

# The Selectivity of Task-Dependent Attention Varies with Surrounding Context

Yee-Joon Kim (김 이준) and Preeti Verghese

Smith-Kettlewell Eye Research Institute, San Francisco, California 94115

Attention is thought to operate by enhancing the target of interest and suppressing the surroundings. We hypothesized that the spatial profile of attention depends on the surround's relationship to the target. Using high-density electroencephalographic measurements, we examined the spatial profile of attention to a grating target surrounded by an annular grating that was either coextensive with the target (unsegmented) or appeared segmented from it due to a gap or phase offset. We directly probed the spread of attention from the central target into the surround by flickering the surround and monitoring frequency-tagged steady-state visual-evoked potentials. Observers were required to detect a contrast increment that occurred only on the target. Successful detection of the increment required selecting the target and suppressing the surround, particularly when the target did not readily segment from the surround. The profile of attention was investigated in five visual regions of interest (ROIs) (V1, V4, V3A, lateral occipital complex, and human middle temporal area), mapped in a separate anatomical magnetic resonance imaging scan. We found that in most ROIs, attention to the target generated smaller responses from the surrounding annulus when it was contiguous compared with when it was clearly segmented. This result shows that the profile of attention depends on task demands and on surrounding context; attention is tightly focused when the target region needs to be isolated but loosely focused when the target region is clearly segmented.

## Introduction

Attention is a top-down process that modulates neural activity to select a feature or location that is relevant to task demands. Here we ask how attention selects a target embedded in a textured background, and how the profile of selectivity depends both on the requirements of the task and on the surrounding context.

The spatial profile of attention to a target among discrete distractors suggests that attention enhances the target and suppresses surrounding distractors (Bahcall and Kowler, 1999; Müller and Kleinschmidt, 2004; Hopf et al., 2006). Single unit studies in extrastriate cortex report that directing attention to one of two stimuli within the receptive field effectively attenuates the effect of the other (Moran and Desimone, 1985; Reynolds et al., 1999; Womelsdorf et al., 2006; Ghose and Maunsell, 2008). However, few studies have addressed the spread of selectivity around the target when it is embedded in a textured surround (Marcus and Van Essen, 2002).

Studies that have examined the interaction between object segmentation and attention indicate that attention spreads within a perceptually linked surface (Duncan, 1984; He and Nakayama, 1995; Valdes-Sosa et al., 2000; Driver et al., 2001). Here

we investigate whether spatial attention can be more selective when observers perform a demanding task at the center of a uniform surface.

We hypothesize that the spatial profile of attention depends on context and task. We predict that in tasks that require discrimination of the properties of a target coextensive with the background, sensitivity to the irrelevant background is attenuated to exclude it from interfering with the target. Figure 1 demonstrates our prediction for a target surrounded by an annulus in a task that requires discrimination of target contrast. When the target is segmented by a gap or phase shift, there is no need to tightly focus attention to the central target, so attention may “leak” to the surrounding region, resulting in a broad spatial profile of attention (Fig. 1*a*, i–iii). In contrast, when the target is contiguous with the surround, it is necessary to either narrow the spatial profile of attention around the target (Fig. 1*a*, iv) or to suppress the surround (Fig. 1*b*, iv).

To test our hypothesis, we combined high-density electroencephalography (EEG) with anatomical and functional magnetic resonance imaging (fMRI) to characterize the population responses to these configurations in different regions of interest (ROIs; Appelbaum et al., 2006; Cottareau et al., 2011*a,b*). Observers were cued to attend to one of two target-surround stimuli that were presented to the left and right of fixation (Fig. 2*b*). Their task was to detect a contrast increment on the target on the cued side. The targets were surrounded by constant-size annuli that flickered at different frequencies on the left and right and generated unique frequency-tagged responses in the evoked activity (Morgan et al., 1996). This frequency-tagging method combined with cortical source localization allowed us to simultaneously monitor population responses in different ROIs from target-

Received Dec. 1, 2011; revised June 1, 2012; accepted June 29, 2012.

Author contributions: Y.J.K. and P.V. designed research; Y.J.K. performed research; Y.J.K. analyzed data; Y.J.K. and P.V. wrote the paper.

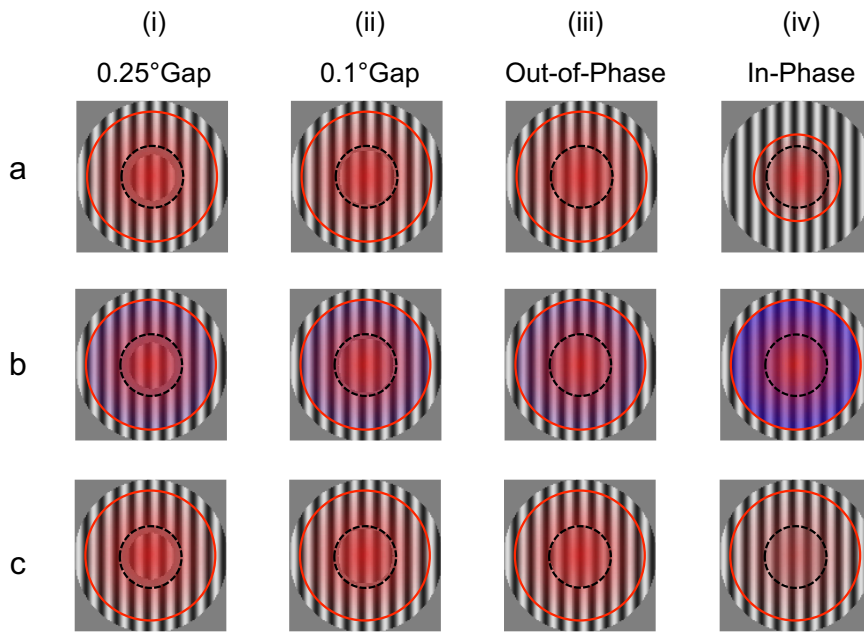
This work was supported by National Science Foundation Grant 0963914 (P.V.). We are grateful to Benoit Cottareau, Suzanne McKee, and Jeff Tsai for valuable comments.

The authors declare no competing financial interests.

Correspondence should be addressed to Yee Joon Kim, Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115. E-mail: joon@ski.org.

DOI:10.1523/JNEUROSCI.5992-11.2012

Copyright © 2012 the authors 0270-6474/12/3212180-12\$15.00/0



**Figure 1.** *a–c*, Three possible models for the spread of spatial attention around the target in each of the configurations tested in our experiments. The transparent red-blue disk represents the hypothetical spread of attention. The black dashed circle shows the inner border of the annulus. These are for purposes of illustration only. Our hypothesis is as follows. When observers attend to an increment on a central target that is clearly segmented from a surrounding annulus by either a gap or a phase offset as in the first three configurations, there may be little cost to having a diffuse attentional window that includes part of the surround, so the spatial profile of attention may spill over into the surround (i–iii). On the other hand, when the central target does not segment easily from an abutting in-phase surround, attenuating the surround grating might isolate the central pedestal and improve contrast discrimination (iv). There are three possible ways in which the spatial profile of attention can change as the degree of segmentation of the target from the surround changes. One model (attentional sharpening) is that target selection in the unsegmented (in-phase) configuration is achieved by narrowing the spatial profile of attention around the target, compared with the segmented configurations. A second model is a variant of the first where spatial selectivity in the unsegmented configuration is achieved by suppressing the surround rather than by restricting the spread of attention. A third model is that the allocation of attention to both center and surround gratings is globally weaker in the unsegmented configuration (global gain modulation). Although the schematic depicts the spatial profile of attention as the same across the first three configurations, it is possible that it could change gradually as the degree of segmentation decreases. For example, attentional sharpening might gradually decrease the width of spatial profile of attention as the degree of segmentation of the target from the surround decreases. Alternatively, the strength of surround suppression may gradually increase without a change in the width of the spatial profile of attention. If attentional selection occurs by global gain modulation, the gain may gradually decrease, without changing the size of the attentional window.

attended and target-ignored surround gratings and determine how attention to the central target affected cortical responses to the surround under segmented and unsegmented configurations (Fig. 2a).

## Materials and Methods

### Observers

A total of 12 observers participated in the recording sessions; the data from all observers were analyzed. In Experiment 1, 10 of these observers (seven men and three women) participated in the high-contrast version of the experiment. Seven of these observers and one additional observer (six men and two women) participated in the mid-contrast version. Responses from medium- and high-contrast stimuli were recorded in two separate sessions. Five of these observers and one additional observer (four men and two women) participated in Experiment 2. All observers had normal or corrected-to-normal visual acuity, gave informed written consent to participate as paid volunteers, and were tested individually in a dark room. The human subjects review committee of Smith-Kettlewell Eye Research Institute approved the study.

### Stimuli

All visual stimuli were composed of a central target surrounded by an annular surround in both experiments 1 and 2.

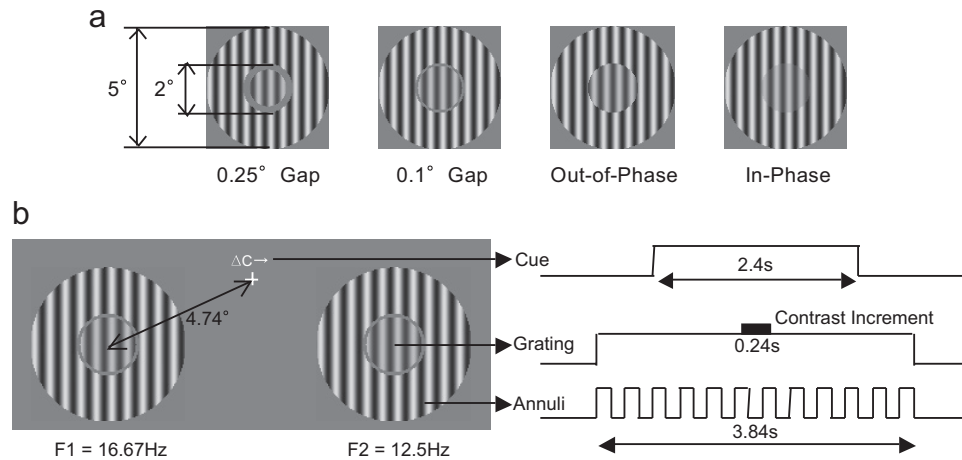
**Experiment 1: spread of attention around the target.** All visual stimuli were displayed on a 19" CRT (LaCie Electron Blue IV) monitor set to a

100 Hz refresh rate. Both temporal and luminance calibrations were performed using a calibrated photodiode, and monitor gamma tables were adjusted to ensure response linearity and a constant mean luminance of 49 cd/m<sup>2</sup>. The center and surround gratings were both vertical and had a spatial frequency of three cycles per degree. The concentric surround grating was either coextensive with the target grating (unsegmented) or segmented from it by a small gap or a phase offset (Fig. 2a). The gap sizes between target and surround were 0.25° (big gap), 0.1° (small gap), and 0° (out of phase and in phase). The size of the surround annulus was constant in all four configurations with inner and outer radii of 1° and 2.5°, respectively. As a result, the size of the inner target grating varied with configuration: it had a diameter of 1.5° in the big-gap condition, 1.8° in the small-gap condition, and 2° in the out-of-phase and in-phase conditions. For the out-of-phase condition, the target had a phase shift of 90° relative to the surround grating. Our choice of gap size and phase offset was based on previous research that shows that the interaction between center and surround configurations varies as a function of gap size and degree of phase offset. This interaction was greatest for abutting gratings and was negligible by a gap size of 0.2° (Xu et al., 2005; Appelbaum et al., 2008). Thus we expect a weak interaction (if any) at a gap size of 0.1° and no significant interaction at a gap size of 0.25°. For the abutting configuration, a 90° relative phase shift is sufficient to substantially eliminate interaction between center and surround (Xu et al., 2005). The in-phase configuration had no gap and no phase shift but may have appeared temporally segmented because the surround flashed on and off (at either 12.5 or 16.7 Hz). Therefore it is important to note that the configurations varied in their degree of segmentation, from most segmented (big gap) to least segmented (in phase).

The center-surround stimulus was presented on both sides of fixation at an eccentricity of 4.74° (1.5° below and 4.5° to the left/right of fixation). The surround gratings on the left and right flickered on and off at 16.67 and 12.5 Hz, respectively. For the high-contrast session, the left and right surround gratings were set at 85 and 75% contrast, respectively, so that they appeared perceptually matched with the static center that had a contrast of 50%. For the medium-contrast session, the left and right surround gratings were set at 25 and 21% contrast, respectively, to match the static center set at 15%.

**Experiment 2: sensitivity to the attended target.** The experimental apparatus was similar to Experiment 1 except that we used a different monitor in Experiment 2, as the monitor that we used for Experiment 1 had stopped working. All visual stimuli were displayed on a 21" CRT (NEC MultiSync FP 2141SB) monitor set to a 100 Hz refresh rate. Both temporal and luminance calibrations were performed using a calibrated photodiode and monitor gamma tables were adjusted to ensure response linearity and a constant mean luminance of 30 cd/m<sup>2</sup>.

In Experiment 2, the center grating was temporally modulated while the surround grating was static. The center gratings on the left and right flickered on and off at 16.67 and 12.5 Hz, respectively. The contrast of the center gratings were set at 72 and 65% on the left and right, respectively, so that they appeared perceptually matched to the static surround that had a contrast of 50%. We compared center responses only in the out-of-phase and in-phase configurations where the center size was identical. We chose to exclude the two gap conditions in this control experiment,



**Figure 2.** Stimuli and trial sequence. *a*, Three segmented (Big Gap, Small Gap, and Out-of-Phase) and one unsegmented (In-Phase) target-background configurations were used. The diameter of the central target was 1.5° (0.25° gap), 1.8° (0.1° gap), and 2° (no gap) in these four configurations. Note that the size of the surround is constant in these configurations; size of the target is reduced for the two gap conditions. *b*, On each trial, two center-surround stimuli were simultaneously presented 4.5° to the left and right of fixation and 1.5° below it. Surround gratings flickered at 16.67 Hz on the left and at 12.5 Hz on the right. Each trial was initiated by a button press. The trial lasted 2.4 s and started with the appearance of the cue at the fixation point indicating the location (left or right) of the increment. The target and grating stimuli came on and went off 720 ms before the start and end of the trial to eliminate onset and offset transients. The contrast increment (if present) appeared 1.2 s after the cue and lasted for 240 ms. Only 40% of the trials had an increment. Observers indicated the presence/absence of the increment with a key press at the end of the trial.

because their smaller center size could lead to confounding effects in response magnitude. The other stimulus parameters were the same as in Experiment 1.

#### Experimental Procedure

The observer initiated each trial by a button press (Fig. 2*b*). A central arrow cue then appeared to direct the observer's attention to the target on the left or the right. The cue indicated (with 100% validity) the side on which a contrast increment could appear on the target. To avoid transients due to stimulus onset, the stimuli came on 720 ms before the cue. During the subsequent 2.4 s period, the observer voluntarily attended to the cued grating to perform contrast discrimination on it, while maintaining fixation at the central cue and attempting to withhold eye blinks. In Experiment 1, a brief contrast increment lasting 240 ms appeared on the static target grating 1.2 s after the start of the trial. In Experiment 2, a brief contrast increment lasting 240 ms appeared on the flickering target grating at a random time (0.96 s, 1.2 s, 1.44 s, or 1.68 s) after the start of the trial. Only 41% of the trials had an increment. The observer indicated whether the increment was present or absent at the end of each trial. Only increment-absent trials were analyzed to avoid transients due to the contrast increment.

In Experiment 1, there were 16 trial types: 2 spatial cues (left or right), 2 states of increment (present or absent), and 4 configurations (big gap, small gap, out of phase, or in phase). We ran the target-absent trial type 30 times and the target-present trial type 21 times while randomly varying and counterbalancing the three factors across trials. We tested each observer for a total of 408 trials in three blocks of 136 trials each. We gave observers breaks within and between blocks as necessary. We initially ran practice trials to equate the difficulty of contrast discrimination across all conditions. This was done by adjusting the magnitude of the contrast increment for each stimulus configuration to yield about 90% accuracy. On average, the contrast increment was 16.1, 16.2, 17.9, and 18.8% for the big gap, small gap, out-of-phase, and in-phase configurations, respectively, for high-contrast session. At medium contrast, the average contrast increment was 9.75, 9.87, 9.75, and 10.8% for the big gap, small gap, out of phase, and in phase, respectively.

In Experiment 2, there were eight trial types: two spatial cues (left or right), two states of increment (present or absent), and two configurations (out of phase or in phase). Similar to Experiment 1, we ran 30 target-absent trials and 21 target-present trials for each type. We tested each observer for a total of 204 trials in three blocks of 68 trials each. We adjusted the magnitude of the contrast increment for each stimulus con-

figuration to yield about 85% accuracy. The performance-correct criterion was set slightly lower than in Experiment 1 for technical and perceptual reasons. At a pedestal contrast of 72% we were limited by the largest contrast increment that we could generate on our monitor (95%). In addition, the detection of a contrast increment was more difficult on a flickering target than on a static target and the temporal uncertainty of the contrast increment made the task harder.

#### EEG signal acquisition and source imaging procedure

The EEG data were collected with 128-sensor HydroCel Sensor Nets (Electrical Geodesics) and were bandpass filtered from 0.1 to 200 Hz. Following each experimental session, the 3D locations of all electrodes and three major fiducials (nasion, left, and right peri-auricular points) were digitized using a 3Space Fastrack 3-D digitizer (Polhemus). For all observers, the 3D digitized locations were used to coregister the electrodes to their T1-weighted anatomical MRI scans.

Raw data were evaluated off-line according to a sample-by-sample thresholding procedure to remove noisy sensors that were replaced by the average of the six nearest spatial neighbors. Once noisy sensors were substituted, the EEG was re-referenced to the common average of all the sensors. Additionally, EEG epochs that contained a large percentage of data samples exceeding threshold (30  $\mu$ V) were excluded on a sensor-by-sensor basis including horizontal and vertical eye channels. This threshold has rejected any large eye movements ( $>1.5^\circ$ ) toward the relevant target. The fact that the topographic steady-state visual-evoked potential (SSVEP) patterns revealed no significant deviation from stable fixation suggests that horizontal eye movements were not significant. If the observers had instead looked directly at the stimulus (rather than at the fixation point), the topographic maps would have shown high activity in the representation of the bilateral medial posterior areas, rather than contralateral visual areas. No such artifacts were seen.

Although all our observers were experienced psychophysical observers who were practiced in our task, it is still possible that observers might have made small eye movements toward the cued target on a fraction of the total trials. To investigate this possibility, we tested the stability of eye fixation for four of our observers in a separate session. Observers performed an identical task to that used for both high- and medium-contrast sessions in Experiment 1 while their left eye position was monitored with a ViewPoint Eye Tracker (Arrington Research) sampling at 224 Hz. Viewing was binocular. Head position was maintained with a chin rest. Calibration was performed in two stages. Each block of trials started with the eye tracker's default calibration program that used a 4  $\times$

4 grid that spanned the display, and then with a custom  $5 \times 5$  point grid that spanned the central  $10^\circ$  of the display where our stimuli were presented.

On average, observers maintained fixation (within a  $0.5^\circ$  window around the fixation cross) for 96.6 and 97.8% of the trials in the high- and medium-contrast sessions, respectively. Both measures indicate that the observers were able to perform the covert attention task while maintaining fixation on the central fixation point.

### Structural and fMRI

Structural and fMRI scanning was conducted at 3 T (Siemens) using a 12-channel head coil. We acquired a T1-weighted MRI dataset (3D MP-RAGE sequence,  $0.8 \times 0.8 \times 0.8 \text{ mm}^3$ ) and a 3D T2-weighted dataset (SE sequence at  $1 \times 1 \times 1 \text{ mm}^3$  resolution) for tissue segmentation and registration with the functional scans. For fMRI, we employed a single-shot, gradient-echo echoplanar imaging sequence (TR/TE = 2000/28 ms, flip angle  $80^\circ$ , 126 volumes per run) with a voxel size of  $1.7 \times 1.7 \times 2 \text{ mm}^3$  ( $128 \times 128$  acquisition matrix, 220 mm FOV, bandwidth 1860 Hz/pixel, and echo spacing 0.71 ms). We acquired 30 slices without gaps, positioned in the transverse-to-coronal plane approximately parallel to the corpus callosum and covering the whole cerebrum. Once per session, a 2D SE T1-weighted volume was acquired with the same slice specifications as the functional series to facilitate registration of the fMRI data to the anatomical scan.

The FreeSurfer software package (<http://surfer.nmr.mgh.harvard.edu>) was used to perform gray and white matter segmentation and a mid-gray cortical surface extraction. This cortical surface had 20,484 isotropically spaced vertices and was used both as a source constraint and for defining the visual areas. The FreeSurfer package extracts both gray/white and gray/CSF boundaries, but these surfaces can have different surface orientations. In particular, the gray/white boundary has sharp gyri (the curvature changes rapidly) and smooth sulci (slowly changing surface curvature), while the gray/CSF boundary is the inverse, with smooth gyri and sharp sulci. To avoid these discontinuities, we generated a surface partway between these two boundaries that has gyri and sulci with approximately equal curvature.

Individual boundary element method conductivity models were derived from the T1- and T2-weighted MRI scans of each observer. The FSL toolbox (<http://www.fmrib.ox.ac.uk/fsl/>) was also used to segment contiguous volume regions for the scalp, outer skull, and inner skull and to convert these MRI volumes into inner skull, outer skull, and scalp surfaces (Smith, 2002; Smith et al., 2004).

### Visual area definition

The general procedures for these scans (head stabilization, visual display system, etc.) are standard and have been described in detail previously (Brewer et al., 2005). Retinotopic field mapping produced ROI-defined visual cortical areas V1, V2v, V2d, V3v, V3d, V3a, and V4 in each hemisphere (Tootell and Hadjikhani, 2001; Wang and Wade, 2011). ROIs corresponding to the human middle temporal area (hMT+) were identified using low-contrast motion stimuli similar to those described by Huk et al. (2002).

The lateral occipital complex (LOC) was defined using a block-design fMRI localizer scan. During this scan, the observers viewed alternating blocks containing intact and scrambled images of everyday objects (Kourtzi and Kanwisher, 2000). The regions activated by these scans included an area lying between the V1/V2/V3 foveal confluence and hMT+ that we identified as LOC. This definition covers almost all regions (e.g., V4d, LOC, lateral occipital parietal) that have previously been identified as lying within object-responsive LOC (Kourtzi and Kanwisher, 2000; Tootell and Hadjikhani, 2001).

ROIs V1, V4, V3a, LOC, and hMT+ were chosen to cover single visual area “clusters”: groups of retinotopic fields with a common foveal representation (Wandell et al., 2005). The dorsal cluster (V3a, V3b, V7) and posteromedial cluster (V1, V2, V3) have separate foveal representations. Area hV4 is technically part of the V1 cluster but is relatively distant from the majority of the V1 ROI and, instead, borders a set of foveal representations (VO1 and 2; Brewer et al., 2005) on the ventral surface. It is generally considered to be a classic “ventral stream” area. LOC shares a

foveal cluster with area +hMT, but we consider cortical sources in both of these LOC and hMT+ ROIs because area hMT+ is a classic “dorsal stream” area, and area LOC is thought to play an important role in object recognition.

### Cortically constrained inverse

An L2 minimum norm inverse was computed with sources constrained to the location and orientation of the cortical surface (Hämäläinen et al., 1993). In addition, we modified the source covariance matrix in two ways to decrease the tendency of the minimum norm procedure to place sources outside of the visual areas. These constraints involved the following: (1) increasing the variance allowed within the visual areas by a factor of two relative to other vertices and (2) enforcement of a local smoothness constraint within an area using the first- and second-order neighborhoods on the mesh with a weighting function equal to 0.5 for the first order and 0.25 for the second (Cottereau et al., 2011a). The smoothness constraint therefore respects areal boundaries unlike other smoothing methods such as low-resolution brain electromagnetic tomography that apply the same smoothing rule throughout cortex (Pascual-Marqui et al., 1994).

### ROI-based analysis of the SSVEP

A discrete Fourier transform was used to estimate the average response magnitude associated with each functionally defined ROI for the first- and the second-harmonic components of the steady-state frequencies (16.67 and 12.5 Hz). To take into account the difference in noise levels between the recordings from each of our observers (Vialatte et al., 2010), we computed the signal-to-noise ratio (SNR) by dividing peak amplitudes by the associated noise, which is defined for a given frequency  $f$  by the average amplitude of the two neighbor frequencies (i.e.,  $f - \delta f$  and  $f + \delta f$  where  $\delta f$  gives the frequency resolution of the Fourier analysis, which was 0.5 Hz in our studies). Our analysis focused on the second-harmonic components of the flicker frequencies (i.e.,  $2f$ ) because previous studies have found the second-harmonic response to be particularly robust in the SSVEP (Rager and Singer, 1998; Herrmann, 2001; Pei et al., 2002) and highly sensitive for attentional modulation (Pei et al., 2002; Kim et al., 2007, 2011; Pastor et al., 2007; Saupé et al., 2009). Indeed, SNRs of the first-harmonic components were either very poor ( $\text{SNR} \ll 1.5$ ) or did not show any attentional modulations in this experiment. The averaged second-harmonic responses (33.3 Hz) from right hemispheric ROIs corresponding to 16.67 Hz flicker and (25 Hz) from left hemispheric ROIs corresponding to 12.5 Hz flicker were used for all statistical analyses.

The robust second-harmonic response to an on-off flickering stimulus is probably due to the involvement of frequency-doubling neurons. Magnocellular neurons, which show transient onset and offset response profile, are a likely substrate for generating the second-harmonic response, although they may not generate a robust first harmonic to an on-off stimulus (Kulikowski et al., 1997). It has also been reported that the magnocellular pathway shows a robust modulation due to attention (Di Russo and Spinelli, 1999; Wang and Wade, 2011).

### Statistical analysis of SNRs

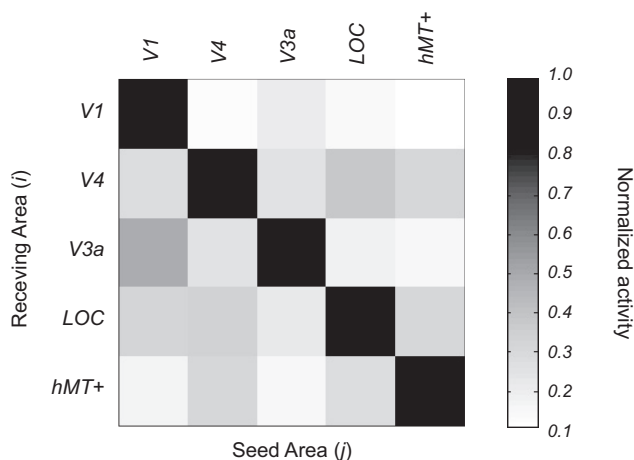
In Experiment 1, we performed a one-way ANOVA on the SNRs of sensor-based data and source-imaged data in the ignored condition to determine if the surround responses were significantly influenced by center-surround configuration. To determine whether the type of center-surround had a significant effect on the spatial profile of attention, we performed statistical tests on the pattern of SNR differences between the attended and ignored conditions in sensor space as well as in source space. We conducted a planned contrast analysis between the responses in the segmented and unsegmented configurations (Howell, 1997; Myers and Well, 2003; Cardinal and Aitken, 2006).

A linear contrast is a linear combination of the mean SNR difference for each configuration. Each mean  $\mu_j$  is weighted by a weight  $w_j$  as follows:

$$L = w_1\mu_1 + w_2\mu_2 + w_3\mu_3 + w_4\mu_4,$$

such that  $\sum w_j = 0$ .

We used two sets of linear contrast values for testing our predictions shown in Figure 1. One set of linear contrast values was for the comparison of “segmented versus unsegmented.” To quantitatively test our hy-



**Figure 3.** Theoretical estimates of cross talk between source-imaged EEG signals in retinotopically defined visual areas. Grayscale values at row  $i$  and column  $j$  represent the relative contribution area  $j$  to the cortical current density estimate in area  $i$ . Estimates are based on five different areas (V1, V4, V3a, LOC, and hMT+) that are defined individually in each of our observers.

pothesis that attention differentially modulates the surround in different configurations, we fit the SNR differences using weight values +1 for each of the segmented conditions (0.25° gap, 0.1° gap, out of phase), and -3 for the unsegmented condition (in phase) as follows:

$$L = \mu_1 + \mu_2 + \mu_3 - 3\mu_4.$$

Another set of linear contrast values was used to test for a “graded” linear trend in SNR difference with decreasing segmentation of the center from the surround. For this linear trend analysis, we fit the SNR difference using weight values +3, +1, -1, and -3 for 0.25° gap, 0.1° gap, out of phase, and in phase, respectively as follows:

$$L = 3\mu_1 + \mu_2 - \mu_3 - 3\mu_4.$$

In Experiment 2, we performed a one-way ANOVA on the SNRs of source-imaged data in the ignored condition to determine if the center responses were significantly influenced by center-surround configuration. To determine whether the type of center-surround had a significant effect on the center of attentional gain field, we performed a  $t$  test on SNR differences between the attended and ignored conditions in each ROI for the out-of-phase and in-phase configurations in each ROI.

### Cross talk

We estimated the theoretical cross talk among visual areas in our EEG study using the methodology described by Cottreau et al. (2011b). Cross talk refers to the neural activity generated in other areas as a result of activity in a particular ROI, due to the smoothing of the electric field by the head volume. In brief, for each observer, we simulated the cross talk by placing sources in one ROI and estimating their contribution to other ROIs, using the same forward and inverse methods described in the sections above. The global cross-talk matrix (i.e., averaged across all the observers who participated in our EEG experiments) is shown in Figure 3 for five ROIs (V1, V4, V3a, LOC, +hMT). For each ROI, this matrix shows how much activity is picked up in a given ROI from activity in each of the other ROIs. The cross-talk magnitude shown in the matrix is proportional to activity originating in the ROI where the cross talk is being estimated. Figure 3 provides the average cross talk between ROIs obtained from our 10 observers from the high-contrast session in Experiment 1.

Values at row  $i$  and column  $j$  represent the relative contribution of area  $j$  to the cortical current density estimate in area  $i$ . The normalization is obtained by dividing by the amplitude obtained in area  $i$  when only area  $i$  was activated in the simulation set. For example, when we estimated the activity in V1, the absolute amplitudes obtained from V4, V3a, LOC, and hMT+ when they were simulated independently (i.e., the second, third, fourth, and fifth columns of the first line of the cross talk matrix) were

respectively 12.43, 21.52, 14.25, and 10.69% of the amplitude in V1 when only V1 was activated. (i.e., first line, first column). An ideal estimation of the cortical current densities would lead to zero cross talk, and the associated matrix would be equal to the identity. In our study however, V4, V3a, LOC, and hMT+ received on average <30% cross talk from other areas. This means that our estimates of activity in each ROI are not influenced strongly by other ROIs. Therefore, our cross-talk matrix indicates that activity in five ROIs is strongly, but not completely, due to activity generated in the corresponding visual area. Note that coactivations from V4, V3a, LOC, and hMT+ together would not equal a linear summation of their individual contributions as strong cancellations could arise between them (Ahlfors et al., 2010).

## Results

### Experiment 1: spread of attention around the target

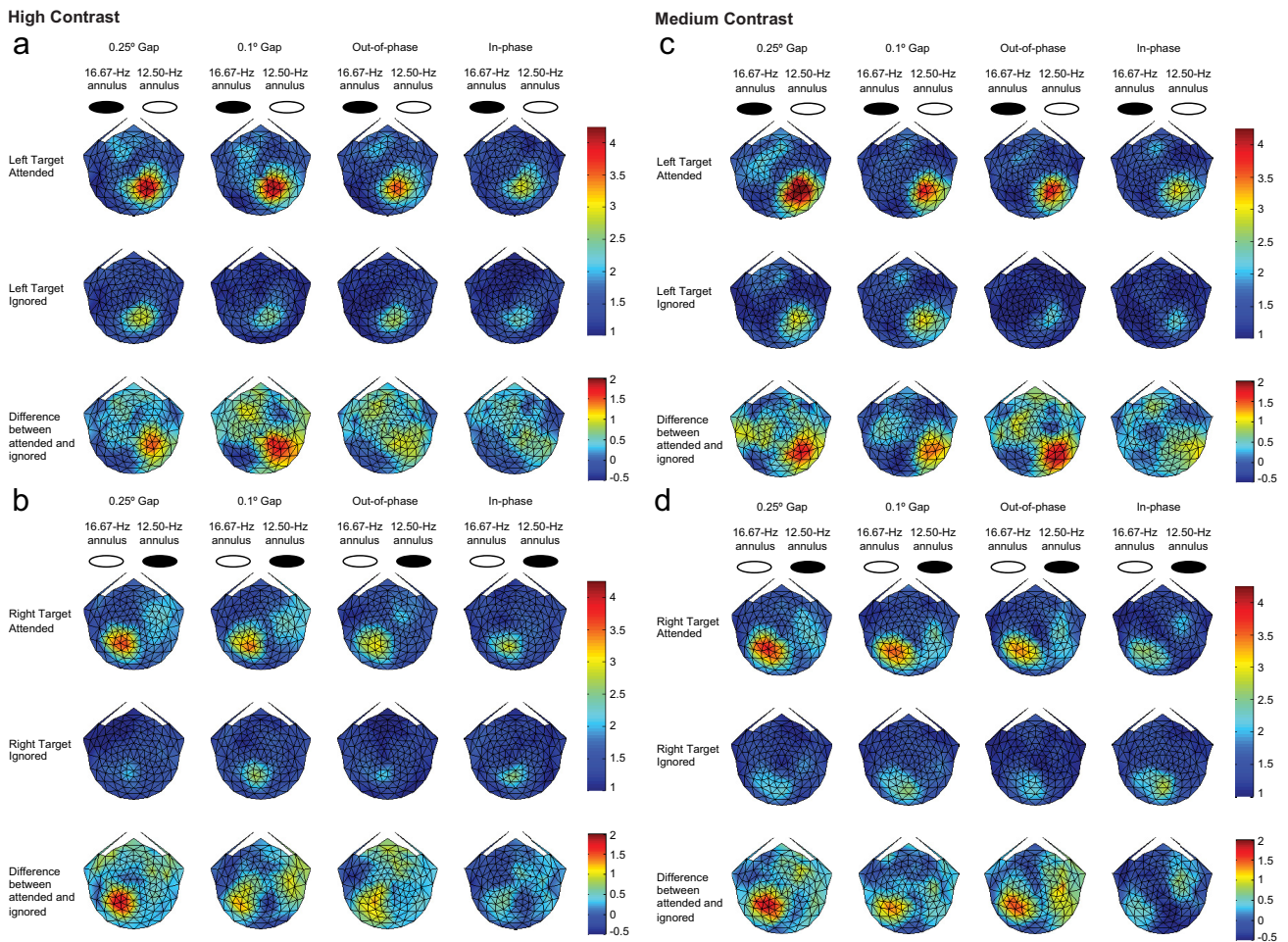
We elicited frequency-tagged SSVEPs from flickering annuli that surrounded a central target, and compared responses in four different center-surround configurations. To obtain a measure of surround modulation due to attention, we contrasted SSVEPs derived from the second-harmonic EEG responses for attended and ignored conditions (Fig. 4). This was done for surrounds that flickered at 16.67 and 12.5 Hz on the left and right of fixation, respectively.

The SSVEP topographies for the responses to the 16.67 Hz surround grating showed maximal amplitudes over contralateral posterior cortex. This pattern was observed at both high- and medium-contrast (Fig. 4*a,c*). The locus of attentional modulation, measured as the difference in SNR between the attended and ignored conditions, showed the same contralateral posterior focus of modulation, indicating that attentional modulation occurred in the same brain regions that were activated by the surround grating. Furthermore, the amplitude of modulation depended on stimulus configuration, but only in the attended condition. A parallel set of results was obtained for SSVEPs synchronized to the 12.50 Hz surround grating (Fig. 4*b,d*).

These findings demonstrate that the surround gratings elicited focal SSVEPs at contralateral posterior scalp regions and that voluntary spatial attention to the central target modulated these localized visual responses. We next evaluated whether the stimulus-evoked population electrophysiological activity due to the surround was consistent with the predictions shown in Figure 1. To evaluate the differential modulations of the surround responses across different stimulus configurations, we used data from 16 scalp locations (eight on each side of the scalp; Fig. 5*a*) that were selected to correspond to the foci of maximal sensory activation and attentional modulation. We averaged SNRs (see Materials and Methods) separately for locations contralateral and ipsilateral to the side of stimulation (combining left and right stimulus presentations).

At both high and low contrast, a one-way ANOVA on SNRs of the ignored condition confirmed that SNRs were similar across all stimulus configurations (Fig. 5*b*, white bars). In contrast, the attentional modulation, measured as the SNR difference between attended and ignored conditions, was greater for the three segmented configurations than for the unsegmented configuration (Fig. 5*c*). We also confirmed that the degree of attentional modulation of the surround showed a pattern of gradual reduction as the segmentation became less clear, especially at high contrast.

Because our hypothesis specifically states that attention interacts with image segmentation, we are primarily interested in whether the attentional modulation of the surround in the unsegmented configuration is smaller than that for the other segmented configurations and whether the degree of this modulation gradually decreases as the target segments less readily from



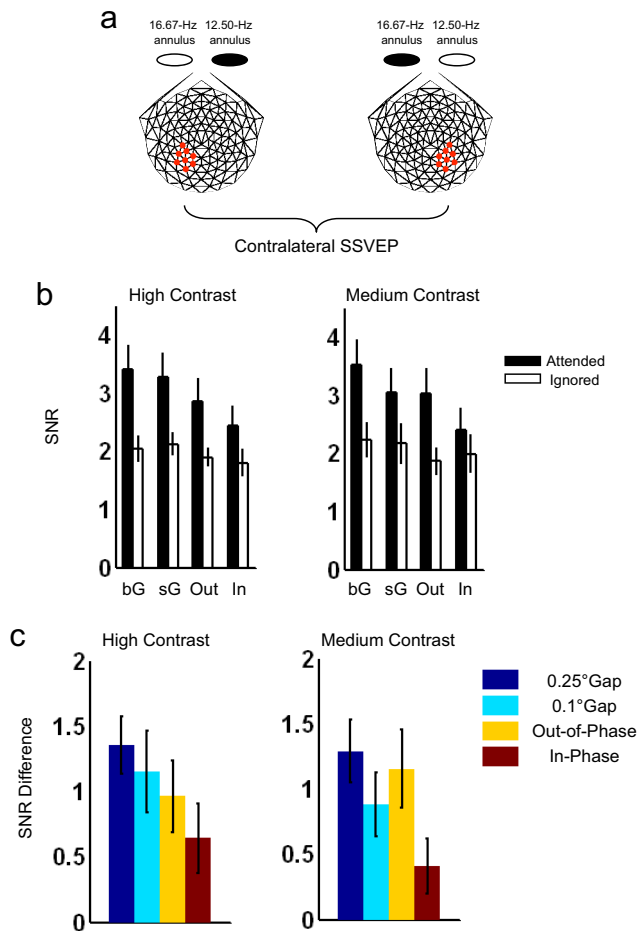
**Figure 4.** Topographic plots of the SNRs of the second-harmonic SSVEPs in response to the two surround gratings at high contrast (**a, b**) and medium contrast (**c, d**), averaged across observers. **a**, Frequency-tagged responses to the 16.67 Hz high-contrast surround grating in the left visual field. **b**, Frequency-tagged responses to 12.50 Hz high-contrast surround grating in the right visual field. **c, d**, Similar to **a** and **b**, but at medium contrast. SNR values smaller than 1 indicate that there was no signal at the driving frequencies. In each figure, the first three columns show the segmented stimulus configurations (0.25° gap, 0.1° gap, and out-of-phase) and the last column shows the unsegmented stimulus configuration (in-phase). The rows show responses when the relevant target was attended (upper), ignored (middle), and the difference between attended and ignored conditions (lower). At both high and medium contrast, the largest responses (SNR) and the largest attention effects occurred at contralateral posterior scalp locations.

the surround (Fig. 1). To quantitatively confirm our predictions, we compared linear combinations of the different configurations using contrast analysis (Howell, 1997; Myers and Well, 2003; Cardinal and Aitken, 2006). This method has the advantage of answering our specific hypotheses (segmented vs unsegmented and graded) whereas an ANOVA compares all configurations to determine if any configuration is different from the others. Another advantage of contrast analysis is that it reduces the chance of type I errors by asking fewer questions relative to comparing all possible configurations. The statistical analysis was performed at both high and medium contrast in each ROI to examine whether the degree of attentional modulations across stimulus configurations depended on stimulus contrast level.

Table 1 summarizes the results of the two planned trend analyses (see Materials and Methods): one trend analysis compared the segmented versus unsegmented configurations (Table 1, second column) and the other trend analysis determined whether there was a gradual reduction in the magnitude of attentional modulation with decreasing segmentation of the surround from the center (Table 1, third column). The trend analyses confirmed that at both high and medium contrast, the attentional modulation of the surround was significantly smaller when it was con-

tiguous with the center compared with when it was segmented. The graded linear trend was also statistically significant showing that the strength of modulation of the surround decreases as the segmentation of the surround decreases. Thus the pattern of results is consistent with our predictions that the attentional modulation of the surround activity is weaker when the target is not readily segmented from the surround (Fig. 1).

We next examined whether this same pattern was apparent within the separate ROIs (see Materials and Methods). Five ROIs (V1, V4, V3a, LOC, and hMT+) were chosen for the statistical analyses as described (see Materials and Methods). To evaluate attentional modulation within each ROI, we averaged SNR values generated by contralateral stimulation, combining left and right stimulus presentations (combining the responses in the right and left hemispheres to the 16.67 and 12.5 Hz surrounds in the left and right visual fields, respectively). Figures 6 and 7 show SNR values for the attended and ignored conditions and their difference. The four configurations did not show significantly different cortical activation in the ignored condition, similar to the result obtained from the topographic map of evoked responses (Fig. 5b, white bars). A one-way ANOVA on SNRs of the ignored condition confirmed that SNRs were indeed statistically indistinguishable



**Figure 5.** *a*, Illustration of the contralateral electrodes used for this analysis. These are the electrode locations that generated a strong stimulus response, as well as robust attention effects in Figure 4. *b*, The average second-harmonic responses to the 16.67 and 12.5 Hz surround gratings in contralateral electrodes. The black and white bars represent the SNRs to the flickering surround grating when the target was attended and ignored, respectively. The SNRs across all stimulus configurations for the ignored condition (white bars) were not statistically different. bG, big (0.25°) gap; sG, small (0.1°) gap; Out, out-of-phase; In, in-phase. *c*, Attentional modulation estimated as the SNR difference between the attended and ignored condition, computed from the 16 scalp electrodes illustrated in *a*. In *b* and *c*, contralateral responses at high and medium contrasts are shown on the left and right, respectively. The trend analyses for the pattern of attentional modulations across the four configurations were significant at both high contrast and medium contrast (Table 1). The SNR and SNR difference was averaged across observers. Error bars indicated  $\pm 1$  SEM.

**Table 1. Trend analysis for sensor-based responses**

EEG session/trend analysis	Segmented versus unsegmented	Graded
High-contrast session	$F_{(1,27)} = 9.188, p < 0.006$	$F_{(1,27)} = 12.49, p < 0.002$
Mid-contrast session	$F_{(1,21)} = 9.912, p < 0.005$	$F_{(1,21)} = 7.591, p < 0.012$

across all stimulus configurations in each ROI at both contrast levels (Fig. 6*a,b*, white bars). This result indicates that small differences in the size of the central target in the different configurations had little effect on the response to the surround gratings when targets are not attended. In contrast, the attended condition evoked stronger cortical activation for the three segmented configurations compared with the unsegmented configuration. Thus the degree of attentional modulation, measured as the difference between the attended and ignored conditions, was greater for the three segmented configurations than for the unsegmented configuration. This was true at both medium and high contrast in

most ROIs (Fig. 7*a,b*). The differential attentional modulation of the surround across stimulus configurations cannot be attributed to differences in the surround as its size was kept constant across all configurations. Nor is it likely due to differences in task difficulty across stimulus configurations, because our behavioral assay, increment detection, was similar across all configurations (Fig. 7*c,d*). These results suggest that when the center is not easily segmented from its surround, attention works to isolate the central target from the surround. On the other hand, when the center is easily segregated from its surround, attention does not need to be more selective to detect a target at the center.

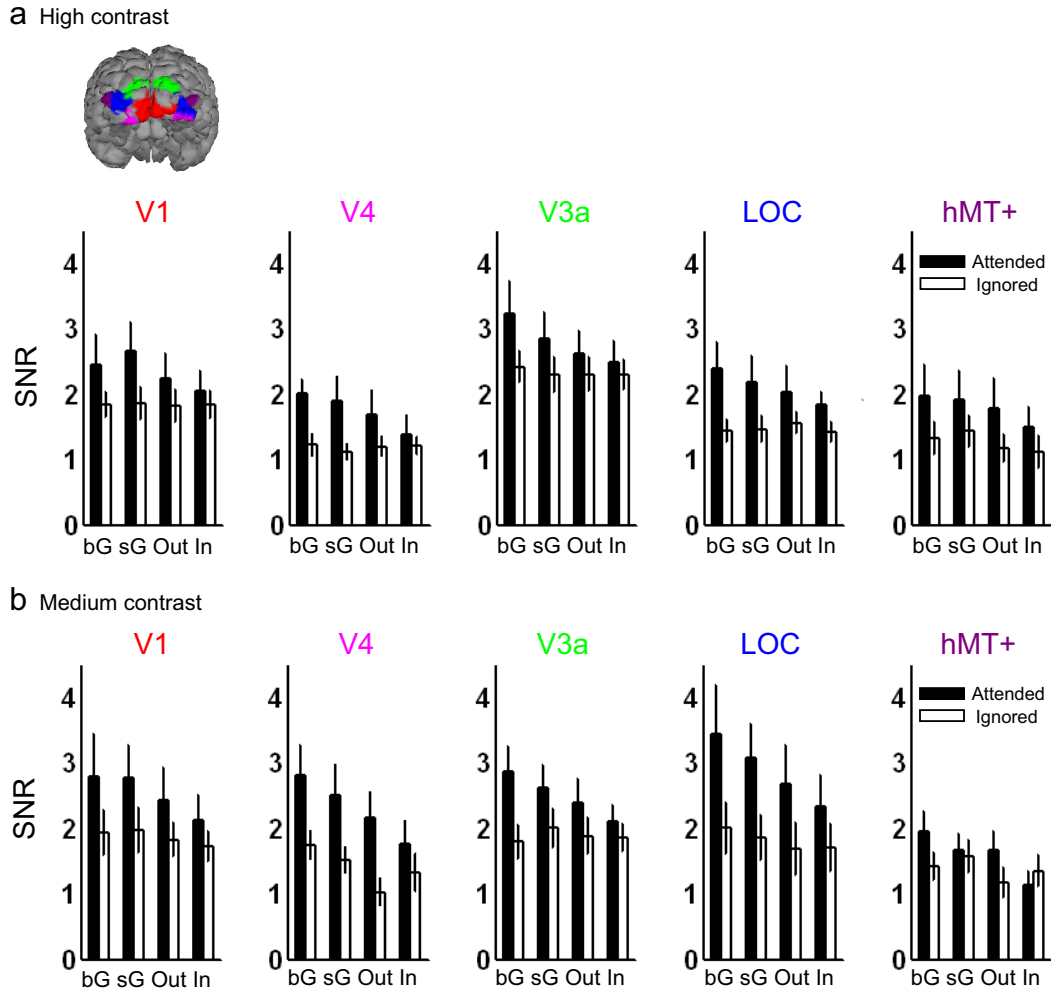
Figure 7*a* shows the pattern of attentional modulation across stimulus configurations at high contrast. The response pattern in V1 and V4 was consistent with segmented versus unsegmented hypothesis that attentional modulation of the surround is more spatially selective in the unsegmented configuration (Table 2). The response pattern in V1, V4, V3a, and LOC was also consistent with the graded hypothesis that the degree of attentional modulation of the surround decreases with the degree of segmentation of the surround from the center (Table 2). In area hMT+, the attentional modulation of the in-phase surround was smaller than for the segmented configurations although the response pattern across configurations was not well fit by our specific weight values.

At medium contrast, most ROIs except V1 showed a significant differential modulation for segmented versus unsegmented as well as for the graded hypotheses (Table 3). In area hMT+, attention suppressed the evoked response in the unsegmented condition, while it enhanced the responses in the segmented configurations to varying degrees (Fig. 7*b*).

Regardless of contrast level, it is important to note that responses in most areas show a more graded pattern that depends on degree of segmentation (size of gap and presence of phase offset), rather than a strict segmented–unsegmented dichotomy. This graded pattern shows the strongest attentional modulation of the surround for the large gap configuration, intermediate modulation for the small gap and out-of-phase configurations, and the smallest modulation for the in-phase, no-gap configuration. It is likely that the areas that show a graded spread of attention to the surround are sensitive to figure-ground segmentation and combine multiple cues that reflect the strength of segmentation. In fact, V4 and LOC are well known for their role in figure-ground segmentations (Altmann et al., 2004; Appelbaum et al., 2006; Chandrasekaran et al., 2007; Cottureau et al., 2011b; Roe et al., 2012); V3a is also sensitive to figure-ground cues generated by depth differences, illusory contours, and segmentation of motion (Mendola et al., 1997; Caplovitz and Tse, 2010; Cottureau et al., 2011b).

Our findings appear consistent with the sharpening model in Figure 1 where the width of the profile of attention becomes more tightly focused in configurations where the target segments weakly from the surround. Our data appear less consistent with a simple version of the suppression model where spatial selectivity in the unsegmented configuration is achieved by suppressing the surround. The only area that shows a clear suppression of the surround response due to attention is area hMT+ for the medium-contrast condition (see Discussion).

However, the global-gain modulation model where the magnitude of the gain field varies in proportion to the strength of segmentation is also consistent with the decreased modulation of the surround under poorly segmented conditions. To evaluate this possibility we need to determine whether the modulation of the target at the center of the configuration also varies in propor-



**Figure 6.** The second-harmonic responses (SNRs) from the surround gratings at high contrast and medium contrast in source space. The second harmonic-responses (SNRs) in five ROIs, averaged across observers, at high contrast and medium contrast (**a, b**). For each stimulus configuration, the second-harmonic responses to the 16.67 and 12.5 Hz surround gratings in contralateral ROIs were averaged. The black and white bars represent the SNRs to the flickering surround grating when the target was attended and ignored, respectively. The SNRs for the ignored condition were similar across all stimulus configurations. bG, big (0.25°) gap; sG, small (0.1°) gap; Out, out-of-phase; In, in-phase.

tion with the surround. To evaluate this, we ran another experiment to measure the strength of modulation at the center of the target.

**Experiment 2: sensitivity to the attended target**

To determine how attention modulates the sensitivity to the central target depending upon the surround context, we repeated Experiment 1, but flickered the center grating instead of the surround grating. We excluded the two gap conditions as their smaller center size could lead to confounding effects in response magnitude and tested only the out-of-phase and in-phase configurations because their center size was identical. Because the SSVEP response to small stimuli (such as the central target) has a very small SNR, only high-contrast stimuli were used.

Observers were cued to attend to either the left or the right center grating while we recorded the SSVEPs elicited by both center gratings. The two center gratings were flickered at 16.67 and 12.5 Hz, respectively, on the left and right of fixation so that we could simultaneously monitor the SSVEPs elicited by the attended and ignored center gratings on the basis of frequency tagging.

Figure 8 shows SNR values from the center gratings for the attended and ignored conditions and their difference (Fig. 8a,b).

A Student’s *t* test on SNRs of the ignored condition showed that SNRs were similar across two configurations in each ROI (Fig. 8a, white bars). As shown in Figure 8b, the attentional modulation of the center gratings, measured as SNR difference between attended and ignored condition, did not show any significant difference between the out-of-phase and in-phase configurations in any ROI. Performance correct on both configurations was similar (Fig. 8c). These results suggest that the responses at the centers of both configurations are enhanced by the same degree due to attention. This in turn implies that the lower modulation of the surround in the contiguous configuration is unlikely to be due to an overall reduction in gain across the stimulus.

In summary, our SSVEP results provided strong evidence supporting our predictions that the selectivity profile of visual spatial attention changes depending on the surrounding context and the task while it does not change the gain at peak of profile.

**Discussion**

To understand how the selectivity profile of visual spatial attention changes around an attended target embedded in a textured background, we considered one specific hypothesis: visual attention is more selective to a target in an unsegmented surround than a segmented surround (Fig. 1). We investigated this hypoth-



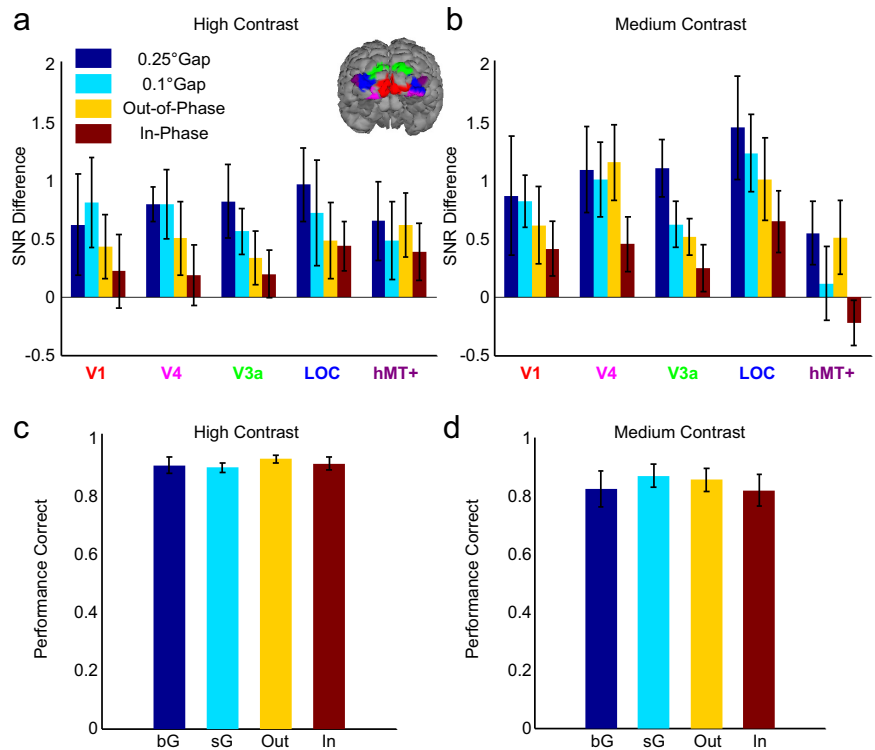
esis at the neural population level by flickering either a surround or a center grating as a “reporter” that monitors either the flanks or the peak of the selectivity profile. We measured the corresponding SSVEPs in five retinotopically defined visual ROIs (V1, V3a, V4, LOC, and hMT+) and demonstrated that attentional modulation of the surround varied with surrounding context whereas that of the center did not, at least for the configurations tested. The spread of attention to the surround was weaker when the target did not readily segment from the surround. Furthermore our data suggest that the extent of the attentional window gradually decreases as the segmentation cue becomes less clear.

What neural mechanisms underlie the configuration-dependent modulation of surround activity due to attention? The spatial profile of attention could change size depending on whether the center-surround configuration is segmented (Fig. 1*a*). Alternatively, attention could enhance the central target region and suppress the irrelevant surround (Fig. 1*b*) with stronger suppression for the unsegmented configuration. Finally, attention could modulate the overall gain profile depending on the surround context, with weaker gain for the contiguous surround (Fig. 1*c*). The latter two models modulate the strength of the component mechanisms without changing their tuning.

Our finding that attentional modulation of the center does not vary with surrounding context runs counter to the global gain model where attention changes the gain of both center and surround depending on the degree of segmentation. But the configuration-dependent modulation of the surround that we observe in most ROIs is consistent with both attention sharpening and surround suppression (Fig. 7*a,b*). The spatial profile of attention can be increasingly focused to better select the target under conditions of decreasing segmentation. Surround suppression can also explain the observed pattern of surround modulation, under the following assumptions: (1) gain of the facilitatory and suppressive regions covary, so that modulation at the peak of the gain field is invariant with center-surround configuration and (2) size of the facilitatory region is always equal or greater than the size of the center grating. If the size of the facilitatory region is equal to the size of the center grating, attention will negatively modulate the surround response. Our results do not rule out the surround suppression hypothesis, although it has to satisfy many constraints to account for our data. This leaves the sharpening hypothesis as the most parsimonious explanation of our data.

### Low-level effects

Could the differential attentional modulation of the surround gratings be due to the well known gain control mechanisms known to operate between adjacent regions? Traditionally neuronal responses to center-surround type stimuli have been explained by a divisive gain control mechanism that divides the



**Figure 7.** The difference of SNRs from the surround gratings difference between the attended and ignored conditions and contrast discrimination performance at both high and medium contrast. The degree of attentional modulation was estimated as the SNR difference between the attended and ignored conditions. *a*, At high contrast, the modulation was significantly higher in areas V1 and V4 for the segmented configurations than for the unsegmented configuration. Areas V1, V4, V3a, and LOC also showed that surround modulation showed a graded pattern that significantly decreased with the degree of segmentation of the surround from the center (Table 2). *b*, At medium contrast, most ROIs except V1 showed a significant differential modulation for segmented versus unsegmented as well as for the graded hypotheses (Table 3). Contrast discrimination performance at high contrast (*c*) and medium contrast (*d*). Performance correct was similar across all four configurations. bG, big (0.25°) gap; sG, small (0.1°) gap; Out, out-of-phase; In, in-phase.

response of a neuron by a weighted sum of the responses of neighboring neurons (Heeger, 1992; Carandini et al., 1997; Schwartz and Simoncelli, 2001). Orientation, size, and contrast level of the stimulus are known to affect gain control. We considered the possibility that low-level stimulus features in the different stimulus configurations might have contributed differentially to the gain control mechanism.

First we consider the effect of target size on surround response. The model predicts that the small central target (in the 0.25° gap configuration) will suppress the surround grating less than bigger central targets (in the 0.1°-gap or no-gap configurations) However, our SSVEP responses from the flickering surround gratings were similar across the four different configurations when targets were ignored. This suggests that the stimulus differences across the four configurations were small enough to leave the strength of suppression from the center to surround unchanged. Our findings are similar to those of Appelbaum et al. (2008) who measured SSVEPs to a center and surround configuration reversing at different frequencies as a function of the gap between them. They showed that the frequency-tagged response to the center, whose size was fixed, was independent of gap size, which parallels our finding that the fixed-size surround evokes similar responses in all configurations regardless of gap size in the ignored condition.

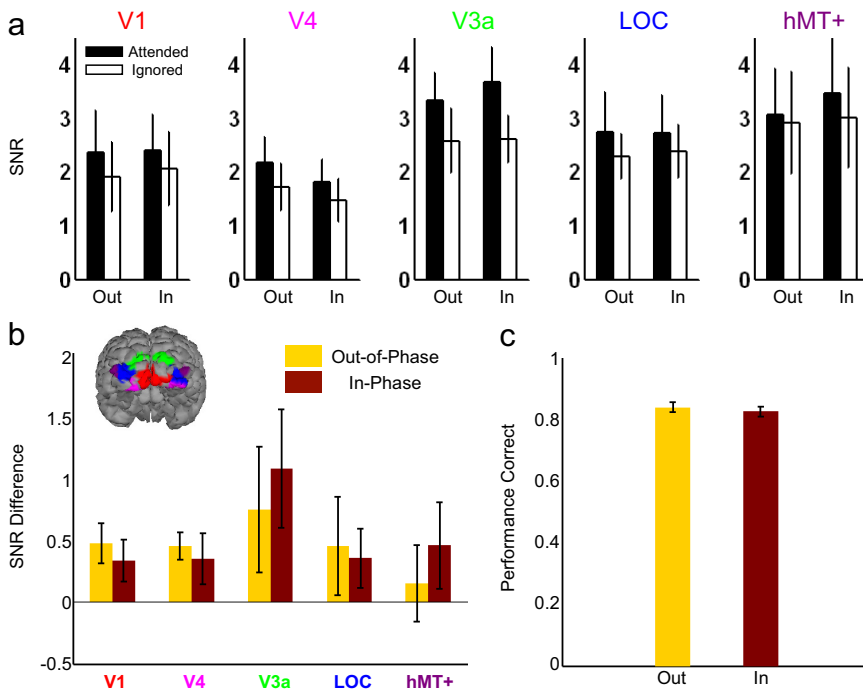
The phase relationship between the center and a collinear surround also affects the neuronal response in areas V1 and V2 of awake behaving monkeys. Maximum and minimum suppression

**Table 2. Trend analysis for source-imaged data at high contrast level**

High contrast	V1	V4	V3a	LOC	hMT+
Segmented versus unsegmented	$F_{(1,27)} = 5.08, p < 0.033$	$F_{(1,27)} = 8.34, p < 0.008$	$F_{(1,27)} = 1.97, n.s.$	$F_{(1,27)} = 2.61, n.s.$	$F_{(1,27)} = 1.23, n.s.$
Graded	$F_{(1,27)} = 5.24, p < 0.033$	$F_{(1,27)} = 5.93, p < 0.022$	$F_{(1,27)} = 9.57, p < 0.005$	$F_{(1,27)} = 6.99, p < 0.014$	$F_{(1,27)} = 0.92, n.s.$

**Table 3. Trend analysis for source-imaged data at medium contrast level**

Medium contrast	V1	V4	V3a	LOC	hMT+
Segmented versus unsegmented	$F_{(1,21)} = 3.13, p = 0.09, n.s.$	$F_{(1,21)} = 10.0, p < 0.005$	$F_{(1,21)} = 6.34, p = 0.02$	$F_{(1,21)} = 8.62, p < 0.008$	$F_{(1,21)} = 9.52, p < 0.006$
Graded	$F_{(1,21)} = 4.10, p = 0.056, n.s.$	$F_{(1,21)} = 5.23, p < 0.033$	$F_{(1,21)} = 12.1, p < 0.0023$	$F_{(1,21)} = 11.7, p < 0.0027$	$F_{(1,21)} = 6.19, p < 0.022$



**Figure 8.** The second-harmonic responses (SNRs) from the center gratings, SNR difference between the attended and ignored conditions, and contrast discrimination performance at high contrast. For each stimulus configuration, the second-harmonic responses to the 16.67 and 12.5 Hz surround gratings in contralateral ROIs were averaged. *a*, The black and white bars represent the SNRs to the flickering center grating when the target was attended and ignored, respectively. The SNRs across all stimulus configurations for the ignored condition (white bars) were not statistically different. Out, out-of-phase; In, in-phase. *b*, Attentional modulation for both configurations in five ROIs, measured as the SNR difference between attended and ignored conditions. Both configurations showed similar levels of attentional modulation for the central target. *c*, Contrast discrimination performance. Performance correct was similar across all both configurations. Out, out-of-phase; In, in-phase.

occurs for relative phase offsets of 0 and 180° (Marcus and Van Essen, 2002; Xu et al., 2005). The phase offsets in our in-phase and out-of-phase conditions were 0 and 90°, respectively, but they generated similar SSVEPs and source responses in the ignored condition. The difference between the single-unit results and our SSVEP results might be due to two factors. First, it is not straightforward to extrapolate from monkey single unit data to human population responses. Second, single-unit studies report the effect of configuration on the center response of a center-surround configuration. The SNR generated by flickering the center in our experiments is low and might be too noisy to reveal the phase-dependent suppression seen in single-cell recording.

In summary, it is unlikely that low-level, center-surround effects can account for the differences in attentional modulation across the segmented versus unsegmented configurations because the responses to the ignored conditions were identical for all four configurations.

### The interaction of attention and segmentation

If attention is implemented as top-down modulation, how does it interact with a demanding task and ongoing mechanisms that are processing the visual image? Neurophysiological research suggests that a response specific to texture segmentation occurs about 60 ms after the initial visual response in area V1, suggesting feedback from higher areas plays a role in texture segmentation (Lamme, 1995; Lamme et al., 1999; Rossi et al., 2001). Specifically, lesions to area V4 impair a monkey's ability to segregate texture (De Weerd et al., 1996; Merigan, 1996). Thus it appears that mid-level areas between V1 and V4 are involved in texture segmentation, and that the delayed activity in V1 is due to feedback from V4. Functional imaging in humans supports the role of these areas in texture segmentation (Kastner et al., 2000; Scholte et al., 2006). Therefore it is not surprising that areas involved in texture processing such as V1 and V4 show a differential modulation of attention depending on surround context. This finding is in accord with previous reports that retinotopically mapped V1 and ventral occipital areas are a physiological basis of the spotlight of visual spatial attention (Tootell et al., 1998; Brefczynski and DeYoe, 1999; Hansen et al., 2007; David et al., 2008). That the differential pattern of

responses appeared when attention interacted with stimulus contexts suggests the role of top-down feedback process such as selective attention may be to improve spatial resolution (Yeshurun and Carrasco, 1998; Boehler et al., 2009, 2011). Thus, when attention is directed to center-surround contexts, a more detailed representation may give rise to a full figure-ground percept (Scholl, 2001).

A recent review also suggests that the unifying function of V4 circuitry is to enable "selective extraction," whether it is by bottom-up feature-specified shape or by attentionally driven spatial or feature-defined selection (Roe et al., 2012). The same V4 network that mediates figure-ground computation may also enable attentional filtering. This perspective is consistent with our findings that show a configuration-dependent modulation of the surround only when observers perform an attentionally demanding task on the cued center, but not when the same target is ignored. Attention is deployed in such a way as to facilitate the

performance of a demanding task at the center by filtering out an irrelevant surround that does not fully segment from the target of interest.

### The effect of task on the spread of attention

One might also wonder why attention did not spread within a perceptually grouped object (Duncan, 1984; He and Nakayama, 1995; Driver and Baylis, 1998; Valdes-Sosa et al., 2000; Driver et al., 2001) as in the in-phase center-surround configuration. Such object-based attention may be useful if the task is to detect a change at an unknown location within the object (He and Nakayama, 1992, 1995). However, as the contrast increment appeared at a fixed location at the center of the stimulus in our studies, an efficient strategy is to exclude the irrelevant surround when it is contiguous with the target. In fact, a study by Boehler et al. (2009) shows the narrowing of the spatial focus of attention occurs only in tasks that require increased spatial resolution. Furthermore, they show that the spatial profile of attention is broad initially and takes about 250 ms to become spatially selective for the target.

Our results indicate that to perform a demanding task efficiently, the spatial profile of attention interacts with scene segmentation to optimize target selection. Our cortical source data strongly suggest that target selection is a distributed process, in which neuronal signals at successive stages of the visual hierarchy, including top-down feedback process such as attention, reflect the global structure of the image as well as the demands of the task.

### References

- Ahlfors SP, Han J, Lin FH, Witzel T, Belliveau JW, Hämäläinen MS, Halgren E (2010) Cancellation of EEG and MEG signals generated by extended and distributed sources. *Hum Brain Mapp* 31:140–149.
- Altmann CF, Deubelius A, Kourtzi Z (2004) Shape saliency modulates contextual processing in the human lateral occipital complex. *J Cogn Neurosci* 16:794–804.
- Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM (2006) Cue-invariant networks for figure and background processing in human visual cortex. *J Neurosci* 26:11695–11708.
- Appelbaum LG, Wade AR, Pettet MW, Vildavski VY, Norcia AM (2008) Figure-ground interaction in the human visual cortex. *J Vis* 8:8 1–19.
- Bahcall DO, Kowler E (1999) Attentional interference at small spatial separations. *Vision Res* 39:71–86.
- Boehler CN, Tsotsos JK, Schoenfeld MA, Heinze HJ, Hopf JM (2009) The center-surround profile of the focus of attention arises from recurrent processing in visual cortex. *Cereb Cortex* 19:982–991.
- Boehler CN, Tsotsos JK, Schoenfeld MA, Heinze HJ, Hopf JM (2011) Neural mechanisms of surround attenuation and distractor competition in visual search. *J Neurosci* 31:5213–5224.
- Brefczynski JA, DeYoe EA (1999) A physiological correlate of the ‘spotlight’ of visual attention. *Nat Neurosci* 2:370–374.
- Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat Neurosci* 8:1102–1109.
- Caplovitz GP, Tse PU (2010) Extrastriate cortical activity reflects segmentation of motion into independent sources. *Neuropsychologia* 48:2699–2708.
- Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17:8621–8644.
- Cardinal RN, Aitken MRF (2006) ANOVA for the behavioural sciences researcher, Chap 6. Mahwah, NJ: Lawrence Erlbaum.
- Chandrasekaran C, Canon V, Dahmen JC, Kourtzi Z, Welchman AE (2007) Neural correlates of disparity-defined shape discrimination in the human brain. *J Neurophysiol* 97:1553–1565.
- Cottareau BR, Ales JM, Norcia AM (2011a) Increasing the accuracy of electromagnetic inverts using functional area source correlation constraints. *Hum Brain Mapp*. Advanced online publication. Retrieved Sept. 21, 2011. doi:10.1002/hbm.21394.
- Cottareau BR, McKee SP, Ales JM, Norcia AM (2011b) Disparity-tuned population responses from human visual cortex. *J Neurosci* 31:954–965.
- David SV, Hayden BY, Mazer JA, Gallant JL (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59:509–521.
- De Weerd P, Desimone R, Ungerleider LG (1996) Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Vis Neurosci* 13:529–538.
- Di Russo F, Spinelli D (1999) Spatial attention has different effects on the magno- and parvocellular pathways. *Neuroreport* 10:2755–2762.
- Driver J, Baylis G (1998) Attention and visual object perception. In: *The attentive brain*, (Parasuraman R, ed), pp 299–326. Cambridge, MA: MIT.
- Driver J, Davis G, Russell C, Turatto M, Freeman E (2001) Segmentation, attention and phenomenal visual objects. *Cognition* 80:61–95.
- Duncan J (1984) Selective attention and the organization of visual information. *J Exp Psychol Gen* 113:501–517.
- 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.
- Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65:413–497.
- Hansen KA, Kay KN, Gallant JL (2007) Topographic organization in and near human visual area V4. *J Neurosci* 27:11896–11911.
- He ZJ, Nakayama K (1992) Surfaces versus features in visual search. *Nature* 359:231–233.
- He ZJ, Nakayama K (1995) Visual attention to surfaces in three-dimensional space. *Proc Natl Acad Sci U S A* 92:11155–11159.
- Heeger DJ (1992) Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9:181–197.
- Herrmann CS (2001) Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Exp Brain Res* 137:346–353.
- Hopf JM, Boehler CN, Luck SJ, Tsotsos JK, Heinze HJ, Schoenfeld MA (2006) Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proc Natl Acad Sci U S A* 103:1053–1058.
- Howell DC (1997) *Statistical methods for psychology*, Ed 4, pp 354–369. Belmont, CA: Duxbury.
- Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 22:7195–7205.
- Kastner S, De Weerd P, Ungerleider LG (2000) Texture segregation in the human visual cortex: a functional MRI study. *J Neurophysiol* 83:2453–2457.
- Kim YJ, Grabowecy M, Paller KA, Muthu K, Suzuki S (2007) Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nat Neurosci* 10:117–125.
- Kim YJ, Grabowecy M, Paller KA, Suzuki S (2011) Differential roles of frequency-following and frequency-doubling visual responses revealed by evoked neural harmonics. *J Cogn Neurosci* 23:1875–1886.
- Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. *J Neurosci* 20:3310–3318.
- Kulikowski JJ, McKeefry DJ, Robson AG (1997) Selective stimulation of colour mechanisms: an empirical perspective. *Spat Vis* 10:379–402.
- Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15:1605–1615.
- Lamme VA, Rodriguez-Rodriguez V, Spekreijse H (1999) Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cereb Cortex* 9:406–413.
- Marcus DS, Van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88:2648–2658.
- Mendola JD, Dale AM, Liu AK, Tootell RBH (1997) The representation of real and illusory contours in human visual cortical areas revealed by {fMRI}. In: *Proceedings of the Society for Neuroscience*, p 550.511. New Orleans, LA.
- Merigan WH (1996) Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. *Vis Neurosci* 13:51–60.
- Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–784.

- Morgan ST, Hansen JC, Hillyard SA (1996) Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci U S A* 93:4770–4774.
- Müller NG, Kleinschmidt A (2004) The attentional ‘spotlight’s’ penumbra: center-surround modulation in striate cortex. *Neuroreport* 15:977–980.
- Myers JL, Well A (2003) *Research design and statistical analysis*, Chap 6, Ed 2. Mahwah, NJ: Lawrence Erlbaum.
- Pascual-Marqui RD, Michel CM, Lehmann D (1994) Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int J Psychophysiol* 18:49–65.
- Pastor MA, Valencia M, Artieda J, Alegre M, Masdeu JC (2007) Topography of cortical activation differs for fundamental and harmonic frequencies of the steady-state visual-evoked responses. An EEG and PET H215O study. *Cereb Cortex* 17:1899–1905.
- Pei F, Pettet MW, Norcia AM (2002) Neural correlates of object-based attention. *J Vis* 2:588–596.
- Rager G, Singer W (1998) The response of cat visual cortex to flicker stimuli of variable frequency. *Eur J Neurosci* 10:1856–1877.
- Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19:1736–1753.
- Roe AW, Chelazzi L, Connor CE, Conway BR, Fujita I, Gallant JL, Lu H, Vanduffel W (2012) Toward a unified theory of visual area V4. *Neuron* 74:12–29.
- Rossi AF, Desimone R, Ungerleider LG (2001) Contextual modulation in primary visual cortex of macaques. *J Neurosci* 21:1698–1709.
- Saupe K, Schröger E, Andersen SK, Müller MM (2009) Neural mechanisms of intermodal sustained selective attention with concurrently presented auditory and visual stimuli. *Front Hum Neurosci* 3:58.
- Scholl BJ (2001) Objects and attention: the state of the art. *Cognition* 80:1–46.
- Scholte HS, Witteveen SC, Spekreijse H, Lamme VA (2006) The influence of inattention on the neural correlates of scene segmentation. *Brain Res* 1076:106–115.
- Schwartz O, Simoncelli EP (2001) Natural signal statistics and sensory gain control. *Nat Neurosci* 4:819–825.
- Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 [Suppl 1]:S208–S219.
- Tootell RB, Hadjikhani N (2001) Where is ‘dorsal V4’ in human visual cortex? Retinotopic, topographic and functional evidence. *Cereb Cortex* 11:298–311.
- Tootell RB, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM (1998) The retinotopy of visual spatial attention. *Neuron* 21:1409–1422.
- Valdes-Sosa M, Cobo A, Pinilla T (2000) Attention to object files defined by transparent motion. *J Exp Psychol Hum Percept Perform* 26:488–505.
- Vialatte FB, Maurice M, Dauwels J, Cichocki A (2010) Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog Neurobiol* 90:418–438.
- Wandell BA, Brewer AA, Dougherty RF (2005) Visual field map clusters in human cortex. *Philos Trans R Soc Lond B Biol Sci* 360:693–707.
- Wang J, Wade AR (2011) Differential attentional modulation of cortical responses to S-cone and luminance stimuli. *J Vis* 11:1.
- Womelsdorf T, Anton-Erxleben K, Pieper F, Treue S (2006) Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat Neurosci* 9:1156–1160.
- Xu WF, Shen ZM, Li CY (2005) Spatial phase sensitivity of V1 neurons in alert monkey. *Cereb Cortex* 15:1697–1702.
- Yeshurun Y, Carrasco M (1998) Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* 396:72–75.