the strength of suppression between orientations was selective for spatial frequency. The threshold for the target + mask condition was found to be significantly higher than those for the other conditions, while the thresholds in other conditions did not differ significantly from each other ( $F_{4,19} = 5.02$ ,  $P = 0.009$ ). Together, these results provide psychophysical evidence of suppression during our main fMRI cross-orientation suppression experiment as well as the expected lack of suppression during our control fMRI experiment.

**Cross-orientation suppression experiment.** A forward model was used to estimate the stimulus-evoked responses in several orientation-selective channels (see MATERIALS AND METHODS for details). In brief, measured responses to full-contrast gratings of six different orientations ( $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ ,  $120^\circ$ , and  $150^\circ$ ) were used to estimate a matrix of weights characterizing the contribution of each orientation channel to each voxel (e.g., the relative number of neurons in each voxel tuned for vertical vs. other orientations). Then the main experiment measured cortical activity as a function of contrast for vertical test gratings (presented at 1.56%, 3.13%, 6.25%, 12.5%, or 50% contrast) and for “plaids” in which vertical test gratings were superimposed with a horizontal mask grating with a fixed contrast of 50% (Fig. 1). The previously estimated weight matrix was used to extract the channel responses from the voxel responses. Each panel of Fig. 5 corresponds to a different channel, i.e., a different subpopulation of neurons with a different orientation preference, with one channel tuned for vertical (Fig. 5A), one tuned for horizontal (Fig. 5D), a pair tuned for  $\pm 30^\circ$  from vertical (Fig. 5B), and a pair tuned for  $\pm 60^\circ$  from vertical (Fig. 5C). The responses of these channels were measured simultaneously to the same set of visual stimuli.

The channel responses showed clear evidence of cross-orientation suppression. For the channel preferring vertical, responses increased with contrast of the (vertical) test grating, saturating at high contrasts (Fig. 5A, open symbols). This contrast-response function shifted rightward on the log-contrast axis in the presence of the (horizontal) mask grating (Fig. 5A, filled symbols). The channel preferring horizontal responded at baseline to test gratings (Fig. 5D, open symbols) and responded strongly when the mask grating was presented in the absence of a test grating, and its responses to the mask grating decreased with increased test contrast (Fig. 5D, filled symbols). The other channels exhibited intermediate effects (Fig. 5, B and C). We confirmed a similar pattern of results for each subject individually (see Fig. 7A).

**Control experiment.** To determine whether the observed channel responses indeed reflected cross-orientation suppression, we performed a control experiment in which we temporally interleaved the mask and target (and doubled the contrasts of both mask and target). Interleaving the mask and target over time kept the net stimulation approximately the same as in the cross-orientation suppression experiment. However, the target and mask stimulus were never presented simultaneously.

There was no evidence for cross-orientation suppression under these stimulus conditions. First, there was no evidence

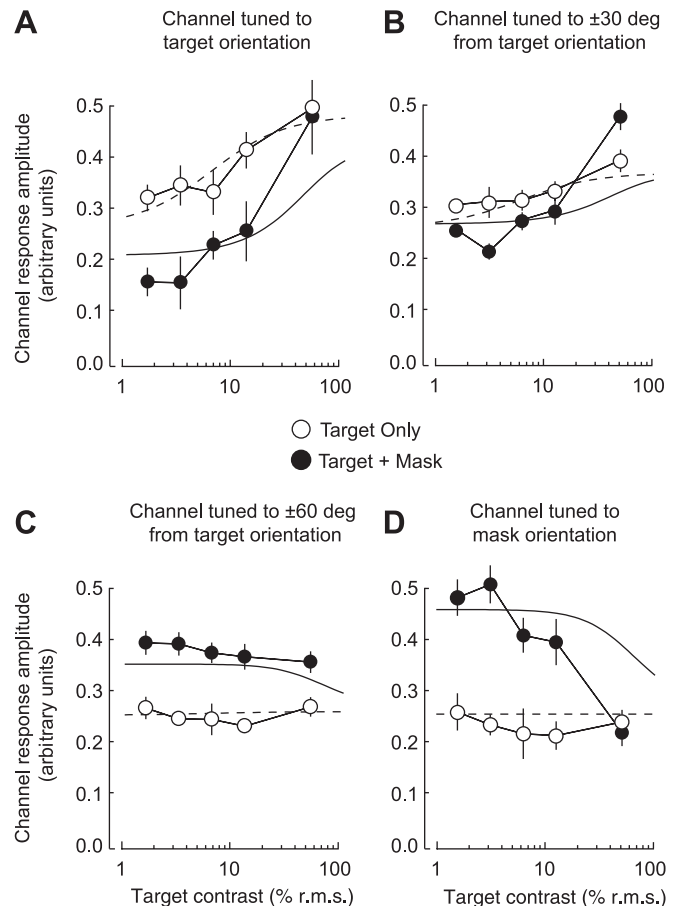


Fig. 5. Cross-orientation suppression in human V1. Each panel plots responses of orientation-selective channels. Solid (target only) and dashed (target + mask) curves depict the best fit of the normalization model. A: channel tuned to target orientation. B: average of the 2 channels adjacent to the target channel ( $\pm 30^\circ$ ). C: average of the 2 channels adjacent to the mask channel ( $\pm 60^\circ$ ). D: channel tuned to the orientation of the mask. Error bars: SE across subjects. r.m.s., root mean square.

for psychophysical cross-orientation masking (Fig. 4). Second, the fMRI contrast-response function of the channel tuned to the target grating did not shift when the mask grating was superimposed (Fig. 6A). Likewise, the channel tuned to the mask grating showed a constant baseline response to the target-only condition and a larger but constant response level to the target + mask condition (Fig. 6D). Similar to the cross-orientation suppression experiment channel responses, the intermediate channels (Fig. 6, B and C) showed intermediate effects.

**Normalization model fit.** To test whether the normalization model could capture these findings, we fitted the data from all channels, both conditions (target only, target + mask), and both the cross-orientation suppression and control experiment simultaneously. For the cross-orientation suppression experiment, the model included cross-orientation suppression terms: the activity in any channel was normalized by the activity across all channels. On the other hand, for the control experiment, the model did not include the cross-orientation suppression terms: the activity of any channel was normalized only by the activity in that channel, ignoring the activity of the other channels, tuned to different orientations. Both models had four free parameters,  $\sigma$ ,  $n$ ,  $r_{\max}$ , and  $b$ , that determined the contrast gain, slope, saturation, and baseline of the contrast-response

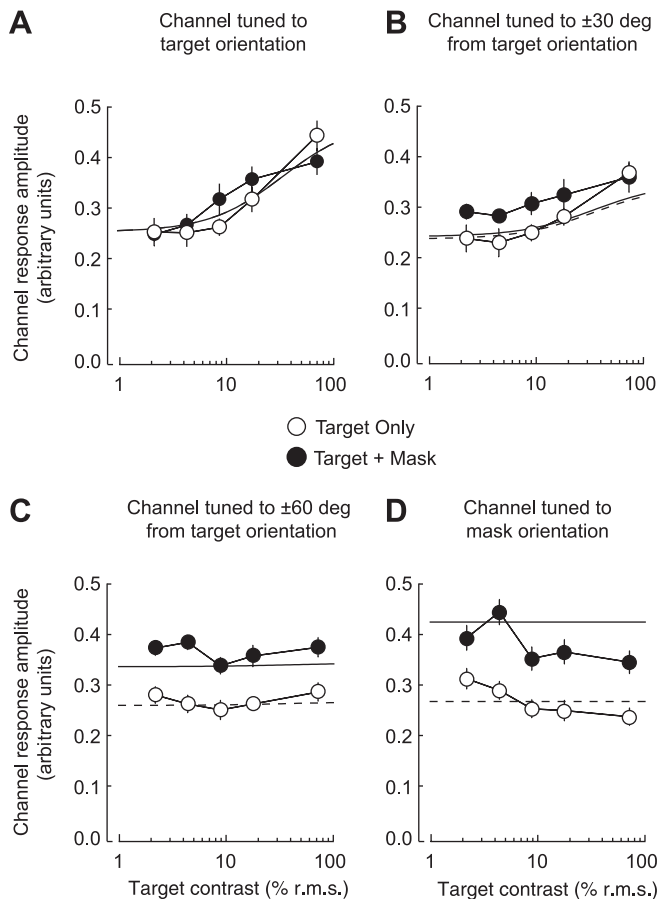


Fig. 6. No cross-orientation suppression for the control experiment. Same format as Fig. 5.

functions, respectively. We used a single  $n$ ,  $r_{\max}$ , and  $b$  parameter for each experiments but allowed  $\sigma$  to differ between experiments. The dashed and solid lines in Figs. 5 and 6 represent the best fits.

The normalization model provided a good approximation of the channel responses for both the cross-orientation suppression and control experiments. The fitted values for the contrast gain parameters were cross-orientation suppression  $\sigma = 0.06$  and control  $\sigma = 0.34$ . These values were similar to values usually obtained with single-unit experiments (see DISCUSSION).

We also fitted both variants of the model to each experimental data set separately, determining which version (suppression vs. no suppression terms) was the more appropriate model for that particular data set. For the cross-orientation suppression experiment, the cross-orientation suppression model (Eq. 4) outperformed the suppression-absent model (Eq. 5). Cross-validation (see MATERIALS AND METHODS) revealed that the former model was a better fit than the latter (cross-orientation suppression model mean  $r^2 = 0.44$ , suppression-absent model mean  $r^2 = 0.26$ ;  $P < 0.001$ ). For the control experiment, the suppression-absent model outperformed the cross-orientation suppression model (cross-orientation suppression model mean  $r^2 = 0.09$ , suppression-absent model mean  $r^2 = 0.27$ ;  $P < 0.001$ ). Thus we found clear evidence for cross-orientation suppression for appropriate stimuli but no evidence for such suppression with stimuli that were identical in extent, orienta-

tion, spatial frequency, and duty cycle but had target and mask gratings temporally interleaved to avoid cross-orientation suppression (Fig. 4).

The suppression in the target channel during the cross-orientation suppression experiment was due to a change in contrast gain (change in  $\sigma$ ). We fitted the responses in this channel to each condition (target only vs. target + mask) separately, using a simplified version of the normalization model (see MATERIALS AND METHODS), allowing  $\sigma$  to differ between conditions (Eq. 6). We found that the change in the target channel response from the target-only to the target + mask condition was associated with a significant increase in the  $\sigma$  parameter (target-only  $\sigma = 0.07$ , target + mask  $\sigma = 0.50$ ; statistical significance of the difference in  $\sigma$  values,  $P < 0.001$ ). For the control experiment, we expected no change in the  $\sigma$  parameter, because of the lack of cross-orientation suppression. Indeed, the curves for the target channel in Fig. 6A appeared to largely overlap and were of a very similar shape. We fitted the responses in this channel to each condition (target only vs. target + mask) separately, identically to the procedure used for the cross-orientation suppression experiment. Cross-validating these fits, we did not find a significant difference in the  $\sigma$  parameter between the target-only and target + mask conditions: (target-only  $\sigma = 0.29$ , target + mask  $\sigma = 0.28$ ;  $P > 0.50$ ). Fitting the data from each subject individually yielded similar results (Fig. 7B).

**Population response amplitudes.** Following previous work with electrode array recordings (Busse et al. 2009), we replotted the data in terms of population responses (Fig. 8), to see how each stimulus combination and strength was encoded by the population of orientation-selective channels. In both the main cross-orientation suppression experiment (Fig. 8A) and the control experiment (Fig. 8B), an isolated target of increasing contrast (left column) increased the responses in the sub-population tuned to the target while evoking only a constant baseline response in the remaining population. When a mask was added, we observed bimodal responses of the population to both the target and mask. However, in the main cross-orientation suppression experiment, lower target contrasts evoked less activity than the control experiment, indicating cross-orientation suppression from the high-contrast mask. The normalization model captured the behavior of population responses (Fig. 8). Note that we would have been unable to observe these effects by measuring the responses of a single neuron, as typically reported in a neurophysiological experiment. Nor would we have been able to observe these results in individual voxel responses or in the average responses across many voxels, the conventional univariate approaches to analyzing fMRI measurements. Nor would we have been able to observe these results in the decoding accuracy of a classification analysis, the typical approach to multivoxel pattern analysis with fMRI data.

**Mean response amplitudes.** The suppression of the responses of the target stimulus were not evident in the mean response amplitudes (averaged across all V1 voxels), without extracting the channel responses. For the cross-orientation suppression experiment, V1 responses increased monotonically with the contrast of the target, both with and without the mask grating (Fig. 9A). In these averaged responses, the main effect of adding a mask (of constant contrast) was a constant shift (increase) in the response amplitudes for all target con-



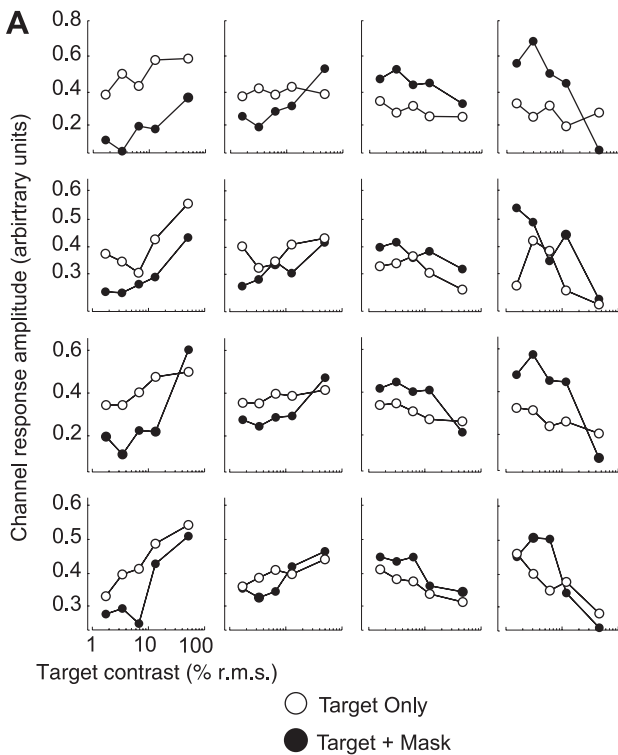


Fig. 7. Cross-orientation suppression in individual subjects. *A*: V1 channel responses for the cross-orientation suppression experiment. Each row corresponds to a different subject (same format as Fig. 5). *B*: best-fit contrast gain ( $\sigma$ ) parameter values. *x*-Axis, target only; *y*-axis, target + mask. Circles, cross-orientation suppression experiment; squares, control experiment; filled symbols, individual subjects; open symbols, fit of the mean responses across subjects. For the cross-orientation suppression experiment, the  $\sigma$  values for target only were lower than those for target + mask, indicating a shift in contrast gain. For the control experiment, the contrast gain was similar with and without the mask.

trasts. There was little, if any, interaction between conditions, such that the presence of the mask did not lead to an observable suppression of the mean responses to the target. If it had, the two contrast-response functions should have converged at high contrasts. To test this formally, we fitted the mean responses of the main cross-orientation suppression experiment and the control experiment to the normalization model with (Eq. 7) and without (Eq. 8) suppression. Both models used the  $\sigma$  and  $n$  parameters obtained in fitting the channels separately, allowing only  $r_{max}$  and  $b$  to vary. For the cross-orientation suppression experiment, the model without cross-orientation suppression provided a better fit to the mean responses (mean  $r^2 = 0.84$ )

than the model with cross-orientation suppression (mean  $r^2 = 0.71$ ); cross-validation of the fits confirmed that this difference was statistically significant ( $P < 0.01$ ). This is in agreement with the observation that there appears to be little interaction between the conditions, as shown in Fig. 9A. A small but significant ( $P < 0.05$ ) difference was found for the fits to the mean responses of the control experiment (normalization model including suppression: mean  $r^2 = 0.94$ ; without suppression: mean  $r^2 = 0.98$ ). Using the  $\sigma$  and  $n$  parameters obtained in fitting the channels separately could have potentially biased our fits of the mean responses. We therefore repeated the fitting procedure, allowing  $\sigma$  and  $n$  to vary freely. This yielded very similar and high  $r^2$  values for both models, to the mean responses of both the main and control experiment ( $r^2 \sim 0.96$ ). However, the low number of data points compared with the number of free parameters (4) resulted in overfitting. Indeed, when unconstrained, the  $\sigma$  and  $n$  parameters varied greatly between permutations during the cross-validation procedure and differed greatly from the parameter values reported for the channel responses.

DISCUSSION

We used fMRI, in combination with a forward modeling analysis (Brouwer and Heeger 2009; Kay et al. 2008), to measure cross-orientation suppression in human primary visual cortex (V1) and test the normalization model. We found that for the channel tuned to the orientation of the target grating, responses to the target grating were suppressed when a second, orthogonal mask grating was added. The remaining channels, tuned to either intermediate orientations or the mask orientation, also showed clear evidence of suppression. We found this suppression to be implemented as a change in the contrast gain of the channel responses. When the target and mask were temporally interleaved, no suppression was observed.

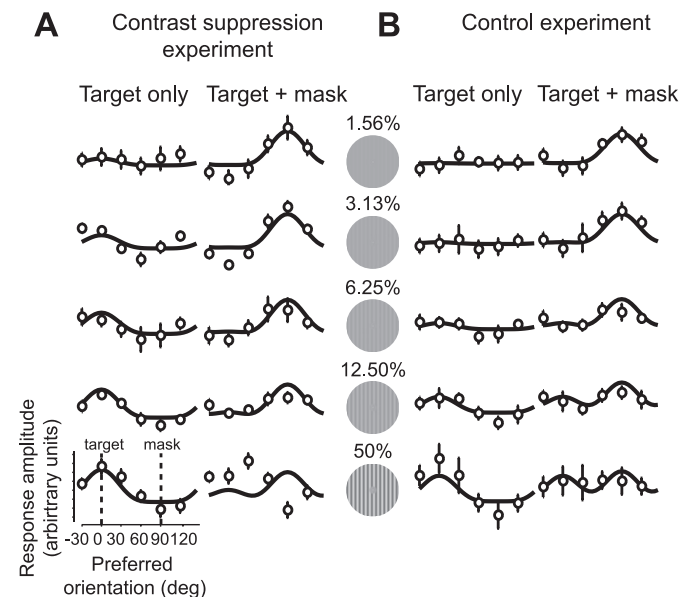


Fig. 8. Population responses. *A*: population responses to the cross-orientation suppression experiment. *Left*: target-only condition. *Right*: target + mask condition. Target contrast increases from *top* to *bottom* in each column. Solid curves depict the best-fit normalization model. *B*: control experiment. Same format as *A*. Error bars: SE across subjects.

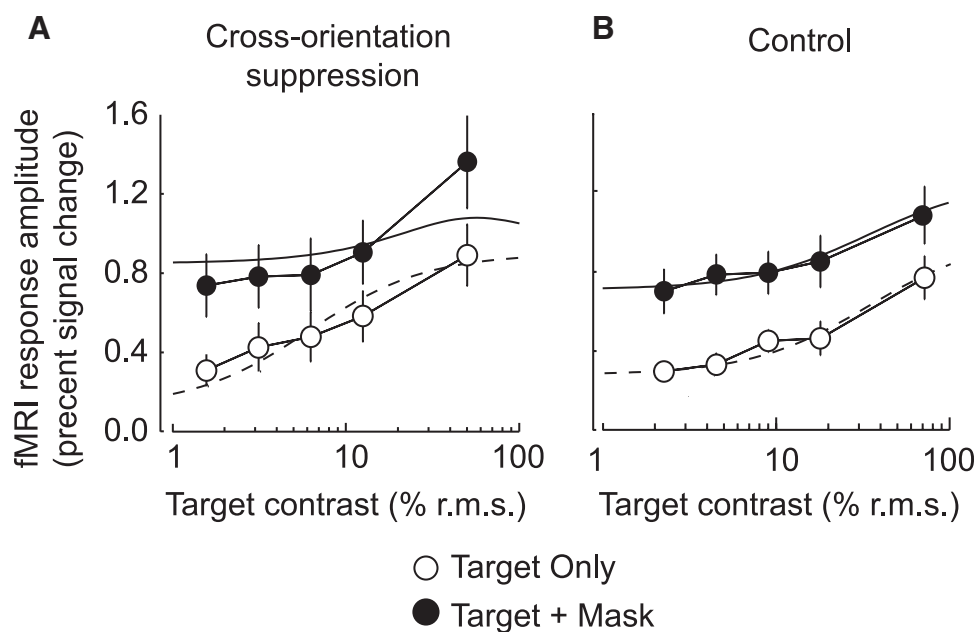


Fig. 9. Mean V1 responses. *A*: mean responses (averaged across V1 voxels) for the cross-orientation suppression experiment. Solid (target only) and dashed (target + mask) curves depict the best fit of a model that includes cross-orientation suppression. *B*: mean responses for the control experiment. Solid (target only) and dashed (target + mask) curves depict the best fit of a model that does not include cross-orientation suppression.

The decomposition of responses into idealized channels with different orientation tunings allowed us to specifically test the predictions of the normalization model (Heeger 1992), a model that has been successful in explaining neural responses in various brain areas (see introduction for references). In most of these previous studies, the normalization model was applied to the responses of individual neurons, with the exception of one study that measured population activity in cat primary visual cortex (measured with electrode arrays) and visually evoked potentials in human subjects (Busse et al. 2009) and another that measured fMRI responses to interocular suppression (Moradi and Heeger 2009). Here we provided additional evidence for the validity of the model by showing its ability to also fit fMRI responses in human visual cortex during cross-orientation suppression.

Decomposition of fMRI responses into channels also allowed us to characterize the inferred responses of an entire orientation-selective population of neurons to both isolated and superimposed stimuli of different contrast strengths. Conventional fMRI analyses average across responses of all differently tuned mechanisms and therefore offer little insight into how stimuli are encoded in the patterns of activity in an entire neural population. At the other extreme, neurophysiology measures the responses of only one neuron at a time. While this provides invaluable knowledge about how neurons integrate incoming signals into spiking activity, it is generally believed that stimuli are represented by the activity of entire neuronal populations. Multivariate pattern classification of fMRI data (Brouwer and Heeger 2009; Brouwer and van Ee 2007; Haynes and Rees 2005b; Kamitani and Tong 2005; Kay et al. 2008) and fMRI-adaptation approaches (Grill-Spector and Malach 2001; Weiner et al. 2010) have demonstrated the potential of fMRI to investigate response selectivity. More recent work, using electrode arrays to simultaneously measure the responses of many neurons, has begun to provide us with a picture of the response profiles of large populations (Busse et al. 2009) to isolated and superimposed stimuli. Here we demonstrated that

our forward model allowed us to infer these population responses from human fMRI data.

The parameter values we obtained by fitting the normalization model to the data deviated from earlier reports based on single-unit measurements. The mean value of the best-fit exponent in single neurons (cat and macaque) is  $n \approx 2$  (Geisler and Albrecht 1997), whereas we found a lower value for the exponent,  $n \approx 1.25$ . However, this was expected because we measured population responses, averaging across large numbers of neurons, each of which probably had a different  $\sigma$  value drawn from a distribution. Although the mean of the distribution of exponents has been found to be  $n \approx 2$ , fitting the mean responses after averaging across many neurons has yielded a best-fit value of  $n \approx 1$  (Busse et al. 2009). Similarly, the  $\sigma$  parameter we found was also somewhat lower than that typically reported from single-unit studies (Geisler and Albrecht 1997).

Why did the mean activity, averaged across all voxels in V1, show no evidence of suppression, unlike the channel responses (compare Fig. 5A with Fig. 9A)? In our analysis, the mean responses averaged across stimulus conditions were removed from each voxel's response before computing the channel responses (baseline removal, see MATERIALS AND METHODS). Hence, the sum of the channel responses, as we computed them, was not equal to the mean responses averaged across voxels. As a consequence, the channel responses revealed phenomena of suppression that were not evident in the mean responses. Even so, a previous study examining dichoptic stimulus presentations did find evidence for cross-orientation suppression in the mean responses averaged across voxels (Moradi and Heeger 2009).

Our main incentive for applying baseline removal to each individual run was that it increased the stability of the weights across runs and sessions. In earlier work we also observed that baseline removal dramatically increased classifier and forward model decoding and reconstruction accuracies (Brouwer and Heeger 2009). What confounding factor does such baseline removal actually remove? One possibility is that it removed

any component of the hemodynamic responses that was time-locked to the stimulus presentations but not differentially responsive to the different stimulus conditions. By removing such a baseline response component, the residual hemodynamic responses have been shown to be tightly linked with spiking activity, such that the time courses of the hemodynamic responses evoked by different stimulus contrasts were well fit as a shift-invariant linear transform of the firing rates (A. Das, personal communication). Alternatively, or in addition, baseline removal might have acted to remove run-specific noise factors (e.g., head motion, gradient heating, physiological noise, attentive state) that influenced the fit of the forward model to the measured response amplitudes. We compared several different methods of baseline removal (subtracting the mean, z-scoring and projecting out the mean, removing the mean across voxels vs. the mean across conditions). These different methods yielded very similar results that supported the same conclusions, but projecting out the mean across conditions gave the cleanest (least noisy) channel responses.

The forward model allowed us to separate responses into channels with different stimulus preferences, and to characterize interactions (e.g., divisive suppression) between neural subpopulations. Conventional fMRI provides only a single measure of activity, which is assumed to be the sum of all underlying responses. In contrast, the forward model separates the responses, allowing us to determine how differently tuned channels respond to various stimuli and how stimuli interact within these channels. We have previously used a very similar forward model to decode colors from the responses in several visual cortical areas (Brouwer and Heeger 2009), showing that this model provides a good approximation of hue responses in visual area V4. An analogous forward model (but with far more basis functions tuned for a variety of orientations, spatial frequencies, and spatial locations) has been used to characterize fMRI responses to natural images (Kay et al. 2008) and to identify a novel image based on the pattern of activity that it evoked. The forward model is related to, but should not be confused with, the classification methods currently prevailing in the neuroimaging literature (Brouwer and van Ee 2007; Dinstein et al. 2008; Haynes and Rees 2005b, 2006; Kamitani and Tong 2005). Both methods rely on the distributed patterns of responses across voxels and the assumption that these differ for different stimuli (e.g., colors, orientations, etc.). However, the forward model specifies a transformation from voxel responses to a much smaller number of channel responses, using a decomposition in terms of physiologically motivated channels. These channels are a mathematical formalization of the tuning curves of the underlying neural subpopulations (see MATERIALS AND METHODS).

The forward model in the present study was limited to our modality of interest: orientation. Trying to use this particular forward model to decode more complex stimuli, like natural images, would inevitably fail. The model would have to be extended to include multiple spatial locations (our model collapsed across space to extract the responses to the 6 orientation channels), spatial frequency bands (our model used only 1 spatial frequency channel because the stimulus components all had the same spatial frequency), etc., as was done by Kay et al. (2008). However, its succinct linear form, based on neurophysiological tuning curves, makes it a powerful model to test modality-specific interactions (specifically, cross-orientation suppression).

The forward model approach used in the present study has the potential for a wide range of applications. By assuming a basis set for neural tuning curves, it allows us to measure simultaneously the responses of each of several subpopulations of neurons (channels) in the human brain that span a particular stimulus or feature dimension (e.g., orientation, color, motion direction, spatial frequency, or even facial expressions). In addition, it allows us to examine the representation of a stimulus across subpopulations of selective neurons within a cortical area.

#### ACKNOWLEDGMENTS

We thank Matteo Carandini for comments on a previous draft of this manuscript.

#### GRANTS

This work was supported by National Eye Institute Grants R01-EY-016752 and R01-EY-019693 and by a grant from the US-Israel Binational Science Foundation (2007224).

#### DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

#### REFERENCES

- Brouwer GJ, Heeger DJ.** Decoding and reconstructing color from responses in human visual cortex. *J Neurosci* 29: 13992–14003, 2009.
- Brouwer GJ, van Ee R.** Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *J Neurosci* 27: 1015–1023, 2007.
- Busse L, Wade AR, Carandini M.** Representation of concurrent stimuli by population activity in visual cortex. *Neuron* 64: 931–942, 2009.
- Carandini M, Heeger DJ.** Summation and division by neurons in primate visual cortex. *Science* 264: 1333–1336, 1994.
- Carandini M, Heeger DJ, Movshon JA.** Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17: 8621–8644, 1997.
- Dale AM.** Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8: 109–114, 1999.
- DeAngelis GC, Robson JG, Ohzawa I, Freeman RD.** Organization of suppression in receptive fields of neurons in cat visual cortex. *J Neurophysiol* 68: 144–163, 1992.
- Dinstein I, Gardner JL, Jazayeri M, Heeger DJ.** Executed and observed movements have different distributed representations in human aIPS. *J Neurosci* 28: 11231–11239, 2008.
- Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN.** fMRI of human visual cortex. *Nature* 369: 525–527, 1994.
- Engel S, Zhang X, Wandell B.** Color tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* 388: 68–71, 1997.
- Foley JM.** Human luminance pattern-vision mechanisms—masking experiments require a new model. *J Opt Soc Am A* 11: 1710–1719, 1994.
- Freeman J, Brouwer GJ, Heeger DJ, Merriam EP.** Orientation decoding depends on maps, not columns. *J Neurosci* 31: 4792–4804, 2011.
- Freeman WT, Adelson EH.** The design and use of steerable filters. *IEEE Trans Pattern Anal Mach Intell* 13: 891–906, 1991.
- Gandhi SP, Heeger DJ, Boynton GM.** Spatial attention affects brain activity in human primary visual cortex. *Proc Natl Acad Sci USA* 96: 3314–3319, 1999.
- Geisler WS, Albrecht DG.** Cortical neurons: isolation of contrast gain control. *Vision Res* 32: 1409–1410, 1992.
- Geisler WS, Albrecht DG.** Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. *Vis Neurosci* 14: 897–919, 1997.
- Grill-Spector K, Malach R.** fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107: 293–321, 2001.
- Haynes JD, Rees G.** Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7: 523–534, 2006.

- Haynes JD, Rees G.** Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8: 686–691, 2005a.
- Haynes JD, Rees G.** Predicting the stream of consciousness from activity in human visual cortex. *Curr Biol* 15: 1301–1307, 2005b.
- Heeger DJ.** Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9: 181–197, 1992.
- Kamitani Y, Tong F.** Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8: 679–685, 2005.
- Kay KN, Naselaris T, Prenger RJ, Gallant JL.** Identifying natural images from human brain activity. *Nature* 452: 352–355, 2008.
- Larsson J, Heeger DJ.** Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26: 13128–13142, 2006.
- Louie K, Glimcher P.** Separating value from choice: delay discounting activity in the lateral intraparietal area. *J Neurosci* 30: 5498–5507, 2010.
- Moradi F, Heeger DJ.** Inter-ocular contrast normalization in human visual cortex. *J Vis* 9: 13 11–22, 2009.
- Morrone MC, Burr DC, Maffei L.** Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proc R Soc Lond B Biol Sci* 216: 335–354, 1982.
- Nestares O, Heeger DJ.** Robust multiresolution alignment of MRI brain volumes. *Magn Reson Med* 43: 705–715, 2000.
- Ogawa S, Lee TM, Kay AR, Tank DW.** Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA* 87: 9868–9872, 1990.
- Ohshiro T, Angelaki DE, DeAngelis GC.** A normalization model of multi-sensory integration. *Nat Neurosci* 14: 775–782, 2011.
- Olsen SR, Bhandawat V, Wilson RI.** Divisive normalization in olfactory population codes. *Neuron* 66: 287–299, 2010.
- Reynolds JH, Chelazzi L.** Attentional modulation of visual processing. *Annu Rev Neurosci* 27: 611–647, 2004.
- Reynolds JH, Heeger DJ.** The normalization model of attention. *Neuron* 61: 168–185, 2009.
- Rust NC, Mante V, Simoncelli EP, Movshon JA.** How MT cells analyze the motion of visual patterns. *Nat Neurosci* 9: 1421–1431, 2006.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB.** Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889–893, 1995.
- Simoncelli EP, Heeger DJ.** A model of neuronal responses in visual area MT. *Vision Res* 38: 743–761, 1998.
- Sirotin YB, Das A.** Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457: 475–479, 2009.
- Weiner KS, Sayres R, Vinberg J, Grill-Spector K.** fMRI-adaptation and category selectivity in human ventral temporal cortex: regional differences across time scales. *J Neurophysiol* 103: 3349–3365, 2010.
- Zoccolan D, Oertelt N, DiCarlo JJ, Cox DD.** A rodent model for the study of invariant visual object recognition. *Proc Natl Acad Sci USA* 106: 8748–8753, 2009.

