MECHANISMS OF VISUAL ATTENTION IN THE HUMAN CORTEX*

Sabine Kastner and Leslie G. Ungerleider
Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892; e-mail: sabine@ln.nimh.nih.gov, lgu@ln.nimh.nih.gov

Key Words vision, selective attention, functional brain imaging, neglect, object recognition

Abstract A typical scene contains many different objects that, because of the limited processing capacity of the visual system, compete for neural representation. The competition among multiple objects in visual cortex can be biased by both bottom-up sensory-driven mechanisms and top-down influences, such as selective attention. Functional brain imaging studies reveal that, both in the absence and in the presence of visual stimulation, biasing signals due to selective attention can modulate neural activity in visual cortex in several ways. Although the competition among stimuli for representation is ultimately resolved within visual cortex, the source of top-down biasing signals derives from a network of areas in frontal and parietal cortex.

INTRODUCTION

In everyday life, the scenes we view are typically cluttered with many different objects. However, the capacity of the visual system to process information about multiple objects at any given moment in time is limited (Broadbent 1958, Neisser 1967, Schneider & Shiffrin 1977, Tsotsos 1990). This limited processing capacity can be exemplified by a simple experiment. If subjects are presented with two different objects and asked to identify two different attributes at the same time (e.g. color of one and orientation of the other), the subjects’ performance is worse than if the task had been performed with only a single object (Treisman 1969; Duncan 1980, 1984). Hence, because of limited processing resources, multiple objects present at the same time in the visual field compete for neural representation.

How can the competition among multiple objects be resolved? One way is by bottom-up, stimulus-driven processes. For example, in Figure 1A, the single vertical line among the multiple distracter lines is effortlessly and quickly detected because of its salience in the display, which biases the competition in favor of the vertical line. Stimulus salience depends on various factors, including simple

---

*The US government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.
figure 1. Cluttered visual scenes. Because of the limited processing capacity of the visual system, multiple stimuli present simultaneously in the visual field compete for neural representation. This competition can be biased in several ways. One way is by bottom-up, stimulus-driven factors, such as the salience of a stimulus (A). In a condition in which the competition is not biased by stimulus salience (B), it can be biased by top-down processes such as directing attention to a particular stimulus location (C, dashed circle). Processing of stimuli occurring at the attended location will be facilitated.

Feature properties, such as line orientation or color of the stimulus (Treisman & Gelade 1980, Treisman & Gormican 1988), perceptual grouping of stimulus features by Gestalt principles (Prinzmetal 1981, Duncan 1984, Driver & Baylis 1989, Lavie & Driver 1996), and the dissimilarity between the stimulus and nearby distracter stimuli (Duncan & Humphreys 1989, 1992; Nothdurft 1993).

In figure 1B, the competition among the multiple lines is not resolved by salience, and one must actively search through the display to identify the vertical line (Treisman & Gelade 1980, Wolfe et al 1989, Wolfe 1994). In such cases, where target salience is low, it is possible to bias the competition among the multiple lines by top-down processes, such as spatially directed attention. For example, if one is spatially cued to attend to the target location, as in Figure 1C, the identification of the vertical line in that location will be facilitated (Posner 1980, Bashinski & Bacharach 1980). This result suggests that spatially directed attention enhances information processing at the attended location. In effect, attention operates to filter out irrelevant information from nearby distracters.

In this review, we discuss mechanisms of selective attention in the human visual cortex in the context of a biased competition account of attention (Harter & Aine 1984; Bundesen 1990; Desimone & Duncan 1995; Duncan 1996, 1998; Desimone 1998). We focus on results from functional brain imaging studies, in particular as they relate to results from monkey neurophysiology. In the following sections, we first describe the evidence for competition among multiple visual stimuli for neural representation. Second, we describe mechanisms of attentional top-down bias in visual cortical areas. We then turn to the potential sources for
generating and controlling attentional top-down bias. Finally, we relate mechanisms of attention to those of working memory.

COMPETITION FOR NEURAL REPRESENTATION IN THE OBJECT VISION PATHWAY

Organization of Visual Cortex

Most of our knowledge about the organization of visual cortex comes from behavioral, anatomical, and physiological studies in monkeys. These studies have shown that monkey cortex contains more than 30 separate visual areas (Felleman & Van Essen 1991), which are organized into two functionally specialized processing pathways (Ungerleider & Mishkin 1982, Desimone & Ungerleider 1989, Ungerleider 1995). Both pathways originate in the primary visual cortex (V1) and both are composed of multiple areas beyond V1. The occipitotemporal pathway, or ventral stream, is crucial for the identification of objects, whereas the occipitoparietal pathway, or dorsal stream, is crucial for the appreciation of the spatial relations among objects as well as for the visual guidance of movements toward objects in space (Ungerleider & Mishkin 1982, Goodale & Milner 1992).

Results from single-cell recordings from areas within the ventral and dorsal streams are consistent with this model of functional specialization. Thus, neurons in areas V4, TEO, and TE of the ventral stream show response selectivities for stimulus attributes that are important for object vision, such as shape, color, and texture (Desimone & Ungerleider 1989). By contrast, neurons in the middle temporal (MT) area and further stations of the dorsal stream are not tuned for these stimulus attributes; rather, they show response selectivity for the speed and direction of stimulus motion, consistent with the role of these areas in visuospatial function (Goldberg & Colby 1989, Newsome & Salzman 1990, Andersen et al 1997).

Within the ventral stream, or object vision pathway, the processing of information is largely hierarchical (Figure 2). For example, the processing of object features begins with simple spatial filtering by cells in V1, but by the time the inferior temporal cortex (area TE) is activated, the cells respond selectively to global object features, such as shape, and some cells are even specialized for the analysis of faces (Desimone & Ungerleider 1989). Likewise, the average receptive field (RF) size increases as one moves along the pathway toward the temporal lobe; at parafoveal eccentricities, RFs of neurons are about 1.5° in V1, and about 4° in V4, whereas neurons in area TE have a median RF size of 26 x 26° (Gattass et al 1981, 1988; Desimone & Ungerleider 1989). It thus appears that large RFs in later areas are built up from smaller ones in earlier areas. Viewed in this way, it is possible to consider much of the neural mechanisms for object vision as a bottom-up process.
Figure 2  The object vision pathway. The occipitotemporal pathway or ventral stream, which is crucial for object recognition, originates in the primary visual cortex (V1) and is composed of multiple areas beyond V1 in occipital (V2, VP, V4) and temporal cortex (TEO, TE). The anatomical organization of ventral stream areas, as outlined on a lateral view of a monkey brain, is largely hierarchical. Connections between successive pairs of areas are reciprocal, such that feed-forward projections from one area to the next are reciprocated by feedback projections from the second area back to the first (arrows).

Anatomical studies reveal that virtually all connections between successive pairs of areas within the ventral stream are reciprocally connected, i.e. projections from one area to the next are reciprocated by projections from the second area back onto the first (Felleman & Van Essen 1991). Additionally, there exist other “feedback” projections to ventral stream areas from both prefrontal and parietal cortex (Cavada & Goldman-Rakic 1989, Ungerleider et al 1989, Webster et al 1994). Feedback projections both within the ventral stream and from areas beyond this processing pathway may form the anatomical basis for top-down influences, such as attention, on object perception.

Functional brain imaging studies have begun to reveal a remarkably similar organization within the human visual cortex. Functional brain imaging measures hemodynamic changes—blood flow in the case of positron emission tomography and blood oxygenation in the case of functional magnetic resonance imaging (fMRI)—and these can be used as indirect measures of neural activity (Fox &
MECHANISMS OF ATTENTION


Functional brain imaging data in humans, like physiological data in monkeys, also argue for an increase in the complexity of processing as activity proceeds anteriorly through the ventral stream into the temporal lobe. Whereas posterior regions in the cortex are preferentially activated during the processing of object attributes, such as colors or scrambled objects and faces, more anterior regions are activated selectively during the processing of intact objects and faces (Haxby et al 1994, Puce et al 1996, Kanwisher et al 1997, Grill-Spector et al 1998). In addition, as described below, the RF sizes of neurons in human visual cortex have been shown to increase progressively from V1 to TEO, an area in the posterior inferior temporal cortex (Kastner et al 1998a, Pinsk et al 1999). Thus, like monkey visual cortex, the visual cortex of humans appears to be organized hierarchically.

Sensory Suppression Among Multiple Visual Stimuli

Competition among objects for representation takes place in ventral stream areas in both monkeys and humans. What are the neural correlates for competitive interactions among multiple objects in the visual field? Single-cell recording studies in monkeys have shed light on this question by comparing responses to a single visual stimulus presented alone in a neuron’s RF with responses to the same stimulus when a second one is presented simultaneously within the same RF (Moran & Desimone 1985, Reynolds et al 1999). The responses to the paired
stimuli were shown to be a weighted average of the responses to the individual stimuli presented alone. For example, if a single good stimulus elicited a high firing rate and a single poor stimulus elicited a low firing rate, the response to the paired stimuli was reduced compared with that elicited by the single good stimulus. This result indicates that two stimuli present at the same time within a neuron’s RF are not processed independently, but rather that they interact with each other in a mutually suppressive way. This sensory suppressive interaction among multiple stimuli has been interpreted as an expression of competition for neural representation.

Sensory suppression among multiple stimuli has also been examined in studies using functional magnetic resonance imaging (Kastner et al 1997, 1998a). In these studies, subjects were presented with images of colorful, complex stimuli in four nearby locations of the upper right quadrant of the visual field while they maintained fixation. Fixation was ensured by having subjects count occurrences of the letter T or L at fixation, an attentionally demanding task. The stimuli were presented under two different conditions, simultaneous and sequential (Figure 3; see color insert). In the sequential condition, a single stimulus appeared in one of the four locations, then another appeared in a different location, and so on, until each of the four stimuli had been presented in the different locations. In the simultaneous condition, the same four stimuli appeared in the same four locations as in the sequential condition, but they were presented at the same time. Thus, integrated over time, the physical stimulation parameters were identical in the two presentation conditions, but sensory suppression among stimuli could take place only in the simultaneous presentation condition. It was therefore predicted that activation in the simultaneous presentation condition would be less than in the sequential presentation condition.

Compared with interleaved blank periods, activation of V1 and ventral stream extrastriate areas V2 to TEO was found under both stimulus presentation conditions (Figure 4A; see color insert). Although the fMRI signal was similar in the two presentation conditions in V1, the activation was reduced in the simultaneous condition compared with the sequential condition in V2, and this reduction was especially pronounced in V4 (Figure 4B) and TEO, as predicted from the sensory suppression hypothesis. The sensory suppression effect appeared to be scaled to the RF size of neurons within visual cortical areas. That is, the small RFs of neurons in V1 and V2 would encompass only a small portion of the visual display, whereas the larger RFs of neurons in V4 and TEO would encompass all four stimuli. Therefore, suppressive interactions among the stimuli within RFs could take place most effectively in these more anterior visual areas. It is likely that surround inhibition from regions outside the classical RF contributed to the small sensory suppression effects found in V1 and V2 (Knierim & Van Essen 1992).

The idea that sensory suppressive interactions are scaled to RF size was directly tested in a second study, in which the spatial separation between the four stimuli was increased (Kastner et al 1998a, Pinsk et al 1999). According to the RF hypothesis, the magnitude of sensory suppression should be inversely related
to the degree of spatial separation among the stimuli. In agreement with this idea, 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° 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 areas in the human visual cortex. From these experiments, at an eccentricity of about 5°, RF sizes were estimated to be less than 2° in V1, in the range of 2–4° in V2, and approximately 6° in V4. In TEO, the RFs were larger than in V4, but they still confined to a single quadrant of the contralateral hemifield (Kastner et al. 1998a, Pinsk et al. 1999). These estimates of RF sizes in human visual cortex are strikingly similar to those measured in the homologous visual areas of monkeys (Gattass et al. 1981, 1988; Boussaoud et al. 1991).

EFFECTS OF ATTENTIONAL TOP-DOWN BIAS IN VISUAL CORTEX

Convergent evidence from single-cell recording studies in monkeys and functional brain imaging and event-related potential studies in humans indicate that selective attention can modulate neural processing in visual cortex. This modulation of visually evoked activity has been interpreted as a top-down bias facilitating information processing of stimuli at attended locations or of attributes of attended stimuli (Desimone & Duncan 1995, Desimone 1998). We argue that top-down biasing signals due to visual attention affect neural processing in several ways. These include the following: enhancement of neural responses to an attended stimulus; the filtering of unwanted information by counteracting the suppression induced by nearby distracters; the biasing of signals in favor of an attended location by increases of baseline activity in the absence of visual stimulation; and the increase of stimulus salience by enhancing the neuron’s sensitivity to stimulus contrast. Below, we review the evidence for each of these effects.

Response Enhancement

In single-cell recording studies, neural responses to visual stimuli presented within a neuron’s RF have been studied under conditions in which the animal covertly (i.e. without executing eye movements) directs its attention to a stimulus within the RF or when the animal directs its attention away from the RF to another location in the visual field. Several studies have shown that neural responses to a single stimulus presented within the RF are enhanced when the animal directs its attention within the RF compared with when the animal attends outside the RF. This effect, which increases with task difficulty (Spitzer et al. 1988, Spitzer & Richmond 1991), has been demonstrated in V1 (Motter 1993), in V2 (Motter 1993, Luck et al. 1997), and in V4 (Haenny et al. 1988; Spitzer et al. 1988; Motter
Response enhancement has also been shown in dorsal stream areas, such as MT (Treue & Maunsell 1996, Treue & Martinez 1999) and the lateral intraparietal (LIP) area (Bushnell et al 1981, Colby et al 1996). This finding suggests that top-down attentional mechanisms operate by enhancing neural responses to stimuli at attended locations, thereby biasing information processing in favor of stimuli appearing at that location.

Similar results have been found in functional brain imaging and event-related potential studies in human visual cortex. In these experiments, identical visual stimuli were presented simultaneously to corresponding peripheral field locations to the right and left of fixation while subjects were asked to direct attention covertly to the right or the left. Directing attention to the left hemifield led to increased stimulus-evoked activity in extrastriate visual areas of the right hemisphere, whereas directed attention to the right hemifield led to increased activity in extrastriate visual areas of the left hemisphere (Heinze et al 1994, Vandenberghe et al 1997). Thus, responses to the stimuli were enhanced on the side of extrastriate cortex containing the representations of the attended hemifield. Response enhancement in extrastriate cortex due to spatially directed attention may occur as early as 80–130 ms after stimulus onset (Heinze et al 1994, Mangun 1995, Hillyard et al 1998, Mangun et al 1998).

Attentional top-down feedback can bias neural responses not only in favor of a stimulus at an attended location but also in favor of an attended stimulus attribute. This has been shown in experiments that compared neural activity evoked by identical stimuli, but when different stimulus attributes were selectively attended. Motter (1994) studied single-cell responses in area V4 when animals were cued to select visual stimuli on the basis of either their color or their luminance. In the majority of neurons, responses to stimuli matching the selected feature were enhanced, whereas responses to stimuli that did not match the selected feature were attenuated. Similar attentional modulatory effects have been shown with other stimulus attributes, such as orientation (Haenny & Schiller 1988, Maunsell et al 1991) and direction of motion (Ferrera et al 1994).

Functional brain imaging studies have investigated attentional top-down bias in favor of stimulus attributes by comparing activity evoked within different visual cortical areas while subjects performed a task requiring selective attention to particular features of identical visual stimuli. In one such study, Corbetta et al (1991a) found that selective attention to either shape, color, or speed enhanced activity in the regions of extrastriate visual cortex that selectively process these same attributes. Attention to shape and color led to response enhancement in regions of the posterior portion of the fusiform gyrus, including area V4. Attention to speed led to response enhancement in areas MT. More recent studies have replicated the attentional effects on color and speed processing (Clark et al 1997, Beauchamp et al 1997, O’Craven et al 1997, Buechel et al 1998). Other investigations have shown that attention to faces or houses led to response enhancement in areas of the mid-anterior portion of the fusiform gyrus, areas specialized for
the processing of faces and objects (Haxby et al 1994, Clark et al 1997, Wojciulik et al 1998, O’Craven et al 1999). Taken together, these results support the idea that selective attention to a particular stimulus attribute biases neural activity in those extrastriate areas that preferentially process the selected attribute. It is interesting that the effects of attention to stimulus attributes, such as color, occur on the order of 60 ms later than those reported for selection based on spatial attention (Hillyard & Vento 1998, Vento et al 1998). Based on this latency difference, Hillyard and colleagues have argued for a hierarchical model of attention, with the selective processing of stimulus attributes dependent on the prior selection for location. Alternatively, selection for location and selection for attributes may take place in parallel (Desimone & Duncan 1995), but the selection for location may be accomplished more rapidly.

Filtering of Unwanted Information

Thus far, attentional top-down bias has been shown to operate by enhancing neural responses to a stimulus at an attended location or to an attended stimulus attribute. However, a typical visual scene contains multiple stimuli, each competing for processing resources. As described above, competition among multiple stimuli for representation is evidenced by mutually suppressive sensory interactions; such interactions were demonstrated in both single-cell recording (Reynolds et al 1999) and fMRI studies (Kastner et al 1997, 1998a). Single-cell recording studies have also shown that sensory suppressive interactions can be modulated by directed attention. In particular, in extrastriate areas V2 and V4, spatially directed attention to an effective stimulus within a neuron’s RF eliminated the suppressive influence of a second stimulus presented within the same RF (Reynolds et al 1999). Attentional effects were less pronounced when the second stimulus was presented outside the RF, which suggests that competition for processing resources within visual cortical areas takes place most strongly at the level of the RF. These findings imply that attention may resolve the competition among multiple stimuli by counteracting the suppressive influences of nearby stimuli, thereby enhancing information processing at the attended location. This may be an important mechanism by which attention filters out unwanted information from cluttered visual scenes (Desimone & Duncan 1995, Desimone 1998).

Recent fMRI studies suggest that a similar mechanism operates in human visual cortex (Kastner et al 1998a). The effects of spatially directed attention on multiple competing visual stimuli were studied in a variation of the paradigm used to examine sensory suppressive interactions among simultaneously presented stimuli (described above and illustrated in Figure 3). In addition to the two different visual presentation conditions, sequential and simultaneous, two different attentional conditions were tested, unattended and attended. During the unattended condition, attention was directed away from the visual display by having subjects count the letters T or L at fixation, exactly as in the original study. In the attended condition, subjects were instructed to attend covertly to the stimulus
location closest to fixation in the display and to count the occurrences of one of the four stimuli, which was indicated before the scan started. Based on the results from monkey physiology, it was predicted that attention should reduce sensory suppression among stimuli. Thus, responses evoked by the competing, simultaneously presented stimuli should be enhanced more strongly than responses evoked by the noncompeting sequentially presented stimuli (Moran & Desimone 1985; Chelazzi et al 1993, 1998; Treue & Maunsell 1996; Luck et al 1997; Reynolds et al 1999).

As illustrated in Figure 4C, directed attention to the display enhanced responses to both the sequentially and the simultaneously presented stimuli. This finding confirmed the effects of attentional response enhancement shown in numerous previous studies in monkeys and humans, as cited previously. More important, and in accordance with the prediction from monkey physiology, directed attention led to greater increases of fMRI signals to simultaneously presented stimuli than to sequentially presented stimuli. Thus, attention partially cancelled out the suppressive interactions among competing stimuli. The magnitude of the attentional effect scaled with the magnitude of the suppressive interactions among stimuli, with the strongest reduction of suppression occurring in areas V4 and TEO. These findings support the idea that directed attention enhances information processing of stimuli at the attended location by counteracting suppression induced by nearby stimuli, which compete for limited processing resources. In essence, unwanted distracting information is effectively filtered out. The degree to which distracting information can be eliminated depends on the load of the target task. For example, Rees et al (1997b) demonstrated that activation in area MT evoked by distracting moving stimuli was totally abolished when subjects performed a high-load linguistic task at fixation, compared with a low-load version of the task. Thus, the greater the attentional resources devoted to the target, the less the processing of irrelevant distracting stimuli.

The attentional effects were retinotopically organized, inasmuch as they were seen only in visual areas with a representation of the attended location (i.e. the upper right quadrant). These areas included those in ventral visual cortex (shown in Figure 4A), as well as areas V3A and MT in dorsal visual cortex. The retinotopic specificity of spatial attention effects was suggested in earlier studies (Heinze et al 1994, Woldorff et al 1997) and has been elegantly demonstrated in recent investigations (Tootell et al 1998, Breftczynski & DeYoe 1999).

It is important to note that the attentional response enhancement to both simultaneously and sequentially presented stimuli appeared to increase from early to later stages of visual processing. Attentional effects were absent or small in V1 and V2, respectively, and much stronger in more anterior extrastriate areas V4 and TEO, which suggests that the latter areas were the primary target of the top-down “feedback.” Single-cell recording studies have shown that neural responses can be modulated by attention as early as in V1 (Motter 1993, Roelfsema et al 1998), and functional brain imaging studies have demonstrated attentional response modulation in V1 with moving (Watanabe et al 1998a, 1998b; Somers
et al 1999, Ghandi et al 1999) and stationary stimuli (Martinez et al 1999). Yet, in all these studies, the magnitude of the attentional response modulation in V1 was smaller than that in more anterior extrastriate areas, which suggests that attentional effects in V1 may be caused by reactivation from higher-order extrastriate areas. This idea is supported by single-cell recording studies, which have shown that attentional effects in area TE of inferior temporal cortex have a latency of approximately 150 ms (Chelazzi et al 1998), whereas attentional effects in V1 have a longer latency, approximately 230 ms (Roelfsema et al 1998).

Increases of Baseline Activity

There is evidence that attentional top-down biasing signals can be obtained not only by the modulation of visually driven activity but also in the absence of any visual stimulation whatsoever. Single-cell recording studies have shown that spontaneous (baseline) firing rates were 30–40% higher for neurons in areas V2 and V4 when the animal was cued to attend covertly to a location within the neuron’s RF before the stimulus was presented there, i.e. in the absence of visual stimulation (Luck et al 1997). A similar effect was demonstrated in dorsal stream area LIP (Colby et al 1996). This increased baseline activity, termed the baseline shift, has been interpreted as a direct demonstration of top-down feedback, biasing neurons representing the attended location and thereby favoring stimuli that will appear there at the expense of those appearing at unattended locations. Thus, stimuli at attended locations are biased to “win” the competition for processing resources.

Attentional top-down biasing signals in human visual cortex in the absence of visual stimulation were studied by adding a third experimental condition to the design used to investigate sensory suppressive interactions and their modulation by attention, as illustrated in Figure 5A (Kastner et al 1999; see color insert). In addition to the two visual presentation conditions, sequential and simultaneous, and the two attentional conditions, unattended and attended, an expectation period preceding the attended presentations was introduced. The expectation period, during which subjects were required to direct attention covertly to the target location and instructed to expect the occurrences of the stimulus presentations, was initiated by a marker presented briefly next to the fixation point 11 s before the onset of the stimuli. In this way, the effects of attention in the presence and absence of visual stimulation could be studied.

As illustrated for area V4 in Figure 5B, the fMRI signals increased during the expectation period, before any stimuli were present on the screen. This increase of baseline activity was followed by a further increase of activity evoked by the onset of the stimulus presentations. The baseline increase was found in all visual areas with a representation of the attended location, indicating the retinotopic specificity of the effect. The increase of baseline activity was strongest in V4, but it was also seen in early visual areas. It is noteworthy that baseline increases were found in V1, even though no attentional modulation of visually evoked activity
was seen in this area. This dissociation suggests either that different mechanisms underlie the effects of attention on visually evoked activity and on baseline activity or that the attentional effects previously reported with visual stimulation in V1 actually derive from sustained shifts in baseline activity rather than from increases in the stimulus-evoked response, per se.

The baseline increases found in human visual cortex may be subserved by increases in spontaneous firing rate similar to those found in the single-cell recording studies (Colby et al 1996, Luck et al 1997) but summed over large populations of neurons. The increases evoked by directing attention to a target location in anticipation of a behaviorally relevant stimulus at that attended location are thus likely to reflect a top-down feedback bias in favor of the attended location in human visual cortex.

Increases in Response Sensitivity

Recent physiological studies suggest that another role for attention is to increase sensitivity. In V4 neurons, attention increased the sensitivity to single grating stimuli presented at different stimulus contrasts by approximately 30%. This effect was especially pronounced with low-contrast stimuli (Reynolds et al 1996). Thus, neurons in V4 responded to an attended stimulus as if its contrast, i.e. its salience, had been increased. Furthermore, when the animal directed its attention away from the RF, a high-contrast stimulus, which was presented with a low-contrast stimulus within the RF, tended to dominate the neuron’s responses. Thus, the neuronal responses were strongly biased by bottom-up processes. When the animal directed its attention to the high-contrast stimulus, the neuron’s responses were further increased. However, when the animal directed its attention to the low-contrast stimulus, the neuron’s responses were reduced, indicating that top-down attentional mechanisms were counteracting the bottom-up stimulus-driven mechanisms (Reynolds & Desimone 1997). A similar result was recently reported for MT neurons (Treue & Martinez 1998). Whether attention operates to increase sensitivity in human visual cortex has not yet been explored.

source areas generating attentional top-down bias

Thus far, we have argued that there is competition among objects within visual cortical areas for neural representation. Further, we have proposed that this competition can be biased in favor of a particular object either by its salience, i.e. by bottom-up, sensory-driven input, or by mechanisms of selective attention, i.e. through top-down inputs. Although it is likely that the competition is ultimately resolved within visual cortex, we propose that the top-down biasing signals derive from areas outside visual cortex and are transmitted via feedback projections to visual cortex. What areas might be the source of these top-down signals?
Both studies of patients suffering from attentional deficits due to brain damage and functional brain imaging studies of healthy subjects performing attention tasks have given insights into a distributed network of higher-order areas in frontal and parietal cortex that appear to be involved in the generation and control of attentional top-down feedback signals.

**Lesion Studies**

There is a long history demonstrating that unilateral brain lesions in humans often cause an impairment in spatially directing attention to the contralateral hemifield, a syndrome known as visuospatial neglect. In severe cases, patients suffering from neglect will completely disregard the visual hemifield contralateral to the side of the lesion (Bisiach & Vallar 1988, Heilman et al 1993, Rafal 1994). For example, they will read from only one side of a book, apply make-up to only one half of their face, or eat from only one side of a plate. In less severe cases, the deficit is more subtle and becomes apparent only if the patient is confronted with competing stimuli, as in the case of visual extinction. In visual extinction, patients are able to orient attention to a single visual object presented to their impaired visual hemifield; however, if two stimuli are presented simultaneously, one in the impaired and the other in the intact hemifield, the patients will only detect the one presented to the intact side, denying that any other object had been presented. These findings suggest that visual extinction reflects an attentional bias toward the intact hemifield in the presence of competing objects (Kinsbourne 1993, Duncan 1998).

Visuospatial neglect may follow unilateral lesions at different sites, including the parietal lobe, especially its inferior part and the temporo-parietal junction (Vallar & Perani 1987), regions of the frontal lobe (Heilman & Valenstein 1972, Damasio et al 1980), the anterior cingulate cortex (Janer & Pardo 1991), the basal ganglia (Damasio et al 1980), and the thalamus, in particular the pulvinar (Watson & Heilman 1979). Studies with monkeys have implicated the same brain regions (Welch & Stuteville 1958; Latto & Cowey 1971; Watson et al 1973, 1974; Petersen et al 1987; Lynch & McLaren 1989; Gaffan & Hornak 1997). The finding that lesions of many different areas may cause visuospatial neglect has led to the notion that these areas form a distributed network for directed attention (Mesulam 1981, Posner & Petersen 1990).

Neglect occurs more often with right-sided parietal lesions than with left-sided parietal lesions, which suggests a specialized role for the right hemisphere in directed attention (Vallar 1993). Based on this hemispheric asymmetry, it has been proposed that the right hemisphere mediates directed attention to both sides of visual space, whereas the left hemisphere mediates directed attention only to the contralateral, right side of visual space (Heilman & Van Den Abell 1980, Mesulam 1981). According to this view, in the case of a left-hemisphere lesion, the intact right hemisphere would take over the attentional function of the damaged left hemisphere, whereas a right-hemisphere lesion would result in a left-
sided hemispatial neglect because of the bias of the intact left hemisphere for the right hemifield. This right-hemispheric dominance of parietal cortex has been demonstrated only in cases of severe neglect; visual extinction appears to result as frequently from left- as from right-sided lesions (Rafal 1994).

It is important to note that stimulus-driven, bottom-up mechanisms within visual cortex, such as figure-ground segmentation or perceptual grouping, which determine the salience of a stimulus, are preserved in the neglected hemifield and may influence a patient’s behavior (Driver et al 1992, Grabowecky et al 1993, Marshall & Halligan 1994, Driver 1995, Mattingley et al 1997, Driver & Mattingley 1998). For example, Mattingley et al (1997) reported a patient with parietal damage whose extinction was less severe when bilateral stimuli were arranged to form an illusory Kanizsa square, a percept based on automatic filling-in of illusory boundaries. This result shows that the patient could use the information from his neglected left hemifield to form the percept of a common surface. It therefore appears that, following parietal damage, the competition among multiple stimuli can be biased equally well across the entire visual field by bottom-up processes, whereas mechanisms under top-down control, such as directing attention to a particular location, are biased toward the intact hemifield.

Functional Brain Imaging Studies

Results from functional brain imaging studies support the idea that top-down signals related to directed attention are generated by a distributed network of areas in frontal and parietal cortex. A network consisting of areas in the superior parietal lobule (SPL), the frontal eye field (FEF), and the supplementary eye field (SEF) extending into the anterior cingulate cortex has been found to be activated in a variety of visuospatial tasks, as illustrated in Figure 6 (see color insert) (Corbetta et al 1993, 1998; Fink et al 1997; Nobre et al 1997; Vandenberghe et al 1997; Culham et al 1998; Kastner et al 1998b, 1999; Rosen et al 1999). In addition, but less consistently, activations in the inferior parietal lobule (IPL) and the lateral prefrontal cortex in the region of the middle frontal gyrus (MFG) have been reported. A common feature among these visuospatial tasks is that subjects were asked to maintain fixation at a central fixation point and to direct attention covertly to peripheral target locations in order (a) to detect a stimulus (Corbetta et al 1993, 1998; Nobre et al 1997; Rosen et al 1999), (b) to discriminate it (Fink et al 1997; Vandenberghe et al 1997; Kastner et al 1998b, 1999), or (c) to track its movement (Culham et al 1998). Thus, there appears to be a general attention network that operates independently of the specific requirements of the visuospatial task.

There are two notable differences in the results from patient and from functional brain imaging studies. First, the patient studies suggest a right parietal dominance in visuospatial attention. That is, directing attention to the left hemifield is presumed to be exclusively subserved by the right parietal cortex, whereas directing attention to the right hemifield is presumed to be subserved by both the
left and the right parietal cortex. This notion has not been unequivocally supported by functional brain imaging studies. Although some investigators have found a stronger or an even exclusive activation of areas in the right parietal lobe (Corbetta et al 1993, Nobre et al 1997, Vandenberghe et al 1997), others have found largely symmetrical activations in the right and left parietal lobes (Fink et al 1997, Kastner et al 1998b). Moreover, these symmetrical activations appeared to be independent of the visual hemifield attended (Vandenberghe et al 1997, Kastner et al 1998b). A second difference between the results from patient and from functional brain imaging studies concerns which portion of the parietal lobe plays a key role in attention. The patient literature has consistently identified the IPL, including the temporo-parietal junction, as the critical lesion site in neglect patients (Vallar 1993). By contrast, the majority of functional brain imaging studies points to the SPL rather than the IPL as the part of the parietal lobe that is involved in visuospatial attention. One possible explanation for this discrepancy is that many tasks used in the imaging studies involved a cue to indicate the location at which the visual stimulus would appear; these tasks thus had an expectancy component. Results from lesion studies suggest that the ability to maintain expectancy depends on the SPL rather than the IPL (Posner et al 1984, Friedrich et al 1998).

A distributed network subserving selective attention to stimulus attributes has not yet been established. Thus far, activations in selective attention tasks to stimulus attributes, such as speed or color, have revealed (a) a network of brain regions similar to the one activated during spatially directed attention (Buechel et al 1998), (b) a partially overlapping network consisting of areas in right and left SPL and the cerebellum (Le et al 1998), or (c) a different network consisting of areas in the globus pallidus, caudate nucleus, posterior thalamus, inferior pre-motor cortex, and lateral orbitofrontal cortex (Corbetta et al 1991b). Additional studies are needed to determine to what extent the network for spatial attention, consisting of areas SPL, FEF, and SEF, also subserves attention to stimulus attributes.

The anatomical connections of SPL, FEF, and SEF put them in a position to serve as sources of top-down biasing signals within visual cortex. In monkeys, FEF and SEF are reciprocally connected with ventral stream areas (Ungerleider et al 1989, Webster et al 1994) and posterior parietal cortex (Cavada & Goldman-Rakic 1989), and the posterior parietal cortex is connected with ventral stream areas via area LIP (Webster et al 1994). Further, single-cell recording studies in monkeys have shown that neural activity can be modulated by attention in these parietal and frontal areas. In regions of parietal cortex, enhancement of neural responses was demonstrated during covert shifts of attention to peripheral visual stimuli (Robinson et al 1978, Bushnell et al 1981, Colby et al 1996). The strongest determinant of neural responsiveness in parietal cortex turned out to be the salience of the stimulus (Colby & Goldberg 1998). In the FEF and SEF, such response enhancement was originally shown only in the context of activity related to the preparation of saccadic eye movements (Wurtz & Mohler 1976, Goldberg & Bushnell 1981). Recent recording studies suggest, however, that the response
enhancement in these frontal areas during covert shifts of attention to peripheral visual stimuli does not depend on the subsequent execution of saccades (Kodaka et al. 1997, Bon & Lucchetti 1997). Thus, results from single-cell recording studies support the idea that areas in parietal and frontal cortex are potential sources for generating and controlling attentional top-down bias.

Recent fMRI studies suggest that the attention-related activity in parietal and frontal areas does not reflect attentional modulation of visual responses; instead, the activity is largely due to the attentional operations themselves. In a study conducted by Rees et al. (1997a), attentional modulation of visually evoked activity was found to be rate dependent in the inferior temporal cortex but rate independent in prefrontal cortex. This result thus demonstrates two distinct effects of attention: one, in the frontal lobe, which generates modulatory influences, and another, in the temporal lobe, in which the visually evoked responses themselves are modulated. In a more recent study, Kastner et al. (1999) investigated activations in the SPL, FEF, and SEF in the presence and in the absence of visual stimulation in the paradigm shown in Figure 5A. During directed attention in the absence of visual stimulation, the increase in activity was stronger in SPL, FEF, and SEF than the increase in activity seen in visual cortex (as exemplified for SPL in Figure 5C), which suggests that these parietal and frontal areas were the sources of feedback that generated the top-down biasing signals seen in visual cortex. In addition, there was no further increase in activity evoked by the attended stimulus presentations in these parietal and frontal areas. Rather, there was sustained activity throughout the expectation period and the attended presentations, demonstrating that the activity reflected the attentional operations of the task and not visual processing (Figure 5C). Because the magnitude of the activity in the parietal and frontal areas was the same during directed attention in the absence and in the presence of visual stimulation, it appears that this activity is independent of the particular visual task, whether detection or discrimination. This would explain the finding that functional brain imaging studies using different visuospatial attention tasks have described similar attentional networks.

ATTENTION AND WORKING MEMORY

Attention and working memory are closely related cognitive processes. For example, if we search for a familiar face in a crowd, we are holding the information required to identify the face “on-line,” i.e. in working memory, while we selectively attend to different people in the crowd until we find the face matching our internal template. Or if we are cued to attend to a particular location in anticipation of stimuli to be presented in that location, as in the fMRI experiments described above (Kastner et al. 1999), we need to hold the information about the spatial location in working memory to accomplish the task. The intimate relationship between attentional selection and working memory has led to the idea that these
cognitive processes may share common neural mechanisms and circuits (Desimone 1998). There are several lines of evidence supporting this notion.

One line of evidence comes from the finding that both attention and working memory can induce top-down bias in visual cortex in the absence of sensory input. Chelazzi et al (1993, 1998) studied neural responses in the inferior temporal (IT) cortex while the monkey performed a visual search task. The task required the monkey to hold a cue stimulus, which was presented at the beginning of a trial, in working memory for a delay period of several seconds and to find a matching target stimulus from a multiple stimulus array after that delay. During the delay period, most cells showed a higher firing rate when the cue stimulus was a preferred stimulus for the neuron than when the cue stimulus was a non-preferred stimulus. This increased delay activity can be interpreted as a top-down bias, favoring those neurons that are involved in the processing of the behaviorally relevant stimulus. Increased activity during delay periods in working memory tasks have been found in IT cortex in both single-cell recording (e.g. Fuster & Jervey 1981, Miller et al 1993) and functional brain imaging studies (Courtney et al 1997). The biasing signals found in IT cortex during working memory tasks are strikingly similar to the biasing signals found in extrastriate areas during directed attention in the absence of visual stimulation (Luck et al 1997, Kastner et al 1999).

A second line of evidence supporting the close relationship between working memory and attention comes from the finding that top-down signals related to both processes may be generated from common sources. Lesion and deactivation studies in monkeys suggest that top-down signals related to working memory are generated in prefrontal cortex. Monkeys with prefrontal lesions show impaired performance in working memory tasks (e.g. Mishkin 1957, Bauer & Fuster 1976, Funahashi et al 1993), and cooling of prefrontal cortex reduces the selectivity of delay activity in IT neurons (Fuster et al 1985). The crucial role of prefrontal cortex in working memory is supported by the response properties of its neurons, many of which show stimulus-specific delay activity (for reviews, see Goldman-Rakic 1995, Fuster 1995) that reflects behaviorally relevant information (Rainer et al 1998). Neurons in ventral prefrontal areas tend to have delay activity that is primarily related to object information, whereas neurons in dorsal prefrontal areas tend to have delay activity that is primarily related to location information (Wilson et al 1993), although many neurons show specific delay activity for both objects and their locations (Rao et al 1997). A similar distinction between areas specialized for object and location working memory has been demonstrated in human frontal cortex in functional brain imaging studies (for reviews, see Courtney et al 1998a, Ungerleider et al 1998, Smith & Jonides 1999). Working memory for objects activated predominantly inferior prefrontal cortex (Courtney et al 1996, 1997), whereas working memory for spatial locations activated predominantly a frontal region dorsal and posterior to it (Courtney et al 1998b). Some of the areas activated in spatial working memory tasks were found to be in close spatial proximity to areas activated in visuospatial attention tasks, such as the FEF and the
SEF (Courtney et al 1998b, Petit et al 1998). It is important to note that activations attributed to the FEF and SEF in visuospatial attention tasks have been extensive (see Figure 6A) and, thus, may have included additional functional areas beyond the FEF and SEF, such as areas involved in spatial working memory. If so, then top-down biasing signals related to attention and working memory would derive from partially overlapping source areas in frontal cortex.

SUMMARY AND CONCLUSIONS

In this review, we have considered the mechanisms of selective attention in human visual cortex in the context of a biased competition account of attention (Figure 7). Evidence from functional brain imaging studies in humans, supported by results from single-cell recording studies in monkeys, indicates that, first, there is competition among multiple stimuli for representation in visual cortex. Thus, multiple stimuli presented at the same time are not processed independently but rather interact with each other in a mutually suppressive way. Such sensory suppressive interactions are scaled to the RF size of neurons within visual cortical areas of the object vision pathway. Second, competition among multiple stimuli can be biased by both bottom-up, sensory-driven mechanisms and top-down feedback mechanisms. For example, stimulus salience, such as a stimulus of high contrast, provides a bottom-up bias favoring neurons that represent the salient stimulus at the expense of less-salient stimuli. Top-down influences on visual cortex, as in the case of selective attention, can also bias the competition and even override sensory-driven inputs. Biasing signals due to selective attention can affect neural processing in several ways, including the following: (a) the enhancement of neural responses to attended stimuli; (b) the filtering of unwanted information by countering the suppression induced by nearby distracters; (c) the biasing of signals in favor of an attended location by increases of baseline activity in the absence of visual stimulation; and (d) the increase of stimulus salience by enhancing the neuron’s sensitivity to stimulus contrast. Thus, attentional modulation of activity in visual cortex can occur not only in the presence but also in the absence of visual stimulation. Third, although competition is ultimately resolved within visual cortex, the source of top-down biasing signals derives from a network of areas outside visual cortex. For spatially directed attention, these areas include the SPL, the FEF, the SEF, and, less consistently, areas in the IPL, the MFG, and the anterior cingulate cortex. Attention-related activity in frontal and parietal areas does not reflect attentional modulation of visually evoked responses, rather it reflects the attentional operations themselves. The involvement of similar areas in spatial working memory tasks suggests that spatially directed attention and spatial working memory may be intimately related cognitive processes, sharing common neural circuits. Fourth, and finally, the stimulus that wins the competition for representation in visual cortex will gain further access to
Biased competition in visual cortex. In a typical visual scene that contains many different objects, because of the limited processing resources of the visual system, competition exists among those objects for neural representation. The competition among stimuli can be biased in several ways. One way is by bottom-up, sensory-driven mechanisms, such as stimulus salience. Another way is by attentional top-down feedback, which is generated in areas outside the visual cortex. For example, directed attention to a particular location facilitates processing of stimuli presented at that location. The stimulus that wins the competition for neural representation will have further access to memory systems for mnemonic encoding and retrieval and to motor systems for guiding action and behavior.

memory systems for mnemonic encoding and retrieval and to motor systems for guiding action and behavior.

ACKNOWLEDGMENTS

We thank Drs. Carol Colby, Robert Desimone, Peter De Weerd, John Duncan, and John Reynolds for valuable discussions, and Mark Pinsk and Christine Rey-Hipolito for help with the preparation of the manuscript.

LITERATURE CITED


Damasio AR, Damasio H, Chang CH. 1980. Neglect following damage to frontal lobe or basal ganglia. Neuropsychologia 18:123–32


Goldberg ME, Colby CL. 1989. The neurophysiology of spatial vision. See Boller & Grafman, pp. 301–15


Halgren E, Dale AM, Sereno MI, Tootell RBH, Marinkovic K, et al. 1999. Location of human face-selective cortex with respect to
Retinotopic areas. Human Brain Map. 7:29–37


Treue S, Maunsell JHR. 1996. Attentional modulation of visual motion processing in cortical areas MT and MST. Nature 382:539–41


Ungerleider LG, Gaffan D, Pelak VS. 1989. Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys. Exp. Brain Res. 76:473–84


Figure 3 Experimental design. Four complex images (2 × 2°) were presented in nearby locations at an eccentricity of 6–10° from a fixation point (FP) to the upper right quadrant in two presentation conditions: sequential (A) and simultaneous (B). Presentation time was 250 ms followed by a blank period of 750 ms, on average, in each location. A stimulation period of 1 s is shown, which was repeated in blocks of 18 s. Integrated over time, the physical stimulation parameters in the two presentation conditions were identical in each location. But suppressive sensory interactions among stimuli could take place only in the simultaneous, not in the sequential, condition. (From Kastner et al 1998a.)
Figure 4 Sensory suppression and attentional modulation in human visual cortex. (A) Brain areas activated by the complex images compared with blank presentations. Coronal brain slices of a single subject at a distance of 25 mm (left) and 40 mm (right) from the posterior pole. The complex images activated the upper visual field representations of areas V1, V2, VP, V4, and TEO of the left hemisphere. R, right hemisphere. (B) Sensory suppression in V1 and V4. As shown by the time series of functional magnetic resonance imaging signals, simultaneously presented stimuli (SIM) evoked less activity than sequentially presented stimuli (SEQ) in V4 but not in V1. This finding suggests that sensory suppressive interactions were scaled to the receptive field size of neurons in visual cortex. Presentation blocks were 18 s. (C) Attentional modulation of sensory suppression. The sensory suppression effect in V4 was replicated in the unattended condition of this experiment, when the subjects’ attention was directed away from the stimulus display (unshaded). Spatially directed attention (blue) increased responses to simultaneously presented stimuli to a larger degree than to sequentially presented ones in V4. Presentation blocks were 15 s. (From Kastner et al 1998a.)
Figure 5 Increases of baseline activity in the absence of visual stimulation. (A) Block design. Visual stimuli were presented sequentially (SEQ) and simultaneously (SIM) in blocks of 10 s interleaved with blank periods (BL). For other parameters, see Figure 3. In addition to the two different presentation conditions, two attentional conditions were tested: an unattended (UNATT) and an attended (ATT). During the unattended condition, subjects maintained fixation and ignored the peripheral stimulus presentations. During the attended condition, subjects maintained fixation and covertly directed attention to the stimulus location closest to fixation in order to count the occurrences of one of the stimuli operating there. The attended presentations were indicated by a marker next to fixation, which was presented 11 s before the onset of the presentations. During this expectation period (EXP), subjects covertly directed attention to the periphery in anticipation of the onset of the stimuli. (B) Time series of functional magnetic resonance imaging (fMRI) signals in V4. Directing attention to the peripheral target location in the absence of visual stimulation led to an increase of baseline activity (gray blocks), which was followed by a further increase after the onset of the stimuli (blue blocks). Baseline increases were found in both striate and extrastriate visual cortex. (C) Time series of fMRI signals in superior parietal lobule (SPL). Directing attention to the peripheral target location in the absence of visual stimulation led to a stronger increase in baseline activity than in visual cortex, which was not followed by a significant further increase of activity after the onset of the stimuli. Sustained activity was seen in a distributed network of areas outside the visual cortex, including SPL, frontal eye field, and supplementary eye field, which suggests that these areas provided the source for the attentional top-down signals seen in visual cortex. (Adapted from Kastner et al 1999.)