



Published in final edited form as:

Nat Neurosci. 2005 August ; 8(8): 1110–1116.

Stimulus context modulates competition in human extrastriate cortex

Diane M. Beck^{1,2} and Sabine Kastner^{1,2}

¹ Department of Psychology, Princeton University, Green Hall, Princeton, New Jersey 08544, USA

² Center for the Study of Brain, Mind, and Behavior, Princeton University, Green Hall, Princeton, New Jersey 08544, USA

When multiple stimuli appear simultaneously in the visual field, they are not processed independently, but interact in a mutually suppressive way suggesting that they compete for neural representation in visual cortex. The biased competition model of selective attention predicts that the competition can be influenced by both top-down and bottom-up mechanisms. Directed attention has been shown to bias competition in favor of the attended stimulus in extrastriate cortex. Here, we show that suppressive interactions among multiple stimuli are eliminated in extrastriate cortex when they are presented in the context of pop-out displays, in which a single item differs from the others, but not in heterogeneous displays, in which all items differ from each other. The pop-out effects appeared to originate in early visual cortex and were independent of attentional top-down control suggesting that stimulus context may provide a powerful influence on neural competition in human visual cortex.

Natural visual scenes are cluttered and contain many different objects that cannot all be processed at once due to limited processing capacity of the visual system¹, suggesting that multiple objects present at the same time in the visual field compete for neural representation²⁻³. Neural correlates for competitive interactions among multiple stimuli have been found in visual cortex in single-cell physiology and functional brain imaging studies, showing that multiple stimuli presented in nearby locations are not processed independently from each other but interact in a mutually suppressive way⁴⁻⁸. These sensory suppressive interactions occur most strongly at the level of the receptive field (RF)^{5,9} and are therefore prominent in extrastriate areas where RFs are large enough to encompass multiple stimuli⁴⁻⁸.

According to the “biased competition model“ of selective attention^{2-3,11}, competition among multiple stimuli can be influenced both by means of top-down processes related to the selection of information that is relevant to current behavioral goals and by bottom-up, stimulus driven processes. For example, if one directs attention to a particular location in a cluttered scene, processing of information at the attended location will be facilitated and processing of unwanted information from nearby distracters will be suppressed¹², suggesting that competition is biased in favor of the attended stimulus. On the other hand, if a salient stimulus is present in a cluttered scene, it will be effortlessly and quickly detected independent of the number of distracters¹³, suggesting that competition is biased in favor of the salient stimulus. At the neural level, evidence in support of the biased competition model was found in studies showing that spatially directing attention to one of multiple stimuli eliminates or reduces the suppressive influences of nearby stimuli in extrastriate cortical areas, consistent with the idea that selective attention biases the competition among multiple stimuli in favor of the attended stimulus by counteracting suppressive interactions^{5-7,9-10}. These mechanisms that operate in

visual cortex appear to be controlled by a distributed network of higher-order areas in frontal and parietal cortex, which generate top-down signals that are transmitted via feedback connections to the visual system¹⁴⁻¹⁶. Here, we asked how bottom-up influences related to stimulus context of a visual display in which a single salient stimulus pops out from a homogeneous background affect suppressive interactions among multiple stimuli competing for neural representation in human visual cortex using functional magnetic resonance imaging (fMRI).

Unlike selective attention, which relies on top-down signals from frontoparietal sources¹⁴⁻¹⁶, a contextual effect like pop-out depends on factors present in the display, including simple feature properties such as the color of the stimulus¹³, perceptual grouping of stimulus features by Gestalt principles¹⁷⁻¹⁹, and the dissimilarity between the stimulus and nearby distracters²⁰⁻²¹. Neural correlates of pop-out have been found as early as in area V1. Responses of V1 neurons to a single item presented in a RF surrounded by a homogeneous array of items presented outside the RF are stronger when the surround differs from the RF stimulus than when it is identical to it²²⁻²⁴, suggesting that neural responses depend on the context in which the stimuli are shown. These context-dependent effects do not appear to rely on top-down control, since they are not only demonstrated in awake, but also in anesthetized animals²³⁻²⁴.

In the present study, suppressive interactions among multiple stimuli present at the same time in nearby locations were assessed across human visual cortex using two display types: pop-out displays, in which a single item differed from the others (Fig. 1a), and heterogeneous displays, in which all items differed from each other (Fig. 1b). We predicted that similar to the way in which top-down attention can counteract suppressive interactions among multiple stimuli^{5-7,9-10}, bottom-up signals related to pop-out can weaken suppressive interactions among stimuli appearing in the context of pop-out relative to heterogeneous displays. However, in accordance with biased competition theory, although neural signals related to the encoding of pop-out may originate early in visual cortex, we predicted that these signals will affect the outcome of competitive interactions among multiple stimuli that typically takes place in later extrastriate areas such as V2 and V4, where RFs are sufficiently large to encompass multiple stimuli⁴⁻⁸.

RESULTS

Four colored Gabor stimuli were presented in randomized order in four nearby locations within the upper right quadrant of the visual field under two presentation conditions: sequential and simultaneous. In the sequential condition (SEQ), each of the stimuli was presented alone in one of the four locations (Fig. 1c). In the simultaneous condition (SIM), the stimuli appeared together in all four locations (Fig. 1d). Integrated over time, the physical stimulation parameters in each of the four locations were identical under the two presentation conditions. However, as shown previously⁷⁻⁸, suppressive interactions among the stimuli could take place only in the simultaneous but not in the sequential presentation condition. The influence of pop-out on competitively interacting stimuli was studied by probing two different display type conditions, heterogeneous (HET) and pop-out (POP), in addition to the SEQ and SIM presentation conditions. In the heterogeneous display condition, all four stimuli differed in orientation and color (Fig. 1b). In the pop-out display condition, one stimulus differed in color and orientation from the other three (Fig. 1a). The display type conditions were equated such that integrated over time, the physical stimulation parameters in each location were identical and only the context, in which the four stimuli were presented, was varied. The subject's task was to detect target letters presented in a rapid stream of letters, digits, and keyboard symbols at fixation during all conditions. The fixation task ensured proper fixation and effectively prevented subjects from covertly attending to the peripheral stimuli. Two versions of the fixation task

were tested in the fMRI experiments: one in which subjects ($n = 6$) made no overt motor response and simply counted the number of targets, and one in which subjects ($n = 6$) pressed a button as soon as they detected a target letter. These two experiments yielded very similar results (see Supplementary Fig. 1 online) and therefore the data from the two experiments were combined for the following fMRI analyses.

Gabor stimuli, as compared to blank presentations, evoked significant activity in areas V1, V2, VP, and V4, as determined on the basis of retinotopic mapping, in all subjects. As the border between V2 and VP could not be distinguished unequivocally in some of the subjects, the two areas were combined for all analyses. The locations of the activations were in the ventral parts of these areas in the left hemisphere, consistent with the locations of stimuli in the upper right visual field.

Experiment 1: heterogeneous versus pop-out displays

For the heterogeneous display condition, we predicted that the fMRI signals evoked by simultaneously presented stimuli would be smaller than those evoked by sequentially presented stimuli in extrastriate cortex due to the mutual suppression induced by competitively interacting stimuli⁷⁻⁸. In support of our hypothesis, an analysis of the fMRI time series and the mean signal changes averaged across all subjects revealed that simultaneous presentations evoked less response than sequential presentations in areas V2/VP and V4 (V2/VP: $t_9 = 5.33$, $P < 0.001$; V4: $t_9 = 6.98$, $P < 0.001$; **Figs. 2a** and **3a**). The difference in activations between sequential and simultaneous presentations increased gradually from V1 to V4 (interaction of area and presentation condition: $F_{2,18} = 30.37$, $P < 0.001$); response differences in area V1 were not significant ($t < 1$). This effect is also reflected in the sensory suppression index (SSI), which quantifies the differences in responses to sequential and simultaneous presentations (main effect of area on SSI_{HET} : $F_{2,18} = 43.89$, $P < 0.001$; V1 vs. V2/VP, $t_9 = 6.46$, $P < 0.001$; V2/VP vs. V4, $t_9 = 2.64$, $P < 0.05$). The gradual increase in magnitude of the SSI_{HET} from V1 to V4 (**Fig. 3b**) suggests that suppressive interactions were scaled to the increasing RF sizes of neurons in areas along the ventral visual pathway in accordance with previous results⁷⁻⁸.

For the pop-out display condition, we predicted that the differences in responses evoked by sequential and simultaneous presentations would be smaller than those obtained with the heterogeneous displays due to a presumed bottom-up contextual effect related to pop-out. In support of our hypothesis, a repeated measures ANOVA revealed a significant interaction of presentation (SEQ vs. SIM) and display type condition (POP vs. HET) in areas V2/VP ($F_{1,9} = 18.34$, $P < 0.01$) and V4 ($F_{1,9} = 18.63$, $P < 0.01$), such that the response differences evoked by sequential and simultaneous presentations were indeed smaller for pop-out displays relative to heterogeneous displays (**Figs. 2** and **3a**). In fact, in areas V2/VP and V4, there was no significant difference between activity evoked by simultaneous and sequential presentations in the pop-out condition (**Fig. 2b**). The interaction of presentation and display type condition can be seen most clearly in comparing the SSI computed for heterogeneous and pop-out display conditions (**Fig. 3b**), which differed significantly in areas V1 ($t_9 = 2.35$, $P < 0.05$), V2/VP ($t_9 = 3.97$, $P < 0.01$), and V4 ($t_9 = 5.03$, $P < 0.001$). Indeed, in V2/VP and V4, the SSI_{POP} was not different from zero. In V1, the SSI_{POP} was significantly different from zero ($t_9 = 3.75$, $P < 0.01$), but it was reversed (negative), indicating that simultaneous presentations evoked more activity than sequential presentations (**Fig. 3b**). The reversal of the presentation condition effect with the pop-out, but not with the heterogeneous display condition, is consistent with single-cell physiology studies showing that neural correlates of pop-out can be found as early as in area V1²²⁻²⁴. Indeed, such a result is suggestive that V1 may be the source of the signal that modulates the suppressive interactions among multiple stimuli at subsequent stages of processing, consistent with the idea that competitive interactions in extrastriate cortex can be modulated by stimulus context in a bottom-up manner.

Experiment 2: homogeneous versus pop-out displays

Due to the spatial resolution limits of fMRI, we were unable to isolate the activity of any one item in the display, and instead, the activity evoked in the pop-out condition represents the summed activity evoked by all items in the display integrated over time. Therefore, we asked whether the effects on suppressive interactions obtained with pop-out displays were due to the salient item, the surrounding homogeneous items, or a combination of both. We performed a second experiment in which pop-out displays were compared to homogeneous displays instead of heterogeneous displays. Neither pop-out nor homogeneous displays induced a significant suppression effect in areas V1, V2/VP, or V4, suggesting that the homogeneous surround did indeed contribute to the smaller sensory suppression found with the pop-out displays. This result is compatible with predictions from biased competition theory² and behavioral data¹⁹⁻²⁰ suggesting that competitive interactions should occur between rather than within perceptual groups. However, importantly, simultaneous pop-out displays evoked significantly more activity than simultaneous homogeneous displays in area V4 [$t_5 = 2.68$, $P < 0.05$; Fig. 4], indicating that the neural responses evoked by pop-out displays were not driven entirely by the homogeneous surround but also depended on the presence of the salient stimulus in the display. This result renders the possibility unlikely that the observed stimulus display effects on suppressive interactions resulted from the relative homogeneity or heterogeneity of the displays. According to such an account, one would predict the pop-out displays (containing two item types) to produce a suppressive effect somewhere in between those produced by the homogeneous and heterogeneous displays. Yet this was not the case. A similar pattern of results was observed in areas V2/VP but the difference was not significant. These findings suggest that the effects on sensory suppression associated with pop-out displays were a function of both the salient item and the surrounding homogeneous items in the display, consistent with the fact that pop-out is a contextual effect, and the perceptual salience of an item is determined by the surrounding items in the display.

Bottom-up versus top-down modulation

Thus far, we have presented evidence that pop-out displays induced less sensory suppression among multiple competing stimuli than heterogeneous displays in extrastriate cortex and that both the salient item and the surrounding items contributed to this effect. Because subject's attention was engaged in a demanding task at fixation, this effect on sensory suppression presumably occurred in a stimulus-driven, or bottom-up fashion. However, it is possible that pop-out displays captured attention²⁵ more than the heterogeneous displays, which would imply that the effects were mediated by top-down rather than bottom-up factors related to visual salience. This issue was addressed in two ways. First, we assessed whether performance on the fixation task differed as a function of the peripheral stimulus condition. If attention was drawn towards the pop-out displays and away from the fixation task, then performance on the fixation task should be poorer during the pop-out condition than during the other conditions. Second, we compared activity evoked by simultaneously presented pop-out and heterogeneous displays to identify brain regions outside visual cortex that were more activated during the pop-out display than the heterogeneous display condition across subjects. If attention was disproportionately captured by the pop-out displays, then we might expect this comparison to result in greater activation in parietal areas associated with attentional capture²⁶⁻²⁸ or spatial shifts of attention¹⁵⁻¹⁶.

Behavioral data was acquired in the scanner by requiring subjects to press a button upon detection of a target letter at fixation. Subjects missed 13% of the targets on average, but no differences in misses across the four block types were obtained ($F_{3,12} = 1.30$, *n.s.*; Table 1). An analysis of subjects' reaction times to correctly detected targets also showed no differences across the four block types ($F_{3,12} = 1.39$, *n.s.*; Table 1). Because the simultaneous heterogeneous and simultaneous pop-out conditions evoked different neural responses in the

fMRI experiment, behavioral performance in these conditions was of particular interest. There were no differences in miss rates or RTs between these two conditions ($t_4 = 0.14$, *n.s.*, $t_4 = 1.77$, *n.s.*, respectively). Thus, the behavioral results did not support the idea that the pop-out displays attracted more attention than the heterogeneous displays.

A similar conclusion can be drawn from the fMRI data analyses of parietal areas. Using the same statistical procedures applied to identify visual areas, a comparison of simultaneous pop-out and simultaneous heterogeneous displays did not reveal any significantly activated voxels anywhere in the parietal cortex, including the superior parietal areas or the temporal parietal junction, which have been previously associated with attentional capture and spatial shifts of attention^{15-16,26-28}. Moreover, a group analysis of the 6 subjects who were tested in the version of the fixation task requiring motor responses did not reveal any significant parietal activity (see Supplementary Methods for more details on parietal analysis). Together, our results from behavioral and fMRI studies suggest that the effects on sensory suppression observed for pop-out displays was not mediated by top-down processes but instead reflected a bottom-up effect of stimulus context related to visual salience.

Finally, we compared the top-down effects on sensory suppression described previously⁷ with the bottom-up effects found in the present study directly by plotting the SSIs from both studies (Fig. 5). Closed symbols refer to SSIs obtained previously⁷; open symbols to those obtained in the present study. In both studies, sensory suppression among four heterogeneous stimuli was assessed across visual cortex when attention was directed away from the display (horizontal axis) and in the presence of either a top-down or a bottom-up factor (attention or pop-out; vertical axis). The SSIs from both studies fall below the dashed line indicating that both pop-out and directed attention conditions led to weaker suppressive interactions relative to the unattended condition. However, while the data probing top-down effects on sensory suppression all fall significantly above zero on the vertical axis, indicating that some suppressive interactions remained when attention was directed to a stimulus, suggesting that competition was not fully resolved by directed attention, the data from the current study fall on or below zero, consistent with the possibility that competitive interactions were eliminated with the pop-out displays. However, it should be noted that this difference is only suggestive because the data from the attention study⁷ also fall to the right of the data from the present study, indicating that the complex stimuli used in that study induced stronger suppressive interactions compared to those induced by the more simple Gabor stimuli used here. Taken together, in accord with the central tenets of biased competition theory, this comparison suggests that the competition among multiple stimuli for neural representation can be influenced by means of both bottom-up and top-down mechanisms operating at intermediate processing stages of human visual cortex.

DISCUSSION

We report evidence for stimulus context modulating competitive interactions among multiple stimuli in human extrastriate cortex. Sensory suppression among multiple stimuli was observed in areas V2/VP and V4 when the stimuli were presented in the context of heterogeneous displays, in accordance with previous studies⁷⁻⁸, but was eliminated when the same stimuli were presented in the context of pop-out displays.

Our results complement previous studies suggesting that sensory suppressive interactions reflect competition among multiple stimuli for neural representation, using the same experimental paradigm of sequential and simultaneous stimulus presentations. As in previous studies⁷⁻⁸, the suppressive interactions obtained in the heterogeneous display condition increased gradually from V1 to V4, suggesting that they were scaled to the increasing RF sizes of neurons in these areas, and supporting the notion that suppressive interactions among

multiple stimuli occur most strongly at the level of the RF. This idea is further corroborated by previous findings that when the spatial separation among the competing stimuli is increased, suppressive interactions are found in more anterior extrastriate areas with larger RFs⁸. The effects of spatial separation on the outcome of competitive interactions, together with the effects of display type found in the present study, rule out the possibility that fewer stimulus onsets in the simultaneous versus sequential presentation condition may account for the smaller activity evoked by simultaneously presented heterogeneous displays. Suppressive interactions among multiple stimuli depend either on the distance between the stimuli⁸, or on the context (i.e. display type) in which the stimuli appeared, despite the fact that the relative number of onsets was unchanged. Taken together, these findings are consistent with the predictions that competition among multiple stimuli for neural representation can be affected by several factors including the spatial layout of stimuli in a display and the context of stimulus presentations.

Our present findings constitute important evidence in support of the biased competition model of selective visual attention, which postulates that competitive interactions among multiple stimuli for neural representation can be biased, not only by top-down allocation of attention, but also by bottom-up stimulus driven influences. Evidence for top-down modulation of competitive interactions has been provided by single cell physiology^{5-6,9-10} and functional brain imaging studies⁷ in which directing attention to one of multiple heterogeneous stimuli presented at the same time results in weaker suppressive interactions in areas V4 and TEO, relative to a condition in which the same stimuli are unattended. Here, we demonstrated a similar effect on suppressive interactions among multiple simultaneously presented stimuli that occurred when attention was directed away from the peripheral stimuli and, instead, stimulus context rendered one of the stimuli salient. This context-dependent effect eliminated suppressive interactions among the stimuli in extrastriate cortex. Taken together, these findings suggest that both top-down mechanisms related to spatially selective attention and bottom-up mechanisms related to stimulus context operate by resolving competitive interactions at intermediate processing stages in visual cortex, although in keeping with single cell recording²²⁻²⁴ and computational models²⁹ of pop-out it appears that these stimulus context effects may have their origin in early visual cortex.

It should be noted that given the spatial resolution of fMRI it is possible that our results obtained with the pop-out displays were the sum of two separate neural processes being generated within the same area (e.g. V4), but from different subpopulations of neurons that did not interact with each other: one subpopulation coding suppressive interactions due to the ongoing competition among the stimuli and another subpopulation coding signals related to pop-out. It is possible, for example, that the reduction of suppression depended entirely on the homogeneous surround and that the increased activity associated with the pop-out displays simply reflected a separate but additive influence of visual salience. However, there is evidence for an interaction of visual salience and competitive processes at the level of single neurons³⁰ that is consistent with the interpretation that competitive interactions may depend on the entire display, including the salient item. Suppressive interactions in V4 neurons are biased towards the more salient (high contrast) of a pair of stimuli presented in the neuron's RF. Such a conception is also consistent with winner-takes-all models of visual salience³¹, in which the more salient stimulus dominates neural responses and thereby wins the competition.

We considered the possibility that the effects on sensory suppression demonstrated with pop-out displays were mediated to some degree by spatially directed attention, given that several models of pop-out have assumed that attention is automatically directed to salient objects in the visual field^{28,32-33}. Although visually salient items do not necessarily capture attention³⁴, it is possible that attention was captured to a greater degree by the pop-out displays than by the heterogeneous displays in our study. If so, the effects on suppressive interactions among the stimuli were not mediated by the context of the display but rather by directed

attention, similar to those found previously⁷. However, our behavioral data and additional analyses of the fMRI data did not favor such an interpretation. There was no effect of display type on subjects' ability to rapidly detect target letters, suggesting that the different display types did not differ in their ability to capture attention, and simultaneously presented pop-out displays evoked no more activity than heterogeneous displays in regions of the parietal cortex known to be activated by displays capturing attention²⁶⁻²⁸ or by spatial shifts of attention¹⁵⁻¹⁶. Finally, the contextual effects of pop-out on sensory suppression appeared to be stronger than the top-down influences of directed attention (see Fig. 5), making it unlikely that the effects observed with the pop-out displays resulted from some small misdirection of attention to the salient stimuli that we were unable to detect in our behavioral studies. Unlike the pop-out displays, directed attention reduces but does not eliminate the suppression induced by nearby stimuli⁷. Rather, it appears that pop-out is a powerful bottom-up process that overcomes competitive interactions among multiple stimuli for neural representation and operates independently of attentional top-down control, consistent with the classical notion that visual salience is processed in a preattentive mode³⁵.

The conception of pop-out as a similar but separate mechanism than top-down attention for modulating competitive interactions among multiple stimuli at intermediate processing stages is in agreement with lesion studies in humans and monkeys. A patient with an isolated V4 lesion³⁶ and monkeys with lesions of areas V4 and TEO show discrimination deficits when targets must be selected among distracters³⁷, suggesting that the filtering mechanisms associated with top-down attention may critically depend on the integrity of extrastriate areas such as V4. Notably, however, the deficit associated with V4 lesions can be ameliorated by increasing the salience of the target stimulus³⁶⁻³⁸, suggesting that visual salience constitutes a separate filtering mechanism than the one mediated by top-down signals from attention.

Although in our study, subjects' attention was drawn away from the peripheral stimuli and engaged in a demanding task at fixation, under natural viewing conditions bottom-up and top-down mechanisms are free to interact, allowing the biasing mechanisms to reinforce each other³⁹⁻⁴¹. Moreover, visual salience may be just one example of a number of bottom-up contextual effects, instrumental in scene segmentation and guiding attention to object-based selections^{17-18,42-44}, that may operate by influencing competition for neural representation in visual cortex.

METHODS

Subjects, visual stimuli, and experimental design

Twelve subjects (7 females; age: 21-34 yrs; normal or corrected-to-normal visual acuity) gave written informed consent for participation in the study, which was approved by the Institutional Review Panel of Princeton University.

Visual stimuli were four Gabor patches (wavelength, 0.47° standard deviation of gaussian envelope, 0.73° each approximately 2 x 2° in size) presented in four nearby locations (2.5° from the center of one Gabor to its nearest neighbor) in the upper right quadrant of the visual field, with the Gabor closest to and furthest from fixation centered at 9.5° and 13.5° from fixation, respectively. The stimuli were either red, blue, green or cyan, and had an orientation of 0° (vertical), 60°, 90° (horizontal) and 150°, respectively (Fig. 1). All stimuli were presented on a dark background. Stimuli were generated on a Power Mac G4 using Matlab software (Mathworks) and the Psychophysical Toolbox⁴⁵.

The stimuli were shown under two presentation conditions: sequential (SEQ) and simultaneous (SIM). In the sequential presentation condition, each of the four Gabor stimuli was shown alone in one of the four locations for 250 ms (Fig. 1c). In the simultaneous presentation condition,

the same four stimuli appeared together for 250 ms (Fig. 1d). The order of stimuli and of locations was randomized. Stimulation periods (shown in Fig. 1c or d) were repeated in blocks of 18 s. Integrated over time, the physical stimulation parameters in each of the four locations were identical for sequential and simultaneous presentations.

In addition to the two presentation conditions, two display type conditions were probed in the main scanning experiment: heterogeneous (HET) and pop-out (POP). In the heterogeneous display condition, all four stimuli differed in both orientation and color (Fig. 1b). In the pop-out display condition, three of the stimuli were identical and the fourth differed in both orientation and color from the others (Fig. 1a). However, in both display type conditions, the same colors and orientations occurred with equal probability in each location, so that integrated across presentation blocks, the stimulation parameters in each location were identical for pop-out and heterogeneous display conditions, and only the context in which the stimuli appeared changed. Specifically, a particular Gabor stimulus (e.g. green-horizontal, Fig. 1a-b) was designated the singleton in the pop-out displays throughout a block and that singleton was presented in the exact same locations as the identical item (e.g. green-horizontal) in a heterogeneous display block from the same scanning run. For each display within a pop-out presentation block, the homogeneous surround was drawn at random from the remaining three colors, with the constraint that each of the 3 colors was presented exactly six times in a block. The remaining three colors in the heterogeneous displays were also presented randomly in each of the remaining three locations.

During a given scan, presentation (SEQ vs. SIM) and display type conditions (POP vs. HET) were combined to produce four blocks of visual stimulation (SEQ POP, SEQ HET, SIM POP, SIM HET) that were interleaved with blank periods of 18 s each. Each scan began with a block of visual stimulation that was discarded from further analysis, and ended with a blank period of 18 s for an overall scan duration of 180 s. Presentation conditions were presented in the sequence SEQ—SIM—SIM—SEQ, with the sequence of display type conditions counterbalanced across scans.

Subjects were engaged in detecting target letters presented in a rapid stream (4 Hz) of letters, digits, and keyboard symbols (720 per scan; 0.5 deg in size) presented at fixation for 250 ms each. Because it has been shown that motor responses can modulate activity in occipital cortex⁴⁶, the experiment was undertaken with two versions of the letter detection task. In one, subjects ($n = 6$) made no overt motor response and simply counted the number of targets (presented at random with a 17.6% probability), reporting the number at the end of the scan. In the second, subjects ($n = 6$) pressed a button as quickly as possible whenever they detected a target letter, which appeared in 20% of the trials. In this version of the experiment, half of the target letters appeared synchronously with the simultaneous displays and half appeared in the intervening intervals between simultaneous displays but during a simultaneous block. These two versions of the experiment yielded very similar fMRI results (see Supplementary Fig. 1) and therefore the fMRI data were combined yielding 10 data sets since two subjects participated in both versions of the experiment. Before being scanned, subjects participated in a training session outside the scanner to ensure that they were able to perform the tasks while maintaining fixation for several minutes.

Experiment 2 compared the pop-out display condition used in Experiment 1 with a homogeneous display condition. In the homogeneous display condition, four identical stimuli were presented in each of the four locations. Display conditions were equated such that, integrated over time, physical stimulation parameters were identical in pop-out and homogeneous display type conditions and, as in the main experiment, only the context of the four stimuli was varied. Visual stimuli and experimental design were as described for the

counting version of Experiment 1 except for the length of visual presentation blocks, which were 12 s, and that of the interleaved blank periods, which were 16 s.

Data acquisition and analysis

Data were acquired in 18 scan sessions with a 3 Tesla head scanner (Allegra, Siemens) using a standard head coil. In addition, retinotopic mapping was performed for all subjects in a separate scan session. Functional images were taken with a gradient echo, echoplanar sequence (TR = 2s; TE = 30 ms; flip angle = 90 ° matrix: 64 x 64). Twenty coronal slices were acquired in an interleaved fashion starting from the posterior pole (3 mm thickness, 1 mm gap, 2.5 x 2.5 mm in-plane resolution) in 12 series of 90 images each. Echoplanar images were compared with a co-aligned high-resolution anatomical scan of the same partial brain volume (FLASH; TR = 184 ms; TE = 4.6 ms, flip angle = 90 ° matrix: 256 x 256; FOV: 160 x 160 mm) for scan sessions testing the counting version of the fixation task and with a high-resolution anatomical scan of the whole brain (MPRAGE; TR = 2.5 s; TE = 4.38 ms, flip angle = 8 ° matrix: 256 x 256; FOV: 256 x 256 mm) for scan sessions testing the motor response version of the fixation task.

Functional images were motion-corrected⁴⁷. The National Institutes of Health functional imaging data analysis program (FIDAP) software was used to perform a multiple regression⁴⁸. Square-wave functions matching the time course of the experimental design contrasted 1) visual stimulation versus blank periods, and 2) sequential versus simultaneous presentations. These square-wave functions were convolved with a gaussian model of the hemodynamic response (lag: 4.8 s; dispersion: 1.8 s) to generate idealized response functions, which were used as regressors in the multiple regression model. Additional regressors were included in the model to factor out between-run changes in mean intensity and within-run linear drifts. Statistical maps were thresholded at a Z-score of 2.33 ($p < 0.01$, corrected for multiple comparisons). Activated voxels in visual cortex obtained during visual stimulation versus blank periods were subsequently assigned to retinotopically organized areas. For each subject, mean signals were computed by averaging across peak intensity values obtained in a given condition and visual area and are given as percent signal change, which was computed relative to the mean signal obtained during blank presentations. These values were further quantified by defining a sensory suppression index [SSI = $(R_{SEQ} - R_{SIM}) / (R_{SEQ} + R_{SIM})$]; R, response computed as mean signal change; SEQ, sequential presentation condition; SIM, simultaneous presentation condition], which was computed separately for the different display type conditions (SSI_{HET}, SSI_{POP}). The SSI quantifies the differences in responses to sequential and simultaneous presentations. Positive values indicate stronger responses to sequential than to simultaneous presentations; negative values indicate the opposite, and values around 0 indicate the absence of response differences. Statistical significance of SSIs and mean signal changes were assessed using repeated measures ANOVAs and paired t-tests with subject as the random variable.

To investigate whether regions in the parietal cortex were differentially activated by the pop-out or heterogeneous displays, both individual subject and group analyses were carried out by comparing activity evoked by simultaneous pop-out and simultaneous heterogeneous displays using AFNI (<http://afni.nimh.nih.gov/afni>; see Supplementary Methods online for more details).

Mapping visual areas

Retinotopic mapping was performed for each subject in a separate scanning session using established procedures⁴⁹ and was used to assign activated voxels to visual areas (for details see Supplementary Methods online).

Behavioral data analysis

RTs to correctly identify targets via a button press in the scanner were computed relative to the onset of the target stimulus as a function of block type for each subject. Correct responses were defined as responses occurring between 250 and 1000 ms after the onset of the target. The RT analysis was restricted to the four counterbalanced visual stimulation blocks from each run (i.e. the first visual stimulation block and blank blocks were excluded from each run). In one subject detections were likely overestimated due to a computer error. Consequently, the data from this subject were excluded from further analysis.

Supplementary Methods

Analysis of parietal cortex

Data from the version of the experiment requiring motor responses ($N=6$), in which an anatomical scan of the whole brain was acquired allowing for spatial normalization to Talairach space, was submitted to an additional analysis in order to investigate whether parietal cortex was differentially activated by the pop-out and heterogeneous displays. If attention was captured more by the simultaneous pop-out displays than by the simultaneous heterogeneous displays we might expect greater activation in parietal areas associated with attentional capture or spatial shifts of attention (i.e. temporal parietal junction or the superior parietal lobe). Data were analyzed using AFNI (<http://afni.nimh.nih.gov/afni>). Functional images were motion-corrected to a single image acquired nearest in time to the anatomical scan, normalized to the mean intensity of the run, and submitted to a single multiple regression. Four regressors of interest were generated by convolving square-wave functions matching the time course of each of the visual stimulation conditions (SEQ POP, SEQ HET, SIM POP, SIM HET) with a gammavariate function (Cohen, 1997). Additional regressors were included in the model to factor out between-run changes in mean intensity, within-run linear and quadratic drifts, and head motion. The coefficients associated with the SIM POP and SIM HET regressors were contrasted to identify brain regions that responded more to simultaneous pop-out displays than heterogeneous displays. As in the main analysis, the resulting statistical maps were thresholded at $P < 0.01$, correcting for multiple comparisons on a voxel-wise basis across the entire imaged brain. Finally, all six subjects were submitted to a group analysis, using a one-way fixed-effect ANOVA comparing SIM POP and SIM HET. Preprocessing was the same as described above except that prior to normalization, each subject's data was spatially filtered with a 4 mm Gaussian kernel, and statistical maps were transformed into standard stereotactic Talairach space. Neither the individual subject analysis nor the group analysis yielded significant activity anywhere in the parietal cortex, including temporal parietal junction and the superior parietal lobe ($P < 0.01$, corrected for multiple comparisons).

Mapping visual areas

All subjects participated in a separate retinotopic mapping scanning session using standard procedures described in detail elsewhere⁸. Briefly, areas V1, V2, and ventral V3 (referred to as VP) were identified by the alternating representations of the vertical and horizontal meridians, which form the borders of these areas⁴⁹. Area V4 was identified by its characteristic upper (UVF) and lower (LVF) visual field retinotopy. The UVF and LVF are separated in V4 and located medially and laterally, respectively, on the posterior part of the fusiform gyrus⁸. Retinotopic mapping served only to assign voxels that survived statistical thresholding (corrected for multiple comparisons across the entire imaged brain) to specific visual areas and was not used in computing the statistical significance of a voxel.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

We thank Anne Treisman and Hans-Christoph Nothdurft for valuable discussions. This work was supported by grants from the National Institute of Mental Health (RO1 MH-64043, P50 MH-62196) and the Whitehall Foundation.

REFERENCES

1. Broadbent, DE. Perception and Communication. Pergamon, London: 1958.
2. Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annu. Rev. of Neurosci* 1995;18:193–222. [PubMed: 7605061]
3. Duncan J. Converging levels of analysis in the cognitive neuroscience of visual attention. *Philos Trans R Soc Lond B Biol Sci* 1998;353:1307–17. [PubMed: 9770224]
4. Miller EK, Gochin PM, Gross CG. Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Res* 1993;616:25–29. [PubMed: 8358617]
5. Reynolds JH, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci* 1999;19:1736–1753. [PubMed: 10024360]
6. Recanzone GH, Wurtz RH. Effects of attention on MT and MST neuronal activity during pursuit initiation. *J. Neurophysiol* 2000;83:777–790. [PubMed: 10669493]
7. Kastner S, De Weerd P, Desimone R, Ungerleider LG. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 1998;282:108–111. [PubMed: 9756472]
8. Kastner S, De Weerd P, Pinsk MA, Elizondo MI, Desimone R, Ungerleider LG. Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *J. Neurophysiol* 2001;86:1398–1411. [PubMed: 11535686]
9. Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science* 1985;229:782–784. [PubMed: 4023713]
10. Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol* 1997;77:24–42. [PubMed: 9120566]
11. Bundesen C, A theory of visual attention. *Psychol. Res* 1990;97:523–547.
12. Posner MI. Orienting of Attention. *Q. J. Exp. Psychol* 1980;32:3–25. [PubMed: 7367577]
13. Treisman AM, Gelade G. A feature-integration theory of attention. *Cognitive Psychol* 1980;12:97–136.
14. Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 1999;22:751–761. [PubMed: 10230795]
15. Kastner S, Ungerleider LG. The neural basis of biased competition in the human visual cortex. *Neuropsychologia* 2001;39:1263–1276. [PubMed: 11566310]
16. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci* 2002;3:201–215. [PubMed: 11994752]
17. Driver J, Baylis GC. Movement and visual attention: the spotlight metaphor breaks down. *J Exp Psychol Hum Percept Perform* 1989;15:448–56. [PubMed: 2527954]
18. Duncan J. Selective attention and the organization of visual information. *J Exp Psychol Gen* 1984;113:501–517. [PubMed: 6240521]
19. Bundesen C, Pedersen LF. Color segregation and visual search. *Percept. Psychophys* 1983;33:487–493. [PubMed: 6877996]
20. Duncan J, Humphreys GW. Visual search and stimulus similarity. *Psychol Rev* 1989;96:433–58. [PubMed: 2756067]
21. Nothdurft HC. The role of features in preattentive vision: Comparison of orientation, motion, and color cues. *Vision Res* 1993;33:1937–1958. [PubMed: 8249312]
22. Knierim JJ, Van Essen DC. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol* 1992;67:961–980. [PubMed: 1588394]
23. Nothdurft HC, Gallant JL, Van Essen DC. Response under modulation by texture surround in primate area V1: Correlates of “popout” anesthesia. *Vis. Neurosci* 1999;16:15–34. [PubMed: 10022475]

24. Kastner S, Nothdurft HC, Pigarev I. Neuronal responses to motion and orientation contrast in cat striate cortex. *Visual Neurosci* 1999;16:587–600.
25. Yantis, S. Goal-directed and stimulus-driven determinants of attentional control. In: Monsell, S.; Driver, J., editors. *Attention and Performance XVIII*. MIT Press; Cambridge, MA: 2000. p. 73-103.
26. De Fockert JW, Rees G, Frith C, Lavie N. Neural Correlates of Attentional Capture in Visual Search. *J. Cogn. Neurosci* 2004;16:751–759. [PubMed: 15200703]
27. Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, Shulman GL. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat. Neurosci* 2000;3:292–297. [PubMed: 10700263]
28. Constantinidis C, Steinmetz MA. Posterior Parietal Cortex Automatically Encodes the Location of Salient Stimuli. *J. Neurosci* 2005;25:233–238. [PubMed: 15634786]
29. Li Z. Contextual influences in V1 as a basis for pop-out and asymmetry in visual search. *Proc. Natl. Acad. Sci. USA* 1999;96:10530–10535. [PubMed: 10468643]
30. Reynolds JH, Desimone R. Interacting Roles of Attention and Visual Salience in V4. *Neuron* 2003;37:853–863. [PubMed: 12628175]
31. Itti L, Koch C. Computational modeling of visual attention. *Nat. Rev. Neurosci* 2001;2:194–203. [PubMed: 11256080]
32. Cave K, Wolfe J. Modeling the role of parallel processing in visual search. *Cognit. Psychol* 1990;22:225–271. [PubMed: 2331857]
33. Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry. *Hum. Neurobiol* 1985;4:219–227. [PubMed: 3836989]
34. Yantis S, Egeth HE. On the Distinction Between Visual Salience and Stimulus-Driven Attentional Capture. *J. Exp. Psychol. Hum. Percept. Perform* 1999;25:661–676. [PubMed: 10385983]
35. Neisser, U. *Cognitive Psychology*. Appleton-Century-Crofts; New York: 1967.
36. Gallant JL, Shoup RE, Mazer JA. A human extrastriate area functionally homologous to macaque V4. *Neuron* 2000;27:227–235. [PubMed: 10985344]
37. De Weerd P, Peralta MR, Desimone R, Ungerleider LG. Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. *Nat. Neurosci* 1999;2:753–758. [PubMed: 10412066]
38. Schiller PH, Lee K. The role of the primate extrastriate area V4 in vision. *Science* 1991;251:1251–1253. [PubMed: 2006413]
39. Wolfe JM, Cave KR, Franzel SL. Guided search: an alternative to the feature integration model for visual search. *J Exp Psychol Hum Percept Perform* 1989;15:419–433. [PubMed: 2527952]
40. Ogawa T, Komatsu H. Target Selection in Area V4 during a Multidimensional Visual Search Task. *J Neurosci* 2004;24:6371–6382. [PubMed: 15254093]
41. Carrasco M, Ling S, Read S. Attention alters appearance. *Nat Neurosci* 2004;7:308–313. [PubMed: 14966522]
42. Kahneman, D.; Henik, A. Perceptual organization and attention. In: Kubovy, M.; Pomerantz, JR., editors. *Perceptual Organization*. Erlbaum; Hillsdale, NJ: 1981. p. 181-211
43. Banks WP, Prinzmetal W. Configurational effects in visual information processing. *Percept. Psychophys* 1976;19:361–367.
44. Driver J, Baylis GC, Rafal RD. Preserved figure-ground segregation and symmetry perception in visual neglect. *Nature* 1992;360:73–75. [PubMed: 1436078]
45. Brainard DH. The Psychophysics Toolbox. *Spat. Vis* 1997;10:433–436. [PubMed: 9176952]
46. Astafiev SV, Stanley CM, Shulman GL, Corbetta M. Extrastriate body area in human occipital cortex responds to performance of motor actions. *Nat Neurosci* 2004;7:542–548. [PubMed: 15107859]
47. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. *J. Comput. Assist. Tomo* 1993;17:536–546.
48. Friston KJ, et al. Analysis of fMRI time-series revisited. *Neuroimage* 1995;2:45–53. [PubMed: 9343589]
49. Sereno MI, et al. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 1995;268:889–93. [PubMed: 7754376]

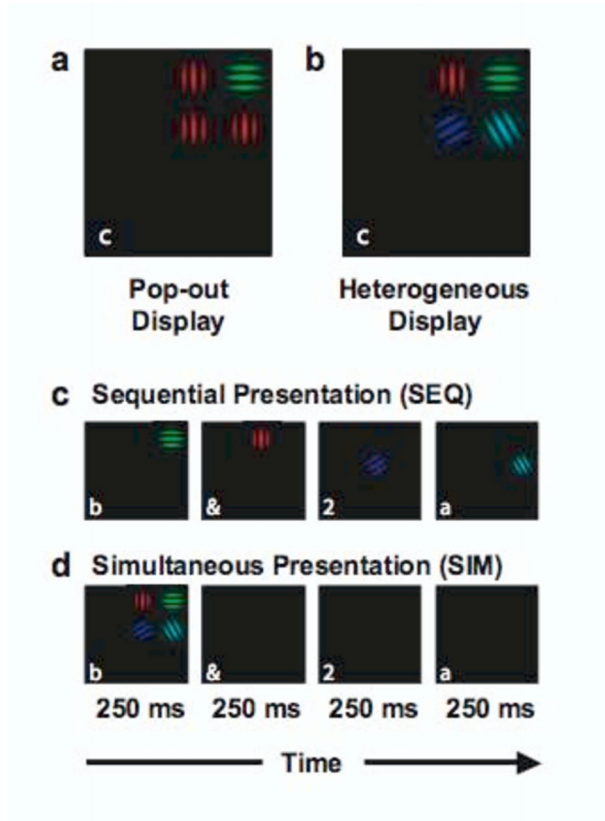


Figure 1.

Experimental design and stimuli. Four Gabor stimuli were presented in four nearby locations in the periphery of the upper right quadrant as (a) pop-out displays, in which a single item differed in color and orientation from the others, or (b) heterogeneous displays, in which all four stimuli differed in color and orientation. These stimuli were presented either (c) sequentially or (d) simultaneously. (c-d) A stimulation period of 1 s, which was repeated in blocks of 18 s, is shown for a heterogeneous display. Stimuli were presented for 250 ms, followed by a blank period of 750 ms, on average, in each location. During all conditions, subjects detected target letters at fixation (illustrated in lower left corner of each display).

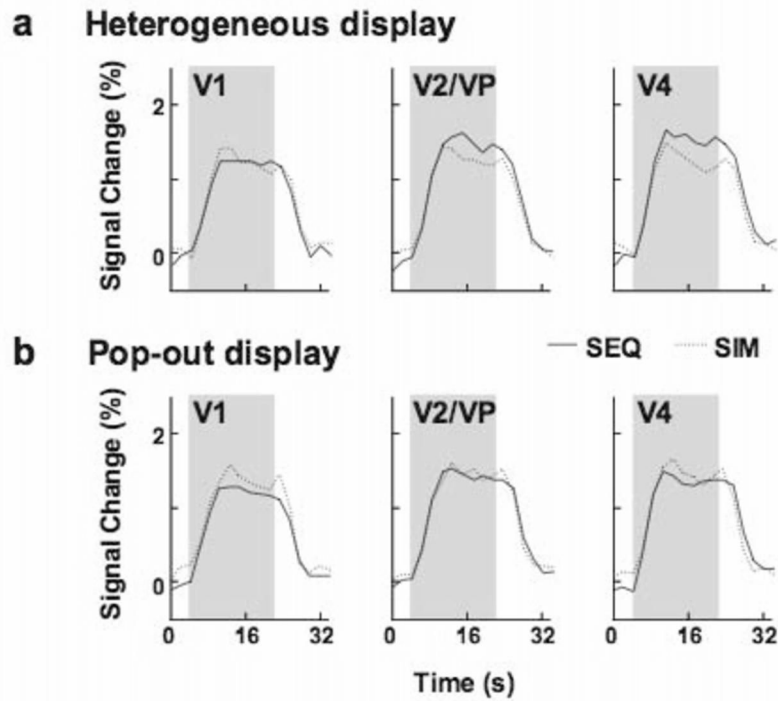


Figure 2. Time series of fMRI signals in visual cortex (Experiment 1). Group analysis ($n = 10$). Solid curves indicate activity evoked by sequential presentations and dashed curves that evoked by simultaneous presentations for (a) heterogeneous displays and (b) pop-out displays.

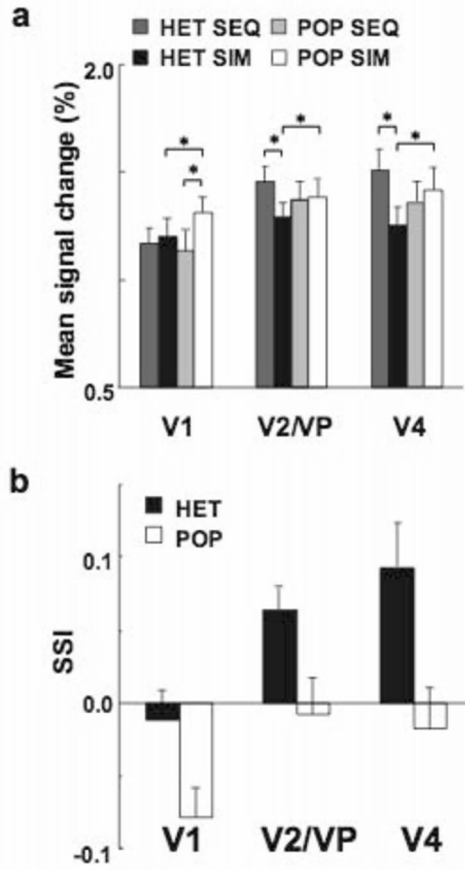


Figure 3. *Mean signal changes and SSIs in visual cortex (Experiment 1).* (a) Mean signal changes for each area and each of the four conditions were averaged across subjects ($n = 10$). For each subject, mean signal change was defined as the average of the nine peak intensities of the fMRI signal obtained during visual presentations. Asterisks indicate significant differences ($p < .05$). (b) Sensory suppression indexes (SSIs) were derived from the data shown in (a). Vertical bars indicate S.E.M.

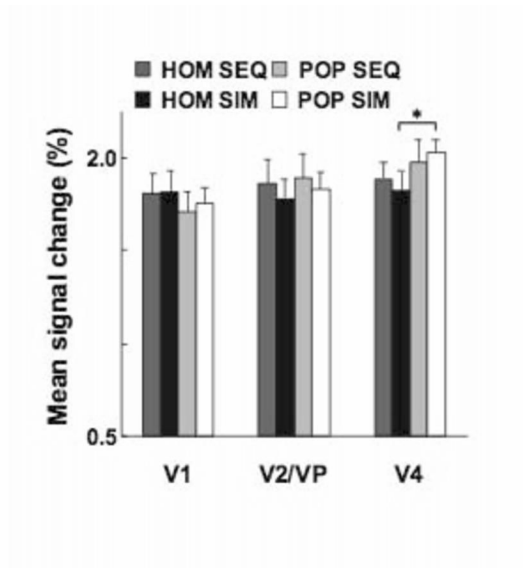


Figure 4.

Mean signal changes in visual cortex (Experiment 2). Mean signal changes, defined as described in Fig. 3, were averaged across subjects ($n = 6$) for each presentation condition of the homogeneous and pop-out displays. Asterisks indicate significant differences ($p < .05$).

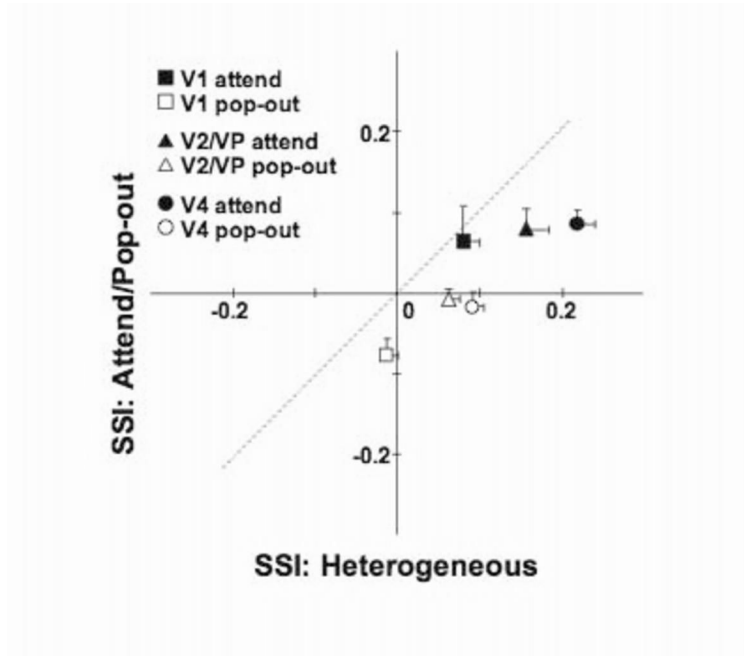
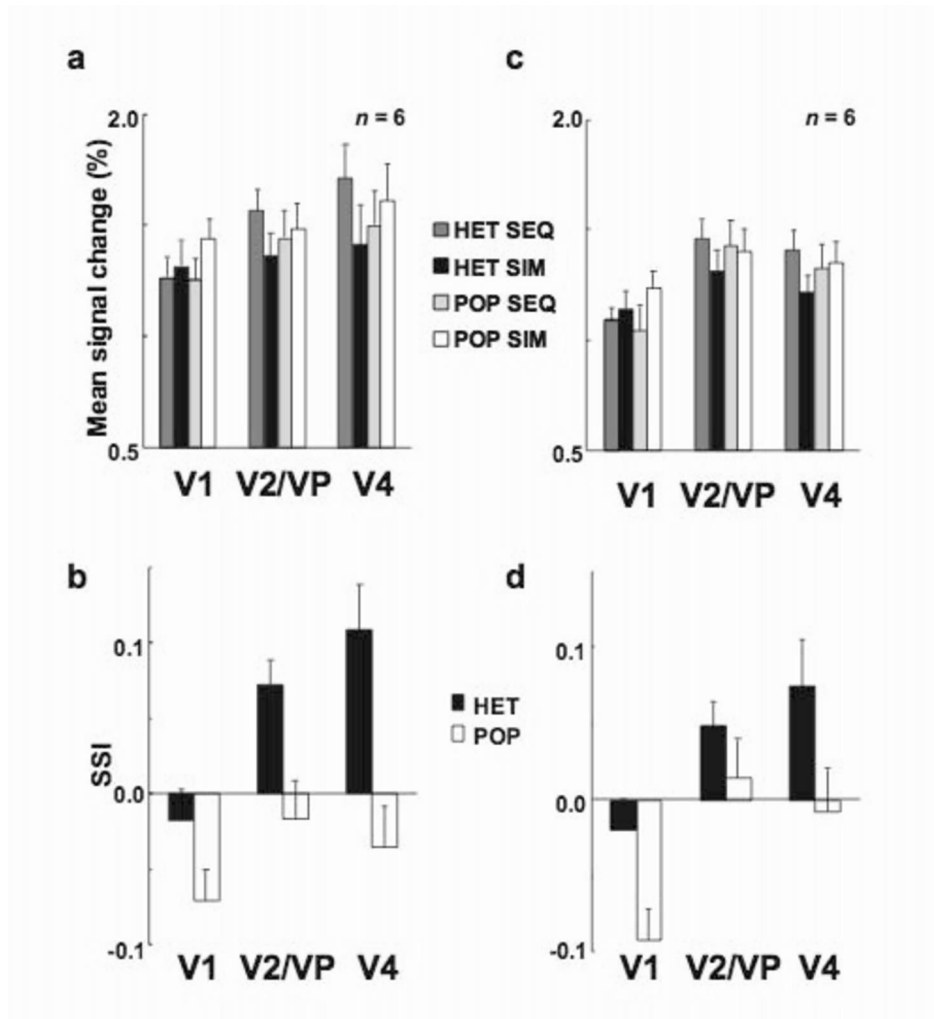


Figure 5.

Effects of pop-out and directed attention on suppressive interactions in human visual cortex. SSIs obtained for areas V1 (squares), V2/VP (triangles), and V4 (circles) are plotted for the current study, probing bottom-up effects of pop-out on suppressive interactions, (filled symbols) and for a study that probed the top-down effects of directed attention on suppressive interactions⁷ (open symbols). The horizontal axis represents the SSIs obtained for heterogeneous display conditions from the two studies, when the peripheral stimuli were unattended. The vertical axis represents the SSIs obtained for the pop-out display condition from the present study and the directed attention condition from the previous study⁷ to directly compare the top-down and bottom-up effects on suppressive interactions. The dashed line represents the points at which the two indices are equal, indicating no modulation of suppressive interactions by top-down or bottom-up influences.



Supplementary Figure 1.

Mean signal changes and SSIs obtained with the two fixation tasks. Mean signal changes for each area and each of the four stimulus conditions were averaged across subjects in the version of the fixation task that required subjects to count target letters (a) and in the version that required them to respond with a button press (b). Conventions are the same as in Figures 3 and 4. (c-d) Sensory suppression indexes (SSIs) were derived from the data shown in (a) and (b), respectively. Vertical bars indicate S.E.M.

Table 1:
Error rates and reaction times for behavioral letter detection tasks (n = 5).

Stimulus condition	Reaction Time (ms)	Percent error
Sequential Pop-out	500 (21)	31 (1)
Simultaneous Pop-out	493 (20)	13 (2)
Simultaneous Heterogeneous	503 (17)	13 (1)
Sequential Heterogeneous	501 (20)	11 (1)

Note: Standard errors of the mean are in parentheses.