The cognitive neuroscience of visual attention
Marlene Behrmann* and Craig Haimson†

In current conceptualizations of visual attention, selection takes place through integrated competition between recurrently connected visual processing networks. Selection, which facilitates the emergence of a ‘winner’ from among many potential targets, can be associated with particular spatial locations or object properties, and it can be modulated by both stimulus-driven and goal-driven factors. Recent neurological data support this account, revealing the activation of striate and extrastriate brain regions during conditions of competition. In addition, parietal and temporal cortices play a role in selection, biasing the ultimate outcome of the competition.

Introduction
The term ‘selective attention’ generally refers to the set of operations that determine which of several possible inputs will be analyzed past the level at which all may be processed in parallel. Standard conceptualizations of selective attention have undergone numerous transformations in the four decades following Broadbent’s [1] initial proposal. Early theories drew analogies between selection and a filtering mechanism that operated in accordance with a set of pre-attentive Gestalt principles (see e.g. [4]). Later theories recast attention as the selective filtering mechanism that selected input associated with a specific set of spatial locations (see e.g. [4]), whereas others pointed to a mechanism that selected input associated with a representation of an object that had already been fully parsed (see e.g. [6]). Most importantly, it was unclear how competition and modulation might be related to processing in the brain.

Computational modelling, however, has provided a means of addressing these issues. Biased competition can be readily understood in terms of interactions between units in ‘winner-take-all’ (WTA) neural networks, and a number of neurally inspired computational models of selective attention have employed WTA network dynamics [7–9]. Moreover, recent proposals (e.g. [10,11]) have interpreted the functioning of the attentional components associated with posterior and anterior neuroanatomical regions [12] in terms of WTA interactions among lower-level representations in early visual cortical and subcortical areas, higher-level representations in the dorsal and ventral visual processing streams, and the frontal lobes (see Figure 1). According to these models, representations within the same processing region that correspond to different portions of input are mutually inhibitory, whereas those in different processing systems that correspond to the same portion of input are mutually excitatory. Thus, selection emerges from local competition and nonlocal cooperation in multiple levels of processing throughout the entire set of interconnected networks.

Competition in lower-level posterior regions will tend to be influenced by exogenous factors such as attribute salience. This will, in turn, affect competition in the higher-level, more anterior regions to which they provide input. However, endogenous factors such as task relevance or goal-driven strategies will tend to bias competition in the more intention-related anterior representations that can, in
Biased competition as a mechanism of selection

Recent support for the biased competition model of visual attention has been provided by both functional magnetic resonance imaging (fMRI) and event-related potential (ERP) studies. In these studies, enhancement of cortical activation or attention-related waveforms is observed under various conditions of competition: for example, when subjects perform more difficult discriminations, when distractors compete with targets and when task demands increase (e.g. when subjects have to saccade to a feature target rather than simply signalling its presence with a key press) [18]. Moreover, when task demands increase, less activation is observed in areas associated with irrelevant task being performed simultaneously, reflecting the decrease in processing of nonselected information [19]. Competitive effects are also observed when, for example, stimuli are presented simultaneously rather than sequentially. In the former condition, greater activation is seen in areas V1, V2, V4, and TEO, possibly reflecting increased receptive field size and greater competition between neurons encoding different objects located within overlapping regions of space. Interestingly, the competition between the stimuli can be cancelled by having subjects attend selectively to one of the possible stimulus locations [20].

Competition is also evident when attending selectively to one of two perceptual features of a stimulus, compared with dividing attention across both. For example, compared with a fixation condition, attending to the colour or shape of a stimulus manifests as increased activation in occipital and inferior temporal regions [21]. Switching attention between the two features of a stimulus, perhaps reflecting the dynamics of the WTA, engages additional cortical regions and results in activation in parietal regions as well as in the cerebellum. Similar effects are observed when subjects attend selectively to the local or global level of a stimulus; this produces an amplitude modulation or increase in the early P1 components of the ERP waveform, in a region corresponding to the fusiform gyrus on PET (positron emission tomography) scans [22]. When attention is divided between local and global levels, later, more hemisphere-specific effects are observed. Interestingly, even when subjects attend to a single dimension such as colour (e.g. to a red or to a blue stimulus), early waveforms associated with attending to one of the two colours may be localized to more anterior fusiform regions and to prefrontal cortex.

Attention-related impairments resulting from damage to parietal cortex can also be interpreted as reflecting competition between targets. When letters are presented to the left and right sides of a patient with a right parietal lesion, report of the left item (‘extinction’) is poor at short SOAs (stimulus onset asynchrony) irrespective of whether the left or right item is presented first [23]. Report of the contralateral item improves, however, with increasing temporal lag between the two presentations, suggesting that when the less powerful contralateral item obtains a temporal advantage, it is sufficiently activated to compete with and surpass the ipsilateral representation. Bottom-up properties of the stimulus can also affect the competition, as the contralateral item is reported more often when it shares orientation with the ipsilateral item and is, therefore, more resistant to extinction [24]. Furthermore, the competition can span more than a single modality: when a visual stimulus is presented on the ipsilateral side, it can inhibit detection of a tactile stimulus presented contralaterally [25].

Figure 1

Reciprocal connections between components of the integrated competition system responsible for attentional modulation.
Selection by spatial location

An efficient means of selecting input is on the basis of spatial location. When spatial regions are selected covertly (i.e. in the absence of an eye movement) by pre-cueing spatial locations in which targets will probably appear, neural regions, including the right anterior cingulate gyrus and right posterior parietal cortex (intraparietal sulcus), are activated [26]. Attentional modulation occurs irrespective of whether the cues are nonsymbolic and appear peripherally [26] or are symbolic and presented centrally [27]. As would be predicted by a competition account, dividing attention between right and left locations results in longer reaction times to detect the target compared with when attention is biased to one side. An early ERP P1 component over the lateral occipital scalp accompanies the attentional benefit and a late positive deflection reveals both the attentional costs and benefits [28]. Moreover, an event-related optical signal can be elicited in early visual areas when stimuli appear at attended, relative to unattended, locations. Whether or not the neural mechanisms associated with covert attentional shifts are identical to mechanisms dedicated to saccadic shifts remains a matter of debate. A direct comparison between a saccadic and a covert attentional shift task resulted in identical regions of activation for both, albeit to a greater extent in the eye movement task [29].

Data from neuropsychological subjects reveal a dissociation, however, between mechanisms subserving attentional and gaze orienting [30]; neglect patients with lesions within the intraparietal sulcus make leftward eye movements without corresponding attentional shifts, whereas neglect patients with parietal lesions, who do not respond to leftward stimuli, show different patterns of covert and overt responses to right-sided targets of differing eccentricities.

Modulation of competition: bottom-up and top-down biases

Competition between input stimuli may be moderated through top-down biases, higher-level representations of spatial location and object structure. In addition to early striate and extrastriate areas, posterior parietal regions are especially well suited to the discrimination of multiple features such as visual search for a target defined by a conjunction of multiple features [18••–31]. Performance on these tasks may be severely impaired in brain-damaged patients with parietal lesions [32]; although see [33]). The critical role of the parietal cortex in the attentional circuit has been confirmed by a study showing that transcranial magnetic stimulation (TMS) applied to the right posterior parietal cortex (intraparietal sulcus) [45]; magnetoencephalography (MEG) alpha rhythms in this region were suppressed most strongly when participants incorrectly labelled objects as noise stimuli and were enhanced even further when stimuli were actually non-objects.

Endogenous biases may also affect the competition in different cortical regions. Even when the visual display is held constant, increased cortical activation in MT–MST was observed when subjects were instructed to attend to moving rather than stationary dots [36]. Under similar conditions, attentional modulation can also be observed in early visual areas such as V1 and/or V2 [37], albeit to a lesser extent [38]. Similarly, enhanced activation is observed in regions of temporal cortex involved in face processing as a function of preferentially attending to faces [39]. Finally, long-latency field potentials in posterior fusiform gyrus, recorded directly from the inferior surface of the temporal lobes in epilepsy patients, are modulated when subjects attend to one of two streams of words [40].

Modulation of competition by object representations: spatially invariant or spatiotopic?

The finding that object representations can alter selection may be readily understood as resulting from interactions between lower-level input representations in early visual cortex and higher-level representations of object structure in the ventral pathway. While some have argued that the object-based attentional effects are mediated by spatially invariant object representations [15], others suggest that they are mediated by more spatiotopic representations (e.g. the ‘grouped array’ proposed by Farah [41]), and susceptible to modulation by perceptual organization [42•]. Both types of representations, however, probably play a role (see Figure 1).

Certainly, it is likely that object perception involves competitive interactions between spatially invariant object representations in later temporal lobe regions (e.g. [43,44]) and that these dynamics play an important role in determining which of several objects ultimately ‘captures’ attention and allows for the generation of a behavioral response. However, given the interconnected nature of the visual system, it is also highly likely that these processes influence competition in lower-level visual areas via feedback connections as well as via connections with spatial representations in the dorsal stream (which, in turn, provide biasing feedback to early visual areas). Evidence to support this highly interactive system comes from a study showing that information about visual shape can even affect activity in the parieto-occipital sulcus [45].
The effect of object properties on single-cell activity can be observed in the earliest cortical visual areas. Roelfsema et al. [46••] found that the firing rates of neurons in V1 corresponding to various segments of a curved line were enhanced relative to responses to a distractor line even when the lines were spatially overlapping. These results suggest that it is the entire object that is modulating the firing rate rather than attention spreading across spatial positions occupied by the curve. The suppression of early ERP waveform components in human subjects also suggests that modulation of early ERP activity may be related to feedback from spatially invariant representations. Because suppression of P1 and N1 ERP components — both of which reflect the spatial distribution of attention — was observed when subjects attended to two different objects that occupied the same spatial position but not when a single object was present, an object-based rather than space-based explanation is more appropriate [47].

Selection appears to involve not only object-based but also space-based representations. Enhanced amplitude of N1 or P1 ERP signals is seen when subjects report attributes of two overlapping objects [48•] or when a single object is observed during trials in which attributes to be reported (e.g., colour and shape) appear in different objects (attributed to a change in the size of the attended region), whereas enhanced P1 activity is observed on same-object trials in which the stimulus often reflecting the increased focus of attention). Importantly, along with these spatial modulations, enhanced P3 waveform (a later, more cognitive component) are observed in occipital, parietal, and temporal sites, suggesting the involvement of object-based representations.

The joint contribution of both space- and object-based attentional systems is also evident in a recent PET study. Fink et al. [49•] found that there is a substantial overlap in activation patterns when subjects perform an object-based (e.g., is dot on the left or right side of a line?) and space-based (e.g., is the dot on the left or right side of the line?) task relative to a control condition. In addition, there were areas that selectively activated in each of the two tasks. Areas of overlap included left and right medial and lateral parietal cortex, as well as left prefrontal cortex; however, the left occipital lobe, the striate and prestriate cortex were activated only in the space-based task.

The interactions between perceptual organization of objects and spatial attention are also evident in studies of patients who extinguish left-sided information following right-parietal lobe damage. Just as a prior temporal entry for a contralateral target can reduce the probability of its being extinguished by a competing ipsilateral target [23••], so can grouping the contralateral item with its ipsilesional counterpart makes it resistant to extinction. Left-sided information can be grouped with right-sided information by bottom-up factors such as colour, proximity, brightness or collinearity [50,51], or even by an illusory contour of a partially occluded figure [52•]. Top-down modulation also plays a role, resulting in less extinction for known, familiar objects or words than for unknown items [53–55].

Computational accounts and integrated competition

Several recent computational accounts of visual attention include an attentional map (also referred to as a priority or saliency map [8,56]) that instantiates the WTA dynamics. The winner, which emerges from the competition, may then be mapped through a smaller 'attentional' window (the selection network [57]) or may be selected as the target because of its high saliency value (both absolute and relative to the other stimuli [58•]). Selected representations may be subsequently inhibited to implement inhibition of return [58•]. These models successfully simulate not only the performance of normal subjects in various experimental paradigms, but also exhibit the behavioral pattern associated with attentional deficits such as hemispatial neglect and extinction. One notable difference between these models and the full integrated competition account is that in the models the competition is more localized to a particular component rather than being widespread throughout the processing circuit. Whether this departure is truly significant remains to be determined.

Conclusions and future directions

Behavioral and neurobiological data are increasingly converging on the notion that selecting a target from a complex array of visual stimuli involves a recurrently connected network of visual processing areas. Competition and cooperation of units in this network facilitate the emergence of a winner from a host of potential targets. While we know that the competition can be modulated by bottom-up factors such as colour, proximity, brightness or collinearity [50,51], or even by an illusory contour of a partially occluded figure [52•]. Top-down modulation also plays a role, resulting in less extinction for known, familiar objects or words than for unknown items [53–55].

Acknowledgements

This work was supported by grants from the National Institute of Mental Health to M. Behrmann (MH57054-07 and MH58246) and a National Defense Science and Engineering Graduate Fellowship to C. Haimson.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as: • of special interest

attended stimulus. Suggesting that focused attention biased competition in favor of the one of the picture locations, however, the suppressive interactions previously activity during simultaneous presentations). When participants attended to participants viewed sequentially presented, as opposed to simultaneously pre-

studies. These conditions include harder-to-discriminate conjunction targets subjects performed a visual discrimination task under various conditions pre-

by functional MRI. The unbalanced competition between ipsilateral and contralateral sides of space is observed in right-hemisphere lesioned patients with tactile extinction even across sensory modalities; when a visual stimulus is presented near a subject’s professional hand, tactile extinction is observed in that the probability of detecting a concomitant tactile stimulus is reduced.

The authors observed an enhanced N2 posterior cortex ERP signal when The authors observed an enhanced N2 posterior cortex ERP signal when subjects attended to the word task.

attended to the word task. As opposed to simple feature targets, presence of distractors in close prox-

patterns while they made an easy or difficult judgement about a centrally pre-


visual-spatial attention produces costs and benefits in response time and evoked neural activity. Neuropsychologia 1998, 36:180-201. ERP and behavioral data recorded while subjects divided attention between right and left hemifields or attended selectively to one revealed response time advantages for the attended regions, accompanied by an early ERP P1 component. A later positive deflection revealed both the attentional costs and benefits.


vida s, zeloni g, soroker n, ring h: Visual extinction and cortical connectivity in human vision, Cogn Brain Res 1997, 6:155-163.


within a block of trials, subjects attended to either stimulus shape or colour (unattended attention) or switched attention across the two features during 4 T fMRI. Relative to a fixation condition, significant bilateral activation in occipital and temporal cortices was observed during sustained attention. Over and above this, in the switching condition, cerebellar, occipital, premotor, and superior parietal lobule activation was observed.


coherent motion 401. The model is consistent with either the global or the local model of extinction whereby the contralateral (left) stimulus loses out in favour of the ipsilateral (right) stimulus.


reflexive saccadic gaze shifts. These data suggest that attention and eye move-


TMS applied over the right parietal cortex of normal subjects disrupts visual search for conjunctions but not simple features. The results are consistent with biased competition in that parietal signals feed back to extrastriate areas, affecting the processing of features from certain spatial positions. Use of TMS provides direct evidence that posterior parietal cortex plays a role in the deployment of attention during effortful visual search.


After subjects received sufficient training on a visual search task (such that search slopes were indicative of roughly parallel search), the application of TMS did not have a disruptive effect. This means that TMS did not transfer to a new conjunction search task, suggesting that practice eliminated the need for highly focused attention only in the detection of a specific, learned combination of features.


Subjects performed an object attentional selection task reporting perceptual attributes for one or two objects while responding to the appearance of probe stimuli that were occasionally presented subsequent to the removal of the object-based attention display. Subcortical performance in this task while visual attention was allocated to objects in the object-based task; this suggests that object-based decisions are made from ‘grouped-arrays’ rather than from a spatially invariant re-representation.


The authors recorded from V1 cells in macaques required to saccade from a central fixation point to a target dot at the end of a curved line originating at fixation (a distractor dot appeared at the end of an additional arc) sometimes intersecting the target. The results revealed enhanced firing rates in those cells whose receptive fields overlapped the curve connecting the target to fixation; simultaneous recording from multiple cells on different parts of the curve revealed that the onset of these enhanced firing rates occurred at the same time across all of the presented positions, suggesting that object-based attention facilitated processing of all portions of the curve simultaneously (as opposed to spreading along the curve from fixation to target).


The authors made ERP recordings while subjects reported the presence or absence of a pair of attributes (a colour and a shape) in displays consisting of two overlapping line drawings;probe dots occasionally appeared subsequent to the removal of these displays. Data revealed enhanced ERP waveform magnitude of early, space-based attention when both objects appeared to overlap, whereas similar waveforms appeared in locations previously occupied by objects containing both target attributes. (P1).


A patient with right parietal damage was better able to detect additional changes to ‘pre-attentive’ stimuli when items formed an illusory square than when they did not. This suggests that pre-attentive construction of illusory contours is mediated by the amodal completion of grouping contextual items with positional items.


59. The authors made ERP recordings while subjects reported the presence or absence of a pair of attributes (a colour and a shape) in displays consisting of two overlapping line drawings;probe dots occasionally appeared subsequent to the removal of these displays. Data revealed enhanced ERP waveform magnitude of early, space-based attention when both objects appeared to overlap, whereas similar waveforms appeared in locations previously occupied by objects containing both target attributes. (P1).

