Behavioral/Systems/Cognitive

Interactions of Top-Down and Bottom-Up Mechanisms in Human Visual Cortex

Stephanie McMains and Sabine Kastner

Princeton Neuroscience Institute, Department of Psychology, Princeton University, Princeton, New Jersey 08544

Multiple stimuli present in the visual field at the same time compete for neural representation by mutually suppressing their evoked activity throughout visual cortex, providing a neural correlate for the limited processing capacity of the visual system. Competitive interactions among stimuli can be counteracted by top-down, goal-directed mechanisms such as attention, and by bottom-up, stimulus-driven mechanisms. Because these two processes cooperate in everyday life to bias processing toward behaviorally relevant or particularly salient stimuli, it has proven difficult to study interactions between top-down and bottom-up mechanisms. Here, we used an experimental paradigm in which we first isolated the effects of a bottom-up influence on neural competition by parametrically varying the degree of perceptual grouping in displays that were not attended. Second, we probed the effects of directed attention on the competitive interactions induced with the parametric design. We found that the amount of attentional modulation varied linearly with the degree of competition left unresolved by bottom-up processes, such that attentional modulation was greatest when neural competition was little influenced by bottom-up mechanisms and smallest when competition was strongly influenced by bottom-up mechanisms. These findings suggest that the strength of attentional modulation in the visual system is constrained by the degree to which competitive interactions have been resolved by bottom-up processes related to the segmentation of scenes into candidate objects.

Introduction

Visual scenes are cluttered and contain many different objects. However, the capacity of the visual system to process information about multiple objects at any given moment in time is limited (Broadbent, 1958). Converging evidence from physiology and functional magnetic resonance imaging (fMRI) studies suggest that neural correlates underlying this limited processing capacity are competitive interactions that occur automatically among multiple stimuli present at the same time in the visual field. Multiple stimuli have been shown to compete for neural representation by mutually suppressing their evoked neural activity throughout visual cortex (Miller et al., 1993; Kastner et al., 1998; Reynolds et al., 1999). Competitive interactions among stimuli can be influenced or biased by bottom-up processes that are based on stimulus-driven properties and top-down processes that are determined by the individual's goals (Desimone and Duncan, 1995; Beck and Kastner, 2009).

For instance, salient items in multiple stimuli displays (e.g., pop-out stimuli) have been found to overcome competitive interactions in extrastriate cortex (Reynolds et al., 1999; Beck and Kastner, 2005). Similarly, the perceptual organization of visual items into groups via Gestalt grouping principles has been found to counteract competitive interactions (McMains and Kastner,

2010). Importantly, the influences of visual salience and grouping have been shown to affect neural competition in an automatic fashion, independent of top-down influences such as selective attention. In addition, top-down mechanisms, such as visual spatial attention, can also bias processing in favor of a stimulus occurring at an attended location, counteracting the competitive influences of nearby stimuli in extrastriate visual cortex (Moran and Desimone, 1985; Kastner et al., 1998; Reynolds et al., 1999).

Thus far, the influences of bottom-up and top-down mechanisms on visual processing have been primarily studied separately. However, in real-world scenarios, these two processes dynamically interact to mediate the selection of behaviorally relevant information. Little is known about the nature of these interactive processes because of the difficulty in isolating bottom-up and top-down processes from one another (Folk et al., 1992; Ogawa and Komatsu, 2006). For instance, bottom-up processes are often measured by how much a stimulus interferes with (Theeuwes, 1991; Folk et al., 1992) or facilitates (Treisman and Gelade, 1980; Yeshurun et al., 2009) performance on an attentional task, and salient stimuli are thought to "capture attention."

Here, we sought to isolate bottom-up and top-down processes to ask how they might interact in influencing competitive interactions among multiple stimuli in visual cortex. We hypothesized that first, and consistent with the traditional "spotlight" view of attention (Posner et al., 1980; Eriksen and St James, 1986), the two processes may not be closely dependent on each other. This hypothesis predicts similar attentional modulation regardless of the degree of neural competition that needs to be resolved in any attended display, provided that the size of the spotlight and task difficulty are held constant (Fig. 1A, spotlight hypothesis). Alter-
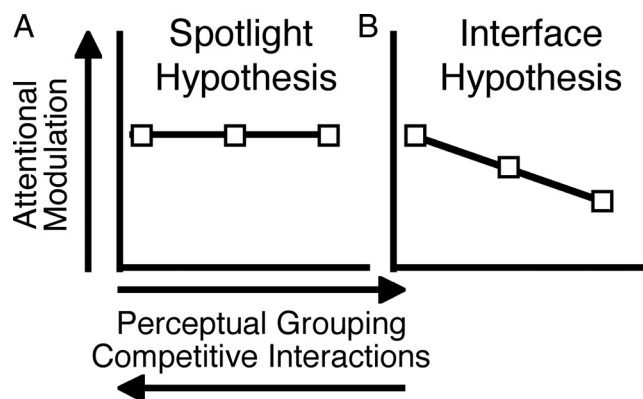
Figure 1. Hypotheses for interactions of top-down and bottom-up processes to resolve neural competition. ***A***, The spotlight hypothesis proposes that attentional modulation will be primarily additive regardless of the amount of neural competition that needs to be resolved. ***B***, The interface hypothesis proposes that the amount of attentional modulation will dependent on the amount of neural competition that has not been resolved through bottom-up processes. Thus, the modulation will be strongest in cases without any perceptual organization and weakest in cases with strong perceptual organization.

natively, attentional modulation may be constrained by the degree to which bottom-up processes have influenced the neural competition. In such an account, top-down operations might use competitive interactions as an interface to counteract residual neural competition that has not been resolved through bottom-up mechanisms (Fig. 1*B*, interface hypothesis) (Qiu et al., 2007). We tested these alternative hypotheses using fMRI while subjects either attended toward or away from multielement displays that differed in the degree of perceptual grouping (McMains and Kastner, 2010).

Materials and Methods

Subjects. Ten subjects (three females; age, 25–37; normal or corrected-to-normal visual acuity) gave informed written consent for participation in the study, which was approved by the Institutional Review Panel of Princeton University. All subjects participated in three scanning sessions each, two sessions for the main experiment, and one session for retinotopic mapping.

Visual display. Visual stimuli were generated on a Macintosh G4 computer (Apple Computer) using Matlab software (The MathWorks) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Stimuli were projected from a PowerLite 7250 liquid crystal display projector (Epson) outside the scanner room onto a translucent screen located at the end of the scanner bore. Subjects viewed the screen at a total pathlength of 60 cm through a mirror attached to the head or surface coil systems. The screen subtended 36° in the horizontal dimension and 28° in the vertical dimension. A trigger pulse from the scanner synchronized the onset of stimulus presentation to the beginning of the image acquisition.

Visual stimuli and experimental design. Visual stimuli consisted of four varin (Varin, 1971) pacman inducers (width of 1.75°) presented in four nearby locations (separated by 0.5°) in the upper right quadrant of the visual field (Fig. 2). The entire stimulus display encompassed 4 × 4° and was centered ~9.5° from fixation. The inducers could be one of several colors and angles (Fig. 2*A–C*; color: red, green, yellow, purple, or cyan; angle: 50, 70, 90, 110, or 130°). Stimuli were presented on a dark gray (1.7 cd/m²) background with an average luminance of 27.9 cd/m².

To investigate how top-down and bottom-up factors interact to influence competition, a 2 × 3 × 2 factorial design was used, in which stimuli were shown under two presentation conditions (sequential and simultaneous) (Fig. 2*D,E*), three levels of grouping (no grouping, intermediate level of grouping, and strong grouping) (Fig. 2*A–C*), and two attentional conditions (attend away and attend to the peripheral stimuli). To probe the amount of competition present within a display, we used the sequential/simultaneous paradigm that has been previously used to assess com-

petitive interactions in visual cortex (Kastner et al., 1998, 2001; Beck and Kastner, 2005, 2007, 2009; McMains and Kastner, 2010; Scalf and Beck, 2010). In the sequential presentation condition of this paradigm (Fig. 2*D*, SEQ), each inducer was presented alone in one of the four locations and in random order for 250 ms. In the simultaneous presentation condition (Fig. 2*E*, SIM), the four inducers were presented in the same locations and for the same time, but they were presented together. In the SIM condition, the stimulus array was randomly presented, either 0, 250, 500, or 750 ms after the beginning of a 1 s presentation period. Integrated over time, the physical stimulation parameters at each of the four locations were identical in the two presentation conditions. However, as shown previously, stimuli could interact with each other in a mutually suppressive way only in the SIM condition (Kastner et al., 1998, 2001; Reynolds et al., 1999; Beck and Kastner, 2005, 2007; McMains and Kastner, 2010; Scalf and Beck, 2010). Such interactions are indicated by a reduction in the evoked response to SIM compared with SEQ presented arrays and are thought to reflect the degree of competition among the stimuli for neural representation.

Previously, competitive interactions have been found to depend on the amount of perceptual grouping among simultaneously presented stimuli. When stimuli formed a perceptual group via illusory contour formation or collinear alignment, competitive interactions were reduced compared with when the stimuli were randomly arranged (McMains and Kastner, 2010). Here, we used the same Kanizsa (1976)-type stimuli as used previously (McMains and Kastner, 2010) to parametrically manipulate the degree of competition present in the stimulus arrays and thus hoped to parametrically manipulate the degree of perceptual grouping. Stimulus arrays contained either a strong perceptual group (Fig. 2*A*, StrongGrp), a weak perceptual group (Fig. 2*B*, WeakGrp), or no perceptual group (Fig. 2*C*, NoGrp). In the StrongGrp condition, the inducers were rotated inward and all consisted of the same angle so that an illusory shape with clearly defined boundaries was present. In the NoGrp condition, the four inducers were randomly rotated outward either 225, 200, 180, 160, or 135° from their positions in the StrongGrp condition; this configuration did not give the appearance of an illusory shape. In the WeakGrp condition, the left two inducers consisted of one angle, whereas the right two inducers consisted of one of the other four possible angles. This resulted in an illusory figure that had ill defined borders between the right and left halves of the figure. Similar configurations have been shown to result in illusory contours perceived somewhere in-between the StrongGrp and NoGrp configurations in terms of perceptual strength (Halko et al., 2008). The stimuli were designed so that, within each run, the same inducer angle and color appeared in a given location for each level of grouping.

Finally, the two presentation conditions and three levels of grouping were presented under two attentional conditions, one in which the peripheral stimuli were ignored (UnATT), and one in which subjects attended to and performed a task on the peripheral stimuli (ATT). This design allowed us to investigate how top-down attentional and bottom-up grouping processes interacted to influence the degree of competition, as indexed by the signal differences evoked by the SEQ and SIM presentations for the two attentional conditions. In the UnATT condition, subjects performed a demanding letter discrimination task at fixation, which was designed to ensure central fixation and to prevent subjects from allocating attention to the peripheral displays. Subjects monitored a rapid serial visual presentation (RSVP) stream, which consisted of letters (KNXPVTHZLAEM) appearing in white, orange, or purple font for the appearance of a target letter (purple K or orange N). Targets occurred on average every 2.3 s, and each letter averaged a height of 0.47° and a width of 0.43°. The duration of each letter presentation was varied for individual subjects to keep behavioral performance at ~75% correct (range, 110–176 ms; average, 130 ms). In the ATT condition, subjects performed a luminance detection task on the lower left inducer closest to fixation. The inducer dimmed on average every 2.5 s, began to dim 50 ms into the inducer presentation, and the dimming lasted for 90 ms. The amount by which the inducer dimmed was evaluated for each subject and presentation condition (SEQ, SIM) in a staircase procedure (two-down, one-up) to keep behavioral performance at ~75% correct

throughout the scan session. Subjects received one and one-half or more hours of training before scanning.

Stimuli were presented in blocks lasting 20 s, thus consisting of 20 different stimulus display presentations. Blocks of peripheral stimulation were interleaved with fixation blocks of 17.5 s during which subjects performed the fixation task. The attentional conditions were blocked by run. To help subjects keep track of which task they were performing within a run, a thin white ring (diameter, 0.9°) was presented around the fixation point, whenever subjects were performing the fixation task. When subjects attended to the peripheral stimuli, the ring disappeared and the RSVP stream turned pink. Before each ATT block, subjects received a cue, a thin white line pointing toward the location to be attended (subtending 0.5°), that lasted for 750 ms. Each run consisted of nine blocks and began and ended with a fixation period for a total run length of 2 min 47.5 s. Within a given run, two levels of grouping were presented both in SEQ and SIM presentation conditions. Presentation conditions were presented in an ABBA block order (SEQ–SIM–SIM–SEQ) with perceptual grouping condition counterbalanced across runs (Kastner et al., 1998). There were 18 runs in a given scan session. Each subjects participated in two scanning sessions for a total of 12 blocks per condition.

Data acquisition and analysis. Data for each subject were acquired in three separate scanning sessions each lasting ~1.5 h using a 3T Siemens Allegra scanner (Allegra; Siemens). An anatomical scan [magnetization-prepared rapid-acquisition gradient echo (MPRAGE) sequence; repetition time (TR), 2.5 s; echo time (TE), 4.38 s; flip angle, 8°; 1 mm³ resolution] was acquired in each session to facilitate cortical surface alignments. For cortical surface reconstructions, two high-resolution structural scans (MPRAGE, same parameters as above) were acquired in a separate session and averaged.

For all studies, functional images were taken with a gradient echo, echo-planar sequence with a 128 square matrix (retinotopic mapping: slice thickness of 2 mm, with a 1 mm gap; interleaved acquisition; field of view, 256 × 256 mm²; TR, 2.5 s; TE, 40 ms; flip angle, 90°; with a partial Fourier factor of 7/8 used to acquire asymmetric fraction of k-space to reduce the acquisition time; main experiment: same parameters, but coronal slices). For retinotopic mapping, 25 axial slices covering the occipital lobe were acquired in three series of 140 volumes using a standard birdcage coil. For the main experiment, 25 contiguous, coronal slices covering occipital and posterior parietal cortex were acquired in 36 series of 67 volumes using a four-channel visual surface coil (Nova Medical). An in-plane magnetic field map image was acquired to perform echo-planar imaging undistortion (TR, 0.5 s; TE, 5.23 or 7.69 s; flip angle, 55°; 2 mm slices; in-plane resolution, 2 × 2 mm).

Data were analyzed using AFNI (including SUMA) (Cox, 1996) (http://afni.nimh.nih.gov/afni/; http://afni.nimh.nih.gov/afni/suma), Matlab (The MathWorks), and FreeSurfer (Dale et al., 1999; Fischl et al., 1999). The functional images were motion corrected to the image acquired closest in time to the anatomical scan and undistorted using the field map scan. Images were spatially smoothed in-plane with a Gaussian filter of 2 mm. The first six images of each scan were excluded from analysis. For the main experiment, statistical analyses were performed using multiple regression in the framework of the general linear model (Friston et al., 1995) with AFNI. Square-wave functions contrasted blocks of UnATT peripheral visual presentations (regardless of presentation or perceptual grouping condition) versus blank periods. Square-wave functions were convolved with a Gaussian model of the hemodynamic response (lag, 2 s; dispersion, 1.8 s) to generate idealized response functions that were used as regressors in the multiple regression model. Additional regressors were used to factor out within-run linear drifts, quadratic drifts, and head movement artifacts. Statistical maps comparing peripheral stimulation blocks versus blank periods were thresholded at $p < 0.0001$ or less (uncorrected for multiple comparisons) such that the comparison revealed only voxels activated by the peripheral stimuli, excluding foveal voxels activated by the RSVP stream at fixation. Functional data were projected onto cortical surface reconstructions created with FreeSurfer that were aligned to each experimental session using AFNI/SUMA. All voxels that fell between the gray- and white-matter boundaries were mapped to the surface.

Activated voxels were assigned to regions of interest (ROIs) in visual cortex, as defined below. Time series of fMRI intensities were extracted from each ROI for each run from unsmoothed data. Time series were normalized to the last two time points of the preceding fixation block. Individual time courses for each subject, area, and condition were investigated for outliers by calculating the sum-squared error. Peripheral presentation blocks in which this value exceeded 2 SDs were excluded, resulting in the exclusion of less than one block per subject, area, and condition (average, 0.81 blocks). Mean time series for each condition were calculated by averaging the activity evoked by the blocks that met this criterion across runs and scan sessions. For each subject, mean signals were computed by averaging across the six peak time points (7.5–20 s) of the average time series, for each condition and visual area. These values were further quantified by defining a sensory suppression index [SSI = $(R_{SEQ} - R_{SIM})/(R_{SEQ} + R_{SIM})$; R represents response computed as mean signal change during the two different presentation conditions] for the three perceptual organization conditions and the two attentional conditions. The SSI quantifies the differences in responses evoked by SEQ and SIM presentations. Positive values indicate stronger responses evoked by SEQ than by SIM presentations (reflecting the mutual inhibition that is observed when nearby stimuli compete for representation), negative values indicate the opposite, and values ~0 indicate the absence of response differences (or no difference in the amount of competition elicited by the two presentation conditions). In addition, an attentional modulation index was computed [AMI = $(R_{ATT} - R_{UnATT})/(R_{ATT} + R_{UnATT})$]. The AMI quantifies the amount of attentional modulation, with positive values indicating enhanced responses when attention was directed to the stimulus, negative values indicating the opposite, and values ~0 indicating no attentional modulation.

Statistical significance of the mean signal changes, SSIs, AMIs, and behavioral accuracy were assessed using within-subject repeated-measures ANOVAs. Linear and quadratic contrasts were used to identify systematic relationships between levels of perceptual grouping and blood oxygen level-dependent (BOLD) modulations. Importantly, a significant linear contrast without a significant quadratic trend implies a strict monotonic relationship between grouping and BOLD signal changes. Whenever a significant linear contrast is reported, a nonsignificant quadratic contrast ($p > 0.05$) was observed unless otherwise stated.

Definitions of ROIs and retinotopic mapping. Retinotopic mapping was performed for each subject in a separate scanning session using color and luminance varying flickering checkerboard stimuli (Swisher et al., 2007; Arcaro et al., 2009). Data were analyzed using standard phase encoding techniques (Sereno et al., 1995; Engel et al., 1997; Schneider et al., 2004; Arcaro et al., 2009). Five retinotopic areas (V1, V2, V3, V3a, V4) (Wade et al., 2002; Arcaro et al., 2009) were identified by the alternating representations of the vertical and horizontal meridians, which form the borders of these areas. Given that the stimuli were presented in the upper right visual quadrant, ROIs were restricted to the upper visual field representation of each visual area in the left hemisphere. After each region was identified, ROIs were created by taking the voxels activated by the contrast all UnATT blocks versus fixation for each region. This resulted in an ROI that encompassed the retinotopic representation of all four inducers. This was done because of the difficulty in localizing responses to a single inducer given the peripheral presentation (9°) and small spatial separation (0.5°) of the inducers. Although ideally we would have liked to measure responses to just the lower left inducer, previous findings suggest that measuring responses to the entire stimulus array are similar to measuring responses to one of the four locations in isolation (Kastner et al., 1998). However, given that the lower left inducer is closest to fixation, and thus presented at a less peripheral eccentricity, it is likely that the responses evoked by the array were dominated by activity evoked by the inducer closest to fixation, which served as the attended inducer in the ATT periphery conditions.

Results

We investigated the interaction of bottom-up and top-down processes, specifically with respect to how they influence competitive interactions that occur among simultaneously presented stimuli. First, competitive interactions among multiple stimuli were as-
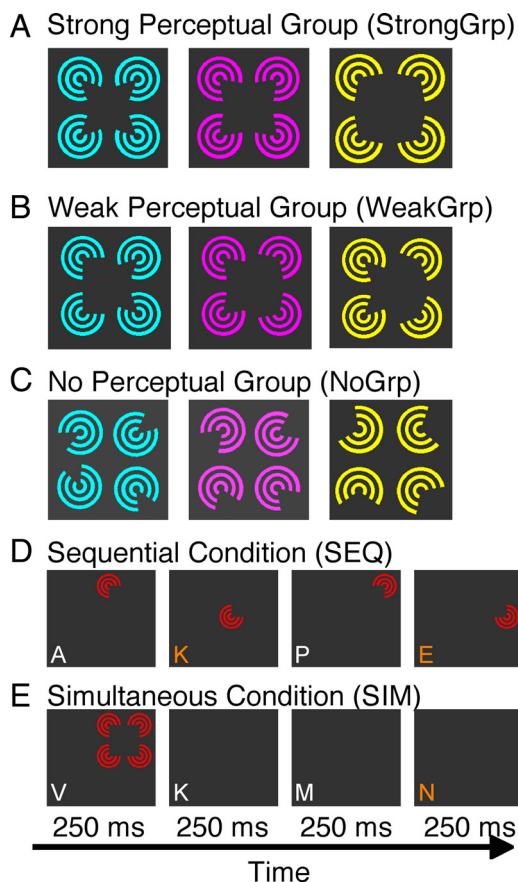
Figure 2. Experimental design and visual stimuli. Four illusory contour inducers were presented under three display conditions: with the inducers rotated inward and with the same angle to form a strong illusory figure (**A**); with the inducers rotated inward, but the left two inducers and right two inducers having different angles, thereby forming a weak percept of an illusory figure (**B**); or with the inducers rotated outward so that no illusory figure was present (**C**). The stimulus display was presented in the upper right visual quadrant using two presentation conditions. **D**, In the sequential condition, each inducer was presented alone for 250 ms each. **E**, In the simultaneous condition, all four inducers were presented at the same time for 250 ms each. Each presentation period lasted 1 s, with 20 presentations in a peripheral stimulation block. On average, a stimulus appeared at each of the four locations every 750 ms. Presentation conditions and levels of perceptual organization were probed while subjects either performed a task at fixation (and ignored the peripheral stimuli), or attended to the peripheral stimuli and performed a task on the lower left inducer closest to fixation.

sessed by using a previously established experimental paradigm, in which multiple stimuli were presented either sequentially or simultaneously to the periphery of the visual field (Kastner et al., 1998, 2001; Beck and Kastner, 2005, 2007). As shown previously (Kastner et al., 1998, 2001; Beck and Kastner, 2005, 2007), competitive interactions among the stimuli could only take place in the simultaneous condition. The degree of competition among stimuli was determined by measuring the differences in responses evoked by the sequential and simultaneous presentations.

Second, the influence of bottom-up perceptual grouping processes on competitive interactions was investigated by parametrically varying the degree of perceptual organization present within the stimulus arrays (Fig. 2A–C). Kanizsa (1976)-type stimuli were used to manipulate the degree of perceptual organization via illusory contour formation present among elements in the stimulus array, while changing the overall arrangement of the individual elements only minimally. In the StrongGrp condition, inducers were aligned such that an illusory figure (e.g., a square) was formed in the SIM condition (Fig. 2A), thus linking the four

inducers together to form a single foreground object by a complex set of processes recruited during illusory contour formation, including visual segmentation, figure–ground assignment, depth perception, and visual interpolation (Palmer, 1999). Previously, stimuli such as these were found to reduce competition in visual cortex compared with when the same inducers were rotated outward, so that no figure was present and the four elements appeared as four single foreground objects with different orientations when simultaneously presented (Fig. 2C, NoGrp) (McMains and Kastner, 2010). To parametrically manipulate the degree of perceptual grouping, a third intermediate condition was created. In the "weak perceptual group" (WeakGrp) condition, the left two inducers were aligned with each other but misaligned with the right two inducers (Fig. 2B), reducing the clarity of the illusory contours and weakening the percept of single foreground object (Halko et al., 2008). To assess the degree of competition, all three perceptual organization conditions were presented simultaneously and sequentially.

To investigate the effects of top-down attentional modulation, the stimulus arrays were presented under two attentional conditions. In the unattended condition (UnATT), subjects ignored the peripheral displays and performed a demanding RSVP task at fixation. In the attended condition (ATT), subjects attended to the lower left inducer closest to fixation and performed a luminance detection task. To investigate whether top-down and bottom-up processes interacted with each other, we tested the interface hypothesis (Fig. 1B), which assumes that the degree of top-down attentional modulation should be dependent on the amount of competition left unresolved by bottom-up processes against the spotlight hypothesis (Fig. 1A), which proposes that attentional modulation should be independent of bottom-up processes, and instead will depend on factors such as the location of the attentional spotlight and task difficulty.

Unattended conditions

The unattended stimulus array evoked robust activity throughout visual cortex, including early visual areas V1, V2, V3, and area V4 of the ventral stream. Notably, the lateral occipital complex (LOC), an area involved in object and contour processing (Malach et al., 1995), was not robustly activated by the illusory contour stimuli. Although this area has been implicated in the processing of illusory contours (Mendola et al., 1999), it has also been shown to prefer object stimuli presented at the fovea (Sayres and Grill-Spector, 2008). Our stimuli, which were presented in the more peripheral parts of the visual field, were therefore not ideal for activating LOC.

Based on previous findings (Kastner et al., 1998), we predicted that activity evoked by the stimulus display in the SIM condition would be smaller than that evoked by the SEQ condition throughout visual cortex, particularly in intermediate areas such as V4, reflecting the suppressive interactions that occur mainly at the level of the receptive field (RF) when multiple objects compete for neural representation (Kastner et al., 1998, 2001). In support of our hypothesis, the average mean signal changes (Fig. 3A) obtained in the SIM conditions were reduced compared with those obtained in the SEQ conditions for all levels of grouping and in all visual areas under investigation (V1, V2, V3, V4: main effect of presentation: all $F_{(1,9)} > 19.79$, $p < 0.01$). Next, we investigated whether bottom-up perceptual grouping processes interacted with competition. Previously, it has been shown that ungrouped elements competed more than grouped elements (McMains and Kastner, 2010). This difference in the amount of competition was driven by activity during the SIM conditions, in

A

B
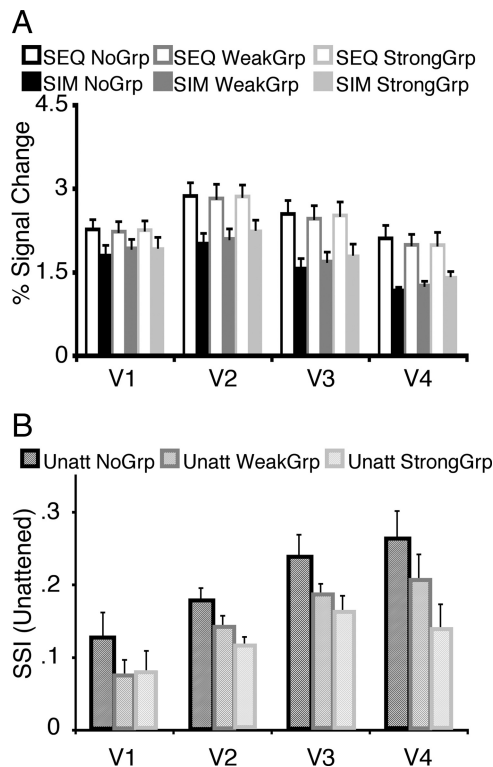
Figure 3. Mean signal changes and sensory suppression indexes for the unattended conditions. ***A***, For each subject, area, and condition, mean signal changes were computed by averaging across the six peak time points of the fMRI time series. Group data were yielded by averaging across subjects ($n = 10$). In general, visually evoked activity was reduced for the SIM (filled bars) compared with the SEQ (open bars) conditions, reflecting the mutual suppression that occurs when multiple stimuli compete. In addition, modulations for just the UnATT SIM conditions varied monotonically with the amount of perceptual organization present in V2, V3, and V4. ***B***, Sensory suppression indexes were calculated on the basis of mean signal changes [(SEQ − SIM)/(SEQ + SIM)]. SSIs in early visual cortex (V1, V2, V3) and area V4 of the ventral stream were largest when the stimuli were randomly arranged as in the NoGrp condition. As predicted, SSIs varied linearly with the strength of the perceptual grouping among the elements in the display in V2, V3, and V4. The vertical bars indicate SEM.

which grouping among the elements could take place, whereas activity was similar for the SEQ conditions, in which only one element was present at a time. Based on these previous results, we made two predictions. First, activity should not differ for the different levels of perceptual grouping in the SEQ conditions, given that only one inducer was present at a time, and varied only minimally in terms of low-level features. In support of this hypothesis, there was no significant effect of grouping in the SEQ condition in any area (all $F_{(2,9)} < 1.43$, $p > 0.27$). Second, we predicted that activity evoked by the SIM conditions should depend on the amount of grouping among the elements, as indexed by an interaction between perceptual grouping and presentation condition. Consistent with this prediction, a significant interaction of perceptual grouping (NoGrp, WeakGrp, StrongGrp) and presentation condition (SEQ, SIM) was found in V2, V3, and V4 (all $F_{(2,18)} > 6.26$, $p < 0.01$). Importantly, when investigating BOLD signal modulations for just the UnATT SIM conditions, a significant linear contrast was found in V2, V3, and V4 (all $F_{(1,9)} > 8.87$, $p < 0.05$), reflecting a monotonic relationship between the amount of perceptual organization and sensory suppression, such that the activity evoked by the NoGrp display was smallest, followed by the WeakGrp display, with the StrongGrp display evoking the most activity. The lack of any differences among the UnATT SEQ conditions suggests that the modulations observed for the SIM con-

ditions were not attributable to low-level changes in the stimulus that were also present in the SEQ conditions (e.g., inducer rotated inward or outward), but instead to the perceptual grouping that occurred among the simultaneously presented inducers.

To quantify the differences in responses evoked by SIM and SEQ presentations further, a SSI was calculated (Fig. 3*B*). The index permits a comparison of the degree of competition effects both across different visual areas and perceptual organization conditions. Positive values indicate stronger responses evoked by SEQ than by SIM presentations (reflecting the mutual suppression that occurs during the simultaneous presentation), negative values indicate the opposite, and values ∼0 indicate the absence of response differences (or no difference in the amount of competition elicited by the two presentation conditions). The sensory suppression indexes varied depending on the amount of perceptual grouping present in the displays in V2, V3, and V4 (all $F_{(2,9)} > 9.58$, $p < 0.01$), such that SSIs varied linearly with the amount of perceptual organization among elements in the display (all $F_{(1,9)} > 16.55$, $p < 0.01$). The largest sensory suppression indexes were found for the NoGrp displays ($SSI_{UnATTNoGrp}$), followed by the WeakGrp displays ($SSI_{UnATTWeakGrp}$), whereas the smallest indexes were found for the StrongGrp displays ($SSI_{UnATTStrongGrp}$). This pattern of activity is consistent with our prediction that competition varies parametrically with the amount of perceptual grouping present in the SIM displays. These results confirm the close relationship between perceptual grouping and competition, further supporting the finding that competitive interactions can be partially overcome when the stimuli form a perceptual group via illusory contour formation (McMains and Kastner, 2010).
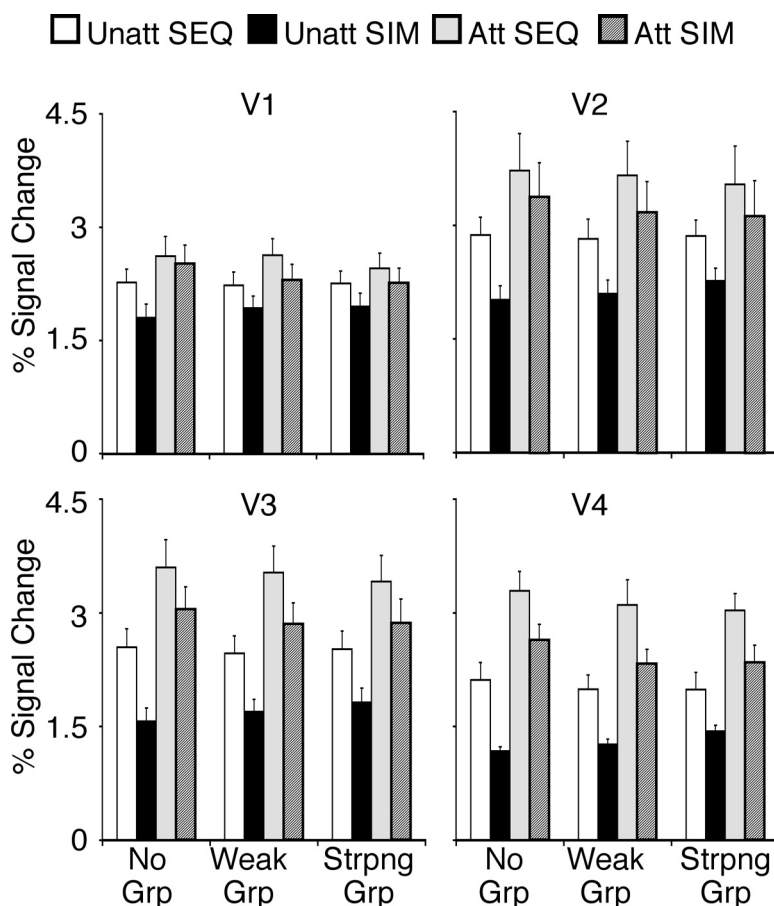
Attended conditions

After establishing the parametric manipulation of bottom-up processes related to grouping and competition, we next investigated the interaction of top-down attention effects and bottom-up processes. When subjects attended to the peripheral displays, we predicted based on previous findings (Kastner et al., 1998; Brefczynski and DeYoe, 1999) that BOLD signal modulations should increase throughout visual cortex in the retinotopic location corresponding to the peripheral stimulus. Accordingly, we found greater activity for the attended compared with the unattended conditions (Fig. 4), reflected as a significant main effect of attention in all areas (V1, V2, V3, V4: all $F_{(1,9)} > 9.35$, $p < 0.05$).

To investigate the amount of attentional enhancement, activity for the ATT conditions was compared with that evoked by the same stimulus when subjects attended to the fixation task and ignored the peripheral stimuli. For that purpose, an AMI was calculated that compares the activity obtained in the attended condition to that obtained in the unattended condition (Fig. 5). The index permits a comparison of the degree of attentional modulation across different visual areas, levels of perceptual grouping, and presentation conditions. Positive values indicate stronger responses evoked by the attended than by the unattended presentations (reflecting the enhancement often associated with directed attention), negative values indicate the opposite, and values ∼0 indicate the absence of response differences (or no effect of attention). Significant attentional modulation in terms of response enhancement was observed in all areas and for all conditions except in area V1 for the SEQ StrongPO condition (all $t > 2.6$, $p < 0.05$). Using the AMI analyses, we tested our specific hypotheses regarding the interaction of top-down attention and bottom-up grouping processes (Fig. 1). The critical conditions to investigate were the SIM conditions, in which grouping among

the elements occurred. When comparing the AMIs for the SIM conditions (AMIs$_{SIM}$), a significant main effect of perceptual grouping was found in all areas (Fig. 5, gray lines) (all areas $F_{(2,9)} > 6.57$, $p < 0.01$). In fact, the AMIs$_{SIM}$ varied linearly with the degree of perceptual grouping, as reflected by a significant linear contrast (all $F_{(1,9)} > 8.62$, $p < 0.05$), suggesting a close relationship between the amount of attentional modulation in visual cortex and the amount of bottom-up grouping processes. Importantly, AMIs for the SEQ conditions (AMIs$_{SEQ}$) did not vary (all $F_{(2,9)} < 1.72$, $p > 0.21$), suggesting that the differences in the AMIs$_{SIM}$ found above were not attributable to low-level differences in the stimulus displays (Fig. 5, black lines). These results are consistent with the interface hypothesis (Fig. 1B), which predicts that attentional modulation varies depending on bottom-up grouping processes. However, given the dependency of visual responses on perceptual grouping for the UnATT SIM conditions, one alternative hypothesis for the observed relationship between perceptual grouping and attentional modulation might be that it resulted from a ceiling effect of the BOLD responses for the ATT SIM conditions given the similar absolute values of the BOLD signal modulations for the ATT SIM conditions (Fig. 4). However, it is clear from Figure 4 that the ATT SEQ condition was greater than the ATT SIM condition, suggesting that larger BOLD signal modulations were possible in visual cortex. In addition, evidence from previous fMRI studies investigating the effects of attention on visual responses over a wide range of stimulus contrast have generally found that attention serves to increase responses additively, that is, additive attention effects can even be found on stimuli of maximum contrast that provide a strong bottom-up drive (Buracas and Boynton, 2007; Li et al., 2008; Murray, 2008). Together, this evidence is opposed to the idea that attentional modulation as measured with fMRI BOLD is constrained by some upper limit or ceiling in absolute evoked response.

Next, we investigated whether attention interacted with bottom-up competitive processes. Consistent with the idea that attentional modulation is dependent on the degree of competition, AMIs were found to be greater for SIM conditions, in which competition among the elements in the display could take place compared with the SEQ conditions in which only one element was present at a time (main effect of presentation: V2, V3, V4: all $F_{(1,9)} > 7.05$, $p < 0.05$; V1: $F_{(1,9)} = 4.3$, $p = 0.07$). However, to investigate in more detail how attention modulated competition, we compared the sensory suppression effects for the different levels of grouping. To do this, SSIs were computed for the ATT (Fig. 6, black dashed lines) and UnATT conditions (Fig. 6, gray dashed lines). The SSIs$_{ATT}$ were reduced compared with the SSIs$_{UnATT}$ (main effect of attention: V2, V3, V4: all $F_{(1,9)} > 6.87$, $p < 0.05$; V1: $F_{(1,9)} = 4.32$, $p < 0.07$), suggesting that directing


Figure 4. Mean signal changes for the unattended and attended conditions. For each subject, area, and condition, mean signal changes were computed by averaging across the six peak time points of the fMRI time series. Group data were yielded by averaging across subjects ($n = 10$). Mean signals are presented for V1, V2, V3, and V4. When subjects attended to the peripheral stimuli (gray and striped bars), activity throughout visual cortex increased compared with when subjects attended away from the peripheral stimuli and performed a task at fixation (white and black bars). The vertical bars indicate SEM.

attention to the stimulus array counteracted competition. In addition, there was a significant interaction of attention and perceptual grouping in V2, V3, and V4 (all $F_{(2,18)} = 4.28$, $p < 0.05$), reflecting that sensory suppression effects were reduced by the largest amount in the NoGrp condition and the smallest amount in the StrongGrp condition. Consistent with this finding, SSIs$_{ATT}$ were significantly reduced compared with SSIs$_{UnATT}$ for the NoGrp displays in V1, V2, V3, V4 (all $t > 3.77$, $p < 0.01$), for the WeakGrp displays in V2 and V3 (both $t > 4.34$, $p < 0.01$), but not for the StrongGrp displays in any area (all $t < 2.09$, $p > 0.07$). These results are in agreement with the hypothesis from biased competition theory that attention resolves competition and thus has the largest effects when there is a large amount of competition left unresolved by bottom-up grouping processes. Interestingly, the SSIs$_{ATT}$ were similar for all levels of perceptual grouping (all areas: $F_{(1,9)} < 1.64$, $p > 0.22$), suggesting that bottom-up and top-down processes interact dynamically to resolve as much competition in each display as possible. In addition, these results suggest a dependency between bottom-up competitive processes and top-down attention, in support of the interface hypothesis.

To investigate the relationship between bottom-up competitive processes and top-down attention further, attentional modulation for the SIM conditions (AMI$_{SIM}$) was correlated with
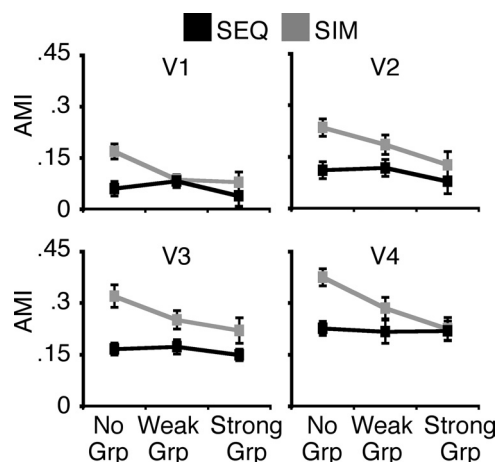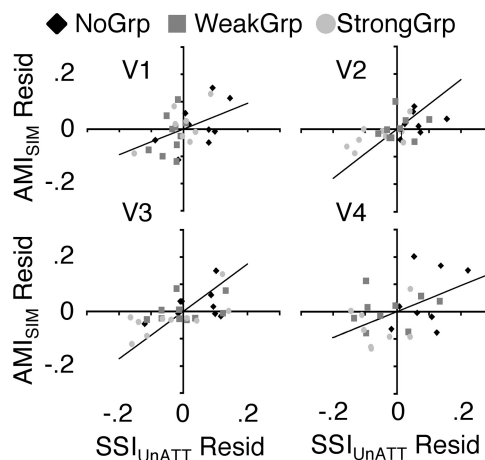
Figure 5. Attentional modulation indexes for SEQ and SIM presentation conditions as a function of perceptual organization. Attentional modulation indexes [(ATT − UnATT)/(ATT + UnATT)] were calculated for each level of grouping for the SEQ (black lines) and SIM (gray lines) conditions. Consistent with the interface hypothesis, attentional modulation depended on display condition for SIM presented displays where competition among elements could occur, but not for the SEQ displays. For all visual areas, there was a significant linear relationship between attentional modulation and perceptual grouping for SIM displays. In addition, attentional modulation was generally larger for the SIM compared with the SEQ conditions. The vertical bars indicate SEM.
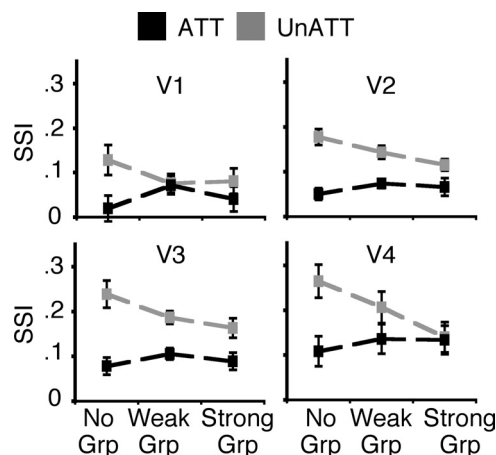
Figure 6. Sensory suppression indexes for the attended and unattended conditions as a function of perceptual organization. Sensory suppression indexes [(SEQ − SIM)/(SEQ + SIM)] were computed for attended (black dashed lines) and unattended (gray dashed lines) conditions for each level of grouping. Sensory suppression was reduced when subjects attended to the display compared with when subjects attended away from the displays. SSIs were reduced by the largest amount for the NoGrp displays and the smallest amount for the StrongGrp displays. The final amount of competition present when top-down attention and bottom-up perceptual grouping combined to counteract competition did not differ for the different display conditions (SSIs$_{ATT}$).

sensory suppression effects for the UnATT conditions (SSI$_{UnATT}$) for each subject and each perceptual grouping condition. Given that both indexes were calculated using the UnATT SIM condition, the two indexes might be inherently correlated. To correct for this, each index was first correlated with the UnATT SIM activity and then the residuals were correlated against each other (Fig. 7). A significant correlation between attentional modulation was observed for all visual areas (V1–V3: $t > 3.03$, $p < 0.01$; V4: $t = 2.35$, $p < 0.05$). This result suggests that the degree of attentional modulation within individual subjects is closely tied to the degree of competition for all levels of grouping.

Figure 7. Correlations between sensory suppression and attention effects. Attentional modulation indexes from the SIM conditions (AMI$_{SIM}$) and sensory suppression indexes from the UnATT conditions (SSI$_{UnATT}$) for each subject and level of grouping were correlated with the respective UnATT SIM to eliminate any inherent correlation because of the fact that each index contains the UnATT SIM condition. Next, the residuals were correlated to investigate the relationship between attention and sensory suppression effects. Values were plotted for the NoGrp (light gray circles), WeakGrp (gray squares), and StrongGrp (black diamonds) for all subjects, excluding outliers that fell >2 SDs away from the mean. A significant correlation between attentional modulation and sensory suppression effects was observed for all visual areas (V1–V3: $t > 3.03$, $p < 0.01$; V4: $t = 2.35$, $p < 0.05$). This suggests that the degree of attentional modulation within individual subjects is closely tied to the degree of competition for all levels of grouping.

Behavioral results

During the UnATT conditions, when subjects were engaged in a demanding task at fixation and the stimulus array was presented in the periphery of the visual field, the influence of perceptual organization on competition presumably occurred in a highly automatic stimulus-driven fashion (McMains and Kastner, 2010). However, it is possible that stimulus arrays with a perceptual group capture attention more strongly than arrays without a perceptual group, resulting in a redeployment of attention to the periphery (Yantis, 2000). If so, it might be argued that the interaction of competition and perceptual organization observed during the UnATT conditions was attributable to attentional modulation instead of bottom-up stimulus-driven effects. However, if attention was redirected to the periphery when a perceptual group was present, attention would be withdrawn from the demanding fixation task at the same time, resulting in a decrease in behavioral performance when the stimulus array contained a perceptual group (i.e., the SIM StrongGrp and WeakGrp conditions). To investigate whether the interaction of perceptual organization and competition was uniquely affected by attentional processes, behavioral performance for the three SIM$_{UNATT}$ conditions was compared. There were no differences in accuracy (Table 1: StrongGrp, 72%; WeakGrp, 73%; NoGrp, 72%; $F_{(2,18)} = 0.4$, $p = 0.67$) or reaction times (Table 1: StrongGrp, 554 ms; WeakGrp, 553 ms; NoGrp, 553 ms; $F_{(2,18)} = 0.04$, $p = 0.96$) for the three UnATT SIM conditions. As a more stringent test of a redeployment of attention, behavioral performance was investigated for the UnATT conditions for trials in which the grouped stimulus array appeared in the periphery at the exact same time that a RSVP target appeared at fixation. Similarly, there were no differences among the conditions in terms of accuracy (StrongGrp, 77%; WeakGrp, 73%; NoGrp, 71%; $F_{(2,18)} = 0.97$, $p = 0.4$) or reaction times (StrongGrp, 554 ms; WeakGrp, 550 ms; NoGrp, 557 ms; $F_{(2,18)} = 0.3$, $p = 0.74$). These behavioral results suggest

Table 1. Accuracy and reaction time for behavioral task

Condition	Accuracy	RT (ms)
UnATT SEQ NoGrp	75 (3)	548 (16)
UnATT SIM NoGrp	72 (4)	553 (13)
UnATT SEQ WeakGrp	76 (3)	544 (14)
UnATT SIM WeakGrp	73 (4)	553 (14)
UnATT SEQ StrongGrp	74 (3)	546 (15)
UnATT SIM StrongGrp	72 (4)	554 (13)
ATT SEQ NoGrp	73 (1)	489 (13)
ATT SIM NoGrp	70 (1)	512 (19)
ATT SEQ WeakGrp	77 (2)	485 (8)
ATT SIM WeakGrp	71 (1)	507 (18)
ATT SEQ StrongGrp	79 (3)	476 (8)
ATT SIM StrongGrp	73 (1)	496 (17)

RT, Reaction time.

that the observed influences of perceptual organization on competition were not attributable to attentional modulation resulting from attentional capture. These results are consistent with previous findings using similar illusory contour stimuli and showing that the influences of illusory contour formation on competitive interactions were mediated by bottom-up stimulus-driven processes that occurred in a highly automatic fashion (McMains and Kastner, 2010).

When subjects attended to the peripheral displays, differential activity was also observed for the different levels of grouping, consistent with the hypothesis that attention operates on neural competition dependent on the level of competition that needs to be resolved. However, it is possible that the presence of a perceptual group in the stimulus array influenced the difficulty of the luminance detection task. To keep behavioral performance similar for the different stimulus arrays, task difficulty was adjusted throughout each scan session using a staircase procedure (two-down one-up). To confirm that behavioral performance was equated, we investigated accuracy for the three ATT SIM conditions. No difference in performance was found (Table 1: StrongGrp, 73%; WeakGrp, 71%; NoGrp, 70%; $F_{(2,9)} = 0.91$, $p = 0.42$). These results suggest that the observed interaction of attentional modulation and perceptual grouping was attributable to the differential amount of competition present in the displays and not to a difference in task difficulty. For the ATT SEQ conditions, there was a trend for subjects to perform worse on the luminance detection task in the NoGrp condition (Table 1: StrongGrp, 79%; WeakGrp, 77%; NoGrp, 73%; $F_{(2,9)} = 3.34$, $p = 0.06$). Importantly, this difference was not attributable to the presence or absence of a perceptual group as only one inducer was present during the SEQ conditions, and instead may reflect the fact that the inducer was rotated outward in the NoGrp condition, resulting in the missing "wedge" piece being rotated toward the fixation point. However, this trend did not result in differential attentional modulation as reflected in the AMIs$_{SEQ}$ (all areas: $F_{(2,9)} < 1.72$, $p > 0.21$).

When subjects ignored the peripheral stimulus arrays, a reduction in competition was found when the elements formed a perceptual group via illusory contour formation. These results suggest that perceptual groups or candidate objects receive an advantage over unorganized elements (McMains and Kastner, 2010). From this result, one might expect a perceptual advantage for grouped stimuli when subjects performed a task on the grouped elements. The staircase procedure used during the peripheral attention task prevented any differences in accuracy between the different display conditions; however, reaction time was not controlled for. Here, we did find an advantage for

grouped stimuli, in that reaction times varied for the ATT SIM conditions (Table 1: StrongGrp, 496 ms; WeakGrp, 507 ms; NoGrp, 512 ms; $F_{(2,9)} = 3.89$, $p < 0.05$). A significant linear contrast was found ($F_{(1,9)} = 10.38$; $p < 0.05$), suggesting that reaction times varied monotonically with the degree of perceptual grouping, such that the largest advantage was found for the StrongGrp condition. These results support the hypothesis that grouped elements receive a processing advantage.

Discussion

By parametrically manipulating the degree of bottom-up perceptual grouping present within multielement displays, a monotonic relationship between the degree of competition and perceptual grouping was found in extrastriate visual cortex, such that grouped stimuli induced less competition. Importantly, these findings build on a previous study, which found that stimuli grouped via illusory contour formation (as used here) or collinear alignment (with oriented gabors) competed less than the same stimuli randomly oriented (McMains and Kastner, 2010), suggesting that perceptual grouping processes in general counteract competition throughout the visual field. When subjects attended to the peripheral stimulus displays, there was an inverse relationship between attentional modulation and the degree of perceptual grouping present in the multielement display, such that attentional modulation was greatest when neural competition was little influenced by bottom-up mechanisms and smallest when competition was strongly influenced by bottom-up mechanisms. These results are consistent with an account that assumes top-down processes to operate on local neural networks that mediate grouping and competition, thereby providing interfaces to constrain and guide attentional selection (interface hypothesis). Our findings suggest that selective attention counteracts competitive interactions among multiple stimuli that have not been resolved by bottom-up grouping processes.

As noted previously (Kastner et al., 1998, 2001; Beck and Kastner, 2005), there were several differences between the SEQ and SIM presentation conditions, in addition to the level of competition they induced. For instance, the visual presentation period of the SEQ condition extended over 1 s, whereas the presentation period for the SIM condition was 250 ms. In addition, the SEQ condition contained four visual onsets, whereas the SIM condition contained only one. However, if the stimulus duration, number of onsets, or any other inherent low-level differences between the SEQ and SIM conditions were solely driving the observed difference in activation for the two conditions, then the difference between conditions should be constant across different stimulus configurations. Previously, Kastner et al. (2001) have found that, within a visual area, the difference between the SEQ and SIM conditions decreased with increasing spatial separation. Similarly, the difference between SEQ and SIM conditions decreased as the perceptual grouping among elements within unattended SIM arrays increased in the current and in a previous (McMains and Kastner, 2010) experiment, whereas the number of onsets and the stimulus duration were held constant. Importantly, the differences in competition as measured by the SSI observed here were attributable to changes in activity evoked by the SIM conditions in which such factors do not vary.

Previous studies have shown that top-down attentional modulation is greater when stimuli are present simultaneously in the visual field, thereby competing for neural representation, as opposed to when the same stimuli are presented in isolation (Kastner et al., 1998; Reynolds et al., 1999). These findings support the biased competition theory, which proposes that top-down processes such as selec-

tive attention operate by counteracting competitive interactions among stimuli (Desimone and Duncan, 1995; Beck and Kastner, 2009). Here, using a parametric modulation of competition, we extend previous findings by demonstrating that it is not simply the presence or absence of competition that influences the amount of attentional effects, but rather that the degree of attentional modulation is closely tied to the degree of competition left unresolved after automatic bottom-up perceptual grouping processes have occurred and influenced competitive processes. In fact, when the stimulus array was attended to, we found that competitive interactions were similar for all levels of perceptual grouping, suggesting that bottom-up and top-down processes interact dynamically to resolve neural competition to the greatest possible extent.

These results are consistent with several recent physiology studies suggesting that top-down and bottom-up processes interact dynamically in visual cortex (Mazer and Gallant, 2003; Reynolds and Desimone, 2003; Bichot et al., 2005; Ogawa and Komatsu, 2006). For instance, a study investigating the interaction of bottom-up stimulus contrast and top-down attention in modulating competitive interactions in monkey V4 (Reynolds and Desimone, 2003), which found a similar inverse relationship between attentional modulation and bottom-up processes. When attention was directed away from a V4 RF, increasing the stimulus contrast of a preferred stimulus within the RF counteracted competition from a neighboring nonpreferred stimulus in the RF. When attention was directed toward the RF, the greatest modulation was observed when competition was little influenced by bottom-up stimulus salience, whereas the least modulation was observed when competition was already counteracted by bottom-up visual salience. In addition, different V4 cells have been found to represent bottom-up visual salience and top-down behavioral relevance (Ogawa and Komatsu, 2006). Interestingly, in the neural population, the initial visual response was dominated by the bottom-up salience (i.e., the singleton type of the RF stimulus), whereas the late pretarget selection response was dominated by the behavioral significance of the stimulus in the RF, suggesting that top-down and bottom-up signals are dynamically represented within the same population of cells.

Attentional selection within the spatial domain has often been conceptualized with the help of a spotlight metaphor. According to the spotlight hypothesis, top-down processes only operate on neurons that represent the attended locations within the spotlight. The strength of attentional enhancement is thought to be related to the size of the spotlight and difficulty of the task being performed at the attended location (Eriksen and St. James, 1986). In general, the attentional spotlight does not consider visual information outside the focus of attention such as nearby distracter stimuli that may influence competition or provide contextual information for grouping processes. Alternatively, the interface hypothesis suggests that selective attention operates on interfaces constituted by local networks. The interface hypothesis was first proposed by von der Heydt and colleagues (Qiu et al., 2007) based on a physiology study that investigated the relationship between bottom-up figure–ground segmentation processes and top-down attention in area V2. They found that, when a monkey directed attention toward the RF of a neuron, attentional modulation was larger when the preferred figure–ground configuration was present within the RF. These findings were interpreted in terms of an interface hypothesis of attention, which proposed that the circuit mediating figure–ground assignment provided an interface within early visual cortex for attentional mechanisms to operate on. According to such an account, the degree of atten-

tional modulation within a region can be predicted by how much the attended stimulus engages local circuits. These results are consistent with the finding that local interneurons that subserve intrinsic circuits receive stronger attentional modulation than other cell classes (Mitchell et al., 2007).

Here, we propose that local networks subserving competitive interactions among groups of neurons coding for multiple nearby stimuli may also provide an interface for attentional mechanisms to operate on, although at a different level of processing compared with the studies on figure–ground segmentation (Qiu et al., 2007). The finding that the amount of attentional modulation was dependent on the output of automatic bottom-up competitive processes may be taken as evidence that the same neural circuits underlie both processes, and that the circuit that mediates competition provides an interface for attentional modulation (Qiu et al., 2007). Thus, top-down mechanisms might operate on local circuits within V4, an area thought to be important in computing competitive interactions among multiple stimuli (De Weerd et al., 1999; Gallant et al., 2000), or on a feedback network that includes V4. Importantly, with fMRI it is often difficult to reveal underlying neural mechanisms. For instance, how might automatic perceptual grouping and competitive processes interact at the neural level? There are two main possibilities, which are not mutually exclusive (McMains and Kastner, 2010). First, competition in intermediate visual areas, such as V4, may be influenced by perceptual organization processes that occur in early visual cortex (von der Heydt and Peterhans, 1989; Sheth et al., 1996; Lee and Nguyen, 2001; Maertens and Pollmann, 2005, 2007; Montaser-Kouhsari et al., 2007). These mechanisms could boost the activity related to the set of stimuli as it enters intermediate visual areas such as area V4, which has the larger RF needed to read out biases resulting from perceptual organization computed in early visual cortex and integrate them with parallel competitive processes. A second possibility is that perceptual grouping and competition may rely on the same set of neural mechanisms implemented at intermediate processing stages (Kastner et al., 2001; Behrmann and Kimchi, 2003). Regardless of the underlying mechanisms, top-down attention appears to modulate local networks involved in computing competitive processes, given that the largest effects of attention were found when competition among array elements was greatest. Together, the results of Qiu et al. (2007) and our present results suggest that attention may operate on several different interfaces, which may be recruited at different levels of processing.

The interface hypothesis is similar to a recent proposal by Gilbert and Sigman (2007), which suggests that the magnitude of attentional enhancement within a visual area is determined by the degree to which the local circuits computing contextual information are involved in the ongoing computation. Both proposals challenge the traditional view that attention acts in a hierarchical manner, a view based on the finding that attention effects are generally larger in extrastriate areas like V4 than in early visual cortex (Kastner et al., 1999; McMains and Somers, 2004). The proposal by Gilbert and colleagues (Li et al., 2004; Gilbert and Sigman, 2007) extends the interface hypothesis by including top-down processes other than attention, such as task set, suggesting that a single area and neuron may perform many different functions depending on the demands of the current behavioral task. The interface hypothesis may help resolve the ongoing debate about the role of attentional modulation in primary visual cortex. Isolated stimuli placed in the RFs of V1 neurons often fail to be modulated by attention (McAdams and

Maunsell, 1999), whereas attention to complex stimuli that provide a contextual framework results in modulation of V1 RFs (Motter, 1993; Ito and Gilbert, 1999). The interface hypothesis would predict this discrepancy, arguing that only when V1 neurons are engaged as part of a local network, such as the neural circuit representing contextual information, would attentional modulation be large.

The current findings extend the biased competition theory of attention, suggesting that top-down attention is constrained by the output of bottom-up processes, such that the degree of attentional modulation varies monotonically with the degree of competition left unresolved after bottom-up grouping processes have occurred. The interaction of top-down and bottom-up processes provides a mechanism by which attention can select low-salient, but task-relevant stimuli. In addition, the finding that the degree of attentional modulation is inversely related to the output of bottom-up processes challenges the traditional spotlight view of attention, which argues that all task-relevant stimuli within the attentional spotlight will be similarly enhanced, regardless of any distractor or contextual stimuli located outside of the spotlight. Alternatively, the inverse relationship between attention and perceptual grouping can be interpreted within the interface hypothesis originally proposed to explain the relationship between figure–ground processes and attention (Qiu et al., 2007). The current results extend the interface hypothesis by providing an example of an interface at a different level of processing on which attention can operate, the network containing V4 involved in competitive processes. Together, converging evidence amounts to suggest that attention may operate on several different interfaces that can be recruited at different levels of processing.

References

Arcaro MJ, McMains SA, Singer BD, Kastner S (2009) Retinotopic organization of human ventral visual cortex. J Neurosci 29:10638–10652.

Beck DM, Kastner S (2005) Stimulus context modulates competition in human extrastriate cortex. Nat Neurosci 8:1110–1116.

Beck DM, Kastner S (2007) Stimulus similarity modulates competitive interactions in human visual cortex. J Vis 7:19.11–19.12.

Beck DM, Kastner S (2009) Top-down and bottom-up mechanisms in biasing competition in the human brain. Vision Res 49:1154–1165.

Behrmann M, Kimchi R (2003) What does visual agnosia tell us about perceptual organization and its relationship to object perception? J Exp Psychol Hum Percept Perform 29:19–42.

Bichot NP, Rossi AF, Desimone R (2005) Parallel and serial neural mechanisms for visual search in macaque area V4. Science 308:529–534.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

Brefczynski JA, DeYoe EA (1999) A physiological correlate of the "spotlight" of visual attention. Nat Neurosci 2:370–374.

Broadbent DE (1958) Perception and communication. London: Pergamon.

Buracas GT, Boynton GM (2007) The effect of spatial attention on contrast response functions in human visual cortex. J Neurosci 27:93–97.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

De Weerd P, Peralta MR 3rd, Desimone R, Ungerleider LG (1999) Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. Nat Neurosci 2:753–758.

Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7:181–192.

Eriksen CW, St James JD (1986) Visual attention within and around the field of focal attention: a zoom lens model. Percept Psychophys 40:225–240.

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II:

Inflation, flattening, and a surface-based coordinate system. Neuroimage 9:195–207.

Folk CL, Remington RW, Johnston JC (1992) Involuntary covert orienting is contingent on attentional control settings. J Exp Psychol Hum Percept Perform 18:1030–1044.

Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, Turner R (1995) Analysis of fMRI time-series revisited. Neuroimage 2:45–53.

Gallant JL, Shoup RE, Mazer JA (2000) A human extrastriate area functionally homologous to macaque V4. Neuron 27:227–235.

Gilbert CD, Sigman M (2007) Brain states: top-down influences in sensory processing. Neuron 54:677–696.

Halko MA, Mingolla E, Somers DC (2008) Multiple mechanisms of illusory contour perception. J Vis 8:17.1–17.17.

Ito M, Gilbert CD (1999) Attention modulates contextual influences in the primary visual cortex of alert monkeys. Neuron 22:593–604.

Kanizsa G (1976) Subjective contours. Sci Am 234:48–52.

Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science 282:108–111.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22:751–761.

Kastner S, De Weerd P, Pinsk MA, Elizondo MI, Desimone R, Ungerleider LG (2001) Modulation of sensory suppression: implications for receptive field sizes in the human visual cortex. J Neurophysiol 86:1398–1411.

Lee TS, Nguyen M (2001) Dynamics of subjective contour formation in the early visual cortex. Proc Natl Acad Sci U S A 98:1907–1911.

Li W, Piëch V, Gilbert CD (2004) Perceptual learning and top-down influences in primary visual cortex. Nat Neurosci 7:651–657.

Li X, Lu ZL, Tjan BS, Dosher BA, Chu W (2008) Blood oxygenation level-dependent contrast response functions identify mechanisms of covert attention in early visual areas. Proc Natl Acad Sci U S A 105:6202–6207.

Maertens M, Pollmann S (2005) fMRI reveals a common neural substrate of illusory and real contours in V1 after perceptual learning. J Cogn Neurosci 17:1553–1564.

Maertens M, Pollmann S (2007) Illusory contours do not pass through the "blind spot." J Cogn Neurosci 19:91–101.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A 92:8135–8139.

Mazer JA, Gallant JL (2003) Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. Neuron 40:1241–1250.

McAdams CJ, Maunsell JH (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci 19:431–441.

McMains SA, Kastner S (2010) Defining the units of competition: influences of perceptual organization on competitive interactions in human visual cortex. J Cogn Neurosci 22:2417–2426.

McMains SA, Somers DC (2004) Multiple spotlights of attentional selection in human visual cortex. Neuron 42:677–686.

Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RB (1999) The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. J Neurosci 19:8560–8572.

Miller EK, Gochin PM, Gross CG (1993) Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. Brain Res 616:25–29.

Mitchell JF, Sundberg KA, Reynolds JH (2007) Differential attention-dependent response modulation across cell classes in macaque visual area V4. Neuron 55:131–141.

Montaser-Kouhsari L, Landy MS, Heeger DJ, Larsson J (2007) Orientation-selective adaptation to illusory contours in human visual cortex. J Neurosci 27:2186–2195.

Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784.

Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol 70:909–919.

Murray SO (2008) The effects of spatial attention in early human visual cortex are stimulus independent. J Vis 8:1–11.

Ogawa T, Komatsu H (2006) Neuronal dynamics of bottom-up and top-down processes in area V4 of macaque monkeys performing a visual search. Exp Brain Res 173:1–13.

Palmer SE (1999) Vision science: photons to phenomenology. Cambridge, MA: MIT.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442.

Posner MI, Snyder CR, Davidson BJ (1980) Attention and the detection of signals. J Exp Psychol 109:160–174.

Qiu FT, Sugihara T, von der Heydt R (2007) Figure-ground mechanisms provide structure for selective attention. Nat Neurosci 10:1492–1499.

Reynolds JH, Desimone R (2003) Interacting roles of attention and visual salience in V4. Neuron 37:853–863.

Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci 19:1736–1753.

Sayres R, Grill-Spector K (2008) Relating retinotopic and object-selective responses in human lateral occipital cortex. J Neurophysiol 100:249–267.

Scalf PE, Beck DM (2010) Competition in visual cortex impedes attention to multiple items. J Neurosci 30:161–169.

Schneider KA, Richter MC, Kastner S (2004) Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. J Neurosci 24:8975–8985.

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.

Sheth BR, Sharma J, Rao SC, Sur M (1996) Orientation maps of subjective contours in visual cortex. Science 274:2110–2115.

Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. J Neurosci 27:5326–5337.

Theeuwes J (1991) Cross-dimensional perceptual selectivity. Percept Psychophys 50:184–193.

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12:97–136.

Varin D (1971) Fenomeni di contrasto e diffusione cromatica nell' organizzazione spaziale del campo percettivo. Revista di Psicologiz 65:101–128.

von der Heydt R, Peterhans E (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. J Neurosci 9:1731–1748.

Wade AR, Brewer AA, Rieger JW, Wandell BA (2002) Functional measurements of human ventral occipital cortex: retinotopy and colour. Philos Trans R Soc Lond B Biol Sci 357:963–973.

Yantis S (2000) Goal-directed and stimulus-driven determinants of attentional control. In: Attention and performance XVII (Monsell S, Driver J, eds), pp 73–103. Cambridge, MA: MIT.

Yeshurun Y, Kimchi R, Sha'shoua G, Carmel T (2009) Perceptual objects capture attention. Vision Res 49:1329–1335.