

Optimal Deployment of Attentional Gain during Fine Discriminations

Miranda Scolari,¹ Anna Byers,¹ and John T. Serences^{1,2}

¹Department of Psychology and ²Neuroscience Graduate Program, University of California, San Diego, La Jolla, California 92093

Most models assume that top-down attention enhances the gain of sensory neurons tuned to behaviorally relevant stimuli (on-target gain). However, theoretical work suggests that when targets and distracters are highly similar, attention should enhance the gain of neurons that are tuned away from the target, because these neurons better discriminate neighboring features (off-target gain). While it is established that off-target neurons support difficult fine discriminations, it is unclear if top-down attentional gain can be optimally applied to informative off-target sensory neurons or if gain is always applied to on-target neurons, regardless of task demands. To test the optimality of attentional gain in human visual cortex, we used functional magnetic resonance imaging and an encoding model to estimate the response profile across a set of hypothetical orientation-selective channels during a difficult discrimination task. The results suggest that top-down attention can adaptively modulate off-target neural populations, but only when the discriminanda are precisely specified in advance. Furthermore, logistic regression revealed that activation levels in off-target orientation channels predicted behavioral accuracy on a trial-by-trial basis. Overall, these data suggest that attention does not only increase the gain of sensory-evoked responses, but may bias population response profiles in an optimal manner that respects both the tuning properties of sensory neurons and the physical characteristics of the stimulus array.

Introduction

When making a difficult perceptual decision, selective attention ensures that relevant features are preferentially processed over distracters (Desimone and Duncan, 1995; Reynolds and Desimone, 1999). Most accounts, such as the feature similarity gain model, suggest that attention increases the firing rates of neurons tuned to relevant features and suppresses neurons tuned to irrelevant features (Treue and Martínez Trujillo, 1999; Martínez-Trujillo and Treue, 2004; Boynton, 2005; Maunsell and Treue, 2006; Lee and Maunsell, 2009; Reynolds and Heeger, 2009). Enhancing the gain of the neurons tuned to a relevant feature (on-target neurons) is optimal when discriminating highly dissimilar features (Fig. 1*a*). However, on-target enhancement is not always ideal; neurons tuned away from the target (off-target neurons) better signal differences between similar features (Fig. 1*b*) (Regan and Beverley, 1985; Seung and Sompolinsky, 1993; Itti et al., 2000; Hol and Treue, 2001; Pouget et al., 2001; Schoups et al., 2001; Purushothaman and Bradley, 2005; Jazayeri and Movshon, 2006, 2007; Navalpakkam and Itti, 2006, 2007; Raiguel et al., 2006; Moore, 2008; Scolari and Serences, 2009, 2010).

Recent psychophysical studies suggest that top-down attentional gain can be optimally deployed to early sensory neurons in accordance with perceptual demands (Navalpakkam and Itti, 2007; Scolari and Serences, 2009). However, these behavioral results are equally consistent with the notion that off-target sensory responses are selectively “readout” by downstream regions (Law and Gold, 2008, 2009), independent from any changes in sensory gain. Both possibilities can account for the behavioral data. One recent functional magnetic resonance imaging (fMRI) study demonstrated a predictive relationship between activation levels in off-target populations in V1 and behavioral performance during a difficult discrimination task (Scolari and Serences, 2010). However, because the data were postsorted based on behavioral accuracy and attention was not directly manipulated, the observed activation differences cannot be conclusively attributed to top-down attentional gain.

Here we first used fMRI and a forward encoding model (Kay et al., 2008; Brouwer and Heeger, 2009, 2011; Naselaris et al., 2011; Serences and Saproo, 2012) to determine whether top-down attentional gain is optimally directed to off-target neural populations during a challenging discrimination. Subjects judged whether two simultaneously presented oriented gratings matched, where a “mismatch” grating was rotated slightly away from a precued orientation; the rotational direction was randomly selected (Experiment 1), or precued (Experiment 2). In addition, we used logistic regression to determine whether behavioral performance was predicted by trial-by-trial changes in off-target responses.

When the rotational offset of the two stimuli was not precued (Experiment 1), top-down attentional gain was applied to on-target neural populations. However, when precise information

Received Nov. 4, 2011; revised March 9, 2012; accepted April 10, 2012.

Author contributions: M.S., A.B., and J.T.S. designed research; M.S. and J.T.S. performed research; M.S. and J.T.S. analyzed data; M.S. and J.T.S. wrote the paper.

This work was funded by NIMH R01-092345 to J.T.S. We thank Soren K. Andersen, John H. Reynolds, Gijs Brouwer, and Thomas Sprague for helpful discussions related to experimental design and data analysis methods.

Correspondence should be addressed to either Miranda Scolari or John T. Serences, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093. E-mail: mscolari@princeton.edu or jserences@ucsd.edu.

DOI:10.1523/JNEUROSCI.5558-11.2012

Copyright © 2012 the authors 0270-6474/12/327723-11\$15.00/0

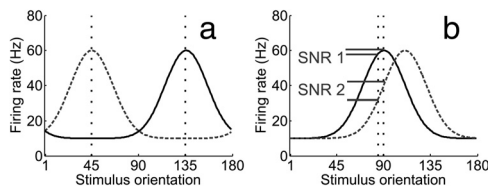


Figure 1. *a*, Target (135°) and distracter (45°) orientations are marked by vertical dashed lines. During a coarse discrimination with low target/distracter similarity, neurons tuned to the behaviorally relevant feature should be the target of attentional gain, given their large signal-to-noise ratio (SNR). However, (*b*) during a fine discrimination, off-target neurons that flank the behaviorally relevant feature theoretically carry more information, yielding a higher SNR. Target (90°) and distracter (92°) orientations are marked by vertical dashed lines. (Adapted with permission from Navalpakkam and Itti, 2007).

about the rotational offset was given (Experiment 2), off-target neural populations were selectively enhanced. In both cases, we observed a significant correlation between off-target activation and behavioral performance, suggesting that decision-making mechanisms rely primarily on input from the most informative populations of sensory neurons.

Materials and Methods

Experiment 1

Subjects. Ten neurologically healthy subjects (three males and seven females) with an age range of 22 to 48 years ($M = 28$, $SD = 7.36$) were recruited from the University of California, San Diego (UCSD) community (one additional subject enrolled in the study, but did not complete both scan sessions). Data from two subjects were discarded: one subject made eye movements to fixate each peripheral stimulus (which was discovered upon debriefing) and the other fell asleep in the scanner. All subjects gave written informed consent as per the Institutional Review Board requirements at UCSD. Subjects completed 1 h of training outside the scanner, followed by two 2 h scan sessions; compensation included \$10 per hour for training and \$20 per hour for scanning.

Experimental paradigm. All visual stimuli were generated using the MATLAB programming language (v.7.7; MathWorks) with the Psychophysics Toolbox (v.3; Brainard, 1997, Pelli, 1997) running on a PC laptop with the Windows XP operating system. The luminance output of both the computer monitors used during training in the lab and the projector used during scanning were measured with a Minolta LS-110 photometer and linearized in the stimulus presentation software. Following training (see below), subjects completed one of three attention tasks during each fMRI scan; all tasks were comprised of identical displays, and task type depended only on the attended aspect of the display (attend-orientation, attend-contrast, or attend central letter stream; see Fig. 2). Throughout each trial, observers maintained fixation on a central white ring (diameter: 1.16° of visual angle) at the center of the screen. An attention cue consisting of a white oriented line (length: 2.52° of visual angle), whose center was occluded by a gray disk demarcated by the ring, was rendered at an orientation that was selected from a set of 10 values spanning the 180° of orientation space in 18° steps (0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, and 162°). One second after cue onset, two spatially anti-aliased square wave gratings (0.74 cycles per degree) appeared simultaneously for 3 s, one centered in each upper quadrant (horizontally and vertically offset from fixation by 4.63 and 2.73° of visual angle, respectively, with the diameter of 6.79° for each grating). Each grating was flickered at a rate of 5 Hz for the duration of each 3 s trial, with the spatial phase of each grating randomly selected on each 100 ms presentation. The orientation of one stimulus (left or right grating) closely matched the orientation of the central attention cue, differing only by a small amount of jitter (a pseudorandomly selected value uniformly selected from a range spanning $\pm 3^\circ$). The orientation of the remaining stimulus was the same as the other stimulus, or was offset by an average of $\pm 7.01^\circ$ (SEM across subjects = 1.3°, range across subjects: 3–12°; henceforth termed the “deviant grating”). The subject’s task on attend-orientation blocks was to make a button-press response indicating if the two gratings were

rendered at the same orientation (match trials) or at different orientations (mismatch trials). The orientation offset of the gratings on mismatch trials was determined separately for each subject in the training session held outside of the scanner (see description below) and then titrated as needed on a scan-by-scan basis to maintain accuracy at $\sim 70\%$ in the two main peripheral attention conditions. For four of the subjects, the general orientation of the cue was selected pseudorandomly on a trial-by-trial basis from the set of 10 possible values. For the remaining four subjects, the same general orientation was presented for four consecutive trials (forming “mini-blocks”) in an effort to maximize orientation-selective responses. However, since this mini-block manipulation had no effect on the magnitude of the orientation-selective response or on the attentional modulations, we do not discuss it further. The contrast of the left or the right stimulus was always selected pseudorandomly from a uniform range extending from 65 to 75%, and the contrast of the remaining stimulus matched the first on half the trials or was offset on average by $\pm 6.8\%$ (SEM across subjects = 0.92%, range across subjects = 3–16%). On attend-contrast blocks, the subject’s task was to ignore differences in orientation and to report whether the contrast of two gratings matched or did not match. On all trials, a central rapid serial visual presentation (RSVP) stream was also presented inside the fixation ring at an average exposure duration of 67 ms per letter (letter height: 0.28° of visual angle); in the attend-letter stream condition, subjects reported on each trial whether an “X” or a “Y” was presented. The attention condition (attend-orientation, attend-contrast, or attend-letter) was blocked and the order of the blocks was counterbalanced across subjects, with all subjects completing an equal number of each block type.

Each scan contained 40 trials (with four trials containing gratings rendered at each of the 10 possible orientations), and every trial was separated by a 2.5 s response window in which only the central fixation ring was present on the screen. Eight blank null-trials, consisting only of the fixation ring for the length of a standard trial (6.5 s), were pseudorandomly intermixed with the 40 task trials with the constraint that a block never started with a null-trial. Subjects completed at least nine 321 s scans in each scanning session (and four subjects completed 12 scans in the first scan session), for a total of six or seven scans per attention condition across both sessions.

Prescanning behavioral training session. Subjects were trained on the task outside of the scanner at least one day before they were scanned. All subjects completed 12 blocks of training on the experimental task (four blocks of each attention condition). Trial-by-trial auditory error feedback was provided via a high- or low-frequency beep; however, auditory feedback was omitted during scan sessions, where feedback on behavioral performance was only given at the end of each scan. For the attend-orientation condition, the offset between two mismatched gratings was randomly selected from a set of four values (1, 4, 8, and 16°); similarly, in the attend-contrast condition, the percentage change in contrast between two mismatched gratings was randomly selected from a set of four values (1, 4, 8, and 16%). The orientation and contrast offsets at which subjects responded correctly on $\sim 70\%$ of trials were selected for subsequent scan sessions.

Independent functional localizer scans. Two 312 s functional localizer scans were run in each scan session to independently identify voxels within retinotopically organized visual areas that responded to the spatial positions occupied by the stimuli used in the attention scans. On half of the trials, a full contrast flickering checkerboard stimulus (4.31 cycles per degree, 5 Hz reversal frequency) was presented in the same central spatial position that was occupied by the fixation circle in the main attention scans for 10 s. On the remaining trials, two flickering checkerboard stimuli were presented in the spatial locations occupied by the peripheral gratings in the main attention scans (0.74 cycles per degree, 5 Hz reversal time). On all trials, regardless of the spatial position of the checkerboards, subjects were instructed to make a button-press response within 1 s when the contrast of the stimulus decreased slightly. Each contrast-change target was presented for two video frames (33.33 ms), and there were three targets presented on each trial. The timing of each target was pseudorandomly determined with the following constraints: (1) each target was separated from the previous one by at least 1 s and (2) targets were

restricted to a window of 1–9 s following the onset of the trial. Each of the 15 stimulus epochs was followed by a blank 10 s intertrial interval. A general linear model (GLM) was then used to identify voxels that responded more during epochs of peripheral stimulation compared with epochs of central stimulation (see below, fMRI data acquisition and analysis).

Retinotopic mapping procedure. Each subject participated in at least one retinotopic mapping scan, which required passive fixation of a full-contrast checkerboard stimulus and standard presentation parameters (stimulus flickering at 8 Hz and subtending 60° of polar angle; Engel et al., 1994; Sereno et al., 1995). This procedure was used to identify ventral visual areas V1, V2v, and V3v (given that the stimuli were centered in the upper quadrants, we did not map dorsal visual areas). We projected the retinotopic mapping data onto a computationally inflated representation of each subject's gray/white matter boundary to better visualize these cortical areas.

fMRI data acquisition and analysis. MRI scanning was performed on a 3 T GE MR750 scanner equipped with an eight-channel head coil at the Keck Center for Functional MRI, UCSD. Anatomical images were acquired using a T1-weighted sequence that yielded images with a 1 mm³ resolution (TR/TE = 11/3.3 ms, TI = 1100 ms, 172 slices, flip angle = 18°). Functional images were acquired using a gradient echo EPI pulse sequence which covered the occipital lobe with 26 transverse slices. Slices were acquired in ascending interleaved order with 3 mm thickness (TR = 1500 ms, TE = 30 ms, flip angle = 90°, image matrix = 64 (anteroposterior) × 64 (right lateral), with FOV = 192 mm (AP) × 192 mm (RL), 0 mm gap, voxel size = 3 mm × 3 mm × 3 mm). Data analysis was performed using BrainVoyager QX (v 1.86; Brain Innovation) and custom time series analysis routines written in MATLAB (version 7.1; MathWorks). EPI images were slice-time corrected, motion corrected (both within and between scans), and high-pass filtered (3 cycles/run) to remove low-frequency temporal components from the time series. The time series from each voxel in each observer was then z-transformed on a scan-by-scan basis to normalize the mean response intensity across time to zero. This normalization was done to correct for differences in mean signal intensity across voxels (e.g., differences related to the distance of a given voxel from the coil elements). In addition, we removed the mean activation level across all voxels on each trial to ensure that orientation selectivity resulted from the pattern of activation as opposed to a spatially global change in the blood oxygenation level-dependent (BOLD) response that was evoked by different orientations. We then added the mean BOLD response for a given attention condition to each respective orientation-tuning function (see below) to account for any differences in mean amplitude across the attention conditions. All reported results are qualitatively unchanged if this spatial normalization step is not performed.

Voxel selection using independent localizer scans. To identify voxels that responded to the retinotopic positions of the two peripheral stimulus apertures, data from the functional localizer scans were analyzed using a GLM that contained two regressors, each one marking the time course of one stimulus type (peripheral or central checkerboard), along with a third regressor to account for shifts in the overall magnitude of the signal in each voxel across each scan (the constant term in the GLM). This boxcar model of each stimulus type was then convolved with the standard difference-of-two gamma function model of the hemodynamic response function (HRF) that is implemented in Brain Voyager (time to peak of positive response, 5 s; time to maximum negative response, 15 s; ratio of positive and negative responses, 6; positive and negative response dispersion, 1). Voxels within each identified visual area (V1, V2v, and V3v) that responded more to epochs of peripheral stimulation, compared with central stimulation, were used to analyze data from the attention scans if they passed a threshold of $p < 0.01$, corrected for multiple comparisons using the false discovery rate algorithm implemented in Brain Voyager.

Estimating trial-by-trial BOLD responses on attention scans. To estimate the trial-by-trial magnitude of the BOLD response during the main attention scans, we shifted the time series from each voxel identified in the functional localizer scans by 6 s to account for the temporal lag in the HRF, and then extracted and averaged data from the two consecutive

1.5 s TRs that corresponded to the duration of each 3 s trial (Kamitani and Tong, 2005; Serences and Boynton, 2007a,b; Serences et al., 2009). These trial-by-trial estimates of the BOLD response amplitude were subsequently used as input to the forward encoding model described below to estimate orientation tuning curves in each visual area.

Estimating orientation-selective BOLD response profiles. We assumed that the BOLD response in a given voxel represents the pooled activity across a large population of orientation-selective neurons, and that the distribution of neural tuning preference is biased within a given voxel due to large-scale feature maps (Sasaki et al., 2006; Freeman et al., 2011; Jia et al., 2011) and/or to subvoxel anisotropies in the number of orientation-selective cortical columns with a specific preference (Kamitani and Tong, 2005; Swisher et al., 2010). Thus, the BOLD response measured from many of the voxels in primary visual cortex exhibit a modest but robust orientation preference (Haynes and Rees, 2005; Kamitani and Tong, 2005; Serences et al., 2009; Freeman et al., 2011). Given this response bias, we inferred changes in the allocation of attentional gain by modeling the response of each voxel as a linearly weighted combination of a set of half-sinusoidal basis functions (Brouwer and Heeger, 2009, 2011). We used half-sinusoidal functions that were raised to the sixth power to approximate the shape of single-unit tuning functions in V1, where the $1/\sqrt{2}$ half-bandwidth of orientation-tuned cells is $\sim 20^\circ$ (although there is a considerable amount of variability in bandwidth; Schiller et al., 1976; Ringach et al., 2002a,b; Gur et al., 2005). Given that the half-sinusoids were raised to the sixth power, a minimum of seven linearly independent basis functions was required to adequately cover orientation space (Freeman and Adelson, 1991); however, since we presented 10 unique orientations in the experiment, we used a set of 10 evenly distributed functions (see Fig. 3). Given the width of the selected functions, the use of more than the required seven basis functions is not problematic as long as the number of functions does not exceed the number of measured stimulus values. While we selected the bandwidth of the basis functions based on physiology studies, all results that we report are robust to reasonable variations in this value (i.e., raising the half-sinusoids to the fifth-eighth power, all of which are reasonable choices based on the documented variability of single-unit bandwidths).

We then used this basis set to estimate orientation-selective responses in visual cortex using a forward model developed by Brouwer and Heeger (2009, 2011). (Note that we adopt their terminology and variable naming convention for consistency.) Let m be the number of voxels in a given visual area, n be the number of trials, and k be the number of hypothetical orientation channels (10 in this case). These data were first segmented into a training dataset (an $m \times n$ matrix B_1 consisting of data from all but one scan of each of the three task types) and a test dataset (B_2 , consisting of the remaining data from one scan of each of the three task types). The training data in B_1 were then mapped onto the full rank matrix of channel outputs (C_1 , $k \times n$) by the weight matrix (W , $m \times k$) that was estimated using a GLM of the following form:

$$B_1 = WC_1, \quad (1)$$

where the ordinary least-squares estimate of W is computed as follows:

$$\widehat{W} = B_1 C_1^T (C_1 C_1^T)^{-1}. \quad (2)$$

The channel responses (C_2) on each trial were then estimated for the test data (B_2) using the weights estimated in Equation 2 as follows:

$$\widehat{C}_2 = (\widehat{W}^T \widehat{W})^{-1} \widehat{W}^T B_2. \quad (3)$$

Next, the 10-point channel response function estimated for each trial (in C_2) was circularly shifted so that the channel matching the orientation of the attention cue presented on that trial was positioned in the center of the tuning curve, thereby aligning each channel response profile to a common center. The channel response estimation procedure was then iterated across all unique combinations of holding out one scan of each attention condition for use as a test set (i.e., on each iteration we held a total of three scans out from the attend-letter, attend-orientation, and attend-contrast conditions in Experiment 1, and two scans out from the attend-orientation and attend-contrast conditions in Experiment 2; see

below, Experiment 2). This ensured that the training set had a balanced number of each trial type. Data were then averaged across all trials from a given condition to generate the orientation tuning curves (see Figs. 5, 7, 8; note that even though the magnitude of the channel responses is influenced by the amplitude of the basis functions, which was set to 1 in this study, the relationship between responses is maintained as long as all channels have the same amplitude; therefore, the y -axes of all data plots read “BOLD response: relative magnitude”).

Relating behavioral responses to orientation-selective BOLD modulations. To determine whether the trial-by-trial fluctuations in response magnitude in each orientation channel predicted the behavioral performance of the observer, we used logistic regression to map the continuous BOLD response in each orientation channel onto the binary accuracy value on mismatch trials (i.e., correct/incorrect). The bilateral stimulus presentation format ensured that the left and right stimuli projected into opposite cortical hemispheres, at least in the early stages of occipital cortex that we measured in the present study (i.e., V1–V3v). To map these spatially lateralized responses onto accuracy, we reasoned that accuracy on mismatch trials should be predicted by the extent to which responses in left and right visual cortex differed. Thus, for each channel, we computed the difference between the responses in contralateral and ipsilateral visual areas with respect to the deviant grating on a trial-by-trial basis. These difference scores were then entered into a logistic regression analysis on a channel-by-channel basis to map them onto the probability of a subject responding correctly on mismatch trials (see Figs. 6, 9).

Assessing statistical significance using a randomization procedure. Due to the nonindependence of the channels in both the main data analyses and the logistic regression (because the half-sinusoidal basis functions overlapped, see Fig. 3), the assumption of independence in the standard GLM/ANOVA framework was violated. Therefore, we used a randomization procedure to assess the significance of the attentional modulations (reported in Figs. 5, 7, 8) and the significance of the logistic fit coefficients (reported in Figs. 6, 9). To evaluate the significance of the reported effects, we first ran a standard t test or ANOVA (as appropriate) to obtain an observed p value. Next, we randomly shuffled the data associated with the channel response profile on each trial. This relabeling procedure thus generated a new dataset that was the same size as the original dataset. We then ran the same statistical test (t test or ANOVA) on the new dataset, and repeated this procedure 10,000 times to generate a distribution of p values given randomized data. Finally, the probability of obtaining the observed p value (or lower) was computed against the distribution of p values generated from the randomized datasets. All reported p values reflect this probability.

Experiment 2

Subjects. Twelve subjects (six males and six females) with an age range of 20 to 34 ($M = 25.83$, $SD = 4.69$) years were recruited to participate in a single fMRI session. Data from one subject were discarded due to excessive motion, so 11 subjects were included in the final analyses. Three of the subjects had previously participated in Experiment 1, and two others had participated in a different study that used similar stimuli, so they were experienced fMRI subjects and experienced with this stimulus paradigm. Before the scan session, each of the five experienced subjects participated in a 1 h training session consisting of eight blocks of the task (see below) to acquaint them with the attentional cueing procedure; the remaining six inexperienced subjects participated in two 1 h training sessions held on separate days. Therefore, more extensive training was used in Experiment 2 compared with Experiment 1. Following training, all subjects participated in a single 2 h scan session that was held on a separate day for most subjects (for two experienced subjects, training and scanning occurred on the same day). All subjects gave written informed consent as per the Institutional Review Board requirements at UCSD. Compensation for participating was \$10 per hour for the training session(s) and \$20 per hour for the scanning session.

Attention scans. The parameters of the attention scans were nearly identical to those of Experiment 1, except here the letter stream task was omitted and the color of the cue (green or red) informed the subjects of the rotational offset of the mismatched grating in the event of a mismatch

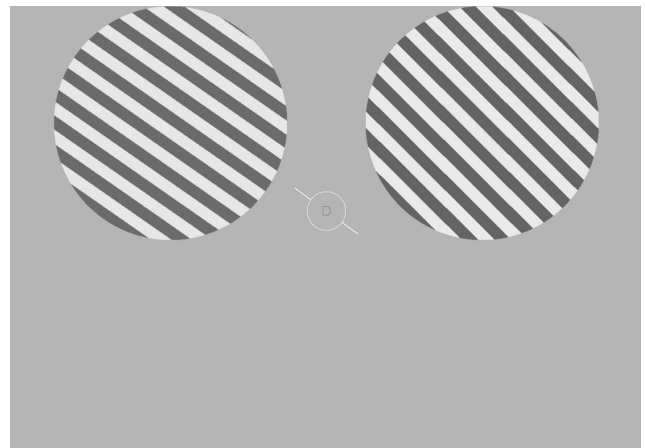


Figure 2. Schematic of the general experimental paradigm. Experiment 1, While maintaining fixation on the central ring, subjects attended to the orientation of the gratings, the contrast of the gratings, or a central letter RSVP stream. The orientation of one grating always closely matched the oriented cue line presented at fixation (randomly selected from a set of 10 orientations equally spaced across 180°); the orientation of the remaining grating matched the other or was mismatched by a small CW or a CCW offset. Similarly, the contrast of one grating ranged from 65 to 75%, and the second grating matched the first or mismatched by adding or subtracting a small contrast change. The subject's task was to decide if the stimuli matched or mismatched with respect to the relevant attended feature. Experiment 2, The display was identical to Experiment 1 (shown), with the following changes: (1) the central RSVP stream was removed, so subjects only attended to orientation or contrast on alternating blocks of trials and (2) the central cue was rendered in green or red to indicate with 100% validity either a CW or CCW rotational offset in the event of an orientation-mismatch trial.

trial on attend-orientation scans. For instance, a green cue indicated that a mismatch grating would be rotated clockwise (CW) from the cue, whereas a red cue indicated a counterclockwise (CCW) rotation from the cue; the direction of rotation was counterbalanced on a block-by-block basis. The same orientation was cued for four consecutive trials, and although the green and red color cues were also presented on attend-contrast blocks, they provided no relevant information about the task. The order of the attention blocks (attend-orientation and attend-contrast) was counterbalanced across subjects. The mean orientation offset on mismatch trials was 5.71° (SEM across subjects, 0.54° ; range, 3.5 – 10°) and the mean contrast offset on mismatch trials was 5.87% (SEM across subjects, 0.47% ; range, 2.5 – 8%).

The stimuli in Experiment 2 were presented on a different screen during scan sessions, such that all visual angles were 72.4% the size of the stimuli used in Experiment 1; the remaining data analysis methods for this experiment were otherwise identical. See Experiment 1 for a description of the functional localizer scans, retinotopic mapping procedures, fMRI data acquisitions and analysis, and the derivation of orientation-selective BOLD response profiles.

Results

Experiment 1

Behavioral task

The behavioral task is shown in Figure 2. Importantly, the stimulus display did not vary as a function of attention instructions: all blocks contained letter targets, orientation matches/mismatches, and contrast matches/mismatches, while only the locus of attention changed. Task difficulty, as operationalized by behavioral response accuracy, was closely matched in the attend-orientation and attend-contrast conditions, but was slightly higher in the attend-letter condition (orientation: 71.16% , $SEM = 0.99\%$; contrast: $M = 69.77\%$, $SEM = 3.6\%$; letter: $M = 78.92\%$, $SEM = 3.6\%$).

Feature-selective tuning curves and attentional gain

First, voxels in V1–V3v that responded to the spatial position occupied by the peripheral attention stimuli were identified using

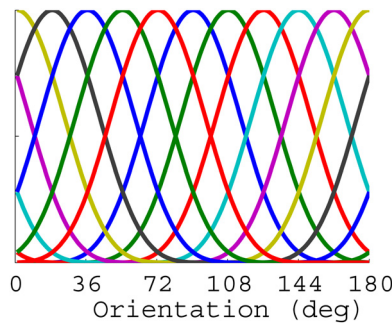


Figure 3. Depiction of the 10 half-cycle sinusoid basis functions evenly distributed across 180° of orientation space.

independent functional localizer scans (see Materials and Methods). Next, we used a forward encoding model to estimate the response profile across 10 hypothetical orientation channels that were equally distributed across 180° (see Materials and Methods, Estimating orientation-selective BOLD response profiles; Fig. 3). By examining the shape of these orientation-selective BOLD response profiles in each attention condition, we indirectly estimated the degree to which top-down attentional gain is applied to underlying neural populations when subjects engage in a fine orientation discrimination task. Importantly, any orientation-selective modulations that we observe using BOLD imaging cannot be directly attributed to changes in neural spiking activity, as the BOLD signal can be modulated by multiple sources such as synaptic input from local and distant sources (Heeger et al., 2000; Logothetis et al., 2001; Heeger and Ress, 2002; Logothetis and Wandell, 2004; Logothetis, 2008; Sirotin and Das, 2009; Kleinschmidt and Müller, 2010; Das and Sirotin, 2011; Handwerker and Bandettini, 2011a,b). However, given that neurons in early sensory areas like V1 are heavily interconnected (Douglas and Martin, 2007), changes in the BOLD signal related to synaptic activity are likely to be tightly correlated with changes in local spiking activity.

Based on optimal models of attentional gain (see Fig. 1*b*), we predicted that the BOLD tuning profiles would follow one of two patterns: Figure 4*a* provides an illustration of BOLD activation changes across orientation space as predicted by on-target gain, where the cued orientation is indicated by 0° on the *x*-axis. Figure 4*b* shows the predicted pattern if off-target gain is applied simultaneously on either side of the cued orientation.

Figure 5 shows the orientation tuning curves from V1, V2v, and V3v when subjects were performing the attend-letter task, the attend-contrast task (control conditions), and when they were performing the main attend-orientation task. The full tuning curves are shown in Figure 5 (positive values indicate CCW rotations and negative values indicate CW rotations). As expected, the magnitude of the orientation-selective tuning curve was smaller when subjects were attending the central RSVP stream compared with when they were attending to contrast or orientation (main effect of attention condition, $p < 0.005$ for both comparisons in V1, V2v, and in V3v). (Note that all statistics were computed using the randomization procedure described in Materials and Methods.) The attend-orientation and attend-contrast conditions evoked a similar overall response amplitude; however, there was a crossover interaction in all tested visual areas such that attending to orientation led to larger responses in channels tuned to the cued orientation and attenuated responses in channels tuned away from the cue, relative to the attend-contrast condition (Fig. 5, $p < 0.01$ in all areas). These results are in close accord

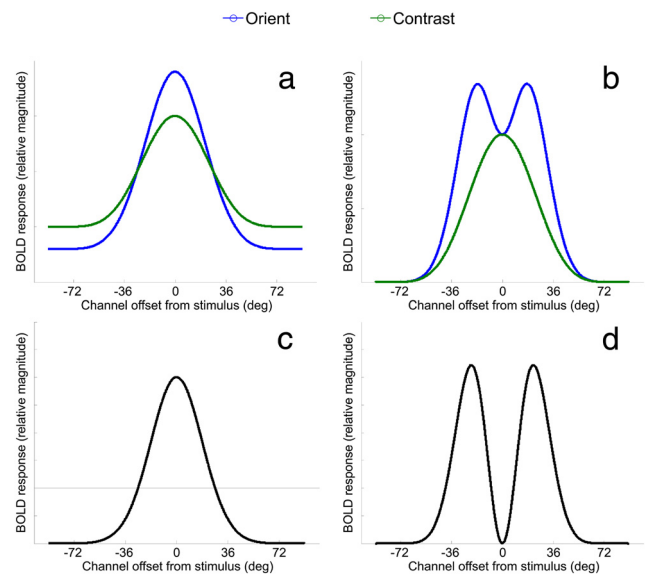


Figure 4. Model predictions of BOLD responses for orientation (in blue) and contrast (in green) conditions based on (*a*) on-target gain models of selective attention and (*b*) the optimal gain hypothesis with bilateral gain. *c, d*, Difference curves between the model predictions for attend-orientation and attend-contrast tasks in *a* and *b*, respectively. Note that the model predictions show the greatest divergence at the cued orientation (0°) where gain is maximized (*c*) or minimized (*d*).

with previous cases documenting on-target attentional gain (as shown in Fig. 4*a*; Treue and Martínez Trujillo, 1999; Martínez-Trujillo and Treue, 2004; Maunsell and Treue, 2006), and are inconsistent with the deployment of off-target attentional gain when performing a fine discrimination.

While the results clearly favor on-target attentional gain, it is still possible that the decision mechanisms selectively readout sensory signals from neuronal populations that carry the most information about a relevant discrimination (Palmer et al., 2000; Eckstein et al., 2002; Gold and Shadlen, 2007; Law and Gold, 2008, 2009; Pestilli et al., 2011). The extent to which such a relationship can be observed is important for two reasons: (1) to establish that fMRI is sensitive enough to measure behaviorally meaningful changes in off-target channels and (2) to establish that our fine discrimination task did induce a reliance on off-channel responses to support behavior. First, we computed the differential response between corresponding channels in areas of visual cortex that were contralateral and ipsilateral to the deviant stimulus on mismatch trials. Next, we used logistic regression to estimate how the differential responses measured on each trial related to behavioral accuracy (see Materials and Methods, Relating behavioral responses to orientation-selective BOLD modulations). The results are shown in Figure 6, where positive values along the *x*-axis refer to offsets in the same rotational direction as the deviant grating, and negative values refer to offsets in the opposite direction. Trial-by-trial fluctuations in V1 channels predicted accuracy on mismatch trials when subjects were performing the orientation discrimination task, but this effect was more prominent during the second scan session (Fig. 6*A, B*; one-way ANOVA across orientation offsets for both scan sessions combined $p = 0.0824$; one-way ANOVA across orientation offsets for second scan session $p = 0.0055$; *t* tests at -18° and $+54^\circ$ on second session $p = 0.031$ and $p = 0.029$, respectively). No predictive relationships were observed in V2v or V3v (one-way ANOVA across orientation offsets for V2v and V3v for both sessions combined $p = 0.5701$ and $p = 0.4482$, respectively, and for

the second scan session alone $p = 0.7284$ and $p = 0.9923$, respectively). This is consistent with the notion that a relatively large differential response in contralateral off-target channels is interpreted as evidence in favor of a correct orientation judgment on mismatch trials, whereas a relatively similar response is interpreted as evidence in favor of an incorrect “match” judgment. Importantly, this comparison was not biased by any physical dissimilarity of the stimuli since only mismatch trials were used (and therefore physical characteristics of the stimuli are held constant across correct and incorrect trials). In contrast, no predictive relationship was observed between channel responses and accuracy on mismatch trials when subjects were performing the contrast discrimination task (one-way ANOVA for both sessions combined $p = 0.9994$ and second session alone $p = 0.1632$; Fig. 6C,D); however, the interaction between attention condition and channel offset did not reach significance (one-way ANOVA on differences between conditions for both scan sessions combined $p = 0.6915$ and second scan session alone $p = 0.097$). Nonetheless, the overall results suggest that even if attentional gain is applied to on-channel neural populations when subjects are attending to the orientation of the gratings (Fig. 5), the success of perceptual decisions appears to be most strongly influenced by the relative activation of channels tuned beyond the target orientation.

Experiment 2

Although we did not observe any evidence for off-target attentional gain in Experiment 1 (Fig. 5), at least two potential issues need to be addressed. First, feature-based attention may be restricted to a single gain field centered at one position in feature space. Given that the direction of the rotational offset on mismatch trials was unknown in advance in Experiment 1, subjects may have thus adopted one of two possible strategies: (1) to focus that single gain field on the orientation indicated by the central cue, as this would slightly elevate the response of informative off-target neural populations relative to a case in which no gain is applied at all or (2) probabilistically deploy attention to one side of the cue or the other, in an attempt to guess the upcoming rotational offset. Finally, it is also possible that if gain was applied to neurons tuned to orientations closely flanking the relevant stimulus (such as the 18° channel), this gain could be effectively smoothed into the estimate of the 0° channel response during the estimation of the forward model. Some smoothing is expected to occur because the basis set of half-cycle sinusoids (Fig. 3) was not orthogonal. Thus, to the extent that there is overlap between adjacent basis functions, the estimated channel responses will not be independent and larger responses in channels tuned $\pm 18^\circ$ from the cue might artificially elevate the estimated response amplitude in the 0° channel. Any of these possibilities may have decreased the probability of observing off-target gain in Experiment 1. We therefore performed a second experiment in which the precue-informed subjects about the rotational offset of the deviant orientation stimulus in the event of a mismatch trial. This manipulation should make it possible to deploy gain to off-target neurons tuned to one side of the cued orientation, and eliminate the symmetry of any smoothing that may occur between adjacent orientation channels, thereby preserving a bias in the shape of the tuning function induced by attention, should one exist.

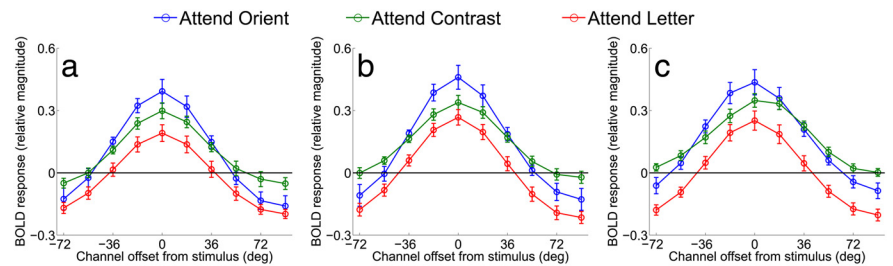


Figure 5. Orientation-selective tuning curves from V1, V2v, and V3v in Experiment 1. *a*, Relative BOLD responses across orientation-selective channels in V1 for each of the three attention conditions averaged across subjects. Orientation tuning curves from each trial were circularly shifted so that 0° indicates the cued orientation by convention, and the x -axis indicates the orientation preference of each channel with respect to the cue (negative values indicate a preferred orientation that is rotated CW from the cue, and positive values indicate a preferred orientation rotated CCW from the cue). *b*, Same as *a*, except that data are from V2v. *c*, Same as *a* and *b*, but data are from V3v. Error bars indicate ± 1 SEM across subjects.

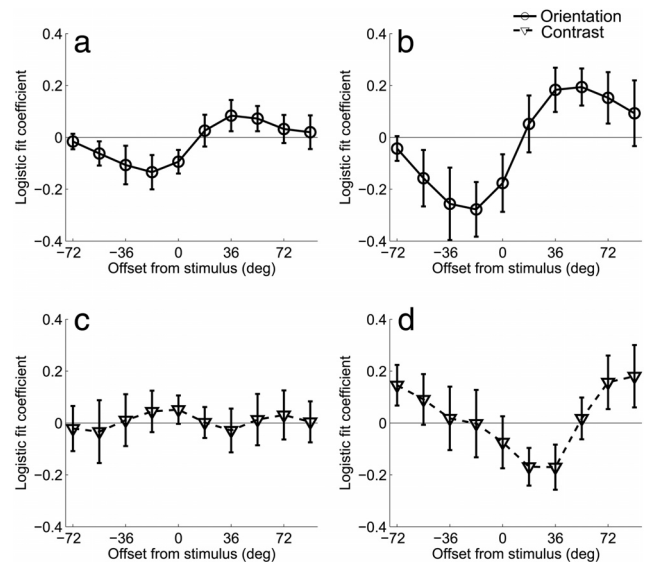


Figure 6. Results of the logistic regression analysis for Experiment 1 that relates behavioral accuracy to the size of the differential BOLD response extracted from regions of V1 that were contralateral and ipsilateral to the deviant grating. Each channel response difference (contralateral–ipsilateral) was entered into a logistic regression as a single predictor; the logistic fit coefficients of each of the 10 analyses are plotted together (*a*) for the attend-orientation task across both sessions and (*b*) day 2 alone. Note that here, positive values refer to offsets in the direction of the deviant grating, and negative values refer to offsets in the opposite direction. On day 2, trial-by-trial fluctuations in orientation channels tuned to -18° and $+54^\circ$ predict subjects’ accuracy on attend-orientation mismatch trials. This pattern was not observed for the attend-contrast task on (*c*) both sessions combined or (*d*) day 2 alone. Error bars indicate ± 1 SEM across subjects.

Behavioral task

In Experiment 2, we used a modified version of Experiment 1 in which we precued the rotational offset (CW or CCW) of the deviant orientation stimulus on mismatch trials with 100% validity. Task difficulty was closely matched across the two conditions, as indicated by behavioral response accuracy: 72.85% (SEM = 2.4%) in the attend-orientation task, and 69.22% (SEM = 1.05%) in the attend-contrast task.

Feature-selective tuning curves and attentional gain

Figure 7*a* shows the orientation tuning curves from V1 when subjects were performing the main orientation discrimination task or the contrast discrimination task. Data were shifted so that the 0° channel indicates the cued orientation and positive values on the x -axis indicate responses in channels that were offset in the cued direction, while negative values indicate responses in chan-

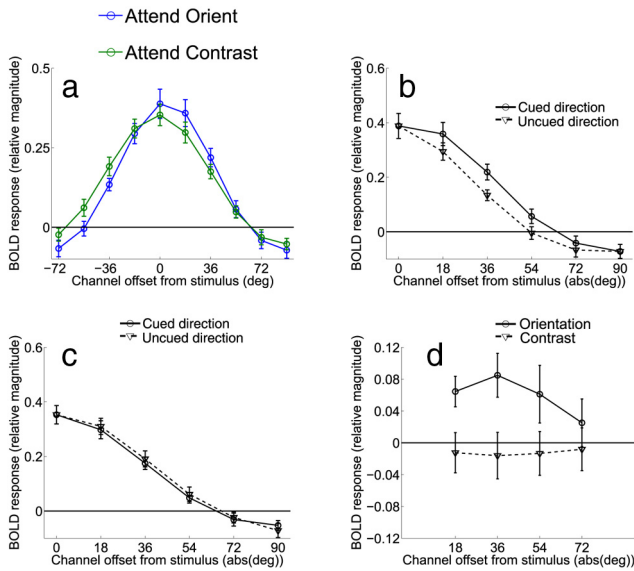


Figure 7. Channel tuning functions from V1 in Experiment 2, averaged across all eight subjects. *a*, Data from the orientation discrimination condition were shifted so that 0° indicates the cued orientation, positive values indicate channels offset in the cued direction, and negative values indicate channels offset in the uncued direction. *b*, Comparison of responses in channels offset in the cued direction and the uncued direction when subjects were attending to orientation. *c*, Comparison of responses in channels offset in the cued direction and the uncued direction when subjects were attending to contrast. *d*, Difference between corresponding channels tuned 18 to 72° in the cued and uncued directions for both the attend-orientation task (solid line) and attend-contrast task (dashed line). Error bars indicate ± 1 SEM across subjects.

nels offset in the uncued direction. Responses in the two attention conditions were similar in overall amplitude. However, there was an interaction such that channel responses were systematically shifted in the cued direction in the attend-orientation condition compared with the attend-contrast condition (attention condition by orientation channel interaction, $p < 0.005$). To highlight this shift in the tuning profiles, Figure 7*b* shows the magnitude of responses on attend-orientation trials in the cued direction relative to each corresponding offset in the uncued direction, and Figure 7*c* shows the corresponding comparison on attend-contrast trials. On attend-orientation trials, responses in channels tuned 18 and 36° in the cued direction were larger than the corresponding channels tuned in the uncued direction (Fig. 7*b*, $p < 0.01$ and $p < 0.025$, respectively). No corresponding differences were found between responses on attend-contrast trials (Fig. 7*c*). A direct comparison of the difference between responses in the cued and uncued directions revealed a larger modulation on attend-orientation trials compared with attend-contrast trials in the 18 and 36° channels (Fig. 7*d*, $p < 0.025$ for both comparisons). Note that this asymmetry cannot be explained by the fact that stimuli on mismatch trials were physically rotated in the cued direction, as the same displays were used in both attend-orientation and attend-contrast conditions, thereby controlling for any sensory-related influences on the channel tuning curves. This asymmetric shift in the response profile was also evident on attend-orientation trials in V2v (Fig. 8*a,c*), as responses in channels tuned 18, 36, and 54° in the cued direction were significantly greater than responses in corresponding channels in the uncued direction on attend-orientation trials (Fig. 8*c*, $p < 0.025$, $p < 0.01$, $p < 0.05$, respectively, see solid line). However, a direct comparison between this response bias on attend-orientation trials and attend-contrast trials did not reach significance in any channel (Fig. 8*c*, all $p > 0.16$, compare solid

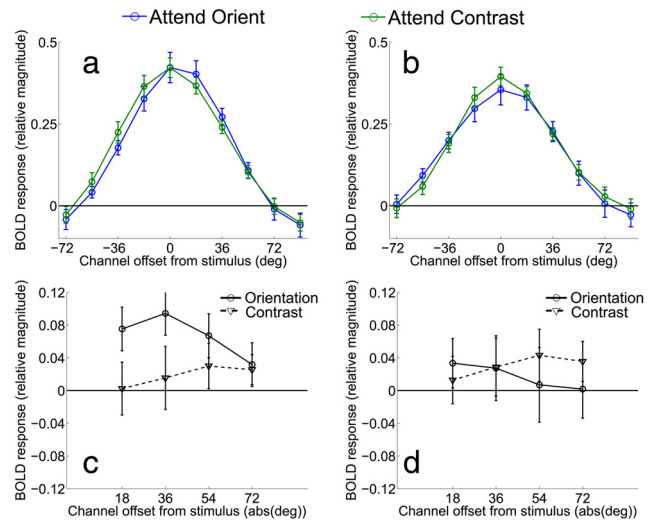


Figure 8. *a, b*, Channel tuning functions from V2v and V3v in Experiment 2, respectively. As in Figure 7*a*, data have been circularly shifted so that 0° indicates the cued orientation, positive values indicate channels offset in the cued direction, and negative values indicate channels offset in the uncued direction. *c*, Difference between corresponding channels tuned 18 to 72° in the cued and uncued directions for both the attend-orientation task (solid line) and attend-contrast task (dashed line) in area V2v. *d*, Difference between corresponding channels tuned 18 to 72° in the cued and uncued directions for both the attend-orientation task (solid line) and attend-contrast task (dashed line) in area V3v. Error bars indicate ± 1 SEM across subjects.

and dashed lines). Finally, no evidence for an attention-related asymmetry was present in V3v (Fig. 8*b,d*).

While these data are consistent with off-target deployment of attentional gain in V1, an alternative explanation is that subjects were actually applying on-target gain to the expected deviant orientation. Because the cue indicated the direction of the rotational offset of mismatched stimuli, treating the cued orientation as the relevant target could potentially contribute to the biased gain pattern that we observed in V1 and in V2v. However, we view this explanation as unlikely given that the average orientation offset on mismatch trials was only 5.71°. Thus, applying on-target gain at this small offset would not be expected to have a large influence on the channel tuned 18° in the cued direction, and certainly would have even less of an impact at 36°, where the observed effect was just as large in both V1 and V2v (i.e., statistically there was no difference between responses in the 18 and 36° channels, although the numerical value at 36° was slightly higher in both areas).

We further explored this possibility by repeating our analysis after recentering the channel responses such that 0° now corresponds to the expected orientation of the deviant stimulus in the event of a mismatch trial (i.e., 0° now indicates the orientation of the expected deviant stimulus). Because the size of the offset was titrated to keep behavioral performance at the desired level, this rotation of the channels was tailored to the offset for each subject on a scan-by-scan basis. If on-target gain was applied to the cued deviant offset in the attend-orientation condition, then we would expect an equivalent response in channels tuned to either side of 0° since there should no longer be any bias in the allocation of attentional gain in the cued direction. Contrary to this prediction, activation in V1 was still greater in the channels in the cued compared with the uncued directions at 18, 36, and 54° (all $p < 0.026$). In addition, these differences were significantly greater in the attend-orientation condition compared with the attend-contrast condition ($p = 0.034$, $p = 0.024$, for the 18° and 36° channels, respectively, while $p = 0.09$ in the 54° channel). These

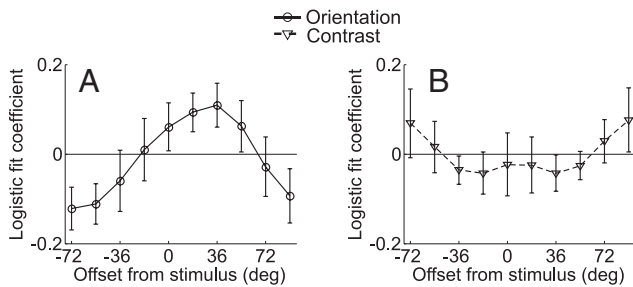


Figure 9. Results of the logistic regression analysis for Experiment 2 that relates behavioral accuracy to the size of the differential BOLD response extracted from regions of V1 and V2v that were contralateral and ipsilateral to the deviant grating (similar to Fig. 6). Note that here, positive values refer to offsets in the cued direction and negative values refer to offsets in the uncued direction. **A**, Increased differential activation in channels tuned to -72° and -54° predicted a lower probability of a correct response, and increased differential activation in the channel tuned to $+36^\circ$ predicted a higher probability of a correct response. **B**, No such effect was observed in the attend-contrast task. Error bars indicate ± 1 SEM across subjects.

results are consistent with off-target gain being directed to channels tuned beyond the deviant orientation.

As in Experiment 1, we assessed the predictive relationship between behavioral accuracy and differential activation levels in orientation channels measured from visual areas contralateral and ipsilateral to the deviant grating on mismatch trials. Here, we averaged corresponding channel responses across V1 and V2v, as both areas exhibited a shifted response function in the attend-orientation condition (Figs. 7, 8 and the logistic regression revealed a qualitatively similar pattern in both areas). The pattern was similar to that observed in Experiment 1: greater activation in the contralateral channels tuned $+36^\circ$ in the cued direction predicted correct responses on mismatch trials when subjects were attending to orientation (Fig. 9; one-way ANOVA across channels, $p = 0.024$; t test at $+36^\circ$, $p = 0.04$, marginally significant effect in the $+18^\circ$ channel, $p = 0.067$). Conversely, greater activation in the contralateral channels tuned to -54° and -72° predicted a higher probability of incorrect responses (t test at -54° , $p = 0.036$, t test at -72° , $p = 0.033$). No predictive relationship was observed between channel responses and accuracy in the attend-contrast condition (one-way ANOVA across channels, $p = 0.81$). Finally, there was a significant two-way interaction between orientation channel and attention condition (Fig. 9, compare $p = 0.031$). No significant effects were observed in V3v.

Discussion

In Experiment 1, we observed on-target gain in the feature-selective BOLD response profiles even though channels tuned to flanking orientations are theoretically more informative in the context of a fine discrimination task. Nonetheless, the logistic regression analysis shows that activation levels in off-target channels predicted subjects' accuracy on mismatch trials (Fig. 6). This suggests that even in the absence of off-target gain, difficult perceptual discriminations may be facilitated by optimizing the readout of sensory information during decision making (Palmer et al., 2000; Eckstein et al., 2002; Gold and Shadlen, 2007; Law and Gold, 2008, 2009; Pestilli et al., 2011). For example, simulations of fine motion discriminations suggest that MT-like neurons tuned away from a motion target ($\sim \pm 40^\circ$ relative to the target direction in this case) should be weighted most strongly by parietal sensorimotor neurons that are thought to integrate sensory evidence over time (Jazayeri and Movshon, 2006; Law and Gold, 2008, 2009). However, the fact that a predictive relationship be-

tween off-target activation and accuracy is not robust until the second scan session suggests that some training may be necessary for optimal readout to emerge (a point that we revisit below).

Given that the rotational offset between the two gratings was unknown in advance on mismatch trials in Experiment 1 (i.e., CW or CCW), we hypothesized two possible factors that may have reduced the likelihood of observing off-target gain: (1) subjects may have only been capable of applying a single gain field, and they centered this single gain field on the cued channel or (2) small amounts of gain in channels tuned $\pm 18^\circ$ from the target may have been smoothed into the estimate of the 0° channel, obscuring any off-target gain. In Experiment 2, these issues were addressed by giving precise advance information about the rotational offset of the deviant grating on mismatch trials. Under these conditions, we observed robust off-target gain, suggesting that feature-based attention can modulate sensory responses in a more complex and optimal manner than previously recognized. Similar off-target modulations have been observed in the context of perceptual learning studies that used a single well trained stimulus discrimination (Schoups et al., 2001; Raiguel et al., 2006); however, the present results suggest that top-down attentional gain can be deployed to informative neural populations in a more flexible manner across a range of possible stimulus values (i.e., the 10 orientations used in the present study as opposed to a single well trained orientation). Furthermore, a logistic regression analysis indicated that the relative amount of off-target activation in the cued direction predicted behavioral accuracy on mismatch trials, just as in Experiment 1. Note that the subjects in Experiment 2 were more experienced with the task than those in Experiment 1 (given either recent exposure to fine orientation discrimination tasks in other fMRI experiments, or an additional training session outside of the scanner; see Materials and Methods, Experiment 2, Subjects), which may account for the emergence of this predictive relationship in a single scan session. Finally, the data collectively suggest that decisions can be based on informative off-target neural populations, either in the absence of corresponding off-target attentional gain (Experiment 1) or in the presence of off-target gain (Experiment 2). This dissociation is consistent with the possibility that optimal sensory enhancement and optimal readout of sensory information can make complementary, and perhaps independent, contributions to the efficiency of perceptual decisions.

An important limitation of the present study is that the precise offset of the most informative sensory neurons cannot be directly inferred based on the point of maximum off-target gain observed in the BOLD signal. This problem arises for two reasons. First, even though changes in the BOLD signal are generally coupled with multiunit neural activity, especially in areas like V1 where most connections are local (Douglas and Martin, 2007), the BOLD signal is also influenced by other factors such as synaptic input from distal sources and by the response properties of astrocytes (Heeger et al., 2000; Logothetis et al., 2001; Logothetis and Wandell, 2004; Logothetis, 2008; Schummers et al., 2008). Second, the encoding model that we used to estimate the response in each orientation channel is a coarse proxy for the activity level of underlying sensory neurons. The loss of resolution arises because each voxel contains a mixture of multiple neural populations that are tuned to different features, so the feature-based tuning curves we derive will always be less precise than their single-unit counterparts. Despite this inability to pinpoint the *exact* orientation offset at which attentional gain is being applied, the observation of higher activation levels in channels tuned away from the

relevant feature (Fig. 7) is consistent with the existence of off-target gain.

The observation that channel responses rotated unilaterally beyond the target predicts behavioral accuracy (Experiments 1 and 2) and show enhanced gain (Experiment 2) are seemingly inconsistent with the observation of bilateral modulations in two previous studies. In one, Scolari and Serences (2009) reported psychophysical evidence suggesting that gain is applied bilaterally to neurons tuned to features on either side of the target. We reasoned that this was an adaptive pattern given that both sets of neurons should undergo a relatively large positive or negative change in firing rate in response to small orientation offsets. A priori, we expected that the same pattern of bilateral gain would be observed in the present study, and the exact reason for this apparent discrepancy is not clear. However, Scolari and Serences (2009) used a search display that contained multiple distracters, and the task was extremely difficult for many observers (based on the distribution of accuracies). Thus, informative off-target neural populations tuned to both sides of the target might be recruited to combat increased competition from additional distracters, particularly in the context of an extremely challenging search task.

Similarly, Scolari and Serences (2010) observed larger off-target BOLD responses bilaterally after postsorting trials based on behavioral accuracy in a task where subjects had to perform a challenging orientation discrimination task. In contrast, an analogous analysis of the relationship between accuracy and BOLD responses in the present experiment only revealed an effect in channels tuned beyond the cued target orientation (Figs. 6, 9). However, in Scolari and Serences (2010), subjects compared sequentially presented “sample” and “test” stimuli that were always offset by a small CW or CCW rotation. The fact that the test stimulus never matched the sample may have thus encouraged a greater reliance on responses in orientation channels tuned on either side of the target, compared with the present paradigm in which 50% of the trials contained matching stimuli. Again, this explanation is highly speculative, but future psychophysical and neuroimaging studies might address the conditions under which bilateral modulations are observed by systematically manipulating both the number of distracters and the type of discrimination that is being performed.

A common thread among most existing models of top-down attention is a focus on the relationship between the tuning properties of a cell and the physical characteristics of the stimulus array. For instance, the biased competition model holds that attention biases activity levels toward the response evoked by a preferred stimulus when presented alone, even when the preferred stimulus is presented along with an unattended distracter that would normally drive down the response of the cell (Moran and Desimone, 1985; Desimone and Duncan, 1995; Reynolds and Desimone, 1999). Likewise, the feature similarity gain model suggests that attention increases the gain of neurons that are tuned to a relevant feature, and suppresses the gain of neurons tuned away from a relevant feature (Treue and Martínez Trujillo, 1999; Martínez-Trujillo and Treue, 2004; Maunsell and Treue, 2006). Critically, both of these accounts focus primarily on the relationship between the relevant stimulus feature and the tuning properties of the recorded neuron, without explicitly factoring in the nature of the perceptual task being performed by the observer. In contrast, optimal gain accounts (Navalpakkam and Itti, 2007; Scolari and Serences, 2009) hold that the magnitude of attentional gain should depend on the relationship between the relevant feature, a neuron’s tuning properties, and the perceptual

task that is currently being performed. However, the addition of a task component poses a mystery: how does the brain “know” which neural populations to modulate given a specific task? In the case of on-target gain, a multiplicative enhancement of all sensory neurons would produce the largest positive modulation in neurons that are already strongly driven by the sensory stimulus (Treue and Maunsell, 1999; McAdams and Maunsell, 1999a,b). However, off-target enhancement could not be implemented via this type of across-the-board increase in sensory gain. While we have no conclusive answer, perceptual learning following extensive practice has been shown to selectively modulate response properties of neurons tuned away from a target (Schoups et al., 2001; Ghose et al., 2002; Raiguel et al., 2006). These off-target modulations might be mediated via a selective reweighting of presynaptic inputs (Ghose and Maunsell, 2008; Ghose, 2009) that might occur either rapidly in response to attentional demands, or after repeated experience with a specific stimulus. While speculative, the experience-based account is at least consistent with our observation that the predictive relationship between modulations in off-target channels and behavioral performance became more pronounced with practice in Experiment 1 (Fig. 6). Future investigations might therefore evaluate potential mechanisms by examining both the flexibility of off-target modulations and the degree to which they grow with practice in the context of a specific task.

Given these findings, existing models might be modified to account for the importance of task demands by stipulating that attention increases the gain of neurons that are most informative given the currently relevant perceptual discrimination (Navalpakkam and Itti, 2007). However, this poses an interesting challenge, as the activity of a sensory neuron is usually thought to convey information about the likelihood that its preferred feature is present in the visual field. Thus, off-target gain might lead to a nonveridical perceptual representation of a relevant feature, which may complicate the decoding of activity patterns during decision making (Navalpakkam and Itti, 2007; Scolari and Serences, 2009). One possible solution is that discriminating between two stimuli does not require the formation of veridical representations. Instead, visual features might be discriminated on the basis of the differential pattern of responses across populations of sensory neurons: a large differential response should count as evidence in favor of a mismatch judgment, and a small differential response should be counted as evidence in favor of a match judgment. In the present task, this would involve comparing the distance between simultaneously evoked response patterns in the left and right visual cortex. If this type of decision rule is used, then increasing off-target gain would lead to a greater separation between response patterns evoked by highly similar stimuli (compared with on-target gain), thereby facilitating accurate behavioral performance.

References

- Boynton GM (2005) Attention and visual perception. *Curr Opin Neurobiol* 15:465–469.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Brouwer GJ, Heeger DJ (2009) Decoding and reconstructing color from responses in human visual cortex. *J Neurosci* 29:13992–14003.
- Brouwer GJ, Heeger DJ (2011) Cross-orientation suppression in human visual cortex. *J Neurophysiol* 106:2108–2119.
- Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114.
- Das A, Sirotin YB (2011) What could underlie the trial-related signal? A response to the commentaries by Drs. Kleinschmidt and Muller, and Drs. Handwerker and Bandettini. *Neuroimage* 55:1413–1418.

- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
- Douglas RJ, Martin KA (2007) Recurrent neuronal circuits in the neocortex. *Curr Biol* 17:R496–R500.
- Eckstein MP, Shimozaki SS, Abbey CK (2002) The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *J Vis* 2:25–45.
- Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994) fMRI of human visual cortex. *Nature* 369:525.
- Freeman J, Brouwer GJ, Heeger DJ, Merriam EP (2011) Orientation decoding depends on maps, not columns. *J Neurosci* 31:4792–4804.
- Freeman WT, Adelson EH (1991) The design and use of steerable filters. *IEEE Trans Pattern Analysis Machine Intelligence* 13:891–906.
- Ghose GM (2009) Attentional modulation of visual responses by flexible input gain. *J Neurophysiol* 101:2089–2106.
- Ghose GM, Maunsell JH (2008) Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. *J Neurosci* 28:5115–5126.
- Ghose GM, Yang T, Maunsell JH (2002) Physiological correlates of perceptual learning in monkey V1 and V2. *J Neurophysiol* 87:1867–1888.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
- Gur M, Kagan I, Snodderly DM (2005) Orientation and direction selectivity of neurons in V1 of alert monkeys: functional relationships and laminar distributions. *Cereb Cortex* 15:1207–1221.
- Handwerker DA, Bandettini PA (2011a) Hemodynamic signals not predicted? Not so: A comment on Sirotin and Das (2009). *Neuroimage* 55:1409–1412.
- Handwerker DA, Bandettini PA (2011b) Simple explanations before complex theories: Alternative interpretations of Sirotin and Das' observations. *Neuroimage* 55:1419–1422.
- Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8:686–691.
- Heeger DJ, Ress D (2002) What does fMRI tell us about neuronal activity? *Nat Rev Neurosci* 3:142–151.
- Heeger DJ, Huk AC, Geisler WS, Albrecht DG (2000) Spikes versus BOLD: what does neuroimaging tell us about neuronal activity? *Nat Neurosci* 3:631–633.
- Hol K, Treue S (2001) Different populations of neurons contribute to the detection and discrimination of visual motion. *Vis Res* 41:685–689.
- Itti L, Koch C, Braun J (2000) Revisiting spatial vision: toward a unifying model. *J Opt Soc Am A Opt Image Sci Vis* 17:1899–1917.
- Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. *Nat Neurosci* 9:690–696.
- Jazayeri M, Movshon JA (2007) Integration of sensory evidence in motion discrimination. *J Vis* 7:7.1–7.7.
- Jia X, Smith MA, Kohn A (2011) Stimulus selectivity and spatial coherence of gamma components of the local field potential. *J Neurosci* 31:9390–9403.
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685.
- Kay KN, Naselaris T, Prenger RJ, Gallant JL (2008) Identifying natural images from human brain activity. *Nature* 452:352–355.
- Kleinschmidt A, Müller NG (2010) The blind, the lame, and the poor signals of brain function—a comment on Sirotin and Das (2009). *Neuroimage* 50:622–625.
- Law CT, Gold JI (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat Neurosci* 11:505–513.
- Law CT, Gold JI (2009) Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat Neurosci* 12:655–663.
- Lee J, Maunsell JH (2009) A normalization model of attentional modulation of single unit responses. *PLoS One* 4:e4651.
- Logothetis NK (2008) What we can do and what we cannot do with fMRI. *Nature* 453:869–878.
- Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. *Annu Rev Physiol* 66:735–769.
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157.
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77:24–42.
- Martinez-Trujillo JC, Treue S (2004) Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14:744–751.
- Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29:317–322.
- McAdams CJ, Maunsell JH (1999a) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441.
- McAdams CJ, Maunsell JH (1999b) Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron* 23:765–773.
- Moore BC (2008) Basic auditory processes involved in the analysis of speech sounds. *Philos Trans R Soc Lond B Biol Sci* 363:947–963.
- Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–784.
- Naselaris T, Kay KN, Nishimoto S, Gallant JL (2011) Encoding and decoding in fMRI. *Neuroimage* 56:400–410.
- Navalpakkam V, Itti L (2006) Top-down attention selection is fine grained. *J Vis* 6:1180–1193.
- Navalpakkam V, Itti L (2007) Search goal tunes visual features optimally. *Neuron* 53:605–617.
- Palmer J, Verghese P, Pavel M (2000) The psychophysics of visual search. *Vis Res* 40:1227–1268.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Pestilli F, Carrasco M, Heeger DJ, Gardner JL (2011) Attentional enhancement via selection and pooling of early sensory responses in human visual cortex. *Neuron* 72:832–846.
- Pouget A, Deneve S, Latham PE (2001) The relevance of Fisher Information for theories of cortical computation and attention. In: *Visual attention and neural circuits* (Braun J, Koch C, Davis JL, eds). Cambridge, MA: MIT.
- Purushothaman G, Bradley DC (2005) Neural population code for fine perceptual decisions in area MT. *Nat Neurosci* 8:99–106.
- Raiguel S, Vogels R, Mysore SG, Orban GA (2006) Learning to see the difference specifically alters the most informative V4 neurons. *J Neurosci* 26:6589–6602.
- Regan D, Beverley KI (1985) Postadaptation orientation discrimination. *J Opt Soc Am A* 2:147–155.
- Reynolds JH, Desimone R (1999) The role of neural mechanisms of attention in solving the binding problem. *Neuron* 24:19–29, 111–125.
- Reynolds JH, Heeger DJ (2009) The normalization model of attention. *Neuron* 61:168–185.
- Ringach DL, Bredfeldt CE, Shapley RM, Hawken MJ (2002a) Suppression of neural responses to nonoptimal stimuli correlates with tuning selectivity in macaque V1. *J Neurophysiol* 87:1018–1027.
- Ringach DL, Shapley RM, Hawken MJ (2002b) Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci* 22:5639–5651.
- Sasaki Y, Rajimehr R, Kim BW, Ekstrom LB, Vanduffel W, Tootell RB (2006) The radial bias: a different slant on visual orientation sensitivity in human and nonhuman primates. *Neuron* 51:661–670.
- Schiller PH, Finlay BL, Volman SF (1976) Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance. *J Neurophysiol* 39:1320–1333.
- Schoups A, Vogels R, Qian N, Orban G (2001) Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412:549–553.
- Schummers J, Yu H, Sur M (2008) Tuned responses of astrocytes and their influence on hemodynamic signals in the visual cortex. *Science* 320:1638–1643.
- Scolari M, Serences JT (2009) Adaptive allocation of attentional gain. *J Neurosci* 29:11933–11942.
- Scolari M, Serences JT (2010) Basing perceptual decisions on the most informative sensory neurons. *J Neurophysiol* 104:2266–2273.
- Serences JT, Boynton GM (2007a) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55:301–312.
- Serences JT, Boynton GM (2007b) The representation of behavioral choice for motion in human visual cortex. *J Neurosci* 27:12893–12899.

- Serences JT, Saproo S (2012) Computational advances towards linking BOLD and behavior. *Neuropsychologia* 50:435–446.
- Serences J, Saproo S, Scolari M, Ho T, Muftuler LT (2009) Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. *Neuroimage* 44:223–231.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Seung HS, Sompolinsky H (1993) Simple models for reading neuronal population codes. *Proc Natl Acad Sci U S A* 90:10749–10753.
- Sirotin YB, Das A (2009) Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457:475–479.
- Swisher JD, Gatenby JC, Gore JC, Wolfe BA, Moon CH, Kim SG, Tong F (2010) Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *J Neurosci* 30:325–330.
- Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–579.
- Treue S, Maunsell JH (1999) Effects of attention on the processing of motion in Macaque middle temporal and medial superior temporal visual cortical areas. *J Neurosci* 19:7591–7602.