

The neural basis of attentional control in visual search

Martin Eimer

Department of Psychological Sciences, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

How do we localise and identify target objects among distractors in visual scenes? The role of selective attention in visual search has been studied for decades and the outlines of a general processing model are now beginning to emerge. Attentional processes unfold in real time and this review describes four temporally and functionally dissociable stages of attention in visual search (preparation, guidance, selection, and identification). Insights from neuroscientific studies of visual attention suggest that our ability to find target objects in visual search is based on processes that operate at each of these four stages, in close association with working memory and recurrent feedback mechanisms.

Introduction

In visual environments where multiple objects compete for attention, the challenge is to find relevant information and to ignore objects and events that are unrelated to current task goals (Figure 1). Many studies of visual spatial attention (see Glossary) have investigated how prior knowledge about the position of target objects in the visual field facilitates behavioural performance and neural processing [1,2]. However, the fundamental problem for visual search is the absence of precise advance information about target locations. In many laboratory-based visual search tasks, target locations are determined randomly on each trial and are therefore completely unpredictable. In real-world search, attention may benefit from contextual spatial information (e.g., kitchen knives are typically found on kitchen counters), but the exact locations of target objects remain unknown. To understand our ability to find known target objects at uncertain locations, it is useful to consider how attentional processes in visual search operate in real time. Based primarily on neuroscientific studies of attention in the human and monkey brain, this review describes four successive stages of attentional selectivity in visual search (Figure 2). Each of these stages performs a specific function and each is characterised by a particular neural signature. Within this framework, ‘attention’ is not seen as a single functionally and anatomically distinct control system, but as emerging from the coordinated operation of a set of neurocognitive mechanisms in real time.

Corresponding author: Eimer, M. (m.eimer@bbk.ac.uk).

Keywords: visual search; selective attention; top-down control; visual cortex; working memory; recurrent feedback.

1364-6613/

© 2014 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tics.2014.05.005>

Preparation

Search starts by deciding which object or feature to look for and representing this search goal in memory. William James [3] believed that such preparatory ‘images in the mind’ are the single most important aspect of selective attention. Representations of search goals in working memory (attentional templates [4]) can be activated before the relevant visual scene is physically present and are assumed to control subsequent stages of the search process in a goal-directed fashion [5,6].

How are James’s images in the mind implemented at the neural level? The sensory recruitment model of working memory (Box 1) suggests that visual target objects for a search task are represented in posterior visual-perceptual brain regions. There is indeed evidence that visual cortical areas are activated in a goal-selective fashion during the preparation phase of visual search. The activity of neurons in the inferior temporal (IT) cortex that selectively respond to a particular object is enhanced in a sustained fashion while monkeys prepare to find this object in an upcoming search display (Figure 3A) and such preparatory ‘baseline shifts’ of neural activity may reflect an activated attentional template [7]. Similar target-selective preparatory activation patterns have been found in human event-related potential (ERP) and functional MRI (fMRI) experiments [8–14]. Some fMRI studies have demonstrated increased activity in colour- or motion-selective visual areas when observers prepare for targets defined in these dimensions [9–11]. Others have found more globally distributed

Glossary

Attentional template: working memory representation of a current search goal that is activated before search and is assumed to control subsequent attentional guidance and selection processes.

Baseline shift: sustained increase in the baseline activity of neurons that are selective for a task-relevant feature or object during the preparation phase of visual search.

Feature-based attention: allocation of selective attention to specific task-relevant visual features. During the guidance phase of visual search, feature-based attention is reflected by a task-dependent modulation of feature-selective activity in the visual cortex that operates in a spatially global fashion across the visual field.

Position-dependent representation: representation of visual information within a spatial coordinate frame that is defined by the position of a visual stimulus on the retina (retinotopic representation) or in the external world (spatiotopic representation).

Position-invariant representation: visual representation that is not sensitive to the position of the represented stimulus on the retina or in the external world.

Spatial attention: allocation of selective attention to specific locations within the visual field. During the selection phase of visual search, spatial attention is reflected by spatially specific processing enhancements for candidate target objects at particular locations.



Figure 1. Finding target objects in crowded visual environments is a complex achievement. When looking for directions to the Central Line in a busy London Underground station, attention has to be allocated to goal-relevant stimuli (such as the distinctive red colour code of the Central Line or the letter string 'Central') while other visual signals have to be ignored.

goal-sensitive activation patterns in the visual cortex during the preparation of search for specific target shapes [12] or for target object categories in real-world visual scenes (e.g., people, cars, houses, faces [13,14]). The exact locus of preparatory activity patterns within the visual processing hierarchy may depend on the nature of the current search goal [12], with lower-level visual areas responsible for representing simple target features [15] and higher-level regions involved in the preparation for more abstractly defined targets [13].

Although the existence of goal-selective activity modulations in the visual cortex during the preparation phase is well documented, it remains unclear whether these modulations causally affect subsequent stages of visual search. Correlations between the target selectivity of preparatory prestimulus activation patterns in the visual cortex and the quality of subsequent target detection performance [10,12–14,16] suggest that preparatory modulations of visual activity might indeed act as attentional templates that are causally involved in the control of visual search. However, this conclusion is by no means universally accepted [17–19]. When the location of targets is uncertain, target-selective visual activation patterns elicited during the preparation for visual search should represent search goals in a position-independent fashion (Box 1). Although spatially global working memory representations do exist

in the visual cortex [15], the question of whether preparatory goal-selective patterns of visual activity are generally position-invariant and whether this is a necessary requirement for their role as attentional templates remains to be systematically addressed.

In line with the involvement of the prefrontal cortex in visual working memory (Box 1), preparatory activation patterns that are sensitive to current task goals have also been observed in prefrontal areas [13,20,21]. These effects may be linked to top-down aspects of attentional preparation, with the prefrontal cortex controlling target-selective preparatory modulations of visual activity, but could also reflect explicit representations of search goals. Because many object representations in the prefrontal cortex are position-invariant, they could act as location-independent preparatory attentional templates that control subsequent stages of visual search when no precise spatial information about target locations is available.

Guidance

Whereas preparation occurs before the arrival of visual input, guidance and selection operate once a search display has been encountered. Models of visual search [22–24] assume that information about the presence of task-relevant features is accumulated in parallel (guidance) and is then used to control the allocation of spatial attention to

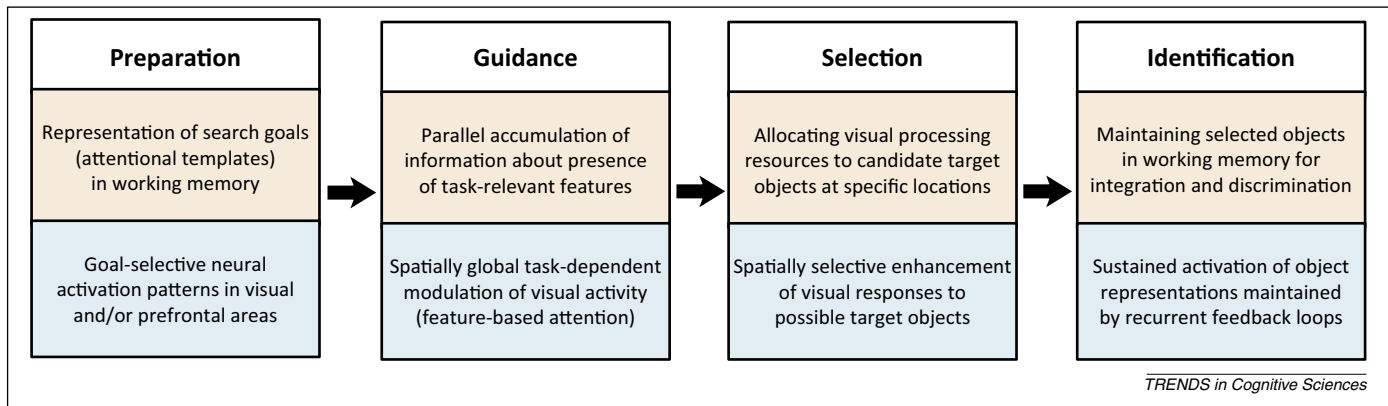


Figure 2. A four-stage model of selective attention in visual search. Preparation, guidance, selection, and identification represent four successive stages of attentional processing. Each of these stages performs a specific cognitive function (red boxes) and each stage is characterised by a particular set of processes at the neural level (blue boxes).

Box 1. The functions of working memory during visual search

Working memory is responsible for the active retention of task-relevant information that is not currently available to sensory perception. Classic accounts of working memory [72] postulate specialised stores for different types of information and a central role of the prefrontal cortex during working memory storage. The sustained activation of prefrontal neurons during memory maintenance [73,74] is in line with this hypothesis. However, it has recently become evident that posterior visual areas are also activated when visual information is memorised [63,75]. In ERP studies, a sustained posterior CDA is observed during visual working memory maintenance and this component is sensitive to memory load and individual differences in working memory capacity [58,76]. The emerging ‘sensory recruitment’ model of visual working memory claims that brain areas involved in visual perception are also the primary locus for the short-term storage of visual information, whereas the prefrontal cortex performs more generic top-down control functions [77,78].

Working memory plays different roles in the preparation and object identification phases of visual search. During preparation, working memory holds a representation of the current search goal (attentional template). During the identification stage, representations of selected objects are maintained in working memory and compared with search goals. These two functions differ markedly in terms of their capacity. Whereas approximately three or four items can be simultaneously maintained in working memory [79], there is evidence that only a single attentional template can be active at any given moment [6,80,81]. This discrepancy suggests qualitative differences between the attentional template and object maintenance functions of visual working memory. Maintenance is based on spatially selective enhancements of object representations in the visual cortex (‘attention directed at internal representations’ [82]) that are sustained by recurrent feedback mechanisms [60]. Individual differences in working memory capacity are therefore closely linked to differences in the ability to select and maintain multiple spatially specific object representations [83]. By contrast, preparatory attentional templates should be position-independent, because target locations are uncertain in visual search, and subsequent feature-based attentional guidance mechanisms operate in parallel across the visual field. Many object representations in prefrontal cortex areas are position-invariant and spatially global representations of memorised visual features also exist in the visual cortex [15]. If attentional templates are based on position-invariant or spatially global representations of search targets, the observation that these templates are strongly capacity-limited suggests that only one position-invariant representation can be active during the preparation stage of visual search.

possible target objects (selection). When target locations are uncertain, guidance processes may operate globally across the entire visual field.

A plausible neural basis of spatially global attentional guidance in visual search has been identified in studies of feature-based attention. When monkeys search for target events defined by a particular feature (such as orientation or motion direction), neurons in visual areas V4 or the middle temporal (MT) cortex that are selective for this task-relevant feature increase their activity, whereas neurons with opposite feature preferences are inhibited [25,26]. Critically, these activity modulations are also elicited in response to stimuli at task-irrelevant unattended locations (Figure 3B), suggesting that feature-based attention is a spatially global phenomenon. During search for colour- or shape-defined target objects, V4 neurons that prefer target-defining features increase their activity when a target object is present in their receptive field, even when monkeys fixate elsewhere, fail to detect the target, and shift gaze to another object [27]. Such observations underline the fact that feature-based attention operates in parallel across the visual field, independent of the current focus of spatial attention. Analogous spatially global feature-based attentional modulations of visual activity have been found in human electrophysiological and fMRI experiments [28–30]. When observers attend to target features in one visual hemifield, objects in the other irrelevant hemifield elicit enhanced visual responses when their features match the features that are currently task relevant.

Because feature-based attentional modulations of visual processing operate in a spatially global fashion, they can provide guidance signals for the subsequent allocation of spatial attention to candidate target objects. Models of visual attention [18,24] postulate that during the parallel analysis of visual input, processing is selectively weighted in favour of target-defining features. The neural mechanisms of feature-based attention can implement such task-dependent attentional biases in a spatially global fashion across the visual field. Feature-based attentional guidance is not necessarily limited to simple visual attributes such as orientation, colour, or movement direction, but can also operate in a spatially global fashion when search goals are

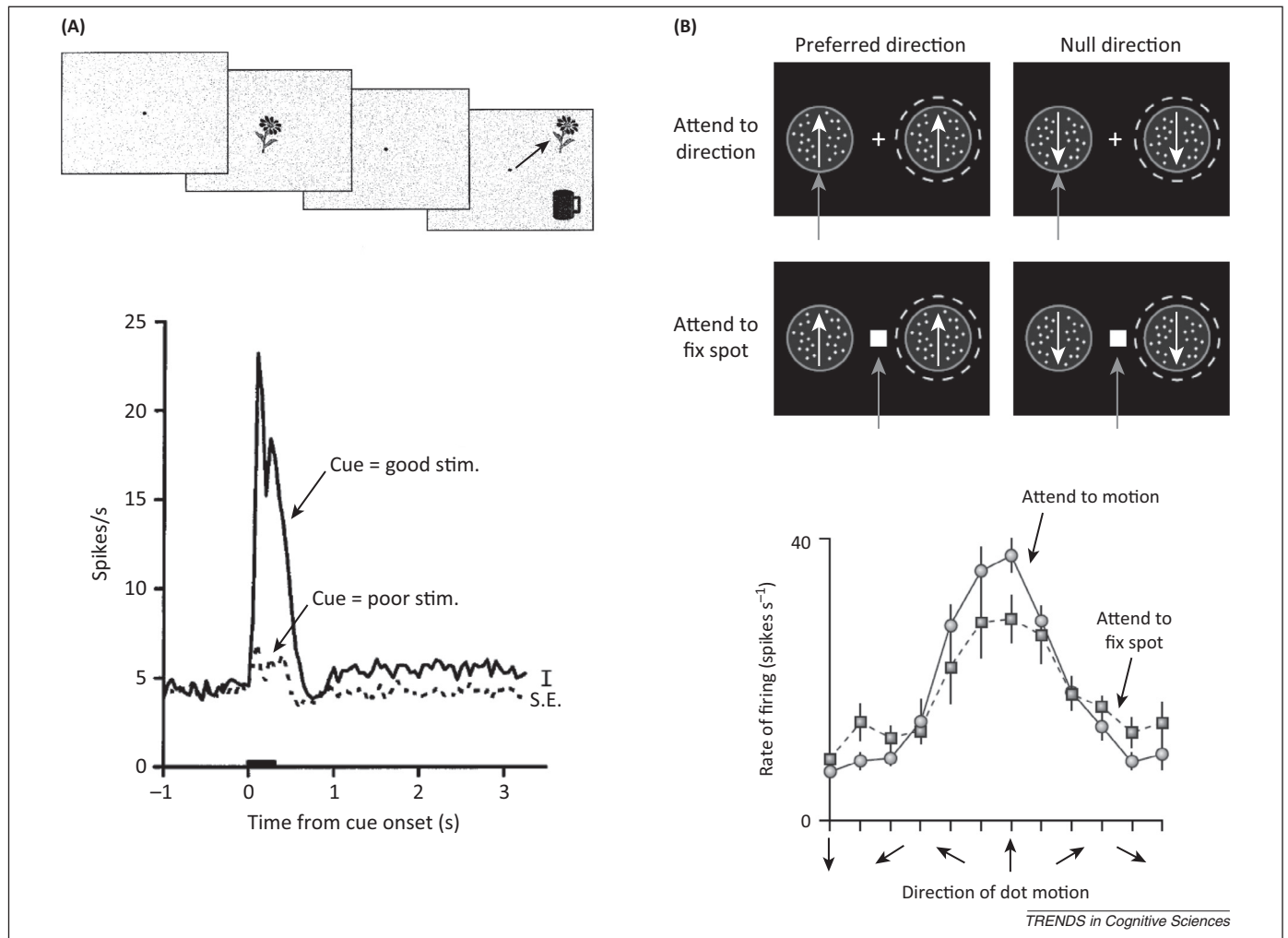


Figure 3. Neural correlates of preparation (A) and guidance (B) during visual search. (A) Stimulus set-up used by Chelazzi *et al.* [7] during a memory-guided visual-search task (top panel) and neural responses recorded in monkey inferior temporal cortex (IT) during the preparation phase (bottom panel). Monkeys remembered a target object presented at the start of each trial during a delay period and moved their eyes to the location of this target in the subsequent search display. When the target object was an effective stimulus for a particular IT neuron, an initial transient sensory response was followed by a sustained increase (baseline shift) of neuronal activity throughout the preparation period. Reproduced, with permission, from [7]. (B) Top: Stimulus set-up used by Martinez-Trujillo and Treue [25]. Monkeys attended to the direction of moving dots in one visual hemifield and ignored another set of dots in the opposite hemifield that moved in the same direction ('attend to direction'). In a baseline condition, they attended to the central fixation spot and ignored both dot arrays ('attend to fix spot'). Bottom: Response of a middle temporal cortex (MT) neuron with a receptive field on the unattended dot array. Attention to a specific movement direction in the opposite visual field increased the neural response to the unattended movement when it matched the preferred movement of this neuron, whereas attention to the opposite ('null') direction produced inhibition. Reproduced from [25], as printed in [33].

more abstractly defined, such as during search for target objects from a particular target category in real-world visual scenes [31]. If guidance depends on the mechanisms of feature-based attention, the fact that visual features differ considerably in their ability to guide spatial attention [32] could be related to systematic differences in the ability of these features to produce task-dependent spatially global modulations of visual activity during the guidance phase of visual search.

Feature-based attentional guidance mechanisms are likely to be closely linked to the processes that operate during the preceding preparation stage. If preparatory goal-selective baseline shifts of visual activity operate in a position-independent fashion, they could be directly responsible for the emergence of spatially global feature-based attention effects during the parallel processing of visual input. For example, spatially global working memory representations of target features [15] may remain

active after search display onset, resulting in feature-selective modulations of sensory responses across the visual field. The observation that feature-based attention effects can spread to currently empty regions of visual space [30] suggests that preparation and guidance might interact in this way. In addition, top-down signals from position-invariant representations of search targets in the prefrontal cortex to visual areas may also play a role in the control of feature-based attention [33]. Despite such close functional links between preparation and guidance, these two stages are not only temporally distinct (i.e., one operates before and the other after the onset of visual stimulation) but are also functionally dissociable. For example, preparation is not followed by feature-selective attentional guidance processes when search displays without any target-matching features are encountered. The reverse scenario (guidance without corresponding preparation) is realised when attention is guided towards salient visual

events irrespective of or even contrary to current selection intentions [34].

Selection

Feature-based attentional guidance highlights the presence of target-defining features and this information can then be employed to select candidate target objects. Because representations in the visual cortex are position-dependent [35], objects compete for representational space in cortical maps [36]. In this context, 'selection' can be defined as the emergence of spatially specific biases in favour of one or more objects at particular locations within these maps. The transition from guidance to selection is therefore marked by the transition from feature-selective activation patterns that are triggered in a spatially global fashion across the visual field to spatially specific modulations of neural responses to potentially task-relevant objects. In contrast to the common assumption that spatial attention is generally faster than attention for features [37,38], feature-based attention should precede spatial

attention during visual search when the location of target objects is not known in advance [39].

How does information accumulated during the parallel guidance phase control the subsequent spatial selection of target objects? The biased competition model of visual attention [5] assumes that feature-selective attentional biases generated at particular levels of the visual processing hierarchy during the guidance phase trigger competitive advantages for possible target objects and that these are then propagated in a spatially selective fashion to lower and higher levels. Other models [18,40] postulate dedicated 'source' areas of attentional control where task-relevant locations are represented, 'sites' of spatially selective processing in visual areas, and recurrent pathways (Box 2) from source to site regions. In these models, the allocation of spatial attention is controlled by priority maps [40,41] in the posterior parietal cortex [42], the frontal eye fields (FEFs) [43], or the thalamus [18]. Priority maps send recurrent top-down control signals to the visual cortex, where they elicit spatially specific enhancements of visual responses to possible target objects. The observation that electrical stimulation of the FEFs triggers activity in spatially corresponding regions of visual area V4 [44] suggests that these two areas are indeed causally linked. An important and controversial

Box 2. The role of recurrent feedback processes in visual search

Visual processing does not operate in a strictly hierarchical bottom-up fashion, with lower-level visual areas representing simple features at specific locations projecting to neurons at higher levels that progressively encode more complex properties of visual objects in a position-invariant fashion. Such feedforward connections are accompanied by parallel feedback projections from higher to lower levels of the visual processing hierarchy [35]. This recurrent architecture allows top-down control signals to modulate visual processing in a flexible task-dependent fashion [84]. According to the Reverse Hierarchy Theory [85], visual input is transmitted in parallel to a high level of processing where complex visual properties and object categories are represented. Information about the presence of potentially task-relevant objects is then fed back to lower visual areas where a detailed analysis of these objects occurs.

Recurrent feedback plays a central role in the object selection and identification stages of visual search [86,87]. During identification, recurrent feedback loops between higher-level control regions and the visual cortex are responsible for the maintenance of spatially selective representations of possible target objects in visual working memory [59,60]. During the preceding selection stage, the initial transient activation of these representations is triggered by recurrent signals from control regions such as the FEFs [43,48] where the locations of task-relevant features are represented in priority maps [41]. The importance of recurrent feedback connections is illustrated by the phenomenon of object substitution masking [88]. Target detection is strongly impaired when targets are immediately replaced by another visual stimulus and therefore can no longer be addressed by recurrent feedback signals. The existence of fast recurrent pathways from category-selective areas to the visual cortex [85] can also explain why category-guided object selection is often remarkably rapid [89].

If high-level control regions (such as the prefrontal cortex) represent task-relevant information in a position-invariant fashion, how can recurrent signals originating from these regions produce spatially selective enhancements of position-dependent object representations in the visual cortex? The existence of parallel and reciprocal feedforward and feedback visual pathways [35,84,85] offers a possible solution to this problem. If only those pathways that were active during the feedforward transmission of task-relevant information to higher areas mediate the flow of recurrent signals to lower areas, information about candidate target objects at particular locations in the visual field can be effectively routed back to spatially corresponding regions of the visual cortex [18].

Box 3. Serial and parallel object selection in visual search

The transition from spatially global guidance to spatially focal selection could coincide with the transition from parallel to serial attentional processing in visual search. In line with this assumption, several models of visual search claim that candidate target objects are selected in a strictly serial fashion. According to Feature Integration Theory [22,71], spatial attention is allocated serially to one object at a time, such that the attentional selection of a new object is always preceded by deallocation of attention from its previous location. In Guided Search [24], guidance and object identification are modelled as parallel processes, but object selection is described as a serial attentional bottleneck. In contrast to such serial-selection accounts, other models of visual attention assume that object selection can operate in parallel at multiple locations in the visual field [5,18]. Along similar lines, the ability to track multiple moving objects has been explained by assuming that spatial attention can be allocated independently and in parallel to different objects in the visual field [90], which has been supported by ERP studies of multiple object tracking [91,92].

Although serial and parallel selection scenarios are often regarded as mutually exclusive accounts of visual search, it is possible that these two types of attentional selection are employed in different task contexts. This is illustrated by two ERP studies [93,94] that used the N2pc component to distinguish serial and parallel selection mechanisms in visual search (see Figure 4 in main text). Woodman and Luck [93] obtained evidence that targets are selected in a serial fashion when observers search for difficult-to-discriminate target objects in crowded visual search displays. By contrast, Eimer and Grubert [94] found that multiple targets are selected in a parallel and independent fashion when the target identification task (letter/digit discrimination) is highly practiced and search displays contain few competing distractor items. The fact that these two N2pc studies draw opposite conclusions about the serial versus parallel nature of visual search suggests a more ecumenical view of attentional object selection. Parallel and serial selection could both be available options in visual search, with the choice between these selection strategies determined by the nature of a particular search task. Serial selection may be the preferred option when a task imposes high demands on object selection and identification mechanisms, whereas a parallel strategy is chosen under conditions where selection and identification requirements are less challenging.

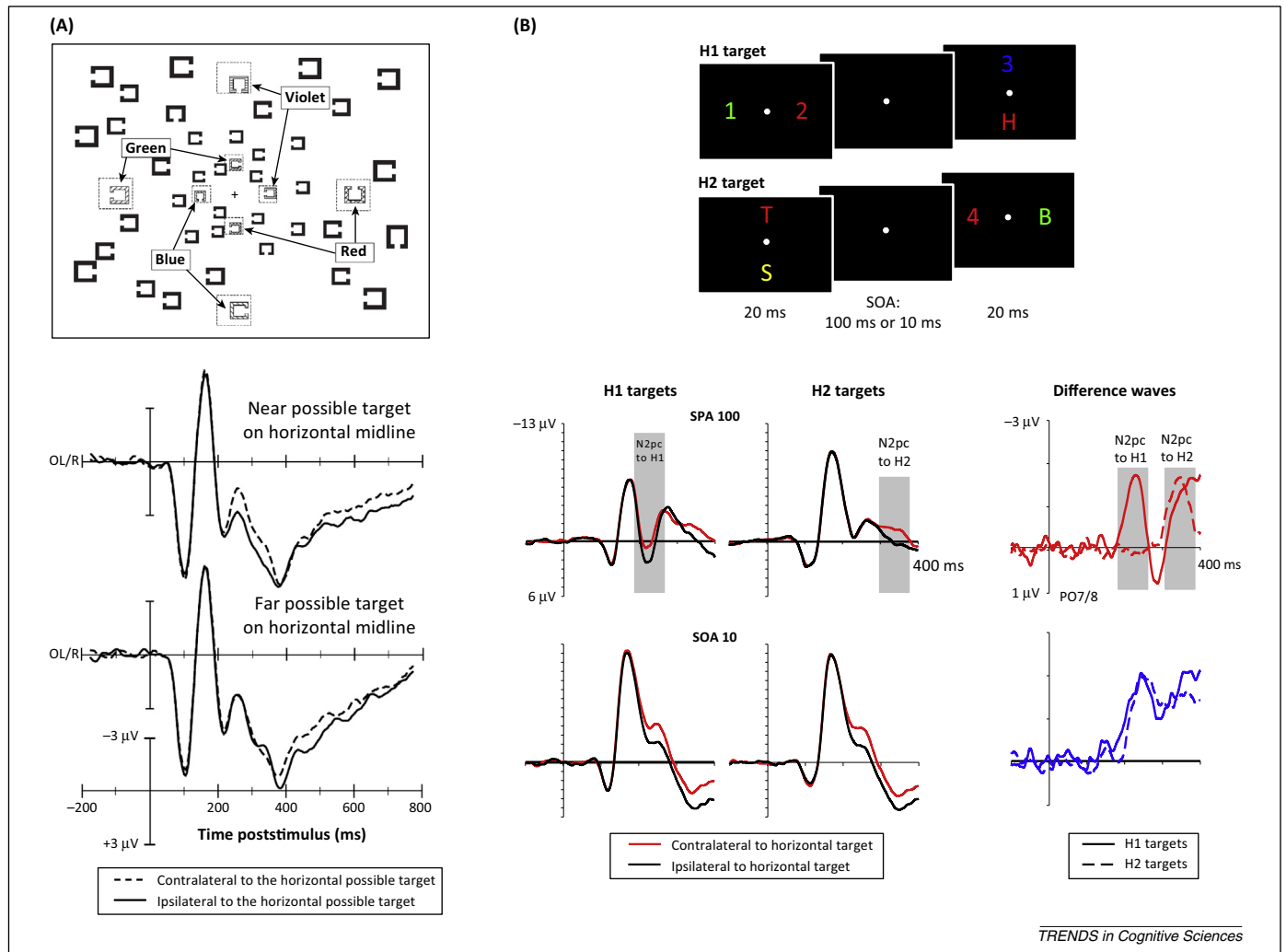


Figure 4. Electrophysiological evidence for serial and parallel object selection in visual search. **(A)** Top: Woodman and Luck [93] instructed participants to search for targets defined by a particular colour (e.g., red) and shape (gap at the top). Each search array contained two target-colour objects. One of these possible targets appeared on the vertical meridian and the other on the horizontal meridian, either near or far from fixation. Because the N2pc is a contralateral event-related potential (ERP) component, it is not triggered by possible targets on the vertical midline and therefore provides a pure measure of the attentional selection of the other (horizontal) target. Bottom: The N2pc measured at lateral posterior electrodes to near possible targets on the horizontal midline emerged 200 ms after stimulus onset and preceded the N2pc to far possible targets by 150 ms. There was no temporal overlap between these two N2pc components, indicating that spatial attention was allocated serially to first the near and then the far target-colour object. Reproduced, with permission, from [93]. **(B)** Top: In the Eimer and Grubert study [94], two search displays that contained a colour-defined (red) target and a distractor on opposite sides were presented in rapid succession with a stimulus onset asynchrony (SOA) of 100 ms or 10 ms. One display contained a horizontal target and the other a target on the vertical midline. Bottom: ERPs at lateral posterior electrodes contralateral and ipsilateral to the horizontal target and N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs. When the SOA between the two targets was 100 ms, the N2pc to a horizontal target in the second array (H2) emerged 100 ms after the N2pc to a horizontal target in the first array (H1). When the SOA was reduced to 10 ms, these two N2pc components were triggered within 10 ms of each other and overlapped in time. These results show that two target objects can be selected in parallel, with each selection process following its own independent time course. Reproduced, with permission, from [94].

question is whether recurrent signals from priority maps to visual areas always address one particular location at a time or whether multiple locations can be addressed simultaneously. Does object selection in visual search operate serially or in a parallel fashion (Box 3)?

The spatial selection of candidate target objects is reflected by enhanced neural responses in the ventral visual cortex that start around 150–200 ms post-stimulus [7,45]. In ERP studies of visual search, target selection is marked by the emergence of the N2pc component at around 180 ms after search display onset. The N2pc is an enhanced negativity at posterior electrodes contralateral to candidate target objects in visual search displays [46,47]. This component is generated during the spatially selective

enhancement of target processing in the ventral visual cortex [39], which is controlled by recurrent signals from higher-level attentional control areas such as the FEFs [48]. Because it tracks the operation of selective attention on a millisecond-by-millisecond basis, the N2pc can provide unique insights into the time course of attentional object selection in visual search (Box 3 and Figure 4).

Identification

The emergence of spatially specific modulations of target processing in the visual cortex does not imply that selected objects are instantly recognised. When spatial attention is employed to track multiple moving objects, access to the features and identity of these objects is remarkably poor

[49], indicating that selection and identification are independent processes [50]. Many models of visual attention distinguish between a selection stage where task-relevant objects are individuated in a spatially specific fashion and a subsequent identification stage where the features of these objects are integrated and object identity becomes accessible to awareness and action control [24,51,52]. The independence of selection and identification is underlined by their sensitivity to different factors. Selection efficiency is determined by the similarity between targets and

competing distractors [4]. By contrast, the efficiency of identification depends on the complexity of target objects [36,51]. Performance impairments observed in tasks where spatial selection demands are minimal and several target objects are presented simultaneously or in rapid succession [53–55] also indicate capacity limitations at an identification stage that follows object selection.

Selection and identification are associated with distinct ERP markers in visual search (Figure 5). The N2pc component reflects target selection, and a later sustained

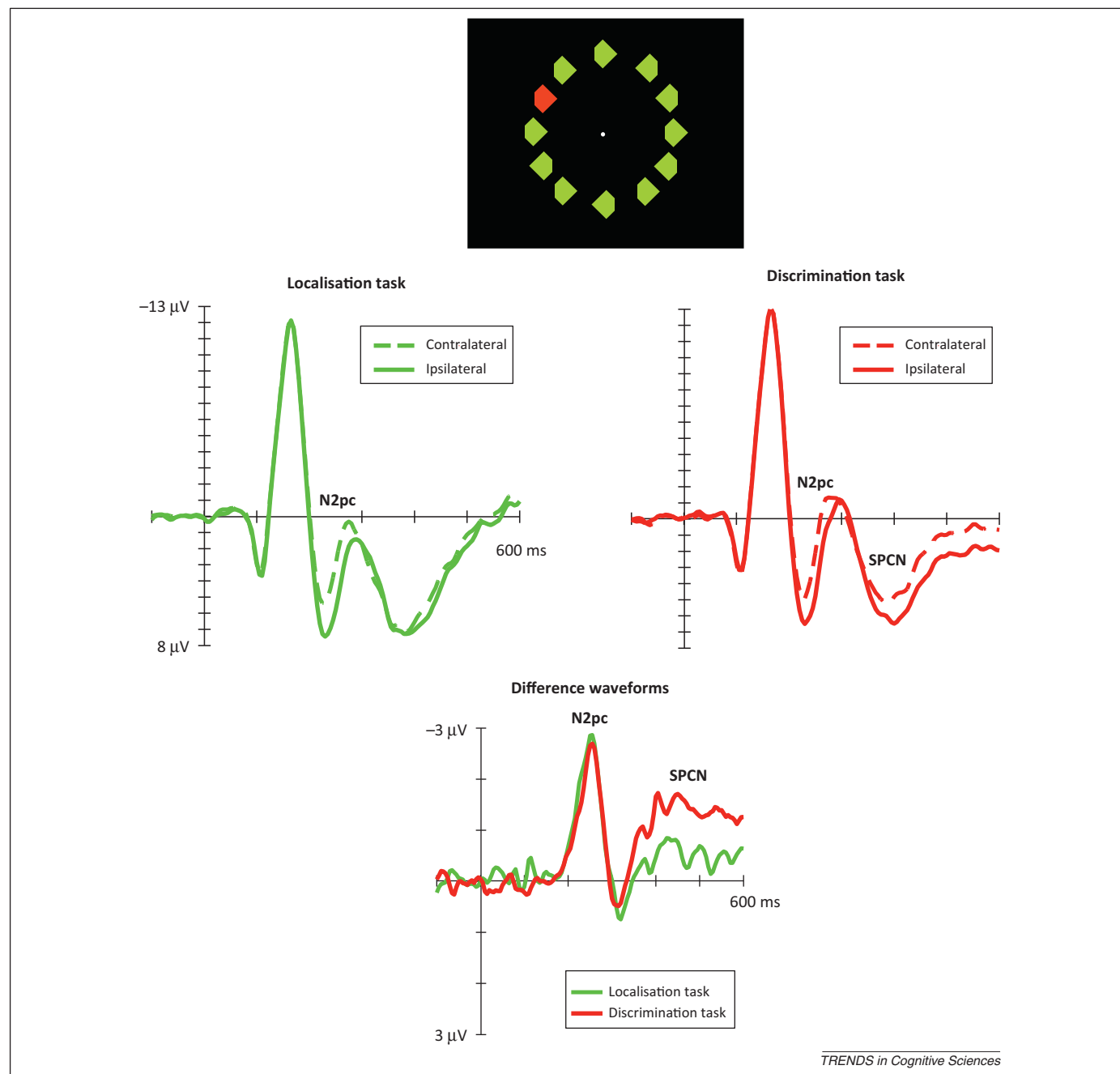


Figure 5. Electrophysiological correlates of successive selection and identification stages in visual search. In the study by Mazza *et al.* [56], search displays containing a colour singleton diamond among uniformly coloured distractors were presented for 150 ms (top panel). Participants had to report either the location of this singleton (localisation task) or its detailed shape (cut on the left or right side; discrimination task). ERP waveforms measured at posterior electrodes contralateral and ipsilateral to the target and corresponding contralateral–ipsilateral N2pc difference waveforms (bottom panels) demonstrate that N2pc components (reflecting target selection) were identical in the two tasks. By contrast, the subsequent sustained posterior contralateral negativity (SPCN) that is generated during the identification stage when selected objects are maintained in working memory [58] was reliably triggered only in the shape discrimination task, demonstrating that selection and identification are separable stages of visual search. Data from [56], reproduced in a different format.

posterior contralateral negativity is elicited during the subsequent object identification stage [56,57]. This sustained negativity is equivalent to the contralateral delay activity (CDA) component observed during the delay period of working memory tasks [58] and its presence during the identification of target objects in visual search demonstrates the involvement of visual working memory [18,59]. Working memory is required for object identification because spatially specific enhancements of target processing that emerge during the preceding selection stage remain transient unless they are sustained by recurrent input from higher-order control areas to the visual cortex [18]. Such sustained feedback loops could represent the neural basis of working memory maintenance [60]. They may be critical for integrating features into object files [61] and for matching perceptual representations of selected objects with stored representations of search targets during the identification stage. Visual areas such as the superior intraparietal sulcus and lateral occipital complex that are sensitive to the number of memorised objects and their complexity [62,63] are likely to be involved in the maintenance of possible target objects in visual search.

Concluding remarks

Our ability to find task-relevant objects in visual scenes depends on attentional processes that unfold in real time. In this review, preparation, guidance, selection, and identification are described as temporally and functionally distinct stages of visual search. This four-stage model can be useful to interpret psychological and neuroscientific findings within a general processing framework of visual search and to clarify the roles of working memory and recurrent feedback mechanisms at different stages of the search process.

This model can provide only a basic outline of the attentional control processes that are active during visual search in complex real-world scenes and needs to be qualified in several important respects. Although Figure 2 may suggest that the four stages are organised in a strictly sequential fashion, it is possible that at least some of these stages overlap in time (Box 4). For example,

the identification of selected objects may be a relatively slow process that could operate in parallel with the attentional selection of other candidate target objects [24]. Identification involves the comparison of visual object representations with representations of current search goals (Box 1), which implies that attentional templates in working memory that are set up during the preparation stage remain active throughout the search process. Furthermore, when targets are not detected on the basis of a single guidance/selection/identification cycle, search will become iterative, with each new iteration of this cycle initiated by a mismatch between selected object representations and current search goals at the identification stage. Search is likely to be based on complex interactions between serial and parallel mechanisms and the simple serial architecture illustrated in Figure 2 is not intended to be a fully realistic representation of these mechanisms. However, because visual search unfolds in real time, and because each of the stages described here depends on the output of operations that occur at preceding stages, the attentional control processes that contribute to successful search performance retain important serial characteristics, which were highlighted in this review.

Another important aspect of visual search that was not discussed here concerns the role of eye movements. Spatial attention and saccade programming are known to be closely linked [41,64]. Selection does not only control the access of task-relevant objects to working memory, but also provides spatial coordinates for upcoming eye movements. Research on saccades and microsaccades [65–67] has provided important insights into the control of selective attention during search in real-world visual scenes. For example, semantic and spatial expectations linked to particular scene contexts strongly constrain which parts of a scene will be visually examined [65,68], which demonstrates that high-level world knowledge plays an important role in attentional guidance and selection. The implicit acquisition of contextual information can guide spatial attention even in simple search displays [69]. Such observations show that visual search does not always operate in the complete absence of prior information about likely target locations and raise important questions for the control of attentional guidance and selection processes. Instead of operating in a spatially unconstrained global fashion across the entire visual field, these processes may sometimes be confined to spatially restricted attentional windows [34] that are linked to context-dependent expectations about target locations. If this is the case, the concept of preparatory attentional templates may have to be extended to include representations not only of target features or objects, but also of scene contexts and likely target positions.

Visual search performance varies greatly across task contexts [70]. It is unlikely that such differences can all be attributed to one particular stage (such as serial selection [71]). In this review, search is described as a process that unfolds in real time and involves successive attentional mechanisms at multiple stages of processing. The efficiency of visual search for known targets at uncertain locations is determined by the complex interplay of all of these mechanisms.

Box 4. Outstanding questions

- How do goal-selective modulations of neural activity during the preparation stage affect attentional processes at subsequent stages of visual search? Do preparatory baseline shifts play a causal role for the selection and identification of search targets?
- How are spatially global feature-based attention effects that emerge during the guidance stage of visual search set up and controlled?
- Are both serial and parallel selection strategies available for the control of visual search? Which factors determine whether the spatial selection of candidate target objects operates in a serial or parallel fashion?
- Do preparation, guidance, selection, and identification operate in a sequential fashion during visual search or can some of these stages be activated in parallel?
- Can other processing bottlenecks identified in the attention literature be linked to the current four-stage model of selective attention? For example, are dual-task interference effects in psychological refractory period (PRP) experiments, which have been attributed to capacity limitations at a central response selection stage [95], generated during the object identification stage described here?

Acknowledgements

This work was supported by the Economic and Social Research Council (ESRC), UK. The author thanks Anna Grubert and John Towler for comments and Joanna Parketny for creating Figure 1.

References

- Posner, M.I. *et al.* (1980) Attention and the detection of signals. *J. Exp. Psychol. Gen.* 109, 160–174
- Moran, J. and Desimone, R. (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229, 782–784
- James, W. (1890) *The Principles of Psychology*, Harvard University Press
- Duncan, J. and Humphreys, G.W. (1989) Visual search and stimulus similarity. *Psychol. Rev.* 96, 433
- Desimone, R. and Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222
- Olivers, C.N.L. *et al.* (2011) Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15, 327–334
- Chelazzi, L. *et al.* (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* 80, 2918–2940
- Carlisle, N.B. *et al.* (2011) Attentional templates in visual working memory. *J. Neurosci.* 31, 9315–9322
- Chawla, D. *et al.* (1999) The physiological basis of attentional modulation in extrastriate visual areas. *Nat. Neurosci.* 2, 671–676
- Giesbrecht, B. *et al.* (2006) Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Res.* 1080, 63–72
- Shibata, K. *et al.* (2008) The effects of feature attention on prestimulus cortical activity in the human visual system. *Cereb. Cortex* 18, 1664–1675
- Stokes, M. *et al.* (2009) Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19569–19574
- Peelen, M.V. and Kastner, S. (2011) A neural basis for real-world visual search in human occipitotemporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12125–12130
- Soon, C.S. *et al.* (2013) Preparatory patterns of neural activity predict visual category search speed. *Neuroimage* 66, 215–222
- Ester, E.F. *et al.* (2009) Spatially global representations in human primary visual cortex during working memory maintenance. *J. Neurosci.* 29, 15258–15265
- Puri, A.M. *et al.* (2009) Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res.* 1301, 89–99
- Fannon, S.P. *et al.* (2007) Baseline shifts do not predict attentional modulation of target processing during feature-based visual attention. *Front. Hum. Neurosci.* 2, 7
- Bundesen, C. *et al.* (2005) A neural theory of visual attention: bridging cognition and neurophysiology. *Psychol. Rev.* 112, 291
- McMains, S.A. *et al.* (2007) Mechanisms of feature- and space-based attention: response modulation and baseline increases. *J. Neurophysiol.* 98, 2110
- Warden, M.R. and Miller, E.K. (2010) Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30, 15801–15810
- Stokes, M.G. *et al.* (2013) Dynamic coding for cognitive control in prefrontal cortex. *Neuron* 78, 364–375
- Treisman, A. and Sato, S. (1990) Conjunction search revisited. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 459
- Wolfe, J.M. (1994) Guided Search 2.0: a revised model of visual search. *Psychon. Bull. Rev.* 1, 202–238
- Wolfe, J.M. (2007) Guided Search 4.0: Current Progress with a model of visual search. In *Integrated Models of Cognitive Systems* (Gray, W., ed.), pp. 99–119, Oxford University Press
- Martínez-Trujillo, J.C. and Treue, S. (2004) Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr. Biol.* 14, 744–751
- Cohen, M.R. and Maunsell, J.H. (2011) Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron* 70, 1192–1204
- Bichot, N.P. (2005) Parallel and serial neural mechanisms for visual search in macaque area V4. *Science* 308, 529–534
- Zhang, W. and Luck, S.J. (2009) Feature-based attention modulates feedforward visual processing. *Nat. Neurosci.* 12, 24–25
- Saenz, M. *et al.* (2002) Global effects of feature-based attention in human visual cortex. *Nat. Neurosci.* 5, 631–632
- Serences, J.T. and Boynton, G.M. (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55, 301–312
- Peelen, M.V. *et al.* (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460, 94–97
- Wolfe, J.M. and Horowitz, T.S. (2004) What attributes guide the deployment of visual attention and how do they do it? *Nat. Rev. Neurosci.* 5, 495–501
- Maunsell, J.H. and Treue, S. (2006) Feature-based attention in visual cortex. *Trends Neurosci.* 29, 317–322
- Theeuwes, J. (2010) Top-down and bottom-up control of visual selection. *Acta Psychol.* 135, 77–99
- Kravitz, D.J. *et al.* (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* 17, 26–49
- Franconeri, S.L. *et al.* (2013) Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn. Sci.* 17, 134–141
- Anllo-Vento, L. and Hillyard, S.A. (1996) Selective attention to the color and direction of moving stimuli: electrophysiological correlates of hierarchical feature selection. *Percept. Psychophys.* 58, 191–206
- Liu, T. *et al.* (2007) Comparing the time course and efficacy of spatial and feature-based attention. *Vision Res.* 47, 108–113
- Hopf, J.M. *et al.* (2004) Attention to features precedes attention to locations in visual search: evidence from electromagnetic brain responses in humans. *J. Neurosci.* 24, 1822–1832
- Itti, L. and Koch, C. (2001) Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203
- Fecteau, J.H. and Munoz, D.P. (2006) Salience, relevance, and firing: a priority map for target selection. *Trends Cogn. Sci.* 10, 382–390
- Gottlieb, J.P. *et al.* (1998) The representation of visual salience in monkey parietal cortex. *Nature* 391, 481–484
- Schall, J.D. (2004) On the role of frontal eye field in guiding attention and saccades. *Vision Res.* 44, 1453–1467
- Moore, T. and Armstrong, K.M. (2003) Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421, 370–373
- Chelazzi, L. *et al.* (2001) Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb. Cortex* 11, 761–772
- Luck, S.J. and Hillyard, S.A. (1994) Spatial filtering during visual search: evidence from human electrophysiology. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 1000–1014
- Eimer, M. (1996) The N2pc component as an indicator of attentional selectivity. *Electroencephalogr. Clin. Neurophysiol.* 99, 225–234
- Cohen, J.Y. *et al.* (2009) On the origin of event-related potentials indexing covert attentional selection during visual search. *J. Neurophysiol.* 102, 2375
- Horowitz, T.S. *et al.* (2007) Tracking unique objects. *Percept. Psychophys.* 69, 172–184
- Ghorashi, S. *et al.* (2010) Spatial selection and target identification are separable processes in visual search. *J. Vis.* 10, 7
- Xu, Y. and Chun, M.M. (2009) Selecting and perceiving multiple visual objects. *Trends Cogn. Sci.* 13, 167–174
- Huang, L. and Pashler, H. (2007) A Boolean map theory of visual attention. *Psychol. Rev.* 114, 599
- Duncan, J. (1980) The locus of interference in the perception of simultaneous stimuli. *Psychol. Rev.* 87, 272
- Raymond, J.E. *et al.* (1992) Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849
- Duncan, J. *et al.* (1994) Direct measurement of attentional dwell time in human vision. *Nature* 369, 313–315
- Mazza, V. *et al.* (2007) Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Exp. Brain Res.* 181, 531–536
- Joliceur, P. *et al.* (2008) Dissociation of the N2pc and sustained posterior contralateral negativity in a choice response task. *Brain Res.* 1215, 160–172
- Vogel, E.K. and Machizawa, M.G. (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751

- 59 Chun, M.M. and Johnson, M.K. (2011) Memory: enduring traces of perceptual and reflective attention. *Neuron* 72, 520–535
- 60 Luck, S.J. and Vogel, E.K. (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400
- 61 Kahneman, D. *et al.* (1992) The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219
- 62 Todd, J.J. and Marois, R. (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754
- 63 Xu, Y. and Chun, M.M. (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440, 91–95
- 64 Krauzlis, R.J. *et al.* (2013) Superior colliculus and visual spatial attention. *Annu. Rev. Neurosci.* 36, 165–182
- 65 Henderson, J.M. (2003) Human gaze control during real-world scene perception. *Trends Cogn. Sci.* 7, 498–504
- 66 Rayner, K. (2009) Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol. (Hove)* 62, 1457–1506
- 67 Martinez-Conde, S. *et al.* (2013) The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nat. Rev. Neurosci.* 14, 83–96
- 68 Hollingworth, A. (2009) Two forms of scene memory guide visual search: memory for scene context and memory for the binding of target object to scene location. *Vis. Cogn.* 17, 273–291
- 69 Chun, M.M. (2000) Contextual cueing of visual attention. *Trends Cogn. Sci.* 4, 170–178
- 70 Wolfe, J.M. (1998) What can 1 million trials tell us about visual search? *Psychol. Sci.* 9, 33–39
- 71 Treisman, A. and Gelade, G. (1980) A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136
- 72 Baddeley, A. (2012) Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29
- 73 Fuster, J.M. and Alexander, G.E. (1971) Neuron activity related to short-term memory. *Science* 173, 652–654
- 74 Curtis, C.E. and D'Esposito, M. (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423
- 75 Harrison, S.A. and Tong, F. (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635
- 76 Ikkai, A. *et al.* (2010) Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *J. Neurophysiol.* 103, 1963
- 77 Postle, B.R. (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38
- 78 D'Esposito, M. (2007) From cognitive to neural models of working memory. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 362, 761–772
- 79 Cowan, N. (2001) The magical number four in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114
- 80 Houtkamp, R. and Roelfsema, P.R. (2009) Matching of visual input to only one item at any one time. *Psychol. Res.* 73, 317–326
- 81 Lewis-Peacock, J.A. *et al.* (2012) Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24, 61–79
- 82 Chun, M.M. *et al.* (2011) A taxonomy of external and internal attention. *Annu. Rev. Psychol.* 62, 73–101
- 83 Anderson, D.E. *et al.* (2013) A common discrete resource for visual working memory and visual search. *Psychol. Sci.* 24, 929–938
- 84 Gilbert, C.D. and Li, W. (2013) Top-down influences on visual processing. *Nat. Rev. Neurosci.* 14, 350–363
- 85 Hochstein, S. and Ahissar, M. (2002) View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804
- 86 Di Lollo, V. *et al.* (2001) The preattentive emperor has no clothes: a dynamic redressing. *J. Exp. Psychol. Gen.* 130, 479
- 87 Hamker, F.H. (2005) The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cereb. Cortex* 15, 431–447
- 88 Enns, J.T. and Di Lollo, V. (2000) What's new in visual masking? *Trends Cogn. Sci.* 4, 345–352
- 89 Nako, R. *et al.* (2014) Rapid guidance of visual search by object categories. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 50–60
- 90 Cavanagh, P. and Alvarez, G.A. (2005) Tracking multiple targets with multifocal attention. *Trends Cogn. Sci.* 9, 349–354
- 91 Drew, T. and Vogel, E.K. (2008) Neural measures of individual differences in selecting and tracking multiple moving objects. *J. Neurosci.* 28, 4183–4191
- 92 Drew, T. *et al.* (2009) Attentional enhancement during multiple-object tracking. *Psychon. Bull. Rev.* 16, 411–417
- 93 Woodman, G.F. and Luck, S.J. (2003) Serial deployment of attention during visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 121–138
- 94 Eimer, M. and Grubert, A. (2014) Spatial attention can be allocated rapidly and in parallel to new visual objects. *Curr. Biol.* 24, 193–198
- 95 Pashler, H. (1994) Dual-task interference in simple tasks: data and theory. *Psychol. Bull.* 16, 220–244