

NEURAL MECHANISMS OF TOP-DOWN SELECTION DURING VISUAL SEARCH

Narcisse P. Bichot

Laboratory of Neuropsychology, National Institute of Mental Health,
National Institutes of Health, Bethesda, MD, USA

Abstract– The brain’s representation of visual information depends greatly on the behavioral relevance of the viewed stimuli. While in some instances behavioral significance is derived from conspicuity, in many situations significance depends on top-down factors such as the viewer’s goals and knowledge. Studies combining neural recordings and behavioral observations have begun to elucidate how the brain selects visual stimuli based on top-down information. While many visual areas of the brain that are selective for visual attributes participate in the selection process, the outcome of the selection process across these areas appears to be represented in structures like the frontal eye field, a key stage in the transformation of visual selection into a command to move the eyes. Evidence shows that the frontal eye field exhibits all the characteristics of a visual salience map in which the behavioral significance of stimuli derived from bottom-up and top-down influences is represented. The patterns of neural modulation in structures like the frontal eye field can be used to design more efficient machine-vision algorithms for target selection.

Keywords– Visual selection, attention, eye movements, saccade, cognition, memory, salience map, frontal eye field

INTRODUCTION

Despite our subjective feeling that we “see” everything within our field of vision, not every part of a visual scene is processed to the same degree [1], [2]. Instead, we attend to objects of interest while ignoring irrelevant ones. Influences on visual selection can be broadly divided into bottom-up and top-down categories. While bottom-up influences are derived from the intrinsic conspicuity of features in a visual scene (e.g., a bright stimulus), top-down influences are derived from the goals and knowledge of the viewer. A large body of literature shows that eye movements reflect cognitive processes during various viewing contexts, which include visual search [3], natural scene perception [4], and reading [5].

Neural mechanisms underlying visual selection have been investigated extensively. Correlates of visual selection appear to be reflected in nearly all visual or visual-association areas (for reviews, see [6], [7]), and perhaps even as early as V1 (e.g., [8]). Except for those at the earliest stages of the visual system, most of these areas project to the frontal eye field (FEF) [9], which appears to be one of the highest points of convergence of the dorsal and ventral visual information streams in the brain [10]. In turn, FEF projects strongly to brainstem oculomotor structures [11], and thus is ideally positioned to transform the outcome of visual processing into a command to move the eyes. Neural correlates of bottom-up selection in FEF

have been reviewed [12]. This paper focuses on studies demonstrating selection in FEF based on top-down factors, completing the argument that this area contains a visual salience map in which behavioral significance, regardless of its source, is represented.

TOP-DOWN INFLUENCES DURING FEATURE SEARCH

A. Expectancy

Cognitive strategies can, in some instances, override the effects of conspicuity (e.g., [13]). For example, experts are more likely than novices to ignore conspicuous, but non-informative features of a visual image from their area of expertise (e.g., [14]). Experience and training also influence the search strategy of monkeys, as well as the concomitant neural selection process in FEF [15]. Monkeys trained to make a saccade to the oddball stimulus in complementary color search arrays adopt a strategy of shifting their gaze according to visual conspicuity. In contrast, monkeys trained exclusively with one of the two complementary search arrays adopt a strategy of ignoring stimuli with the distractor feature, even when a stimulus with that feature becomes the oddball in the visual search array (Fig. 1A).

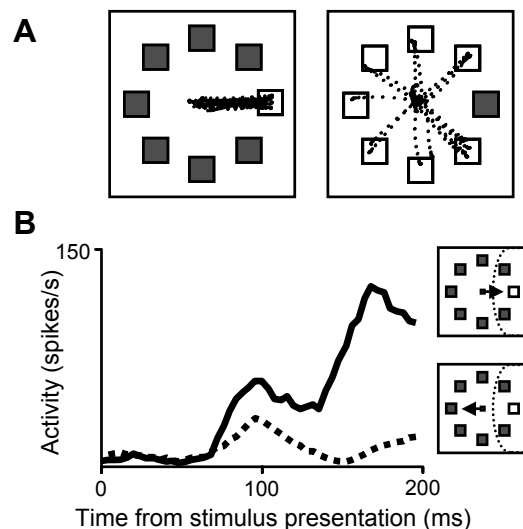


Fig. 1. Effect of feature expectation on gaze behavior and neural modulation in FEF during a color popout search. A. Eye movements of one monkey during search with the learned search array (left) and during viewing of the complementary search array (right). Each dot represents an eye movement sample. B. Response of one FEF visuomovement neuron while this monkey performed search with the learned search array. Spike density functions during correctly performed trials when the target was in the neurons’s response field (solid line) and when distractors were in its response field (dashed line) are shown superimposed. (Modified from [15])

Report Documentation Page

Report Date 25OCT2001	Report Type N/A	Dates Covered (from... to) -
Title and Subtitle Neural Mechanisms of Top-Down Selection During Visual Search	Contract Number	
	Grant Number	
	Program Element Number	
Author(s)	Project Number	
	Task Number	
	Work Unit Number	
Performing Organization Name(s) and Address(es) Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD	Performing Organization Report Number	
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	Sponsor/Monitor's Acronym(s)	
	Sponsor/Monitor's Report Number(s)	
Distribution/Availability Statement Approved for public release, distribution unlimited		
Supplementary Notes Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 25-28 Oct 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.		
Abstract		
Subject Terms		
Report Classification unclassified	Classification of this page unclassified	
Classification of Abstract unclassified	Limitation of Abstract UU	
Number of Pages 4		

In monkeys using the strategy of searching for the learned target color, about half of FEF visuomovement neurons discriminated the target from distractors as soon as they responded (Fig. 1B). In other words, FEF neurons exhibited an apparent feature selectivity in their initial response unlike what had been observed before in this area [16]. Furthermore, this selectivity manifested itself as a suppression of the response to the learned distractor, rather than an enhancement of the response to the learned target. This finding is consistent with the suppression of distractor information during color popout visual search determined psychophysically with human observers [17].

In the study mentioned above, the search process was affected by expectation of a non-spatial stimulus property (i.e., color). Recently, recordings in superior colliculus have shown that the initial activity of buildup neurons in this structure is modulated by the expectation of target location [18]. Similarly, reward (or gain) expectancy has been shown to influence the initial response of neurons in the lateral intraparietal area (LIP) [19]. Both the superior colliculus and area LIP have also been surmised to contain a visual salience map [20], [21].

B. Singleton distractors

Experiments with human observers have shown that when searching for an oddball stimulus in one feature dimension, it is impossible to ignore an oddball stimulus in another, irrelevant feature dimension [22]. Similarly, during a shape feature search, monkeys are adversely affected by the presence of a distractor that differs from all other stimuli in color [23] (Fig. 2A). The presence of the color singleton distractor results in longer saccade latencies to the target, as well as an increase in error rates (Fig. 2B). Furthermore, when monkeys make an error during search with a singleton distractor, most of the errors are accounted for by saccades to the singleton distractor.

Neural recordings in FEF during shape feature search with a singleton distractor show that selection based on both conspicuity and top-down guidance are reflected in this area [23] (Fig. 2C). Similar to previous observations with monkeys that are not biased towards searching for a particular shape or color, the sampled population of FEF visuomovement neurons did not initially discriminate the target from distractors. When the activity of these neurons became selective, the target (i.e., oddball shape) elicited the highest activation, despite not being the most conspicuous stimulus in the display. Furthermore, the color singleton distractor elicited a higher activation than the other, non-salient distractors, presumably due to the strong attentional capture by highly conspicuous stimuli. Thus, FEF in this task reflects both the top-down goal of the search (i.e., the stimulus with the oddball shape), as well as the bottom-up attentional capture by the irrelevant singleton.

While items that differ in one or more features from neighboring items draw attention, items that appear as sudden-onsets capture attention even more strongly [24].

A recent study in area LIP has described the neural correlates of the attentional capture by sudden-onsets [25]. In this study, stimuli were either presented in the receptive field of area LIP neurons as sudden-onsets, or were brought into their receptive field by a saccade. While neurons in this area responded vigorously to the sudden appearance of stimuli in their receptive field, the same neurons exhibited weak or no responses to stable stimuli brought into their receptive field by a saccade. Thus, the strong initial responses of area LIP neurons to stimuli that are presented in their receptive field appear to reflect the strong attentional capture by sudden-onsets, rather than a passive sensory response to stimuli in their receptive field. Further research is needed to determine whether the strong, seemingly visual responses commonly observed in FEF are also due to attentional capture by stimuli appearing suddenly. A study in which visual responses were recorded in FEF while monkeys freely scanned natural images supports this possibility [26]. However, even if this were the case, the results of recordings in the FEF of monkeys exhibiting a bias towards a particular feature [15] (Fig. 1) suggest that the magnitude of the attentional capture by sudden-onsets can be modulated by top-down influences.

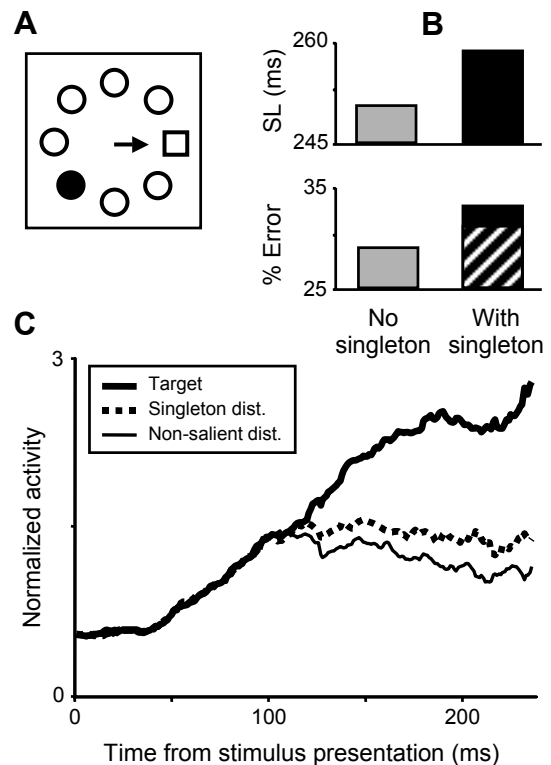


Fig. 2. Effect of a color singleton distractor on performance and neural modulation in FEF during a shape feature search. A. Task display. The arrow illustrates the saccade to the target (i.e., oddball shape). B. Mean saccade latency and error rate during the shape search task without the color singleton (gray bars) and with the color singleton (black bars). The proportion of saccades to the color singleton in the “with singleton” search condition are represented by the white stripes. C. Pooled normalized response of a population of FEF visuomovement neurons during correctly performed feature search trials with a singleton distractor. Responses to the target (thick solid line), to the singleton distractor (dashed line), and to the non-salient distractors (thin solid line) are shown superimposed.

CONJUNCTION SEARCH

In many real-world situations, an object of interest cannot be located based on conspicuity, and a memory of the object is required to locate it (e.g., “searching for a face in the crowd”). An analogous situation is obtained during conjunction search where the target is defined by one combination of possible features (e.g., a red cross), and distractors are formed by other possible combinations (e.g., a green cross, a red circle, or a green circle) (Fig. 3A). Early experiments reported a powerful dichotomy between feature and conjunction search [27], and played an important role in the development of theories of visual attention. Based mainly on measures of reaction time, feature search appeared to be effortless, while conjunction search appeared to be effortful and attentionally-demanding as evidenced by an increase in the time to detect the presence of the target as the number of distractors in the display was increased. However, later experiments showed that conjunction search can be performed efficiently [28], [29], and suggested that the search display can be processed in parallel to identify stimuli with the desired features.

To the extent that attention and eye movements are functionally related, such a parallel search strategy predicts that subjects would be more likely to shift gaze to a distractor that shares a target feature (i.e., “similar distractor”) than to a distractor that has no features in common with the target (i.e., “opposite distractor”). This prediction has been confirmed with both humans [30] and monkeys [31] searching for a target defined by the combination of color and shape (Fig. 3B). Interestingly, when target properties remained the same within an experimental session, but changed across experimental sessions, there was evidence that the history of target properties affected behavior [31]. In addition to being influenced by visual similarity to the target, errant saccades tended to land on the distractor that was the search target during the previous session (Fig. 3B). This tendency manifested itself across sessions at least a day apart and persisted throughout a session. Although of much longer timecourse, this phenomenon may be related to the perceptual priming observed during popout search with both human [32] and monkey [31] subjects, and was thus referred to as long-term priming.

Recordings in FEF during conjunction search further support the hypothesis that this area represents a visual salience map [33]. After an initially non-selective response, FEF neurons not only discriminated the target from distractors, but also discriminated among the distractors based on their visual similarity to the target and the history of target properties across sessions (Fig. 3C). In other words, while the highest activation was associated with the target, distractors similar to the target elicited stronger responses than the distractor that shared no target features. Furthermore, a similar distractor primed by virtue of being the target of the previous session elicited a stronger response than an unprimed similar distractor. In recording sessions

during which the opposite distractor was primed (not shown), there was a relative increase in its neural representation, although it was still weaker than that of distractors similar to the target. This observation is consistent with the fact that, in those sessions, there was a relative increase in erroneous saccades to the opposite distractor, although these saccades were not as frequent as those to distractors similar to the target. Thus, this study shows that neural modulation in FEF reflects a variety of top-down influences, and predicts gaze patterns during a complex visual search.

Neural selection has also been investigated in area LIP during a complex visual search in which no stimulus was conspicuous [25]. In this study mentioned earlier, neurons did not automatically respond to the entry of stable stimuli into their receptive field after a saccade. However, when the stimulus in the receptive field was the cued search target, neurons responded to signal its presence. This study suggests that the representation of visual information in area LIP is normally sparse, with only behaviorally relevant stimuli being strongly represented.

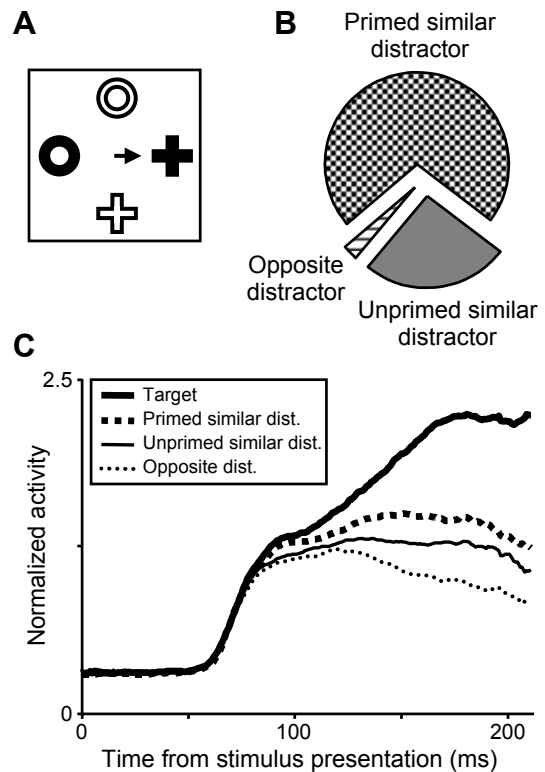


Fig. 3. Gaze behavior and neural modulation in FEF during conjunction search. A. Conjunction search display. The arrow illustrates the saccade to the target. B. Incidence of saccades to distractors representing the probability of shifting gaze to each of the distractor types during error trials. A similar distractor refers to a distractor that shares a target feature (i.e., same color or same shape); a primed distractor refers to a distractor that was the search target during the previous session. C. Pooled normalized response of a population of FEF visuomovement neurons during correctly performed conjunction search trials (i.e., initial saccade was made to the target). The response to the target (thick solid line), to the primed similar distractor (thick dashed line), to the unprimed similar distractor (thin solid line), and to the opposite distractor (thin dotted line) are shown superimposed.

CONCLUSION

Since its proposal by Koch and Ullman [34] nearly two decades ago, the concept of a visual salience map has figured prominently in most models of covert attention and overt saccade production. Evidence strongly supports the view that FEF contains such a salience map. Neurons in this area are not selective for visual attributes, but represent the behavioral relevance of stimuli derived from bottom-up influences, as well as top-down influences involving goals and knowledge. Neural modulation in FEF accurately predicts the gaze patterns observed in a variety of search tasks that involve different degrees of bottom-up and top-down influences. Furthermore, a recent study shows that small populations of FEF neurons can account for reaction time and error rate across a wide range of search efficiency [35]. Thus, FEF represents an ideal area in which to test theories on the mechanisms of visual selection, as well as response production. The mechanisms of visual selection observed in FEF, which appear to reflect processing throughout the visual system, can be used to design machine-vision systems that operate with an efficiency and reliability closer to that of the primate visuomotor system.

ACKNOWLEDGMENT

Thanks to Dr. K.G. Thompson for valuable comments, and to Drs. S.C. Rao, J.D. Schall, and K.G. Thompson for their contributions to this work.

REFERENCES

- [1] R. Rensink, "Change blindness: implications for the role of attention in scene perception," in *Vision and Attention*, M. Jenkin and L. Harris, Eds. New York: Springer-Verlag, 2001.
- [2] D. Simons, "Change blindness and inattention blindness: the role of attention and inattention in perception," in *Vision and Attention*, M. Jenkin and L. Harris, Eds. New York: Springer-Verlag, 2001.
- [3] P. Viviani, "Eye movements in visual search: cognitive, perceptual and motor control aspects," in *Eye Movements and Their Role in Visual and Cognitive Processes*, E. Kowler, Ed. Amsterdam: Elsevier Science Publishers, 1990, pp. 353-393.
- [4] A. L. Yarbus, *Eye Movements and Vision*, New York: Plenum Press, 1967.
- [5] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychol. Bull.*, vol. 124, pp. 372-422, 1998.
- [6] R. Desimone and J. Duncan, "Neural mechanisms of selective visual attention," *Annu. Rev. Neurosci.*, vol. 18, pp. 193-222, 1995.
- [7] J. H. R. Maunsell, "The brain's visual world: representation of visual targets in cerebral cortex," *Science*, vol. 270, pp. 764-769, 1995.
- [8] P. R. Roelfsema, V. A. Lamme, and H. Spekreijse, "Object-based attention in the primary visual cortex of the macaque monkey," *Nature*, vol. 395, pp. 376-381, 1998.
- [9] J. D. Schall, A. Morel, D. J. King, and J. Bullier, "Topography of visual cortical afferents to frontal eye field in macaque: functional convergence and segregation of processing streams," *J. Neurosci.*, vol. 15, pp. 4464-4487, 1995.
- [10] B. Jouve, P. Rosenstiehl, and M. Imbert, "A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey," *Cereb. Cortex*, vol. 8, pp. 28-39, 1998.
- [11] M. A. Segraves and M. E. Goldberg, "Functional properties of corticotectal neurons in the monkey's frontal eye field," *J. Neurophys.*, vol. 58, pp. 1387-1419, 1987.
- [12] K. G. Thompson, "Neural mechanisms of bottom-up selection during visual search," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 23, 2001.
- [13] W. F. Bacon and H. E. Egeth, "Overriding stimulus driven attentional capture," *Percept. Psychophys.*, vol. 55, pp. 485-496, 1994.
- [14] C. F. Nodine, H. L. Kundel, S. C. Lauer, and L. C. Toto, "Nature of expertise in searching mammograms for breast masses," *Acad. Radiol.*, vol. 3, pp. 1000-1006, 1996.
- [15] N. P. Bichot, J. D. Schall, and K. G. Thompson, "Visual feature selectivity in frontal eye fields induced by experience in mature macaques," *Nature*, vol. 381, pp. 697-699, 1996.
- [16] J. D. Schall, D. P. Hanes, K. G. Thompson, and D. J. King, "Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation," *J. Neurosci.*, vol. 15, pp. 6905-6918, 1995.
- [17] N. J. Cepeda, K. R. Cave, N. P. Bichot, and M.-S. Kim, "Spatial selection via feature-driven inhibition of distractor locations," *Percept. Psychophys.*, vol. 60, pp. 727-746, 1998.
- [18] M. A. Basso and R. H. Wurtz, "Modulation of neuronal activity in superior colliculus by changes in target probability," *J. Neurosci.*, vol. 18, pp. 7519-7534, 1998.
- [19] M. L. Platt and P. W. Glimcher, "Neural correlates of decision variables in parietal cortex," *Nature*, vol. 400, pp. 233-238, 1999.
- [20] J. M. Findlay and R. Walker, "A model of saccade generation based on parallel processing and competitive inhibition," *Behav. Brain Sci.*, vol. 22, pp. 661-721, 1999.
- [21] K. D. Powell, C. L. Colby, J. P. Gottlieb, M. Kusunoki, and M. E. Goldberg, "Space and salience in parietal cortex," in *Current Oculomotor Research*, W. Becker, H. Deubel, and T. Mergner, Eds. New York: Plenum Press, 1999, pp. 25-35.
- [22] J. Theeuwes, "Cross-dimensional perceptual selectivity," *Percept. Psychophys.*, vol. 50, pp. 184-193, 1991.
- [23] N. P. Bichot, S. C. Rao, and J. D. Schall, "Continuous processing in macaque frontal cortex during visual search," *Neuropsychologia*, in press.
- [24] S. Yantis, "Attentional capture in vision," in *Converging Operations in the Study of Visual Selective Attention*, A.F. Kramer, M.G.H. Coles, and G.D. Logan, Eds. Washington, DC: American Psychological Association, 1995, pp. 45-76.
- [25] J. P. Gottlieb, M. Kusunoki, and M. E. Goldberg, "The representation of visual salience in monkey parietal cortex," *Nature*, vol. 391, pp. 481-484, 1998.
- [26] D. D. Burman and M. A. Segraves, "Primate frontal eye field activity during natural scanning eye movements," *J. Neurophys.*, vol. 71, pp. 1266-1271, 1994.
- [27] A. Treisman, "Features and objects: the fourteenth Bartlett memorial lecture," *Q. J. Exp. Psychol. A*, vol. 40, pp. 201-237, 1988.
- [28] J. M. Wolfe, K. R. Cave, and S. Franzel, "Guided Search: an alternative to the feature integration model for visual search," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 15, pp. 419-433, 1989.
- [29] A. Treisman and S. Sato, "Conjunction search revisited," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 16, pp. 456-478, 1990.
- [30] J. M. Findlay, "Saccade target selection during visual search," *Vision Res.*, vol. 37, pp. 617-631, 1997.
- [31] N. P. Bichot and J. D. Schall, "Saccade target selection in macaque during feature and conjunction visual search," *Vis. Neurosci.*, vol. 16, pp. 81-89, 1999.
- [32] V. Maljkovic and K. Nakayama, "Priming of pop-out: I. Role of features," *Mem. Cognit.*, vol. 22, pp. 657-672, 1994.
- [33] N. P. Bichot and J. D. Schall, "Effects of similarity and history on neural mechanisms of visual selection," *Nat. Neurosci.*, vol. 2, pp. 549-554, 1999.
- [34] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Hum. Neurobiol.*, vol. 4, pp. 219-227, 1985.
- [35] N. P. Bichot, K. G. Thompson, S. C. Rao, and J. D. Schall, "Reliability of macaque frontal eye field neurons signaling saccade targets during visual search," *J. Neurosci.*, vol. 21, pp. 713-725, 2001.