Visual attention mediated by biased competition in extrastriate visual cortex

Robert Desimone

Laboratory of Neuropsychology, NIMH, 49 Convent Drive, Building 49, Room 1B80, Bethesda, MD 20892-4415, USA

According to conventional neurobiological accounts of visual attention, attention serves to enhance extrastriate neuronal responses to a stimulus at one spatial location in the visual field. However, recent results from recordings in extrastriate cortex of monkeys suggest that any enhancing effect of attention is best understood in the context of competitive interactions among neurons representing all of the stimuli present in the visual field. These interactions can be biased in favour of behaviourally relevant stimuli as a result of many different processes, both spatial and non-spatial, and both bottom-up and top-down. The resolution of this competition results in the suppression of the neuronal representations of behaviourally irrelevant stimuli in extrastriate cortex. A main source of top-down influence may derive from neuronal systems underlying working memory.

Keywords: attention; vision; primates; inferior temporal cortex; V4

1. INTRODUCTION

A typical visual scene contains many different objects, not all of which can be fully processed by the visual system at any given time. Thus, attentional mechanisms are needed to limit processing to items that are currently relevant to behaviour (for examples, see Broadbent 1958; Neisser 1967; Treisman 1969; Bundesen 1990; Tsotsos 1990; Allport 1993; Desimone & Duncan 1995; Duncan 1996). Probably the dominant neurobiological hypothesis to account for attentional selection is that attention serves to enhance the responses of neurons representing stimuli at a single behaviourally relevant location in the visual field (see Colby 1991, for a review). This enhancement model is closely related to older ‘spotlight of attention’ models in psychology, in which visual attention serves to limit processing to a single locus of variable size in the visual field. According to this classical view, a behaviourally relevant object in a cluttered field is found by rapidly shifting the spotlight from one object in the scene to the next, until the sought-for object is found. Attention essentially serves as an internal eye that can shift its focus from one location to another. Because all visual attention is inherently spatial according to this view, even objects defined by their shape or colour must be found by examining candidate objects with the serially scanning spotlight, unless the object is so distinctive that it ‘pops out’ and automatically attracts the attentional spotlight. The neurobiological spotlight hypothesis has the advantage of both simplicity and a clear relation to the neuronal enhancement effects seen in the oculomotor system for stimuli that are the targets of eye movements; indeed, a common corollary to the enhancement hypothesis is that the control signals for attentional selection derive from structures involved in oculomotor control.

On the basis of behavioural, neuropsychological, and neurophysiological data, we have developed an alternative model for attentional selection, which we term ‘biased competition’ (Desimone et al. 1990; Desimone & Duncan 1995; Desimone 1996; Duncan 1996; Luck et al. 1997; Chelazzi et al. 1998). According to this account, any enhancing effect of attention on neuronal responses is best understood in the context of competition among all of the stimuli in the visual field for control over behaviour. The psychological aspects of this model are more fully described elsewhere (Desimone & Duncan 1995; see Duncan, this issue) and will not be considered here. In this review, we will first briefly outline the neurobiological aspects of the model as it applies to attentional selection in the ‘central stream’ of extrastriate cortex, which is the network of cortical visual areas that is important for object recognition in primates (see Ungerleider & Mishkin 1982; Ungerleider & Haxby 1994; Ungerleider 1995). We will then review some of the neurophysiological evidence that led to the development of the model.

There are five main tenets in the biased competition model as it applies to visual processing in cortex. The first is that objects in the visual field compete for the responses of cells in visual cortex. For example, if two stimuli are presented simultaneously within the visual field, they will initially activate their neuronal representations in parallel throughout the extrastriate visual areas. If both are independent objects, and if a local region of cortex receives inputs from both of them, neuronal responses in that region will be determined by a competitive interaction between the two stimuli. On average, these interactions will be mutually suppressive.

The second tenet is that competitive interactions are strongest in a given cortical area when competing stimuli activate cells in the same local region of cortex. Thus, in visuotopically organized areas in which neurons have restricted receptive fields, competitive interactions...
between two stimuli will be strongest when the two stimuli fall within the same receptive field.

The third tenet is that these competitive interactions can be biased in favour of one stimulus in a cluttered field by virtue of many different mechanisms, rather than by a single overall 'attentional control' system. Such mechanisms include both 'bottom-up', or stimulus-driven, influences (e.g. one stimulus has greater novelty or has a higher contrast than another) and 'top-down' feedback mechanisms (e.g. one stimulus has greater behavioural relevance than another). Increases in feedback bias are frequently accompanied by increases in the maintained activity of visual cortical neurons or by an increase in sensory-evoked responses, but these effects are not invariable.

The fourth tenet is that the feedback bias is not purely spatial, i.e. it is not limited to cells with receptive fields at a single locus in the visual field. Thus, processing can be biased in favour of stimuli possessing a specific behaviourally relevant colour, shape, texture, and so on, in parallel throughout the visual field, in addition to biases in favour of stimuli occupying a specific relevant spatial location. In this view, the search for an object with a specific shape in a scene, for example, does not necessarily require a serially scanning neural selection process. This is consistent with recent psychological accounts of attention, which post at least some top-down parallel selection of candidate objects throughout the visual field (see, for example, Duncan & Humphreys 1989; Treisman & Sato 1990; Wolfe et al. 1989; Grossberg et al. 1994).

Finally, the fifth and perhaps most speculative tenet of the model is that a main source of the 'top-down' biasing inputs to ventral stream areas in extrastriate cortex derives from structures involved in working memory, specifically prefrontal cortex. We will now consider some of the neurophysiological evidence that bears on the model.

2. VISUAL SEARCH

Some of the strongest support for the model comes from studies of neurons in inferior temporal (IT) cortex of macaque monkeys, studied while the monkeys performed a visual search task (Chelazzi et al. 1993, 1998). IT cortex is a high-order visual processing area in the ventral stream crucial for object recognition. Monkeys and humans with lesions of IT cortex are severely impaired in object recognition (see Desimone & Ungerleider 1989). Consistent with this role in object recognition, IT neurons have extremely large receptive fields that are typically bilateral and include the centre of gaze, and they have complex stimulus specificities, such as selectivity for shape, colour and texture (Desimone & Gross 1979).

In these studies, the monkeys were rewarded for finding a target object in an extrafoveal array of stimuli, similar to finding a 'face in a crowd' (Chelazzi et al. 1993, 1998). A schematic representation of the task is shown in figure 1a. At the start of each trial, a cue stimulus was presented at the centre of gaze, followed by a blank delay period of either 1.5 s or 3 s. The monkey was required to maintain fixation throughout the cue and delay period. At the end of the delay, an array of one to five stimuli was presented extrafoveally, and the monkey was rewarded for making a saccadic eye movement to the target stimulus matching the cue. The location of the target stimulus on a given trial was random—the monkey had to find it based on non-spatial features such as shape or colour. On some 'target-absent' trials, none of the stimuli in the array matched the cue, and the monkey was rewarded for simply maintaining fixation. The stimuli were digitized pictures of complex objects, which are the sort of stimuli that typically elicit selective responses from IT neurons.

The results are shown in figure 1b, c, which shows the average response of a population of individually recorded IT neurons on target-present trials in which a two-stimulus choice array was presented in the contralateral visual field. As shown in figure 1b, at the time of the presentation of the cue, cells responded better in trials when the cue was the good stimulus for the cell than when it was the poor stimulus. This was expected because the cue stimuli were chosen on the basis of the cells' selective responses to them.

Following the response to the cue, there were three critical time periods in the trial. The first critical period was the delay interval after the cue, where the monkey gazed at a blank screen, awaiting the choice array. During this delay (figure 1b), most cells maintained a higher firing rate when their good stimulus was the cue than when their poor stimulus was the cue. This higher maintained activity was not simply a prolonged sensory afterdischarge to the good stimulus because the difference in delay activity was eliminated in control blocks of trials, where the same stimuli were presented but the monkey did not perform the search task. Furthermore, in some sessions the same cue was used for several consecutive trials in a
block, and, in these blocks, the maintained activity of the cells was higher at the start of the trial, preceding the good cue, than it was preceding the poor cue. This can be seen in the population histogram of figure 1b, which shows the results from the blocked-cue trials. Thus, the animal’s expectation of a behaviourally relevant stimulus caused a chronic elevation of activity in cells representing that stimulus. Together, the results argue for some type of feedback activation, or bias, in favour of cells selective for a feature of the cue-target stimulus on a given trial.

The next critical time period was when the choice array was presented. As shown in figure 1c, the initial population response to the array was the same, regardless of whether the good or poor stimulus was the target. Thus, several stimuli appearing in the visual field initially activate their cortical representations in parallel, consistent with the biased competition model.

The last critical time period began about 200 ms after the onset of the array, when responses to the array diverged dramatically depending on which stimulus was the target. When the good stimulus was the target, the response to the good stimulus in the array remained high until the time of the eye movement, which is indicated by the small vertical bar on the horizontal axis in figure 1c.

Figure 1. Responses of neurons in IT cortex in a visual search task, with search arrays confined to the contralateral visual field.
(a) Schematic representation of the task. A cue stimulus was presented at the start of the trial, followed by a delay, and then an array of stimuli. In target-present trials, the array contained a target matching the cue, and the monkey was rewarded for making a saccade to it. In some trials the cue-target was a good stimulus for the cell (top row), and in other trials it was a poor stimulus for the cell (middle row). In target-absent trials, the array did not contain a target, and the animal was rewarded for maintaining fixation during the presentation of the stimulus array. Relative locations of the good and poor stimulus in the array varied randomly from trial to trial. (b) Response of a population of 88 individually recorded IT neurons at the time of the cue presentation. Trials with a given cue were run in blocks. Cells showed higher maintained activity both before and after the cue presentation on trials when the cue was a good stimulus for the cells. (c) When the choice array was presented, the same cells shown in (b) initially responded well, regardless of which stimulus was the target. By 170 ms after stimulus onset, responses diverged dramatically depending on whether the target was the good or poor stimulus for the cell. This target-selection effect occurred well in advance of the saccade to the target, indicated by the small vertical bar on the horizontal axis. (d) In target-absent trials, the response to the good and poor stimulus together was smaller than the response to the good stimulus alone. Adapted from Chelazzi et al. (1998).
Following the eye movement, activity increased somewhat, owing to the shift of the good stimulus to the foveal part of the retina, which is the most sensitive portion of IT receptive fields. In contrast, when the poor stimulus was the target, the response to the (now irrelevant) good stimulus in the array became strongly suppressed over the course of approximately 100 ms. Similar results were found with larger arrays. In principle, the suppression began long enough in advance of the eye movement (at least 100 ms) to be the signal to the oculomotor system to make the eye movement to the target. With selection of the target, cells responded as though the irrelevant distractors had been filtered from the visual field. Thus, the resolution of competition results in a suppression of cells representing behaviourally irrelevant stimuli. The temporal dynamics of the suppressive effects are consistent, even at a quantitative level, with the predictions of a computational implementation of biased competition during search (Usher & Nieber 1996).

Together, these physiological results illustrate some of the basic components of the biased competition model, including a bias in favour of cells representing the relevant stimulus, an initial parallel activation of cortical representations by several stimuli in the visual field, and the ultimate suppression of response to the behaviourally irrelevant distractors. An examination of the results from trials with a single stimulus in the choice array as well as from trials in which the target was absent from the array allowed us to ask more detailed questions about the nature of the competitive interactions, which are described in the next section.

(a) Competitive interactions

Several lines of evidence strongly suggested that a main effect of target selection is to bias an underlying suppressive interaction between the stimuli in the choice array. First, we compared the response to the two-stimulus arrays when the target was absent (target-absent trials in figure 1a) to the response to the good and poor stimuli presented alone, with stimuli contained within the contralateral visual field. Figure 1d shows the population histograms for this comparison, which reveal a suppressive influence of the poor stimulus in the array on the response to the good stimulus. The response to the two-stimulus array was intermediate between the responses to either stimulus alone. Other studies have also reported that the response to two stimuli in the receptive field (RF) of an IT neuron are smaller than the response to either stimulus alone (Richmond et al. 1983; Sato 1989; Miller et al. 1993a; Rolls & Tovee 1995), supporting the idea of a competitive interaction.

Next, we compared the responses to the individual stimuli presented alone with the response to the two-stimulus array on target-present trials. These data showed that, by the time the eye movement was made, the effect of selecting the good stimulus as the target in the two-stimulus array was to eliminate the suppressive effect of the poor stimulus. That is, by the time of the eye movement, the response to the good stimulus had become about equal to the response to the good stimulus presented alone. Conversely, the effect of selecting the poor stimulus as the target in the two-stimulus array was to nearly eliminate the excitatory influence of the good stimulus. That is, by the time of the eye movement, the response to the array when the poor stimulus was selected as a target had been reduced almost to the level of the response to the poor stimulus presented alone. These results strongly argue against a model for selection based purely on enhancement. Rather, the main effect of attentional selection in this study appears to be a modulation of the underlying competitive interaction between the stimuli in the visual field.

Interestingly, a somewhat different picture emerged when we examined the data from trials where the two stimuli in the choice array were positioned on opposite sides of the vertical meridian, in opposite hemifields. In this case, whichever stimulus was in the contralateral visual field appeared to dominate the cell’s response to the two-stimulus array, and this dominant effect was largely unaffected by selecting either stimulus as a target. Other studies have found significant attentional effects with this configuration, but the effects are smaller than when stimuli compete within the same hemifield (Sato 1988; R. Desimone, unpublished observations). Apparently, anatomical connections that cross the midline are at a great competitive disadvantage compared with those that originate within the same hemisphere. Again, these attentional results are inconsistent with a simple attention-enhancement model; rather, they indicate that attentional selection is strongly linked to the underlying anatomy that subserves competitive interactions between stimuli in the visual field.

The effects of target selection in visual search are not confined to temporal lobe areas. In preliminary experiments, we have found very similar results in area V4 (Chelazzi & Desimone 1994), a visuotopically organized area that projects to IT cortex and forms part of the ventral stream for object recognition (Zeki 1971; Van Essen & Zeki 1978; Desimone et al. 1980; Gattass et al. 1988; Baizer et al. 1991). Neurons in area V4 have restricted receptive fields and respond selectively to stimulus form and colour (Zeki 1973, 1983, 1996; Desimone & Schein 1987; Schein & Desimone 1990; Gallant & Van Essen 1996).

We found that, as in IT cortex, stimuli that are contained within the receptive field of a V4 neuron appear to engage in suppressive interactions, which are strongly modulated by attentional selection (Chelazzi & Desimone 1994). Independent of target selection, the presence of a poor stimulus in the receptive field has a suppressive effect on the response of a V4 neuron to a good stimulus in the receptive field, and this suppressive effect is reduced as the poor stimulus is moved further away, outside the receptive field. If a good and a poor stimulus are paired in a search array contained within the receptive field and the animal selects the good stimulus as the target, the cells give a good initial response and continue to respond well until the time of the eye movement. In contrast, with the same pair of stimuli, the response of the cell to the good stimulus is initially strong but soon becomes greatly suppressed on trials when the poor stimulus is the target. If the poor stimulus is moved well outside the receptive field, these attentional effects are reduced in magnitude. Thus, consistent with the biased competition model, competitive interactions are strongest when competing stimuli activate cells in the same local
region of cortex with similar receptive fields. Because the attentional feedback biases this underlying competitive interaction, the attentional effects are stronger with nearby stimuli as well.

These data on attentional modulation of V4 responses in visual search are also in general agreement with those of Motter (1994), who has also reported attentional effects not limited to a single spatial location in V4. In Motter’s study, the animal was presented with a large array of coloured stimuli and was cued to attend to all stimuli of a given colour, which varied from trial to trial. The responses of V4 cells were larger when the cued colour matched the colour of a stimulus in the receptive field, even though the monkey had not been cued to attend specifically to the receptive field location. Although Motter did not study competitive interactions between the stimuli in the array, such interactions are likely to have been present owing to the density of stimuli in the vicinity of the receptive field.

3. SPATIALLY DIRECTED ATTENTION

From the point of view of the competition model, attentional selection of a stimulus on the basis of its spatial location should involve neuronal mechanisms that are qualitatively similar to those for selection by nonspatial stimulus features. Thus, spatial selection should involve a feedback bias in favour of a stimulus that is behaviorally relevant because of its location, and this bias should then modulate an underlying competitive interaction among the attended and unattended stimuli in the visual field. This is precisely what has been found in studies of spatial attention in ventral stream areas, and a recent study indicates that the same principle may hold in some dorsal stream areas as well (see next paragraph).

Moran & Desimone (1985) and Luck and co-workers (1997) studied spatial attention in ventral stream areas. In these studies, two stimuli were typically presented simultaneously in the visual field, and the animal was rewarded for performing a task (e.g. matching-to-sample or target detection) on the stimulus at one of the locations, while ignoring the stimulus at the irrelevant location. One of the stimuli was placed within the receptive field, and the other was placed at a variable location, either inside or outside the receptive field. The animal was cued as to the relevant location at the start of each block of trials, and the relevant location was switched between blocks. As in the visual search studies, one stimulus in the pair was typically chosen to be a good stimulus for the cell and the other was chosen to be a poor stimulus. By using this strategy, one could measure the neuronal response to the good stimulus in the receptive field when it was attended on one set of trials, compared with when the same stimulus was an irrelevant distractor on a different set of trials.

(a) Feedback bias

With a single stimulus inside the receptive field, the effects of any feedback bias during spatially directed attention may be observed independently of any effects of the bias on competitive interactions. When one stimulus was located inside the receptive field of a cell in either V2 or V4 and another stimulus was located outside, Luck and co-workers (1997) found that the maintained firing rate of the cells was higher when attention was directed to the location inside the receptive field than when attention was directed outside. As shown in figure 2, the increase in firing rate began at the start of the trial, before any stimulus was presented, and thus was owing strictly to the animal expecting a relevant stimulus at the receptive field location. This is very similar to the elevation in maintained firing rate found by Chelazzi and co-workers (1993) during visual search, where IT cells showed higher maintained firing rates if the animal was attending to a stimulus with a preferred, non-spatial, feature for the cell. As in the visual search experiment, this increase in firing rate during spatial attention is consistent with the idea of feedback biasing activity in favour of the behaviourally relevant stimulus.

The spatial resolution of this feedback bias for spatial location is also very high. Luck and co-workers (1997) measured the maintained firing rate when the animal directed its attention to one of two locations within the RF of cells in V4, one closer to the ‘hot spot’ of the receptive field than the other (i.e. elicited higher firing rates to the same stimulus). The maintained firing rate was higher...
when attention was directed to the location closer to the hot spot, indicating that the resolution of the mechanism responsible for the increase in firing rate has a resolution finer than the dimensions of the receptive field. This fine spatial resolution of the feedback is presumably related to the ability of cells to distinguish between attended and unattended stimuli located simultaneously within the receptive field, described here.

In addition to increasing maintained neuronal activity, other studies suggest that the bias in favour of the relevant location also increases the sensitivity of cells to stimuli presented at that location, although this increase in sensitivity may not always result in stronger sensory responses. Spitzer and co-workers (1998) found that attention to a stimulus within the receptive field of cells in V4 increased the cells' selectivity for orientation and colour and also caused an increase in the gain in the response by about 20%, compared with when the animal attended to a stimulus outside the receptive field. However, this increase in the magnitude of the response to the attended stimulus was found only when the animal performed a difficult task. Likewise, a preliminary study has shown that attention to a stimulus in V4 specifically increases the contrast sensitivity of cells (Reynolds et al. 1996; Reynolds & Desimone 1997), i.e. cells responded to the attended stimulus as though its contrast had been increased. This increase in contrast sensitivity with attention was found predominantly with low-contrast stimuli, because the responses of many V4 cells were already saturated at high stimulus contrast even in the absence of attention (Reynolds et al. 1996). An overall increase in response with attention in V4 has recently been confirmed in a preliminary study by McAdams & Maunsell (1997). Differences in task-difficulty, contrast-saturation effects, or both factors together may explain why several other studies have failed to detect the effects of feedback bias on the gain or sensitivity of cells to a single stimulus in the receptive field of V4 cells (Moran & Desimone 1985; Haenny et al. 1988; Maunsell et al. 1991; Motter 1993; Luck et al. 1997).

In summary, a bias in favour of an attended stimulus inside a V4 receptive field causes an increase in sensitivity, or excitability, which may become manifest as a change in maintained activity and/or sensory-evoked responses. However, these attentional effects with a single stimulus inside the receptive field are variable and have, at best, a rather modest effect on the absolute magnitude of firing rate in prestriate areas such as V2 and V4. As described below, strong consistent effects of attention occur when these biases in favour of the relevant stimulus influence the competition between stimuli in the visual field.

(b) Effects of attention on competition

Similar to what was found in the visual search experiment in IT cortex, attention has a large effect on responses when two stimuli compete within the same receptive field. When two stimuli are located within the receptive field of cells in V2 or V4, and the animal attends to one of them, the cell's response is predominantly determined by the attended stimulus (Moran & Desimone 1985; Luck et al. 1997). If the good stimulus for the cell is attended, the response is strong. If the poor stimulus is attended, the response to the good stimulus in the receptive field is greatly suppressed. Figure 3 shows the effects of attention on competition between two stimuli in the receptive field for a population of V4 cells. Comparable effects are found in IT cortex, although the receptive fields of IT cells are much bigger and the attentional effects generalize over a larger portion of the visual field than in V2 and V4 (Moran & Desimone 1985). As predicted by the biased competition model, a preliminary study in areas V2 and V4 indicates that attention serves to modulate the suppressive interaction between two or more stimuli within the receptive field (Reynolds et al. 1994, 1995).

Interestingly, during spatially selective attention, the onset of suppression for an unattended stimulus often occurs earlier in the neuronal response than was found in the visual search task. In the visual search task, the onset of suppression for the unattended stimulus began as early as 170 ms in the population response. However, in one study of spatial attention in V4, attentional effects were found at the very onset of the response (Luck et al. 1997), 50 ms after stimulus onset, and other studies have found an onset of suppression shortly after the onset of the visual response, in the range of 90–120 ms (J. Moran and R. Desimone, unpublished data; J. Reynolds and R. Desimone, unpublished data). It is possible that this difference is owing to the nature of the feedback bias, which may be more accurately targeted to all cells with the same

---

**Figure 3.** Stimulus-evoked responses in V4 vary according to the locus of spatial attention. Top: schematic representation of stimulus configuration. While the animal maintained its gaze on a fixation spot (Fix.), two stimuli were presented simultaneously within the receptive field (RF, dashed box). One of the stimuli was a good stimulus for the cell (i.e. would activate the cell well if presented alone), and the other was a poor stimulus for the cell (i.e. would not activate the cell very well if presented alone). The animal was cued to perform a discrimination task on the stimulus at one location or the other, and the task-relevant location varied across blocks of trials. Bottom: average response of 29 individually recorded V4 neurons to the pair of stimuli when either the good or poor stimulus was attended. The response to the good stimulus was suppressed when the animal attended to the poor stimulus of the pair. Adapted from Luck et al. (1997).
receptive field in a small portion of a visuotopically organized area than to cells that share a common selectivity for shape or colour but which are widely distributed throughout the cortical area.

A somewhat different account of the effects of attention has been advanced by Connor and colleagues (1996) who have studied the effects of spatially directed attention in V4. In their experiment, the animal’s attention was directed to one of several large stimuli in the shape of rings immediately outside the classical receptive field of the cell. They then measured the response to a probe stimulus presented at a variable location within the receptive field. They found that the response to the probe stimulus at a given location in the receptive field was typically greatest when attention was directed to a nearby ring on the same side of the receptive field as the probe than when it was directed to a more distant ring on the opposite side of the receptive field. The authors interpreted this effect as a shift in the receptive field sensitivity profile towards the locus of attention. A related notion that the receptive field shrinks around the locus of attention was originally proposed to account for the effects of attention in V4 (Moran & Desimone 1983; Desimone et al. 1989). Neither idea is strictly incompatible with the competition model because, according to the model, attentional modulation of competition may lead to changes in the apparent size or profile of the receptive field. The empirical question is whether the apparent shift in receptive field profile in the study by Connor and colleagues (1996) is actually owing to an attentional modulation of a competitive interaction between the ring stimulus and the probe. Competition from stimuli outside the receptive field borders is likely to be strongest when the stimuli are both large and numerous, as in the study by Connor and co-workers (1996).

(c) **Biased competition in the dorsal stream**

Evidence that some type of biased competition may operate in the dorsal stream of visual areas important for spatial vision has recently been reported by Treue & Maunsell (1996). They made recordings from cells in MT and MST in monkeys trained to attend to one of two stimuli moving in opposite directions in the visual field. Cells in both MT and MST are often highly selective for the direction of motion of a stimulus, although receptive fields are much larger in MST than MT. Treue & Maunsell used the same strategy that had been used in V4 and IT cortex to study attentional selection of one of two competing stimuli in the receptive field. They chose one stimulus to be a good stimulus for the recorded cell (moving in its preferred direction) and the other to be a poor stimulus (moving in the opposite direction). They compared the response to the good stimulus, when it was attended, to when it was an irrelevant distractor. When both stimuli were in the same receptive field of cells in either MT or MST, the response of the cell was almost completely determined by the attended stimulus, similar to what has been found in V2, V4, and IT cortex. The response was good when the animal attended to the stimulus moving in the preferred direction of the cell and poor when it attended to the stimulus moving in the opposite direction. The competition model predicts that this attentional modulation of responses should be correlated with a competitive interaction between the two stimuli in the receptive field, but this has not yet been tested in MT and MST. Much smaller effects of attention were found when one of the two stimuli was moved outside the receptive field, presumably reducing local competitive interactions in these areas. This inverse dependence of the attentional effects on the spatial separation between the stimuli is inconsistent with any simple enhancement model of attention but is similar to what was found in V2 and V4. Groh and co-workers (1996) have recently reported that they have failed to replicate strong attentional effects in MT using a different task and different stimuli from those used by Treue & Maunsell (1996), raising the possibility that the attentional modulation of a particular visual area will be stimulus- or task-specific (see Richmond & Sato 1987; Spitzer & Richmond 1991; Corbetta et al. 1991).

It seems likely that some form of biased competition works in other dorsal stream areas as well. Several studies have shown that responses to a stimulus in the receptive field of cells in posterior parietal cortex are larger when the animal attends to it than when it ignores it (Lynch et al. 1977; Robinson et al. 1978, 1995; Bushnell et al. 1981; Steinmetz et al. 1994; Steinmetz & Constantinidis 1995; Colby et al. 1996). This is the traditional ‘enhancement’ effect of attention. However, there is at least some suggestive evidence for competitive interactions between targets and distractors in structures associated with the dorsal stream, in addition to the traditional enhancement for stimuli chosen as targets.

In the frontal eye fields (FEF), an oculomotor structure closely associated with the dorsal stream, Schall and colleagues (Schall & Hanes 1993; Schall et al. 1995) have found that the response to a distractor stimulus within the receptive field of visual cells is suppressed when the animal plans to make an eye movement to a target stimulus outside the receptive field. However, this suppression of the distractor response is much greater when the target stimulus is located just beyond the receptive field borders than when it is in a more distant location, consistent with some type of competitive interaction.

Likewise, in the superior colliculus, another oculomotor structure, Basso & Wurtz (1997) have found that the activity of build-up cells in the intermediate layers is suppressed by the presence of several stimuli in the visual field. The suppression ceases when a stimulus inside the cell’s movement field is chosen as a target. Suppression is also reduced by increasing the probability that the stimulus in the movement field of a cell will be chosen as a target. Basso & Wurtz interpreted the suppressive effects of several stimuli in the visual field to mean that collicular responses are reduced by target uncertainty, i.e. that the suppressed responses reflect the reduced likelihood that a stimulus in a given cell’s movement field will be the target of an eye movement. An alternative view is that target uncertainty per se is not computed. Rather, the response suppression may be owing to mutually inhibitory interactions among cells activated throughout the visual field map of the colliculus (or an upstream structure). These interactions would then be modulated by feedback to the colliculus, biasing the interactions in favour of cells at the target location. Competition between stimuli is likely to take place at all levels of the nervous system, from stimulus to response.

Phil. Trans. R. Soc. Lond. B (1998)
4. ATTENTION, MEMORY, AND PREFRONTAL CORTEX

Attention is often thought of as the gateway to memory, as we typically remember little about stimuli we ignore. However, the reverse is also true—mechanisms for learning and memory play a critical role in selecting which stimulus we will attend to in a crowded scene. Attention may be directed preferentially to novel stimuli, to recognized stimuli, to stimuli with learned behavioural relevance, and so on. In fact, some of the neuronal mechanisms for memory and attention are so intertwined that one may question whether they are even distinguishable (Desimone et al. 1994; Ungerleider 1995; Desimone 1996). Here, we will focus on the role of neuronal mechanisms for working memory, as those are the ones most likely to be involved in many of the attentional experiments described so far in this review.

Working memory in monkeys typically refers to any type of short-term memory in which specific information is kept actively ‘on-line’ for up to several seconds and then discarded (see Badeley 1986), for a review of human studies). In one type of working memory task, known as matching-to-sample, the monkey is shown a sample stimulus at the start of the trial, followed by a sequence of one or more test stimuli. Delay intervals intervene between all of the stimulus presentations, and the monkey is rewarded for signalling when a test stimulus matches the sample. Working memory is required to solve such a task when small stimulus sets are used, because the relevance of a particular stimulus changes from trial to trial—a stimulus used as the sample on one trial will temporarily have great behavioural relevance but will become behaviourally irrelevant in another trial in which it is a nonmatching stimulus.

Neurons in IT cortex of monkeys performing the matching-to-sample task often show stimulus-specific activity during the delay interval following the sample (Fuster & Jervey 1981; Miyashita & Chang 1988; Miller et al. 1993a,b, 1996; Vogels & Orban 1994). If the sample stimulus is a good stimulus for the cell, the maintained activity during the delay is high, whereas if the sample is a poor stimulus the maintained delay activity is low. Furthermore, in a version of the task that is particularly demanding of working memory, the response of an IT neuron to a given test stimulus is enhanced if it matches the sample (Miller & Desimone 1994). Both the higher maintained activity and enhanced responses suggest that some IT cells are sensitized, or biased, to respond preferentially to the behaviourally relevant (matching) choice stimulus in a trial.

This biasing of IT neurons in a working memory task is remarkably similar to the biasing effects on neuronal responses found throughout extrastriate cortex during visual search and spatially directed attention. Indeed, there is every reason to expect that they should be similar, because in both types of tasks the animal must use information held flexibly in working memory to guide its selection of stimulus or response when confronted with a choice. The difference is that in many working memory tasks the choice stimuli are distributed in time, whereas in visual search the choice stimuli are distributed in space. Because the functional requirements for this feedback to visual cortex are the same in both working memory and attention tasks, it is reasonable to assume that it derives from the same sources.

There are several reasons to suspect that a main source of top-down feedback to visual cortex during both working memory and attention is prefrontal cortex (see Desimone & Duncan 1995). One main reason is that prefrontal cells seem to have the appropriate properties. In working memory tasks, prefrontal cells show stimulus-specific delay activity (Fuster 1973, 1995; Niki & Watanabe 1976; Funahashi et al. 1989, 1993a,b; Sawaguchi et al. 1988a,b; Wilson et al. 1993, di Pellegrino & Wise 1993; Miller et al. 1996; Rao et al. 1997). Thus, if at the start of a behavioural trial animals are shown a brief cue that specifies the critical information that they need to solve the task on that trial, some prefrontal cells will show high cue-specific activity that persists for as long as that information is important. Cells in the dorsolateral portion of prefrontal cortex tend to have delay activity that is specific for a remembered location in space, whereas cells on the ventral convexity of prefrontal cortex tend to have delay activity that is specific for a particular complex object (Wilson et al. 1993), and many cells in both regions have delay activity that is specific for both object and place (Rao et al. 1997). If this maintained activity were fed back to the appropriate cells in extrastriate cortex, prefrontal cortex would be in a position to bias extrastriate activity in favour of a relevant object at either a specific spatial location or with a specific shape, colour, and so on. Furthermore, delay activity in prefrontal cortex is not disrupted by intervening or distracting stimuli, unlike delay activity in IT cortex (Miller et al. 1996). Brain imaging studies indicate that both prefrontal cortex and extrastriate visual areas are activated during both visual stimulation and the delay periods of working memory tasks, but prefrontal activity is more closely linked to the delay period whereas extrastriate activity is more closely linked to the visual stimulation (Courtney et al. 1997).

Both behavioural and anatomical data also suggest that feedback from prefrontal cortex plays an important role in attention and working memory. Lesions or deactivation of prefrontal cortex impair performance on working memory tasks in monkeys (Mishkin 1973; Bauer & Fuster 1976; Mishkin & Manning 1978; Fuster et al. 1985; Sawaguchi & Goldman-Rakic 1991; Funahashi et al. 1993a; Schindly et al. 1994), and cooling a small part of prefrontal cortex causes delay activity in IT cortex to become less selective (Fuster et al. 1985). Anatomical studies reveal that prefrontal cortex has reciprocal connections with virtually all extrastriate cortical areas and is, therefore, in a position to directly influence visual cells throughout all areas except for primary visual cortex (Barbas 1988; Barbas & Pandya 1989; Ungerleider et al. 1989; Webster et al. 1994). Finally, it makes sense for the feedback to extrastriate visual cortex in working memory and attentional selection tasks to come from structures such as prefrontal cortex that are not purely visual. Behavioural relevance and expectations are often established by the behavioural context, which is itself defined by abstract and multimodal sources of information. A hypothetical representation of the interaction between prefrontal and IT neurons during visual search is shown in figure 4.

Despite this suggestive evidence, the definitive experiments to test the role of prefrontal feedback to extrastriate
cortex have not yet been done. Several cortical regions share at least some key properties with prefrontal cortex (see, for example, Suzuki et al. 1997), and it may be overly simplistic to think of a single cortical region providing most or all of the top-down feedback to extrastriate visual areas, even in well-defined tasks of working memory and attention. Furthermore, the behavioural relevance of a stimulus will often be defined not only by the contents of working memory but also by affective state and long-term learned associations, which probably involve different neural systems. How these many potential sources of feedback target the specific visual cells representing behaviourally relevant stimuli remains perhaps the most mysterious aspect of attention.

I thank my collaborators on the attentional experiments: L. Chelazzi, J. Duncan, S. Hillyard, S. Luck, J. Moran, E. Miller, T. Pasternak, J. Reynolds and H. Spitzer. L. Ungerleider provided valuable comments on the manuscript.

REFERENCES


Figure 4. Possible interaction between prefrontal and IT cortex during visual search. (a, b) Each box represents the activity in a hypothetical population of prefrontal or IT neurons. Each dot represents a population of cells, and the size of the dot represents the level of activity. (c) Representation of stimuli used in the task. When the cue is presented, the prefrontal cells that represent the features of the cue develop a high state of maintained activity which persists for the duration of the trial. This activity is fed back to IT cortex, biasing activity in favour of cells selective for features of the stimulus used as the cue and target. When the choice array is presented, cells in IT cortex are initially activated if they are selective for a feature of any stimulus in the array. Cells representing different stimuli engage in mutually suppressive interactions, which are biased in favour of cells representing the target. By late in the choice period, cells representing non-target stimuli are suppressed. About 100 ms after the onset of suppression, the eye movement to a target (arrows) is initiated.

Colby, C. L. 1991 The neuroanatomy and neurophysiology of attention. J. Child Neurol. 6, 90–118.


Morton, B. C. 1993a Focal attention produces spatially selective processing in visual cortical areas VI, V2 and V4 in the presence of competing stimuli. J. Neurophysiol. 70, 909–919.


Zeki, S. M. 1971 Cortical projections from two prefrontal areas in the monkey. *Brain Res.* **34**, 19–33.


Zeki, S. M. 1996 Are areas TEO and PIT of monkey visual cortex wholly distinct from the fourth visual complex (V4 complex)? *Proc. R. Soc. Lond.* **B** **263**, 1339–1544.