

Modulation of Sensory Suppression: Implications for Receptive Field Sizes in the Human Visual Cortex

SABINE KASTNER,^{1,2} PETER DE WEERD,⁴ MARK A. PINSK,^{1,2} M. IDETTE ELIZONDO,² ROBERT DESIMONE,³ AND LESLIE G. UNGERLEIDER²

¹*Department of Psychology, Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, New Jersey 08544;* ²*Laboratory of Brain and Cognition and* ³*Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892;* and ⁴*College of Social and Behavioral Sciences, Department of Psychology, University of Arizona, Tucson, Arizona 85721*

Received 27 June 2000; accepted in final form 17 May 2001

Kastner, Sabine, Peter De Weerd, Mark A. Pinsk, M. Idette Elizondo, Robert Desimone, and Leslie G. Ungerleider. Modulation of sensory suppression: implications for receptive field sizes in the human visual cortex. *J Neurophysiol* 86: 1398–1411, 2001. Neurophysiological studies in monkeys show that when multiple visual stimuli appear simultaneously in the visual field, they are not processed independently, but rather interact in a mutually suppressive way. This suggests that multiple stimuli compete for neural representation. Consistent with this notion, we have previously found in humans that functional magnetic resonance imaging (fMRI) signals in V1 and ventral extrastriate areas V2, V4, and TEO are smaller for simultaneously presented (i.e., competing) stimuli than for the same stimuli presented sequentially (i.e., not competing). Here we report that suppressive interactions between stimuli are also present in dorsal extrastriate areas V3A and MT, and we compare these interactions to those in areas V1 through TEO. To exclude the possibility that the differences in responses to simultaneously and sequentially presented stimuli were due to differences in the number of transient onsets, we tested for suppressive interactions in area V4, in an experiment that held constant the number of transient onsets. We found that the fMRI response to a stimulus in the upper visual field was suppressed by the presence of nearby stimuli in the lower visual field. Further, we excluded the possibility that the greater fMRI responses to sequential compared with simultaneous presentations were due to exogenous attentional cueing by having our subjects count T's or L's at fixation, an attentionally demanding task. Behavioral testing demonstrated that neither condition interfered with performance of the T/L task. Our previous findings suggested that suppressive interactions among nearby stimuli in areas V1 through TEO were scaled to the receptive field (RF) sizes of neurons in those areas. Here we tested this idea by parametrically varying the spatial separation among stimuli in the display. Display sizes ranged from $2 \times 2^\circ$ to $7 \times 7^\circ$ and were centered at 5.5° eccentricity. Based on the effects of display size on the magnitude of suppressive interactions, we estimated that RF sizes at an eccentricity of 5.5° were $<2^\circ$ in V1, $2\text{--}4^\circ$ in V2, $4\text{--}6^\circ$ in V4, larger than 7° (but still confined to a quadrant) in TEO, and larger than 6° (confined to a quadrant) in V3A. These estimates of RF sizes in human visual cortex are strikingly similar to those measured in physiological mapping studies in the homologous visual areas in monkeys.

INTRODUCTION

The visual scenes that we experience in everyday life are typically cluttered with many different objects. However, only

a limited amount of this information reaches awareness or gets stored in memory, indicating that there is limited processing capacity within the visual system (Broadbent 1958; Duncan 1980; Treisman 1969). Because of this limited capacity, multiple objects in cluttered visual scenes compete for neural representation.

What are the neural correlates for competition among multiple objects? Single-cell recording studies have investigated this question by comparing responses evoked by a single visual stimulus presented within a neuron's receptive field (RF) to those evoked by the same stimulus when a second stimulus is presented simultaneously with it in the RF (Moran and Desimone 1985; Reynolds et al. 1999). It has been shown that the responses to the paired stimuli are a weighted average of the responses to the individual stimuli when presented alone. For example, if a single effective stimulus evoked a high firing rate and a single ineffective stimulus evoked a low firing rate, the responses to the paired stimuli were reduced compared with those evoked by the single effective stimulus. This result indicates that two stimuli presented together within a neuron's RF are not processed independently, but rather interact with each other in a mutually suppressive way. This sensory suppressive interaction among multiple stimuli within RFs has been interpreted as an expression of competition for neural representation, and it has been found in several areas of the visual cortex, including areas V2, V4, the middle temporal (MT) and medial superior temporal (MST) areas, and inferior temporal (IT) cortex (Miller et al. 1993; Moran and Desimone 1985; Recanzone et al. 1997; Reynolds et al. 1999; Rolls and Tovee 1995; Sato 1989).

In a recent short report, we demonstrated sensory suppressive interactions in the human visual system using functional magnetic resonance imaging (fMRI) (Kastner et al. 1998). Complex visual stimuli, known to evoke robust responses in ventral visual areas of the monkey brain, were presented in four nearby locations under two presentation conditions: sequential and simultaneous (Fig. 1, *A* and *B*). In the sequential condition, each stimulus was presented alone in one of the four locations. In the simultaneous condition, the stimuli were

Address for reprint requests: S. Kastner, Dept. of Psychology, Center for the Study of Brain, Mind and Behavior, Princeton University, Green Hall, Princeton, NJ 08544 (E-mail: skastner@princeton.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

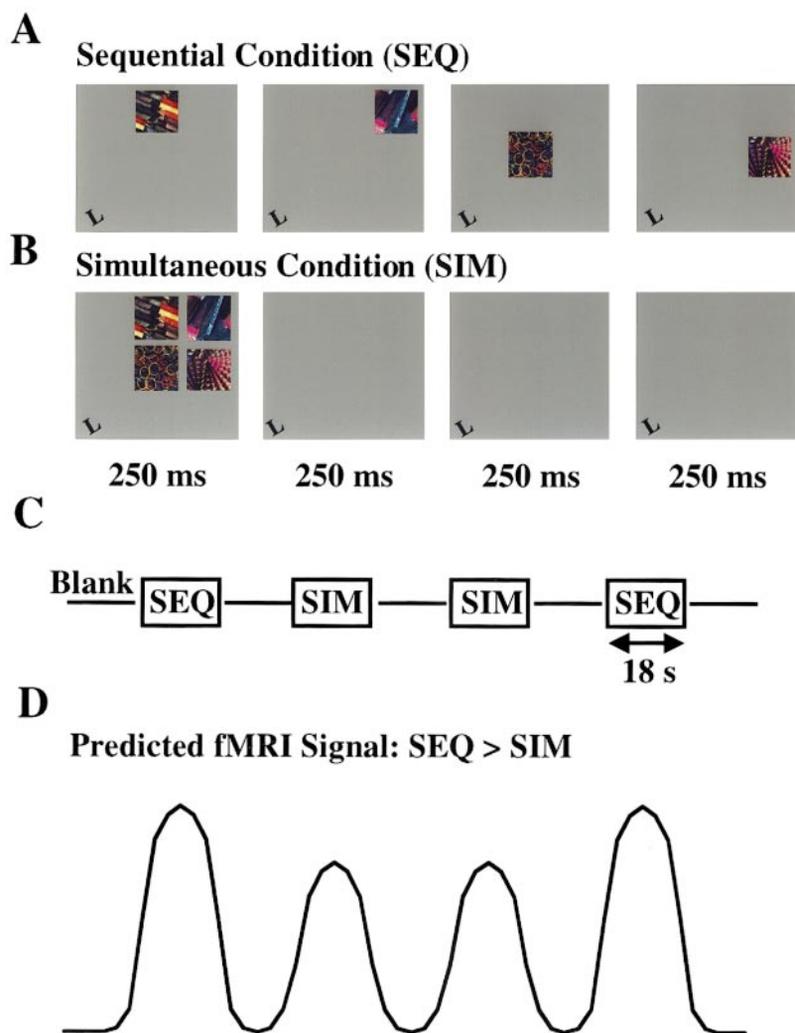


FIG. 1. Experimental design. Four complex images (each $2 \times 2^\circ$ in size) were presented at $6\text{--}10^\circ$ eccentricity from a fixation point, either sequentially (A) or simultaneously (B). Presentation time was 250 ms, followed by a blank period of 750 ms, on average, in each location. Stimulus location and order of presentation were randomized. A stimulation period of 1 s is shown, which was repeated in an ABBA scheme interleaved with equally long blank periods (C). Integrated over time, the physical stimulation parameters were identical in each of the 4 locations. However, sensory suppressive interactions could only take place in the simultaneous but not in the sequential presentation condition. Based on results from monkey physiology, we predicted therefore that the functional magnetic resonance imaging (fMRI) signal to simultaneous presentations would be smaller than to sequential presentations (D).

shown together in the four locations. Integrated over time, the amount of visual stimulation in each of the four locations was identical under the two conditions. However, suppressive interactions among stimuli within RFs could take place only in the simultaneous, not in the sequential one. Based on the results from monkey recordings, we hypothesized that the fMRI signals would be smaller during the simultaneous than during the sequential presentations because of the mutual suppression induced by competitively interacting stimuli (Fig. 1D). As predicted, simultaneous presentations evoked weaker fMRI responses than sequential presentations in V1 and ventral extrastriate areas V2/VP, V4, and TEO. Moreover, the difference in activations between sequential and simultaneous presentations increased from V1 to V4 and TEO, suggesting that the suppressive interactions were scaled to the progressive increase in RF size of neurons across these areas (Kastner and Ungerleider 2000; Kastner et al. 1998).

In the present report, we provide a full description of our previous findings (Kastner et al. 1998), including both group and single subject analyses, and we extend the findings to dorsal extrastriate areas. Further, we test the idea that sensory suppressive interactions are scaled to the RF size of neurons in visual cortex. According to the RF hypothesis, the magnitude of sensory suppression should be inversely related to the degree of spatial separation among the stimuli. If so, it should be

possible to derive an estimate of RF sizes across several areas in the human visual cortex by systematically varying the spatial separation among the stimuli and determining the degree of suppressive interactions. Preliminary reports of these findings have been published (Pinsk et al. 1999a,b).

METHODS

Subjects

Eight subjects (4 females, age: 22–35 yr) participated in the study, which was approved by the National Institute of Mental Health Institutional Review Board. The subjects participated in *experiment 1*, four in *experiment 2*, and three in *experiment 3*. All subjects were in good health with no past history of psychiatric or neurological diseases and gave their informed written consent. Subjects had normal or corrected-to-normal (with contact lenses) visual acuity.

Visual tasks

EXPERIMENT 1: SEQUENTIAL AND SIMULTANEOUS STIMULUS PRESENTATIONS. This experiment was designed to test whether multiple stimuli presented together in nearby locations interact in a mutually suppressive way in human visual cortex. Colorful, complex bitmaps were used as visual stimuli. Examples of stimuli out of a library of about 100 are given in Fig. 1, A and B. Four of these stimuli, each $2 \times 2^\circ$ in size, were presented in four nearby locations to the

upper right quadrant centered at 8° eccentricity from a fixation point. Stimuli were shown in two conditions: sequential (SEQ) and simultaneous (SIM). In the sequential condition, stimuli were presented alone in one of the four locations for 250 ms (Fig. 1A). In the simultaneous condition, the four stimuli appeared together for 250 ms (Fig. 1B). The order of stimuli and of locations was randomized. During a given scan, sequential and simultaneous conditions were presented in blocks of 18 s interleaved with equally long blank periods in the sequence SEQ—SIM—SIM—SEQ (Fig. 1C). Each scan started with a blank period of 36 s and ended with a blank period of 18 s. Different stimuli were used for different scans. T's and L's (0.6° in size) were presented for 250 ms in random order and in different orientations at 4 Hz at a central fixation point. The subjects' task was to count T's or L's at the fixation point throughout the scan. Before being scanned, subjects received three to four training sessions outside the scanner to learn to fixate well over several minutes. Eye movements were monitored during these training sessions.

EXPERIMENT 2: SPATIAL SEPARATION OF STIMULI. The purpose of this experiment was to use sensory suppressive interactions as a way to assess RF sizes in V1 and in extrastriate visual areas. The visual stimulation paradigm for *experiment 2* was the same as for *experiment 1*, except for the size of the stimuli, which was $0.5 \times 0.5^\circ$, and the eccentricity of the display, which was centered at 5.5° . The display size was parametrically varied by spatially separating the four stimuli. In the first series of experiments, display sizes of $2 \times 2^\circ$ and $7 \times 7^\circ$, presented to the upper right quadrant, were tested. In the second series of experiments, display sizes of $2 \times 2^\circ$, $4 \times 4^\circ$, and $6 \times 6^\circ$, presented to the upper right quadrant, were used (Fig. 2). The $6 \times 6^\circ$ display was also presented centered over the horizontal meridian, and thus spanned two quadrants of a hemifield (HF, $6 \times 6^\circ$; Fig. 2). The data from the two series of experiments were pooled in the analysis presented here. The subjects were engaged in the T/L task at fixation.

EXPERIMENT 3: STIMULUS PRESENTATIONS ALONG THE HORIZONTAL MERIDIAN. This experiment was designed to rule out the possibility that the differences between activations evoked by simultaneous and sequential presentation conditions were due to the faster overall presentation rate in the latter condition. That is, across the visual field, there were four stimulus onsets in the sequential condition, but only one in the simultaneous condition. We sought to demonstrate sensory suppressive interactions directly in areas that have the upper visual field (UVF) and the lower visual field (LVF) representations separated by the horizontal meridian (HM). Four complex images of $2 \times 2^\circ$ in size were presented centered at an eccentricity of 6° . One stimulus was presented just above the HM to the UVF, and three stimuli were presented just below the HM to the LVF (see Fig. 11). Stimuli were presented for 250 ms in blocks of 18 s interleaved with equally long blank periods in the following three conditions: 1) one stimulus presented to the UVF, 2) three stimuli presented to the LVF, and 3) all four stimuli presented together (Fig. 11). The order of the stimulus conditions was randomized. The rate of the presentations was 1 Hz in all conditions. Subjects were engaged with the T/L task at fixation.

Retinotopic mapping

For each subject, retinotopic mapping was performed in a separate scanning session. Areas V1, V2, and VP were identified by determining the alternating representations of the vertical and horizontal meridians, which form the borders of these areas (DeYoe et al. 1996; Engel et al. 1997; Grill-Spector et al. 1998; Sereno et al. 1995; Shipp et al. 1995; Tootell et al. 1997). This was accomplished by presenting high-contrast color and luminance checker stimuli along the meridians, flickering at 4 Hz. As it was difficult to separate V2 and VP in some subjects, activity was averaged across the two areas in the group analyses. In the context of the group analyses, the combined region will be referred to as V2. Areas V4 and TEO were identified on the basis of their characteristic UVF and LVF retinotopy. The UVF and the LVF are separated in V4 and located medially and laterally, respectively, on the posterior part of the fusiform gyrus (BA 19; see Fig. 10, Table 1). Area TEO is also located on the fusiform gyrus, just anterior to area V4 (BA 37; Table 1). This area contains a representation of the contralateral hemifield but, in contrast to area V4, without a separation of UVF and LVF (Kastner et al. 1998). Area V4 in this study likely corresponds to area V4 of McKeefry and Zeki (1997) and appears to overlap with V4v and V8 described by Hadjikhani et al. (1998). Mapping the UVF and LVF retinotopy was accomplished by presenting the complex stimuli to either the upper right or the lower right quadrant at $8\text{--}12^\circ$ eccentricity. In contrast to Hadjikhani et al. (1998), we were not able to distinguish V4v, an area with a representation of the contralateral UVF located just anterior to VP, from area V8, which they described as having both UVF and LVF representations. This discrepancy may be due to differences in retinotopic mapping procedures and/or magnetic field strength between their study and ours. Activations in area V3A were identified on the basis of their location in dorsal extrastriate cortex, where the UVF is represented among LVF representations of other visual areas (Tootell et al. 1997). Activations in area MT were identified based on the characteristic anatomical location of this area at the junction of the ascending limb of the inferior temporal sulcus and the lateral occipital sulcus (Tootell et al. 1995; Watson et al. 1993; Zeki et al. 1991). In four of the eight subjects, the locations of areas MT, V4, and TEO were confirmed by performing additional functional scans, which probed the motion or color selectivity of these areas, respectively (e.g., Beauchamp et al. 1999; Hadjikhani et al. 1998; McKeefry and Zeki 1997; Zeki et al. 1991). Talairach coordinates of visual areas are given in Table 1.

Data acquisition

Images were acquired with a 1.5 Tesla GE Signa scanner (Milwaukee, WI) using a standard head coil. Subjects were comfortably placed on their backs with their heads restrained and surrounded by soft foam to reduce head movements. Data were acquired in 26 scan sessions, each lasting 2 h. In addition, retinotopic mapping was performed in all subjects during a separate scan session. Functional images were taken with a gradient echo echo-planar imaging sequence (TR = 3 s, TE = 40 ms, flip angle = 90° , 64×64 matrix). Sixteen contiguous coronal slices were taken starting from the posterior pole (thickness: 5 mm; in

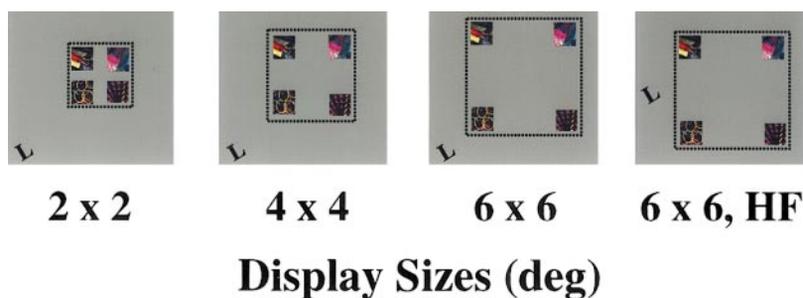


FIG. 2. Spatial separation of stimuli. Examples of display sizes used in *experiment 2*. Four stimuli, $0.5 \times 0.5^\circ$ each, were presented in displays of $2 \times 2^\circ$, $4 \times 4^\circ$, and $6 \times 6^\circ$ presented to the right upper quadrant, or in a $6 \times 6^\circ$ display presented to a hemifield (HF). All displays were centered at 5.5° eccentricity. For each display size, stimuli were presented sequentially or simultaneously in blocks of 18 s interleaved with blank periods, as in *experiment 1*.

TABLE 1. Talairach coordinates of activated areas in visual cortex (experiment 1)

Area	x	y	z	Z Score	n
<i>Visual stimulation versus blank (regressor 1)</i>					
V1	-3 ± 2	-80 ± 7	+8 ± 8	8.1 ± 2.5	8
V2	-3 ± 4	-79 ± 9	-6 ± 8	9.5 ± 2.5	8
VP	-12 ± 4	-76 ± 11	-13 ± 5	9.0 ± 3.0	4
V4	-18 ± 4	-74 ± 6	-17 ± 3	12.3 ± 1.9	8
TEO	-23 ± 6	-59 ± 9	-11 ± 5	8.3 ± 1.5	8
V3A	-21 ± 4	-91 ± 6	+24 ± 7	8.2 ± 2.3	6
MT	-44 ± 5	-75 ± 8	+8 ± 3	5.1 ± 1.8	5
<i>Sequential versus simultaneous presentations (regressor 2)</i>					
V1					
V2	-8 ± 5	-80 ± 2	-13 ± 8	3.1 ± 0.9	2
VP	-14 ± 4	-71 ± 16	-14 ± 6	3.9 ± 1.3	2
V4	-21 ± 3	-74 ± 8	-17 ± 4	5.1 ± 1.2	8
TEO	-22 ± 6	-60 ± 7	-11 ± 6	3.4 ± 0.6	8
V3A	-21 ± 3	-90 ± 2	+23 ± 6	3.3 ± 0.6	3
MT	-45 ± 4	-69 ± 8	+4 ± 10	3.1 ± 0.2	3

Values are means ± SD of peak coordinates in mm; n is number of subjects showing significant clusters of activation.

plane resolution: 2.5 × 2.5 mm). Data for *experiment 1* were acquired in one scanning session for each subject, during which 10–12 scans were taken. Data for the first series of *experiment 2* were acquired in one scanning session for each subject, during which six scans with the 2 × 2° display and six scans with the 7 × 7° display were taken. Data for the second series of *experiment 2* were acquired in three sessions for each subject. In *session 1*, six scans with a display size of 2 × 2° and six scans with a display size of 4 × 4° were taken. In *session 2*, another six scans of the 4 × 4° display size and six scans of the 6 × 6° within-a-quadrant display size were taken. In *session 3*, another six scans of the latter condition and six scans of the 6 × 6° within-a-hemifield display size were acquired. Data for *experiment 3* were acquired in one scanning session for each subject, during which 16–20 scans were taken.

Echo-planar images were compared with a co-aligned high-resolution anatomical scan of the same subject's brain taken in the same session (3D SPGR, TR = 15 ms, TE = 7 ms, flip angle = 30°, 256 × 256 matrix, FOV = 160 × 160 mm, 28 coronal slices, thickness: 5 mm). Another high-resolution anatomical scan of the whole brain (3D SPGR, TE = 5.4 ms, flip angle = 45°, 256 × 256 matrix, FOV = 240 × 240 mm, 124 sagittal slices, thickness: 1.5 mm) was taken in a different scan session to perform spatial normalization in SPM96b and for reconstruction of the cortical surface using BrainVoyager.

Visual stimuli were presented to the subjects as videotapes rear-projected onto a translucent screen placed 40 cm from the subject's feet with a magnetically shielded liquid crystal display (LCD) projector. Stimuli were viewed from inside the bore of the magnet via a mirror system attached to the head coil. Synchronization of the video presentation with the MR data acquisition was accomplished by manually starting the video the same time as the scanner.

Data analysis

Between-scan head movements were corrected by aligning each image to a mean image of one of the scans obtained in the middle of the session using Automatic Image Registration (AIR) software (Woods et al. 1993). Images were spatially smoothed in-plane with a small Gaussian filter (FWHM of 1.2 voxel lengths), and ratio-normalized to the same global mean intensity. Statistical analyses were restricted to brain voxels with adequate signal intensity (average intensity of >20% of the maximum value across voxels) and per-

formed on both smoothed and unsmoothed data. The first six images of each scan were excluded from analysis. Statistical analyses were performed using multiple regression in the framework of the general linear model (Friston et al. 1995a,b) with National Institutes of Health functional imaging data analysis program (FIDAP) software. Square-wave functions matching the time course of the experimental design were defined as effects of interest in the multiple regression model. The square-wave functions contrasted 1) visual stimulation versus blank periods (*regressor 1*), and 2) sequential versus simultaneous presentations (*regressor 2*). For each effect of interest, 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 into the model to partial out variance due to baseline shifts between time series and linear drifts within time series.

To rule out the possibility that the RF size estimates we obtained did not depend on the statistical model described above, we computed a second statistical model, in which the square-wave functions contrasted 1) sequential presentations versus blank periods and 2) simultaneous presentations versus blank periods. The resulting activation maps were then added, and RF sizes were estimated. The estimates obtained were quantitatively very similar and not significantly different from the RF size estimates derived from the original model. Therefore the RF size estimates resulting from the two statistical models indicated that the estimates did not depend on the statistical model. Because our original statistical model was the more conservative approach, the results reported below were based on this model.

Regions of interest (ROI) were located by identifying clusters of seven or more contiguous voxels. Statistical significance (P < 0.01) of these clusters was assessed using random Gaussian field methods based on their spatial extent and peak height (Friston et al. 1994; Poline et al. 1997). All statistical results have a single voxel Z threshold of 2.33 (P < 0.01, *experiment 3*), or 3.07 (P < 0.001, *experiment 1* and 2) (degrees of freedom corrected for correlation between adjacent time points). Statistically significant clusters of voxels were overlaid on structural T1-weighted scans taken in the same session and in the same plane. Activity in visual cortex was assigned to retinotopically organized areas based on meridian mapping and UVF and LVF retinotopy. For three subjects, cortical surface reconstructions, based on three-dimensional (3-D) volumetric data, were performed using BrainVoyager software (version 3.9) (Goebel et al. 1998).

All time course analyses were performed on unsmoothed data. Time series of fMRI intensities were usually averaged over all voxels in a given ROI during visual stimulation versus blank presentations and normalized to the mean intensity obtained during the baseline condition. For *experiment 2*, in which data were pooled from multiple scan sessions, the time course analysis was restricted to voxels that were consistently activated across all conditions. For each subject, the six peak intensities of the fMRI signal obtained during the sequential and simultaneous periods were averaged resulting in mean signal changes. These values were further quantified by defining a sensory suppression index [SSI = (R_{SEQ} - R_{SIM})/(R_{SEQ} + R_{SIM}); R is the averaged responses of the peak MRI intensities obtained during visual presentation blocks for a given presentation condition]. Statistical significance was assessed with repeated measures ANOVAs on the peak intensities of the fMRI signal. Two-way ANOVAs were calculated to assess significance for indexes. For each subject, Z-score maps and structural images were transformed into the standard stereotactic Talairach space (Talairach and Tournoux 1988) using SPM96b. For this purpose, structural and functional partial volumes were aligned to a high-resolution structural whole brain volume from the same subject using AIR software in Medx.

RESULTS

Experiment 1: sensory suppressive interactions among multiple stimuli

In this experiment, epochs of visual presentations alternated with blank presentations as the subjects counted T's or L's at the fixation point. The T/L task had a high attentional load to ensure proper fixation and to prevent participants from covertly attending to the peripheral stimuli. Performance measured outside the scanner in this task (75% correct on average) did not differ during blank, sequential, or simultaneous presentation periods [$F(2, 143) = 1.6, P = 0.21$]. Hence, neither presentation condition interfered with the T/L task, indicating that this task provided sufficient attentional load to preclude exogenous attentional cueing.

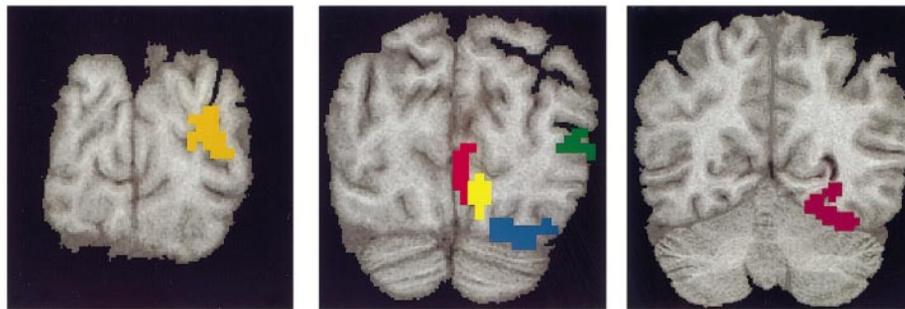
The complex stimuli, as compared with blank intervals, evoked significant activity in visual areas V1, V2, VP, V4, and TEO of the left hemisphere in all eight subjects (see Table 1). In four of the eight subjects, the border between V2 and VP could not be distinguished unequivocally. The locations of 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. In addition, the UVF representations of dorsal extrastriate areas V3A and MT were activated by the complex stimuli in six and five of the eight subjects, respectively (see Table 1). The locations of activations for a single

subject are illustrated in coronal sections at different distances from the occipital pole in Fig. 3, and on a flattened surface reconstruction of that subject's brain in Fig. 4. In Fig. 3A, the assignment of activated voxels to areas V1, V2, V4, TEO, V3A, and MT, based on meridian mapping and on UVF and LVF topography, is also shown. The activation within VP for this subject was on a different coronal section than the ones illustrated here.

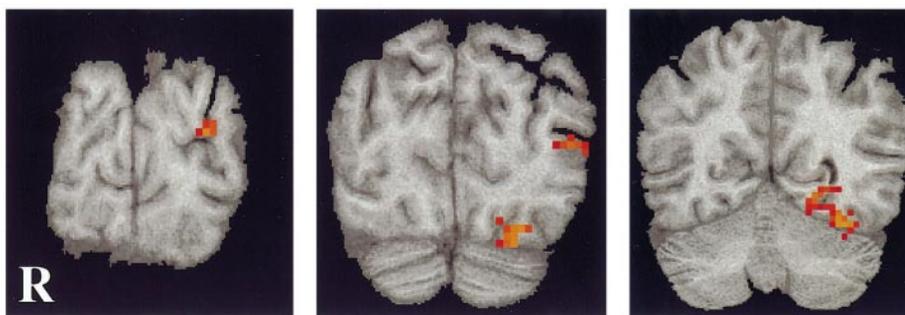
An analysis of the time series of the fMRI signal (Fig. 5) and the mean signal changes (Fig. 6A) averaged across all subjects confirmed and extended these results. Among ventral visual areas, the complex stimuli in the two conditions compared with blank periods evoked strongest responses in V4 [main effect of area: $F(3, 21) = 4.0, P < 0.05$; main effect of visual stimulation: $F(23, 161) = 15.4, P < 0.001$] with a significant interaction of area and visual stimulation [$F(69, 483) = 3.7, P < 0.001$]. There was a nonsignificant trend for the complex stimuli to evoke stronger responses in ventral extrastriate areas V4 and TEO compared with V3A and MT [$F(1, 3) = 8.7, P = 0.06$; Figs. 5 and 6A]. This trend is also apparent in the volume analysis given in Table 2 (*regressor 1*).

As predicted by our hypothesis that stimuli presented together interact in a mutually suppressive way, sequential presentations evoked stronger responses than simultaneous presentations in V4 and TEO of all eight subjects, in V3A and MT of three subjects and in V2 and VP of two subjects. However,

A



B



+ 90 mm

+ 80 mm

+ 60 mm

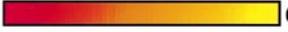
2  6
Z score

FIG. 3. Brain regions activated in human visual cortex. *A*: brain areas activated by the complex images as compared with blank presentations (*regressor 1*). Coronal slices of a single subject at different distances from the posterior pole. Activated voxels were assigned to areas V1, V2, V4, TEO, V3A, and MT based on meridian mapping and upper visual field (UVF) and lower visual field (LVF) topography. *B*: brain regions more strongly activated by sequential than by simultaneous presentations (*regressor 2*). Same subject and coronal slices as in *A*. Sequential presentations evoked significantly more activity than simultaneous presentations in V4, TEO, V3A, and MT. The number below each coronal section indicates the approximate y Talairach coordinate. R indicates right hemisphere.

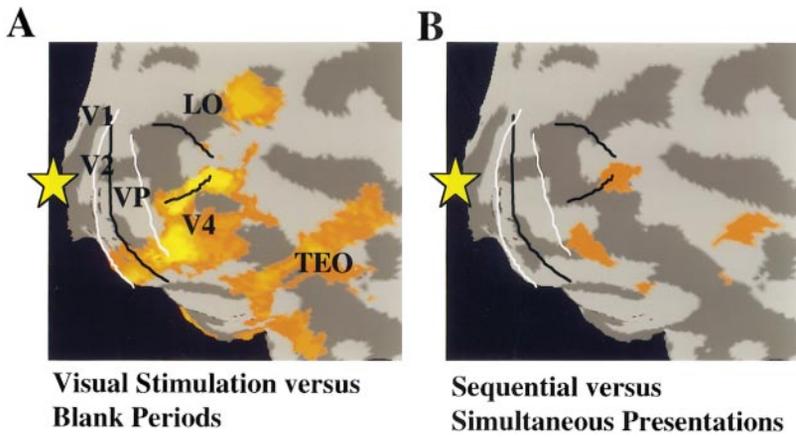


FIG. 4. Activated brain regions on flattened surface reconstructions. Same subject as in Fig. 3. The flattened surface reconstruction shows areas in ventral visual cortex from V1 to TEO. White lines indicate representations of the vertical meridians, which form the areal boundaries of V1/V2 and VP/V4, respectively. Black lines indicate representations of the horizontal meridians (HM), which form the areal boundary of V2/VP and separate the UVF and LVF within V4, respectively. Another HM representation separates V4 from the lateral-occipital complex, LO. A: colored regions indicate activations evoked by visual presentations to the periphery of the visual field compared with blank periods (*regressor 1*). Ventral visual areas from V1 to TEO were activated. In this subject, there is also activity in LO. B: colored regions indicate activations evoked by the sequential compared with the simultaneous presentations (*regressor 2*). The sequential presentations evoked stronger responses than the simultaneous presentations in V4 and TEO. The star indicates the region of foveal representations of the visual field.

no differences in responses were seen in V1 (see Table 1). This pattern of activation can also be seen for the single subject illustrated in Figs. 3 and 4, who showed significantly stronger activations evoked by the sequential presentations as compared with the simultaneous presentations in V4, TEO, V3A, and MT. For this subject, no response differences were seen in V1 or V2 (Figs. 3B and 4B).

The analysis of the time series of the fMRI signal and the mean signal changes averaged across all subjects revealed that sequential presentations evoked stronger responses than simultaneous presentations in all areas [V1: $F(1, 7) = 18.7, P < 0.01$; V2: $F(1, 7) = 30.4, P < 0.001$; V4: $F(1, 7) = 510.3, P <$

0.0001 ; TEO: $F(1, 7) = 50.0, P < 0.001$; V3A: $F(1, 5) = 7.7, P < 0.05$; MT: $F(1, 5) = 42.9, P < 0.01$; Figs. 5 and 6A]. In ventral visual areas, the difference in activations between sequential and simultaneous presentations increased gradually from V1 to V4 and TEO [interaction of area and presentation condition: $F(3, 15) = 25.1, P < 0.001$]. Interestingly, the level of activity to simultaneous presentations was similar in V1, V2, and V4, whereas the responses to sequential presentations increased from V1 to V4. The gradual increase of sensory suppression effects across ventral visual areas is also reflected in the sensory suppression index (SSI; Fig. 6B). 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. The SSI gradually increased from V1 to V4 and TEO, with significantly larger suppression effects in the latter areas [SSI: V1/V2 vs. V4/TEO, $F(1, 30) = 38.4, P < 0.0001$; Fig. 6B]. Sensory suppression effects in dorsal extrastriate areas V3A and MT were similar compared with ventral extrastriate areas V4 and TEO (Fig. 6B), even though these dorsal areas were less activated by the complex stimuli (Figs. 5 and 6A). These results are also reflected in the ratio of volumes activated during sequential versus simultaneous presentations (*regressor 2*) to those acti-

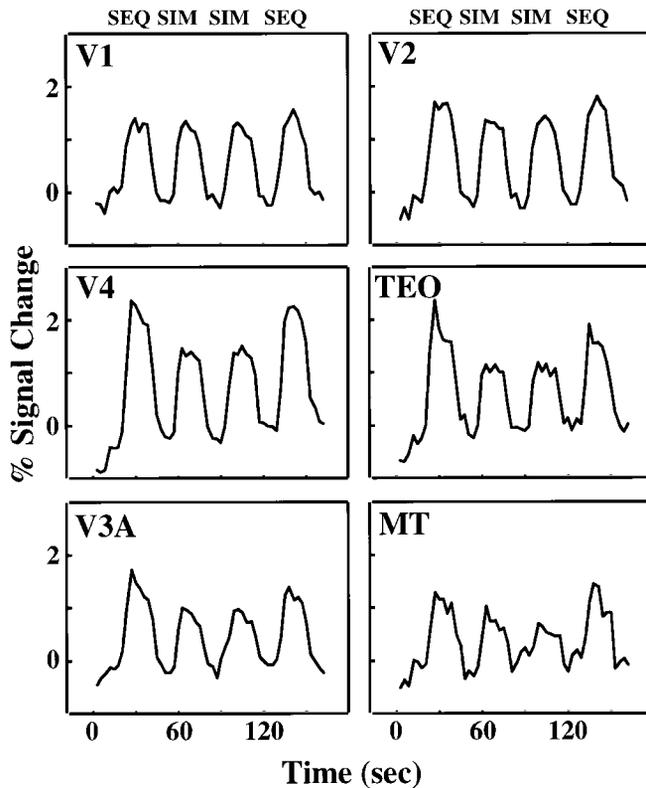


FIG. 5. Time series of fMRI signals in visual cortex. Averaged fMRI signals in V1, V2, V4, TEO, V3A, and MT ($n = 8$). Sequential presentations evoked significantly more activity than simultaneous presentations in all visual areas, but there was a graded increase in response differences in ventral visual areas from V1 to V4 and TEO. Differences in responses between sequential and simultaneous presentations were similar in ventral extrastriate areas V4 and TEO and dorsal extrastriate areas V3A and MT.

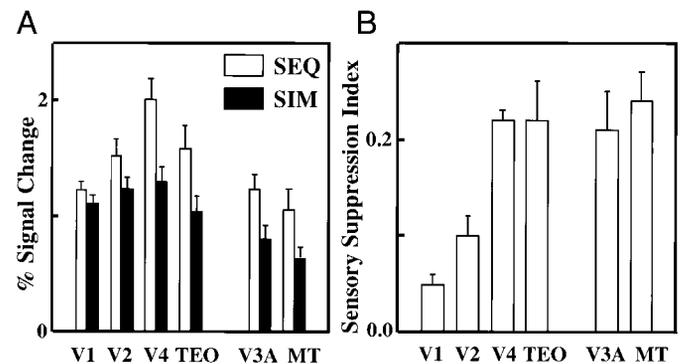


FIG. 6. Mean signal changes and sensory suppression index. A: mean signal changes in V1, V2, V4, TEO, V3A, and MT, averaged across subjects. For each subject, the 6 peak intensities of the fMRI signal obtained during sequential and simultaneous presentations were averaged. B: sensory suppression indexes (SSIs) derived for the data shown in A. SSIs increased from V1 to V4 and TEO, which suggests that the effects were scaled to the increasing receptive field (RF) sizes of neurons in these areas. SSIs were similar in ventral extrastriate areas V4 and TEO and dorsal extrastriate areas V3A and MT. Vertical bars indicate SE.

TABLE 2. Volume analysis of activated areas in visual cortex (experiment 1)

Area	Activated Volume, mm ³		Ratio: Regressor 2/ Regressor 1
	Regressor 1	Regressor 2	
V1	727 ± 155	23 ± 16	0.03
V2	973 ± 135	168 ± 52	0.17
VP	515 ± 87	172 ± 92	0.33
V4	2,613 ± 298	1,210 ± 247	0.46
TEO	1,531 ± 179	613 ± 127	0.40
V3A	938 ± 161	271 ± 93	0.29
MT	1,250 ± 507	425 ± 194	0.34

Values are means ± SE.

vated during visual stimulation versus blank (regressor 1), shown in Table 2.

Experiment 2: an estimate of RF sizes

The increase in the magnitude of the suppression index across ventral visual areas (Fig. 6B) suggests that the suppressive interactions were scaled to the progressive increase in RF size of neurons within these areas. This is illustrated schematically in Fig. 7. Because of their small RFs, individual neurons in V1 and V2 would be capable of processing information only from a very limited portion of the $4 \times 4^\circ$ display, resulting in minimal interaction effects among stimuli. In contrast, neurons in V4 and TEO with their larger RFs would process information from all four stimuli in the display, resulting in greater suppressive interaction effects. According to this interpretation, RFs of neurons in dorsal extrastriate areas V3A and MT would be similar or possibly larger in size compared with those in V4 and TEO. This RF size hypothesis does not preclude suppression arising from the surround outside the classical excitatory RF. Indeed, suppressive interactions from the RF surround have been shown in physiological recording studies (e.g., Allman et al. 1985; Desimone et al. 1985; Kastner et al. 1999; Knierim and Van Essen 1991). The hypothesis simply assumes that suppression is greatest when nearby stimuli are separated by distances that are scaled to the RF size in a given area.

According to the RF hypothesis, sensory suppressive interactions among stimuli falling within RFs should be modulated by the spatial separation of stimuli. Specifically, the

magnitude of sensory suppression should be inversely related to the degree of spatial separation among the stimuli. If so, modulation of sensory suppression by spatial separation of multiple visual stimuli may be used to derive an estimate of RF sizes across multiple areas in the human visual cortex. To test this prediction, we performed two series of experiments, in which the distance between stimuli in the display was parametrically varied. In the first series of experiments, display sizes of $2 \times 2^\circ$ and $7 \times 7^\circ$, presented to the upper right quadrant, were tested. Results will be reported for V1 and ventral extrastriate areas, because areas V3A and MT were not reliably activated in the three subjects tested in this experiment. In the second series of experiments, display sizes of $2 \times 2^\circ$, $4 \times 4^\circ$, $6 \times 6^\circ$, presented to the upper right quadrant, and $6 \times 6^\circ$, presented within a hemifield, were tested (see Fig. 2). Results will be reported for V1, ventral extrastriate areas, and V3A, but not for MT, which was not reliably activated in the four subjects performing this experiment. All displays were centered at 5.5° eccentricity.

The prediction for the first series of experiments was that increasing the display size from $2 \times 2^\circ$ to $7 \times 7^\circ$ would eliminate sensory suppressive interactions in areas V1 and V2, which have small RFs, reduce or eliminate them in area V4, which has RFs of intermediate size, but would not alter them in area TEO, which has large RFs. Time courses of the fMRI signal obtained with the two display sizes in V1, V2, V4, and TEO are shown for a single subject in Fig. 8. In V1, sensory suppressive interactions were absent with both display sizes. In both V2 and V4, the sequential presentations evoked stronger responses than the simultaneous presentations with the $2 \times 2^\circ$ display, but not with the $7 \times 7^\circ$ display. In contrast, in TEO, response differences between sequential and simultaneous presentations were found with both the $2 \times 2^\circ$ and $7 \times 7^\circ$ displays. Similar results were found with the other two subjects tested in this series of experiments. Thus as predicted, suppressive interactions were eliminated in V4, but not in TEO. Hence, these results supported the idea that increasing the distance between the stimuli in the display modulates sensory suppressive interactions.

In the second series of experiments, the display sizes were systematically varied to derive an estimate of RF sizes across multiple areas in the human visual cortex. The SSIs derived for the various display sizes tested are shown in Fig.

Receptive Field Sizes

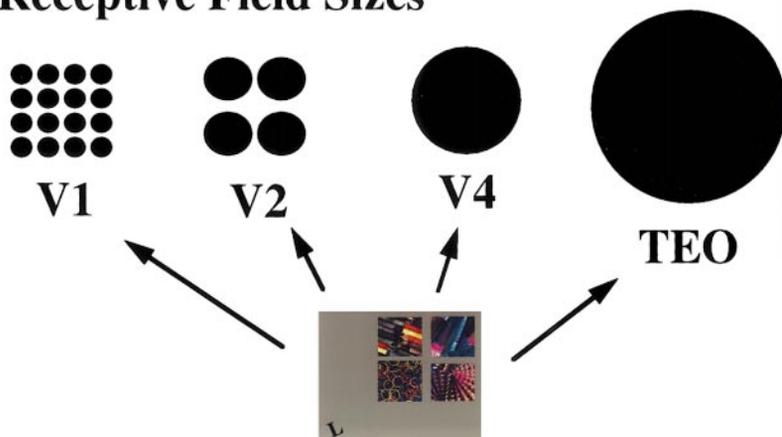


FIG. 7. The RF size hypothesis. RF sizes increase in size from V1 to TEO. The schematic drawing shows RF sizes in ventral visual cortex in relation to the $4 \times 4^\circ$ display used in experiment 1. Sensory suppression effects were likely scaled to the increasing RF sizes of neurons in these areas. Because of their small RFs, individual neurons in V1 and V2 would be capable of processing information only from a very limited portion of the $4 \times 4^\circ$ display, resulting in minimal interaction effects between stimuli; by contrast, neurons in V4 and TEO, with their larger RFs, would process information from all 4 stimuli in the display, resulting in significantly greater suppressive interaction effects.

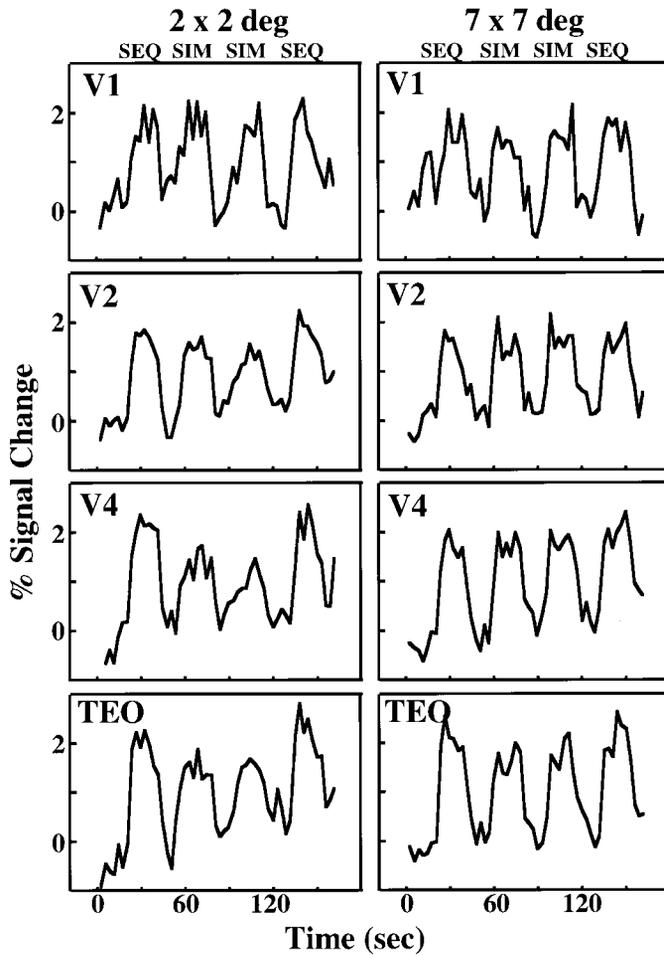


FIG. 8. Time series of fMRI signals in ventral visual areas with display sizes of $2 \times 2^\circ$ and $7 \times 7^\circ$. Four stimuli were presented either in the $2 \times 2^\circ$ display or in the $7 \times 7^\circ$ display within the same quadrant in sequential and simultaneous presentation conditions. Data are from a single subject. When stimuli were presented with the $2 \times 2^\circ$ display, response differences to sequentially and simultaneously presented stimuli were found in V2, V4, and TEO. When stimuli were presented with the $7 \times 7^\circ$ display, the response differences to the sequentially and simultaneously stimuli were abolished in V2 and V4, but unchanged in TEO.

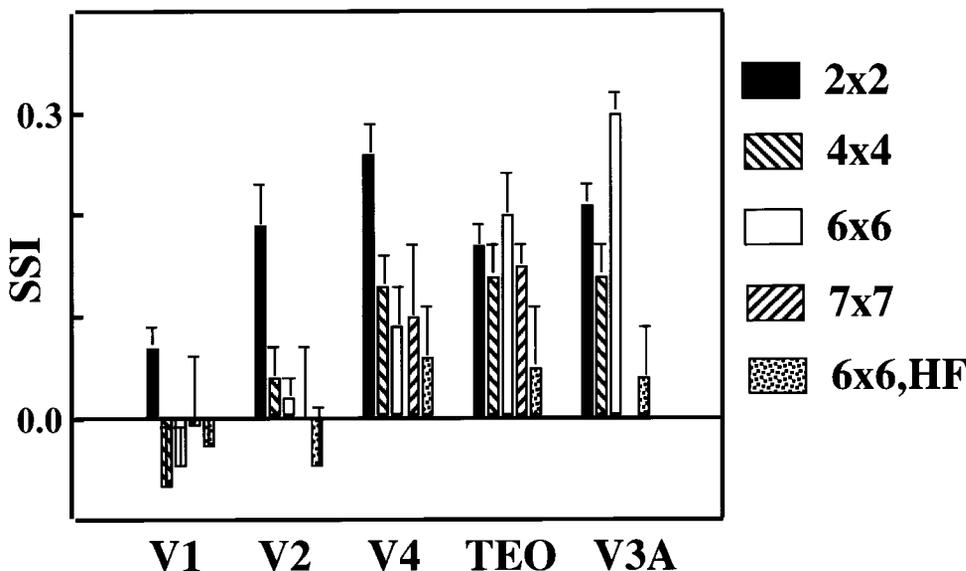


FIG. 9. SSI with various display sizes. SSIs for display sizes of $2 \times 2^\circ$, $4 \times 4^\circ$, $6 \times 6^\circ$, and $7 \times 7^\circ$, presented to a quadrant, and of $6 \times 6^\circ$, presented to a hemifield, for areas V1, V2, V4, TEO, and V3A. With a display of $2 \times 2^\circ$, sensory suppression was significant in all areas, but V1. With a display size of $4 \times 4^\circ$, sensory suppression was significant in V4, TEO, and V3A, but not in V1 or V2. With display sizes of $6 \times 6^\circ$ and $7 \times 7^\circ$ within-a-quadrant, there were significant suppressive interactions in TEO and V3A, but not in the remaining areas. Finally, no sensory suppression was seen with the $6 \times 6^\circ$ display presented to a hemifield in any of these areas (6×6 , HF).

9. The $2 \times 2^\circ$ display size evoked significant sensory suppression in all visual areas, but V1 [V2: $F(1, 7) = 22.7$, $P < 0.01$; V4: $F(1, 7) = 53.8$, $P < 0.001$; TEO: $F(1, 6) = 25.9$, $P < 0.01$; V3A: $F(1, 3) = 54.5$, $P < 0.01$]. The $4 \times 4^\circ$ display induced suppressive interactions in V4, TEO, and V3A [V4: $F(1, 4) = 9.9$, $P < 0.05$; TEO: $F(1, 4) = 26.1$, $P < 0.01$; V3A: $F(1, 3) = 11.8$, $P < 0.05$], but not in V1 or V2. The $6 \times 6^\circ$ within-a-quadrant display evoked significant suppressive interactions in TEO and V3A [TEO: $F(1, 4) = 25.3$, $P < 0.01$; V3A: $F(1, 3) = 24.8$, $P < 0.05$], but not in V1, V2, or V4. Finally, no significant sensory suppressive interactions were seen in any of these areas when the $6 \times 6^\circ$ display spanned two quadrants of a hemifield. A two-way ANOVA of the SSIs revealed a main effect of display size [$F(3, 74) = 13.8$, $P < 0.0001$], a main effect of area [$F(4, 74) = 19.0$, $P < 0.0001$], and a significant interaction of display size and area [$F(12, 74) = 2.0$, $P < 0.05$]. From these experiments, at an eccentricity of 5.5° , RF sizes were estimated to be $< 2^\circ$ in V1, $2\text{--}4^\circ$ in V2, and $4\text{--}6^\circ$ in V4. In TEO and V3A, the RFs were larger than $6\text{--}7^\circ$, but still confined to a single quadrant of the contralateral hemifield.

Experiment 3: a direct demonstration of sensory suppression

In the experiments described thus far, the stimulus presentation rate at any one of the four locations was 1 Hz in both the sequential and simultaneous conditions. However, across the visual field the overall presentation rate in the two conditions differed. To rule out the possibility that the differential responses evoked by the two presentation conditions reflected differences in overall stimulus presentation rate, we designed an experiment to demonstrate suppressive interactions while the presentation rate was held constant. The stimulus display was arranged so that one of the four stimuli was presented just above the HM to the UVF and the other three stimuli were presented just below the HM to the LVF (see outlines in Fig. 11). The idea of this experiment was that nearby stimuli placed on opposite sides of the HM may competitively interact in areas with spatially separated UVF and LVF representations, such as V2 and V4. Although the stimuli were placed on

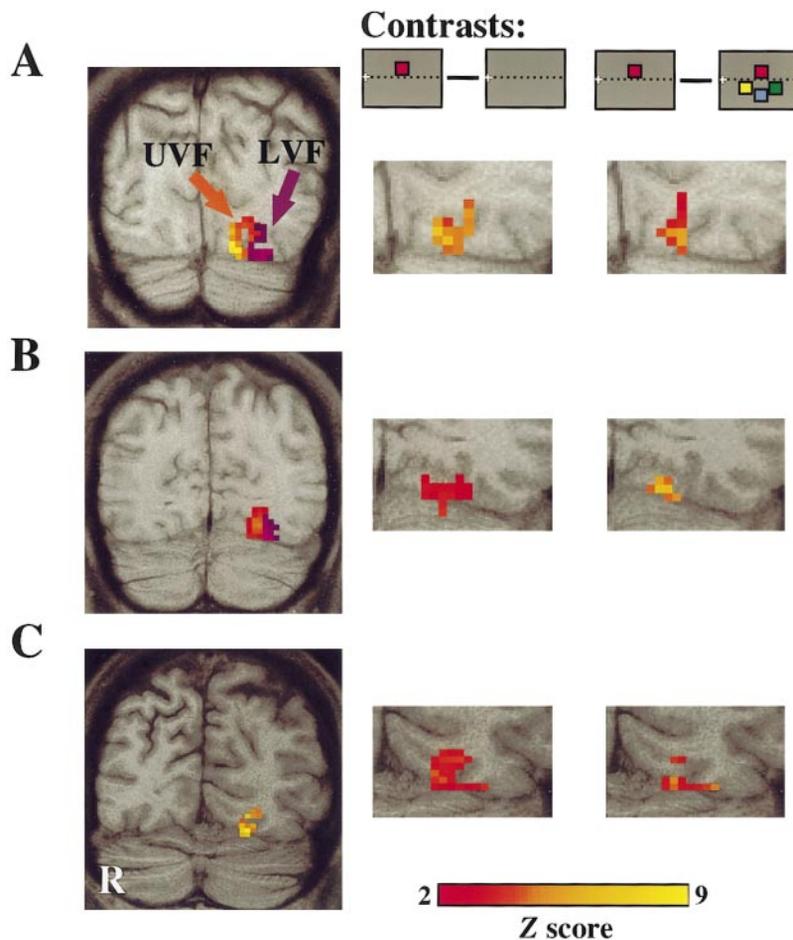


FIG. 10. A direct demonstration of sensory suppression. A–C: data for 3 individual subjects. *Left*: topography of area V4. The representation of V4's UVF and its LVF are located medially and laterally, respectively, in separated but in adjacent locations on the fusiform gyrus. The UVF and the LVF are split along the HM. In the subject shown in C, the LVF was not activated in this coronal plane. *Middle*: the activity evoked by a single stimulus ($2 \times 2^\circ$) presented at 8° eccentricity just above the HM, as compared with blank presentations, was confined to V4's UVF representation. *Right*: more activity was evoked in V4's UVF when the stimulus was presented alone than when it was shown together with 3 stimuli in the LVF, presented just below the HM. In all presentation conditions, stimuli were presented for 250 ms at 1 Hz.

opposite sides of the HM, they presumably fell within the surrounds of cells in the adjacent visual quadrant, close to the classical RFs.

Individual results for the three subjects tested in this experiment are shown for area V4 in Fig. 10. In V4, the UVF and LVF are represented medially and laterally, respectively, on the fusiform gyrus, separated along the HM (see *left panel* in Fig. 10 for V4 topography in the 3 subjects; A–C) (cf. also McKeefry and Zeki 1997). The responses to the single stimulus presented to the UVF compared with blank presentations are shown in the *middle panel* of Fig. 10. As shown in the *right panel* of Fig. 10, these responses were significantly reduced when the same stimulus was presented together with the three stimuli in the LVF. The averaged signal change was significantly different in the two conditions in V4's UVF across the subjects ($P < 0.01$; Fig. 11). It should be noted that there was considerable signal spread into V4's UVF evoked by the three stimuli presented to the LVF. Because of this spread, the actual suppression effect is likely to be larger than that reflected in the difference in responses to the single stimulus and to the four stimuli. Unlike in V4, in V2, the difference in responses to the single stimulus and to the four stimuli was not significant (Fig. 11). Thus with this experimental design, suppressive interactions among nearby stimuli could be demonstrated only in an area with sufficiently large RFs and surrounds to be influenced by all of the stimuli in the display. These findings in V4 rule out stimulus presentation rate as the explanation for the suppressive effect.

DISCUSSION

Using fMRI, we have demonstrated, in multiple areas of human visual cortex, stronger responses evoked by visual stimuli presented sequentially in four nearby locations than by the same stimuli presented simultaneously. Based on evidence from monkey physiology, the reduced responses to simultaneously presented stimuli were interpreted as sensory suppressive interactions among multiple stimuli that compete for neural representation. The suppressive interactions increased progressively in ventral visual processing areas, with smallest effects in V1 and strongest effects in V4 and TEO, suggesting that the suppressive effects were scaled to the increasing RF sizes of neurons in these areas. In addition, sensory suppressive interactions in dorsal extrastriate areas V3A and MT were found to be of similar magnitude to those in ventral extrastriate areas V4 and TEO. Importantly, sensory suppressive interactions were shown to be modulated by parametrically increasing the spatial separation of the stimuli in the display. In this way, an estimate of RF sizes for multiple visual cortical areas was derived.

Relation to monkey physiology

Single-cell recording studies in monkey visual cortex have investigated sensory suppressive interactions among multiple stimuli. In these studies, responses to a single stimulus presented within a neuron's RF have been compared with the responses to that same stimulus presented together with a

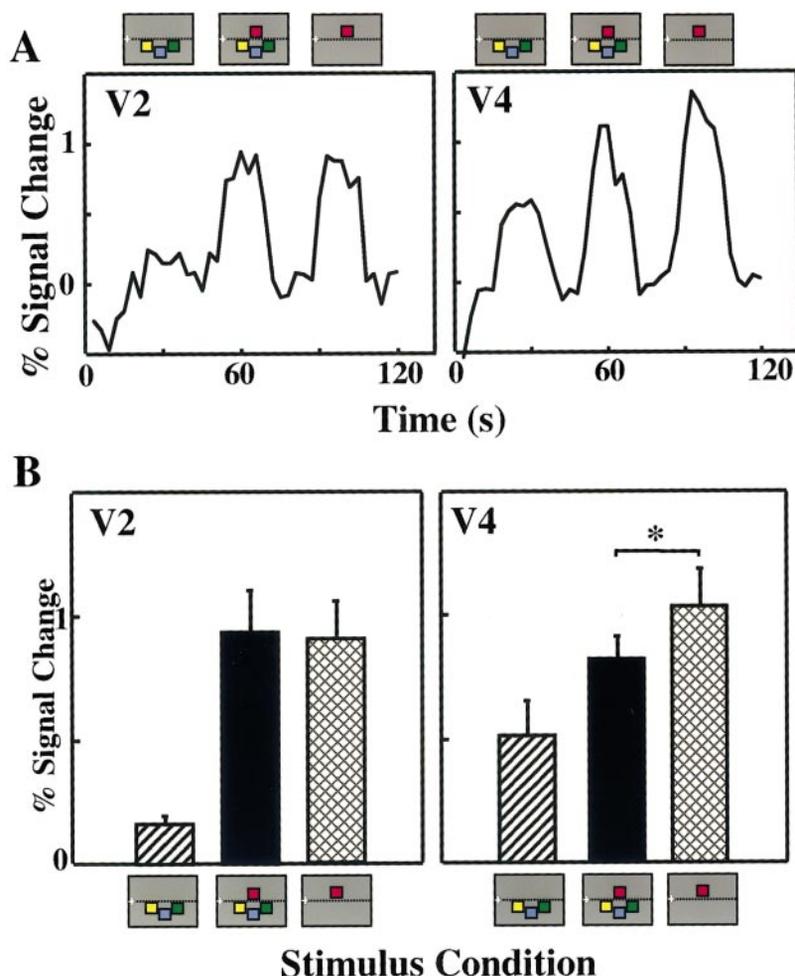


FIG. 11. Time series of fMRI signals and mean signal changes in V2 and V4 (*experiment 3*). *A*: time series of fMRI signals in V2's and V4's UVF, averaged across subjects ($n = 3$). A single stimulus presented just above the HM evoked stronger responses than when the same stimulus was presented together with 3 stimuli shown just below the HM. This effect was found in V4, but not in V2. *B*: mean signal changes to the single stimulus presented to the UVF above the HM, the 3 stimuli presented to the LVF stimulus below the HM, and the 4 stimuli presented together, averaged across subjects. For each subject, the 6 peak intensities of the fMRI signal obtained during the different conditions were averaged. * $P < 0.05$.

second stimulus within the RF. In areas V4 and MT/MST, it has been shown that the addition of an ineffective stimulus, eliciting a low firing rate, to an effective stimulus, eliciting a high firing rate, drove the neuron's firing rate down (Recanzone et al. 1997; Reynolds et al. 1999). Similarly, in IT cortex, a high proportion of neurons exhibited weaker responses to pairs of stimuli relative to the responses to the effective single stimulus of the pair (Miller et al. 1993; Rolls and Tovee 1995; Sato 1989). Because the responses to the paired stimuli did not summate in these studies, these findings suggest that two stimuli present simultaneously in a neuron's RF are not processed independently. Rather, multiple stimuli appear to interact in a mutually suppressive way.

Based on these results from monkey physiology, we hypothesized that fMRI signals evoked by simultaneously presented stimuli would be weaker than those evoked by sequentially presented stimuli, due to the putative suppressive interactions that would take place among the stimuli in the simultaneous, but not in the sequential condition. In accordance with this hypothesis, we found that simultaneously presented stimuli indeed evoked weaker activations than sequentially presented stimuli in multiple visual areas. Moreover, the effects increased gradually from V1 to V4 and TEO, with the strongest effects in V4, TEO, MT, and V3A. As these areas have RFs of intermediate or large size, in which the four stimuli of the $4 \times 4^\circ$ display could interact, we suggest that the suppressive effects occur predominantly among multiple stimuli within RFs.

It is unlikely, however, that sensory suppressive interactions among multiple stimuli within RFs accounted for the suppressive effects found in areas V1 and V2, where only a portion of the display would fit within the neurons' small RFs. Although the suppressive effects were small in these areas, they were significant. It may be that the suppression found in these areas in the simultaneous condition was due to surround inhibition, induced from regions beyond the classical RF. Surround inhibition, a reduction in the response to a stimulus within the RF by stimuli presented outside the classical RF, has been demonstrated for V1 (e.g., Kastner et al. 1999; Knierim and Van Essen 1991) and extrastriate areas MT and V4 (Allman et al. 1985; Desimone and Schein 1987; Desimone et al. 1985). For example, in V1, it has been shown that the responses to a bar stimulus presented in a RF were smaller when that stimulus was surrounded by similar bar stimuli presented outside the RF than when the same bar stimulus was presented in the RF without the surrounding stimuli. Surround inhibition has been shown to operate over large spatial scales, up to $10\text{--}12^\circ$ (Knierim and Van Essen 1991; Lamme 1995; Nothdurft et al. 1999) and likely accounts, at least in part, for the suppressive effects found in V4, when stimuli were placed above or below the HM. The fact that these effects are long ranging may also explain the suppression obtained during simultaneous compared with sequential presentations even in areas with small RFs.

Even in areas beyond V1 and V2, it is difficult to quantita-

tively relate the magnitude of activation in the sequential condition to that in the simultaneous condition. As described above, single-cell recording studies have shown that responses to multiple competing stimuli within RFs are best described as a weighted average of the responses to each of the stimuli presented alone, due to suppressive interactions within the RF (Recanzone et al. 1997; Reynolds et al. 1999). The complex, colorful stimuli that we used were chosen because they have been shown to be effective in driving neurons in ventral visual areas of monkeys (Chelazzi et al. 1993). In the course of our experiments, we used a large library of about 100 different stimuli, which were similar in terms of their general properties, such as colorfulness and texture richness. Therefore we assume that all stimuli were equally effective in driving neural responses in these areas, with the qualification that the most central stimulus in the display probably contributed the most to the integrated response in the sequential condition due to the cortical magnification factor. The activity evoked by the stimuli in the sequential presentation was presumably close to the sum of the responses to each stimulus presented alone, integrated over time. By contrast, the activity evoked by the multiple stimuli presented simultaneously was presumably closer to the weighted average of the responses to the single stimuli presented alone. Furthermore, as indicated above, the four stimuli in our display did not contribute equally to the response, inasmuch as one stimulus in the display was presented closer to the fovea than the others and thus almost certainly dominated the population response. The major contribution of the three more peripheral stimuli in the simultaneous display was probably to reduce the response to the more central stimulus. Given the differential contributions of the central and peripheral stimuli, in conjunction with the limited spatial and temporal resolution of the fMRI method, we are not able to assess the relationship between the responses to the sequential and simultaneous stimulus displays quantitatively, as has been possible using single-cell recordings (Reynolds et al. 1999). However, clearly the physiological results predict that the responses to the simultaneously presented stimuli should be qualitatively smaller than the responses to the sequentially presented stimuli, and that the spatial dependence of this relationship should be closely linked to RF size, as we have found.

RF sizes in human and monkey visual cortex

Single-cell recording studies in the monkey have provided detailed topographical maps of retinotopically organized visual areas. One key characteristic is the increase in RF sizes at successive stages of visual processing. For example, at parafoveal eccentricities, RFs of neurons are about 1.5° in V1, and about 4° in V4, whereas neurons in TE have a median RF size of $26 \times 26^\circ$ (Desimone and Gross 1979; Gattass et al. 1981, 1988; Van Essen et al. 1984). Functional brain imaging studies have begun to reveal a remarkably similar topographical organization within the human visual cortex (for review see Courtney and Ungerleider 1997; Tootell et al. 1996). However, so far, RF sizes in human visual cortex have not been determined. Based on our observation that the sensory suppression effects gradually increased from V1 to V4 and TEO, we hypothesized

that these effects were scaled to the RF sizes of neurons in these areas. If so, we expected that sensory suppression would be modulated by spatially separating the stimuli in the display. Moreover, the magnitude of the suppression effect should be inversely related to the degree of spatial separation among the stimuli. In agreement with these predictions, separating the stimuli by 4° abolished sensory suppressive interactions in V2, reduced them in V4, but did not affect them in TEO. Separating the stimuli by $6\text{--}7^\circ$ led to a further reduction of sensory suppression in V4, but again it had no effect in TEO. Thus by systematically varying the spatial separation among the stimuli and measuring suppressive interactions, it was possible to get an estimate of RF sizes across several visual areas in the human cortex. The RFs were estimated, at an eccentricity of about 5° , to be $<2^\circ$ in V1, in the range of $2\text{--}4^\circ$ in V2, in the range of $4\text{--}6^\circ$ in V4, larger than 7° in TEO, and larger than 6° in V3A, but for both TEO and V3A, still confined to a quadrant.

In monkeys, RF sizes have been defined at the level of single cells. Here, we have measured hemodynamic responses, that is, BOLD contrast, to determine RF sizes in the human visual cortex. It should be noted that there are several important differences between these two methods. First, it is not known how single-unit activity translates into hemodynamic responses. There is evidence that hemodynamic responses best reflect local field potentials rather than single-unit activity (Logothetis et al. 2000). Second, we have investigated the responses of large populations of neurons, that is spatially integrated signals from entire visual areas, rather than localized signals as in single-unit recordings. Population responses integrated over large cortical areas have never been measured using single-cell recordings. In addition, our RF size estimates depend on the assumption that RF sizes and sensory suppression effects scale with a factor of one. Because the true scale factor is not known from physiological studies, these estimates represent approximate, but not absolute values. For example, as discussed above, it is possible that suppressive effects from beyond the classical RF contributed to the overall suppression effect measured with our paradigm. If so, this may have resulted in an overestimation of RF sizes. Finally, it is possible that the integration of neural activity evoked by stimuli presented over extended periods of time (e.g., 18 s with our paradigm) introduces nonlinearities between the neural and the hemodynamic measures. From the studies of Boynton et al. (1996), there is no evidence for such nonlinearities for the time periods used in this study. However, because we did not probe such nonlinearities directly, more research is needed to resolve this particular issue. Given all these caveats, it is remarkable

TABLE 3. *Receptive field sizes in human and monkey visual cortex at 5.5° eccentricity*

Area	Human	Monkey*
V1	<2	1.5
V2/VP	2–4	2.5
V4	4–6	4
TEO	$>7^\dagger$	8
V3A	$>6^\dagger$	

Values are in deg. * From Gattass et al. (1981, 1988), Van Essen et al. (1984), and Boussaoud et al. (1991). \dagger Confined to a quadrant.

that our estimates of RF sizes in human visual cortex turned out to be strikingly similar to those measured in the putative homologous visual areas of monkeys, as shown in Table 3 (Boussaoud et al. 1991; Gattass et al. 1981, 1988; Van Essen et al. 1984). Importantly, our findings indicate that, as in monkey visual cortex, RF sizes of neurons in human visual cortex increase at successive stages of processing, in accordance with preliminary findings from Smith et al. (1999). Our findings strongly support the notion that results from monkey physiology can be used to derive hypotheses for human fMRI studies despite the uncertainties in terms of the translation of single-unit activity into hemodynamic responses, and the integration of signals over space and time in fMRI studies compared with physiology studies.

Transient onset effect, hemodynamic rate effect, exogenous attentional cueing: alternative accounts?

In *experiments 1* and *2*, sensory suppressive interactions among multiple competing stimuli were probed in a design in which the stimuli were presented sequentially and simultaneously in four nearby locations. It would have been ideal to probe sensory suppressive interactions among multiple stimuli more directly by comparing the responses to a single stimulus presented alone to the responses to the same stimulus presented together with multiple competing stimuli. Such a design would have been exactly comparable to those typically used in the physiology studies described above. However, such a design would have required us to spatially resolve responses to single stimuli presented in nearby locations, which is not possible in many areas of visual cortex using conventional fMRI techniques at 1.5T.

We were able to separate activations in some visual areas by presenting stimuli on opposite sides of the horizontal meridian (*experiment 3*); however, the suppressive interactions in this experiment turned out to be small and only existent in areas with sufficiently large RFs, such as V4. Therefore unlike the design used in *experiment 1* and *2*, the design of *experiment 3* neither allowed us to compare sensory suppression effects across multiple visual areas nor to derive an estimate of RF sizes by modulation of sensory suppression. However, the design used in *experiments 1* and *2* raises certain questions regarding our interpretation that sensory suppression accounts for the signal difference found between sequentially and simultaneously presented stimuli.

Although the physical stimulation parameters in each of the four locations were identical in both conditions, there were four transient onsets during sequential presentations compared with one onset during simultaneous presentations. Thus the stronger neural responses to the sequential presentations compared with the simultaneous presentations in areas with RFs of intermediate or large size may be due to differences in transient onsets rather than sensory suppressive interactions among competing stimuli. To rule out this possibility, we conducted *experiment 3*, in which the presentation rate was kept constant. We used a diamond-shaped configuration of stimuli, presented along the HM, and exploited the fact that the UVF and LVF representations are separated along the HM in V2 and V4. This anatomical organization allowed us to distinguish activations evoked by

stimuli presented to the UVF and LVF in locations close to the HM. Thus we were able to investigate the activations evoked by a single stimulus presented to the UVF and compare them with activations evoked when the same stimulus was shown together with three other stimuli presented to the LVF. In both conditions, the presentation rate was the same. The results demonstrated that the activation in V4 evoked by a single stimulus presented in the UVF was reduced when that same stimulus was presented simultaneously with three nearby stimuli in the LVF. Because the stimulus presentation rate and onset transients in the two conditions were identical, sensory suppressive interactions can be the only interpretation of the result. It is interesting to note that the suppressive effect was not seen in V2, which is likely due to the fact that the RFs of neurons in V2 were too small to encompass the four stimuli in the display.

Another issue concerning the design used in *experiments 1* and *2* is whether the differences in presentation rate during sequential and simultaneous conditions could have led to differences in the evoked hemodynamic responses. The dependence of the hemodynamic response on presentation rate is well-known (e.g., Rees et al. 1997; Schneider et al. 1994). Typically, for both striate and extrastriate areas, the hemodynamic response increases with increasing presentation rate. Therefore across several visual areas, one would expect a similar increase in response as stimulation rate increases. However, in contrast to this prediction, we found a graded increase in response differences to sequentially and simultaneously presented stimuli in ventral visual areas. Moreover, there was a modulation of response differences in the two conditions by spatially separating the stimuli. Both findings cannot be explained by a presentation rate account. Further, in an attention study using the same visual paradigm, we found stronger effects of attention on simultaneously than on sequentially presented stimuli (Kastner et al. 1998), even though attentional effects on stimuli differing in rate should be similar (Rees et al. 1997). Taken together, these arguments strongly speak against the possibility that differences in presentation rate and corresponding differences in the evoked hemodynamic response could account for the present findings.

A final issue concerning the design used in *experiments 1* and *2* is the possibility that the sequentially presented stimuli led to stronger exogenous attentional cueing due to the four transient onsets during the sequential presentations as compared with the one transient onset during the simultaneous presentations. If the larger responses in the sequential condition were due to stronger exogenous attentional cueing, then one would expect stronger behavioral interference in the T/L task during the sequential compared with the simultaneous condition. However, in behavioral studies conducted outside the scanner, we showed that the subjects' performance did not differ during blank, sequential, and simultaneous presentations, indicating that the T/L task provided sufficient attentional load to preclude exogenous attentional cueing in either presentation condition. Finally, if the differences in activation between sequential and simultaneous presentations were due to greater exogenous cueing in the sequential condition, then increasing the separation between stimuli should not make any difference.

However, we showed that increasing the spatial separation among stimuli modulated the response differences to sequentially and simultaneously presented stimuli, which cannot be explained by exogenous attentional cueing. Rather, our data are best interpreted in terms of sensory suppressive interactions among multiple visual stimuli that compete for neural representation within RFs. The data presented in this paper cannot be explained in terms of any of the alternative accounts discussed.

This study was supported in part by Deutsche Forschungsgemeinschaft Grant Ka 1284/1-1 to S. Kastner.

REFERENCES

- ALLMAN J, MIEZIN F, AND MCGUINNESS E. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev Neurosci* 8: 407–429, 1985.
- BEAUCHAMP MS, HAXBY JV, JENNINGS JE, AND DEYOE EA. An fMRI version of the Farnsworth Munsell 100-Hue Test reveals multiple color-selective areas in human ventral occipitotemporal cortex. *Cereb Cortex* 9: 257–263, 1999.
- BOUSSAOUD D, DESIMONE R, AND UNGERLEIDER LG. Visual topography of area TEO in the macaque. *J Comp Neurol* 306: 554–575, 1991.
- BOYNTON GM, ENGEL SA, GLOVER GH, AND HEEGER DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207–4221, 1996.
- BROADBENT DE. *Perception and Communication*. London: Pergamon, 1958.
- CHELAZZI L, MILLER EK, DUNCAN J, AND DESIMONE R. A neural basis for visual search in inferior temporal cortex. *Nature* 363: 345–347, 1993.
- COURTNEY SM AND UNGERLEIDER LG. What fMRI has taught us about human vision. *Curr Opin Neurobiol* 7: 554–561, 1997.
- DESIMONE R AND GROSS CG. Visual areas in the temporal cortex of the macaque. *Brain Res* 178: 363–380, 1979.
- DESIMONE R AND SCHEIN SJ. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurosci* 57: 835–868, 1987.
- DESIMONE R, SCHEIN SJ, MORAN J, AND UNGERLEIDER LG. Contour, color, and shape analysis beyond the striate cortex. *Vision Res* 25: 441–452, 1985.
- DEYOE EA, CARMAN GJ, BANDETTINI P, GLICKMAN S, WIESER J, COX R, MILLER D, AND NEITZ J. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci USA* 93: 2382–2386, 1996.
- DUNCAN J. The locus of interference in the perception of simultaneous stimuli. *Psychol Rev* 87: 272–300, 1980.
- ENGEL SA, GLOVER GH, AND WANDELL BA. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 7: 181–192, 1997.
- FRISTON KJ, HOLMES AP, POLINE JB, GRASBY PJ, WILLIAMS SCR, FRACKOWIAK RSJ, AND TURNER R. Analysis of fMRI time-series revisited. *Neuroimage* 2: 45–53, 1995a.
- FRISTON KJ, HOLMES AP, WORSLEY KJ, POLINE JP, FRITH CD, AND FRACKOWIAK RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Map* 2: 189–210, 1995b.
- FRISTON KJ, WORSLEY KJ, FRACKOWIAK RSJ, MAZZIOTTA JC, AND EVANS AC. Assessing the significance of focal activations using their spatial extent. *Hum Brain Map* 1: 210–220, 1994.
- GATTASS R, GROSS CG, AND SANDELL JH. Visual topography of V2 in the macaque. *J Comp Neurol* 201: 519–539, 1981.
- GATTASS R, SOUSA APB, AND GROSS CG. Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8: 1831–1845, 1988.
- GOEBEL R, KHORRAM-SEFAT D, MUCKLI L, HACKER H, AND SINGER W. The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur J Neurosci* 10: 1563–1573, 1998.
- GRILL-SPECTOR K, KUSHNIR T, HENDLER T, EDELMAN S, ITZCHAK Y, AND MALACH R. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum Brain Map* 6: 316–328, 1998.
- HADJIKHANI NK, LIU AK, DALE AM, CAVANAGH P, AND TOOTELL RB. Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neurosci* 1: 235–241, 1998.
- KASTNER S, DE WEERD P, DESIMONE R, AND UNGERLEIDER LG. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282: 108–111, 1998.
- KASTNER S, NOTHDURFT HC, AND PIGAREV IN. Neuronal responses to orientation and motion contrast in cat striate cortex. *Visual Neurosci* 16: 587–600, 1999.
- KASTNER S AND UNGERLEIDER LG. Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci* 23: 315–341, 2000.
- KNIERIM JJ AND VAN ESSEN DC. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67: 961–980, 1991.
- LAMME VAF. The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15: 1605–1615, 1995.
- LOGOTHETIS N, PAULS J, OELTERMANN A, TRINATH T, AND AUGATH M. The relationship of LFPs, MUA, and SUA to the BOLD fMRI signal. *Soc Neurosci Abstr* 26: 820, 2000.
- MCKEEFERY DJ AND ZEKI S. The position and topography of the human colour center as revealed by functional magnetic resonance imaging. *Brain* 120: 2229–2242, 1997.
- MILLER EK, GOCHIN PM, AND GROSS CG. 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, 1993.
- MORAN J AND DESIMONE R. Selective attention gates visual processing in the extrastriate cortex. *Science* 229: 782–784, 1985.
- NOTHDURFT HC, GALLANT JL, AND VAN ESSEN DC. Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia. *Vis Neurosci* 16: 15–34, 1999.
- PINSK M, KASTNER S, DESIMONE R, AND UNGERLEIDER LG. An estimate of receptive field sizes in human visual cortex. *Soc Neurosci Abstr* 25: 916, 1999a.
- PINSK MA, KASTNER S, DESIMONE R, AND UNGERLEIDER LG. An estimate of receptive field sizes in human visual cortex. *Neuroimage* 9: S885, 1999b.
- POLINE JB, WORSLEY KJ, EVANS AC, AND FRISTON KJ. Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage* 5: 83–96, 1997.
- RECANZONE GH, WURTZ RH, AND SCHWARZ U. Responses of MT and MST neurons to one and two moving objects in the receptive field. *J Neurophysiol* 78: 2904–2915, 1997.
- REES G, FRACKOWIAK RSJ, AND FRITH CD. Two modulatory effects of attention that mediate object categorization in human cortex. *Science* 275: 835–838, 1997.
- REYNOLDS JH, CHELAZZI L, AND DESIMONE R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19: 1736–1753, 1999.
- ROLLS ET AND TOVEE MJ. The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Exp Brain Res* 103: 409–420, 1995.
- SATO T. Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Exp Brain Res* 77: 23–30, 1989.
- SCHNEIDER W, CASEY BJ, AND NOLL D. Functional MRI mapping of stimulus rate effects across visual processing stages. *Hum Brain Map* 1: 117–133, 1994.
- SERENO MI, DALE AM, REPPAS JB, KWONG KK, BELLIVEAU JW, BRADY TJ, ROSEN BR, AND TOOTELL RB. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889–893, 1995.
- SHIPP S, WATSON JDG, FRACKOWIAK RSJ, AND ZEKI S. Retinotopic maps in human prestriate visual cortex: the demarcation of areas V2 and V3. *Neuroimage* 2: 125–132, 1995.
- SMITH AT, SINGH KD, AND GREENLEE MW. FMRI mapping of receptive field size in human striate and extrastriate visual cortex. *Soc Neurosci Abstr* 25: 2059, 1999.
- TALAIRACH J AND TOURNOUX P. *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme, 1988.
- TOOTELL RBH, DALE AM, SERENO MI, AND MALACH R. New images from human visual cortex. *Trends Neurosci* 19: 481–489, 1996.
- TOOTELL RBH, MENDOLA JD, HADJIKHANI NK, LEDDEN PJ, LIU AK, REPPAS JB, SERENO MI, AND DALE AM. Functional analysis of V3A and related areas in human visual cortex. *J Neurosci* 17: 7060–7078, 1997.

- TOOTELL RBH, REPPAS JB, KWONG KK, MALACH R, BORN RT, BRADY TJ, ROSEN BR, AND BELLIVEAU JW. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15: 3215–3230, 1995.
- TREISMAN AM. Strategies and models of selective attention. *Psychol Rev* 76: 282–299, 1969.
- VAN ESSEN DC, NEWSOME WT, AND MAUNSELL JHR. The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Res* 24: 429–448, 1984.
- WATSON JDG, MYERS R, FRACKOWIAK RSJ, HAJNAL JV, WOODS RP, MAZZIOTTA JC, SHIPP S, AND ZEKI S. Area V5 of the human brain: evidence from combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex* 3: 79–94, 1993.
- WOODS RP, MAZZIOTTA JC, AND CHERRY SR. MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 17: 536–546, 1993.
- ZEKI S, WATSON JDG, LUECK CJ, FRISTON KJ, KENNARD C, AND FRACKOWIAK RS. A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11: 641–649, 1991.