

Visual Attention: Bottom-Up Versus Top-Down

Dispatch

Charles E. Connor¹, Howard E. Egeth² and Steven Yantis²

Visual attention is attracted by salient stimuli that ‘pop out’ from their surroundings. Attention can also be voluntarily directed to objects of current importance to the observer. What happens in the brain when these two processes interact?

To mangle a well-known Wordsworth line, the world is too much for us — it contains far too much information for us to perceive at once. We typically pay attention to individual items, one after another. But which items? That depends on two distinct types of attentional mechanism. Bottom-up mechanisms are thought to operate on raw sensory input, rapidly and involuntarily shifting attention to salient visual features of potential importance — the spot of red against a field of green that could be a piece of fruit, the sudden movement that could be a predator. Top-down mechanisms implement our longer-term cognitive strategies, biasing attention toward colored spots if we are hungry or toward sudden movements and quadrupedal shapes if we fear a predator.

Psychologists have investigated how these two attentional processes function in human perception, both individually and — as is usual under natural conditions — together [1]. In a typical experiment [2], on some trials subjects were asked to search for and identify a color singleton, such as a single red target in a search display containing several white distractor stimuli, while on other trials the target was an onset singleton — the only item in the display with an abrupt onset. The search display was immediately preceded by a ‘cue’ display, also containing a color or onset singleton. Even though subjects were instructed that the preceding cue display was irrelevant and should be ignored, the cue singleton attracted attention in an involuntary, bottom-up fashion, affecting how quickly subjects could respond to the upcoming targets.

When a color-singleton cue appeared at the same location as the upcoming color target, for example, attention was drawn to the target location in advance and responses were faster. When a color-singleton cue appeared at a different location, attention was temporarily drawn away from the target location and responses were slower. However, this bottom-up attention effect was modulated by the subject’s top-down attentional state: color-singleton cues only affected response times when the subject was

searching for a color target; likewise, onset-singleton cues only mattered during search for onset targets. Thus, bottom-up attention alerts us to salient items in our environment, but top-down attention modulates bottom-up signals when we need to look for something specific.

A new paper by Ogawa and Komatsu [3] provides a glimpse at the neural processes underlying such interactions between bottom-up and top-down attention. These investigators trained monkeys on two top-down behavioral strategies: color-singleton search and shape-singleton search. The search displays (Figure 1, red and green boxes) always contained six stimuli, including one color singleton and one shape singleton. At the beginning of the trial, the monkey had to fix its gaze at the center of the screen. Then the display appeared, and the monkey had to saccade (shift its gaze; arrows in Figure 1) to the correct target to receive a reward. In the color search task (the two boxes on the right), the correct target was the color singleton; in the shape search task (left-hand boxes) the correct target was the shape singleton.

While the monkey performed these tasks, Ogawa and Komatsu [3] recorded the electrical spike activity of individual neurons in area V4, an intermediate stage in the visual object-processing pathway of the brain, where neural activity is known to be affected by attention [4–8]. The display was arranged so that one stimulus fell inside the V4 neuron’s receptive field (gray circles in Figure 1). On a given trial, the receptive field stimulus could be the shape singleton (red boxes in Figure 1), the color singleton (green boxes) or one of the four non-singleton stimuli (not shown). Depending on the search condition, a receptive field singleton stimulus could either be a target (solid outline boxes) or a non-target (dashed outline boxes).

The firing rate *versus* time plots in Figure 1 show how an example V4 neuron responded to a white cross-shaped stimulus during shape search (left) and color search (right). In both conditions, this neuron’s response to the white cross was stronger when that stimulus was the color singleton in the display (green lines). The color-singleton responses began to exceed the shape-singleton (red) and non-singleton (blue) responses at about 125 milliseconds after stimulus onset. Other neurons showed similar response enhancements for color singletons, shape singletons or both, again under both search conditions. Thus, even when the top-down strategy was to focus on one dimension (shape or color), singletons in the other dimension evoked enhanced responses from some neurons.

This obligatory singleton effect at the neural level is neatly consistent with psychophysical evidence reported by Theeuwes [9]. He showed that, during searches for shape-singleton targets, the attention of human observers is ineluctably drawn to irrelevant color singletons. This psychophysical effect now has

¹Department of Neuroscience and Krieger Mind/Brain Institute, ²Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA.
E-mail: connor@jhu.edu

a neural analog: the neural response enhancement for color singletons that occurs even during shape search (Figure 1). Response enhancement is the signature for attention at the neural level [4–8], and it is easy to imagine that the enhanced responses to singletons could cause (or in a sense constitute) the bottom-up attentional capture observed in humans.

Citing his psychophysical evidence, Theeuwes [10] has maintained that bottom-up attention, for example to singletons, can never be overridden by top-down attention. Do the neurophysiological results of Ogawa and Komatsu [3] support this categorical position? Not necessarily. First, in both the psychophysical experiments of Theeuwes [9] and neurophysiological experiments of Ogawa and Komatsu [3], the task itself requires search for singletons (in a particular dimension). Subjects are deliberately looking for unique stimuli that pop out from a field of distractors — a top-down cognitive mode that might permit any singleton to capture bottom-up attention. What if the search target is not a singleton? Bacon and Egeth [11] tested this by requiring search for a specific shape embedded in an array of multiple shapes, rendering the target a non-singleton and discouraging a singleton search strategy. Under this circumstance, color-singleton distractors had no apparent effect, indicating that bottom-up effects can be overridden in some cases. It would be fascinating to know whether the neural singleton effect would likewise disappear under non-singleton search conditions.

Second, while the neural singleton effect is not completely overridden by top-down strategy in the Ogawa and Komatsu [3] experiment, it is nevertheless strongly modulated, especially at later timepoints. The response to the irrelevant color singleton (during shape search; Figure 1, dashed green line) falls to near 20 spikes per second by 200 milliseconds after display onset, while the relevant (color search) color-singleton response (solid green line) stays up near 50 spikes per second. Thus, the enhancement of neural responses to color singletons is stronger during color search and weaker during shape search. Most neurons showed similar task-dependent response differences. This could be a general effect of compatibility between search dimension and singleton type, or it could represent selective enhancement of responses to the target stimulus once it has been identified. In either case, it exemplifies top-down modulation of bottom-up effects, consistent

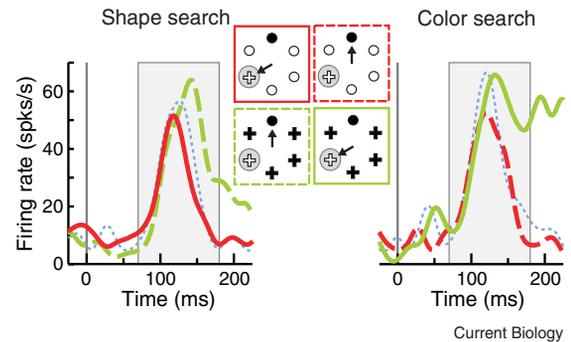


Figure 1. Responses of a typical monkey area V4 neuron during the shape search and color search tasks.

The left and right plots show the neuron’s firing rate as a function of time following display onset. The center boxes show four different task conditions. See text for details.

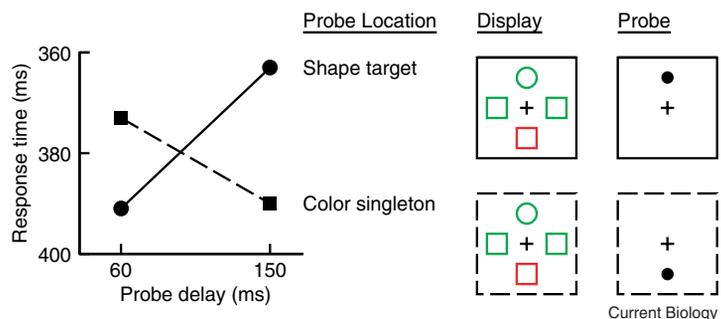
with theories and psychophysical evidence favoring the idea that top-down and bottom-up factors interact to control allocation of attention [1,2,11–15].

The differential time course of the bottom-up and top-down neural effects provides a particularly revealing comparison with psychophysics. In Figure 1, the color-singleton response functions (green lines) begin to separate from shape-singleton (red) and distractor (blue) responses at about 125 milliseconds (in both shape search and color search). The additional top-down effect of search dimension begins to appear at about 175 milliseconds (where the dashed green line continues to fall but the solid green line stays high). At the neural population level, bottom-up singleton response enhancement first became significant at about 120 milliseconds, while top-down search dimension effects first became significant at about 195 milliseconds [3].

A similar time course difference has been demonstrated in human psychophysical experiments [13,14]. Figure 2 is based on an experiment by Lamy *et al.* [14] in which the primary task was to report the presence or absence of a circle target among square distractors (a shape task). In every trial, one of the square distractors was a task-irrelevant color singleton designed to attract bottom-up attention. The allocation of attention was tested at different time points with a concurrent secondary task: a small probe dot could appear at one of the square/circle

Figure 2. Probe reaction time as a function of delay between target display and probe.

The graph on the left shows the response times of human subjects to probe stimuli presented at the location of a shape target (circles, solid lines) or a color singleton (squares, dashed lines) at two different probe delays (time between display onset and probe onset). The vertical response time scale is inverted so that faster responses appear higher. See text for details.



stimulus locations, and the subject had to respond to the probe dot by pressing a key as quickly as possible. (Later, at the end of the trial, the subject reported whether the circle target had been present.) At an early time point — tested with a probe delay of just 60 milliseconds following display onset — subjects responded faster to the probe at the color-singleton location, showing that this irrelevant singleton had captured attention in a bottom-up fashion.

At a later time point — 150 milliseconds probe delay — subjects responded faster to the probe at the shape-target location, showing that top-down attention had taken over. Thus, at both the neural and psychophysical levels, bottom-up attention acts early, and top-down attention takes over within on the order of 100 milliseconds. These parallel time courses help to confirm the functional relationship between the neural events and the behavioral performance measurements. The earlier time course for bottom-up attention makes sense — salience effects based on simple visual properties could be rapidly implemented at early visual processing stages, while top-down cognitive control would require subsequent operations on sensory input by higher-level cortex.

Psychology has shown that the complex dynamic interplay between bottom-up and top-down attention determines what we are aware of from moment to moment [1,2,11–15]. Ogawa and Komatsu [3] have provided the closest look yet at this interplay on the neural level. Parallels are often drawn between neural and psychological events, but it is rare to see a neurophysiological paradigm so well-targeted toward an extensive preexisting set of psychophysical observations. One hopes this is the beginning of a closer interaction between the two literatures that will lead to a far richer understanding of visual attention in all its complexity.

References

1. Egeth, H.E., and Yantis, S. (1997). Visual attention: control, representation, and time course. *Annu. Rev. Psychol.* 48, 269-297.
2. Folk, C.L., Remington, R. and Johnston, J.C. (1992). Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psychol.: Hum. Percept. Perform.* 18, 1030-1044.
3. Ogawa, T. and Komatsu, H. (2004). Target selection in area V4 during a multidimensional visual search task. *J. Neurosci.* 24, 6371-6382.
4. Moran, J. and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science* 229, 782-784.
5. Motter, B.C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J. Neurosci.* 14, 2178-2189.
6. Connor, C.E., Preddie, D.C., Gallant, J.L. and Van Essen, D.C. (1997). Spatial attention effects in macaque area V4. *J. Neurosci.* 17, 3201-3214.
7. McAdams, C.J. and Maunsell, J.H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* 19, 431-441.
8. Reynolds, J.H. and Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron* 37, 853-863.
9. Theeuwes, J. (1992). Perceptual selectivity for color and form. *Percept. Psychophys.* 51, 599-606.
10. Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychon. Bull. Rev.* 11, 65-70.
11. Bacon, W.F. and Egeth, H.E. (1994). Overriding stimulus-driven attentional capture. *Percept. Psychophys.* 55, 485-496.
12. Wolfe, J.M. (1994). Guided search 2.0: a revised model of visual search. *Psychon. Bull. Rev.* 1, 202-238.
13. Kim, M.-S. and Cave, K.R. (1999). Top-down and bottom-up attentional control: on the nature of the interference from a salient distractor. *Percept. Psychophys.* 61, 1009-1023.
14. Lamy, D., Tsai, Y. and Egeth, H.E. (2003). Does a salient distractor capture attention early in processing? *Psychon. Bull. Rev.* 10, 621-629.
15. Yantis, S. and Egeth, H.E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 661-676.