Salience, relevance, and firing: a priority map for target selection

Jillian H. Fecteau and Douglas P. Munoz
Netherlands Institute of Neuroscience, Meibergdreef 47, 1105BA, Amsterdam, The Netherlands

The salience map is a crucial concept for many theories of visual attention. On this map, each object in the scene competes for selection – the more conspicuous the object, the greater its representation, and the more likely it will be chosen. In recent years, the firing patterns of single neurons have been interpreted using this framework. Here, we review evidence showing that the expression of salience is remarkably similar across structures, remarkably different across tasks, and modified in important ways when the salient object is consistent with the goals of the participant. These observations have important ramifications for theories of attention. We conclude that priority – the combined representation of salience and relevance – best describes the firing properties of neurons.

Introduction
The complexity of the visual world exceeds the processing capacity of the human brain [1], which forces us to select one (or a few) object(s) in the scene for more detailed analysis at the expense of other items [2]. This act of selection provides a succinct definition of the term ‘visual attention’. In this article, we explore one basic issue surrounding this selection process – how do we choose the next object of attention [1,3–13]?

The salience map provides one conceptual framework that accounts for how the next object is selected. It consists of a topographical map of space, upon which all objects in the visual scene compete. Only the physical distinctiveness of objects factor into this competitive process and, over time, the most salient object is chosen in a winner-take-all fashion [4].

Many disciplines have converged upon the same basic idea: computational and psychological models have implemented concepts similar to the salience map when describing how an object is selected [4,8,10,11,14–17,25]. Over the last several years, the spiking patterns of single neurons have been likened to the salience map as well [18–29]. It is both an exciting prospect and a mark of true convergence of computational and psychological theories is very specific – salience refers to the physical, bottom-up distinctiveness of an object [4,8,11,13–17]. It is a relative property that depends on the relationship of one object with respect to other objects in the scene [8,10]. This property of salience is highlighted in Figure 1a: it is much easier to detect the presence of a target in a display when it possesses a distinct feature (Figure 1a left, and Figure 1b: search reaction times are flat across set sizes) than when it possesses a less distinct feature (Figure 1a middle) or is characterized by a combination, or ‘conjunction’, of features (Figure 1a right, and Figure 1b, search times increase with increases in set size).

There are many different physical qualities that can make an object more salient than other objects in the display, such as its color, orientation, size, shape, movement or unique onset [9,11]. The salience map allows the most distinct object to be identified, independently of the particular features that it possesses [4], by receiving input from different feature maps that represent specific qualities of the scene and then summing these values so the relative distinctiveness of the object is represented in a featureless manner. In the presence of multiple competing objects, the object possessing the greatest sum of activity is the one that is selected first. When subsequent shifts of attention occur, the representation of this salient object is suppressed through inhibition of return [31–34], allowing the object possessing the second highest level of salience to be selected next.

Where is the salience map?
These properties constrain the neurophysiological evidence that should be taken to support the salience map. First, the task should be designed to encourage the bottom-up processing of items in the scene, as opposed to goal-directed selection. This is an important constraint because many neurophysiological investigations have used the terms ‘salience’ and ‘relevance’ interchangeably [18,19,21,23,24,29,30]. Second, the neurons should be spatially selective, but otherwise encode visual objects in a featureless manner, and lesions to the structures in question should produce obvious deficits in selective attention. These later constraints are important because...
several areas of the brain have been considered the locus of the theoretical salience map (e.g. the primary visual cortex [25] and areas of the ventral visual pathway [26]), but do not meet these criteria. Importantly, the oculomotor network does possess many of these properties [18,19,21,23,24,29] (and see Box 1):

(i) neurons in this network encode visual information in a featureless manner
(ii) lesions involving these structures produce deficits in attentional selection [35]
(iii) electrically stimulating these regions facilitates the selection of objects with attention [36–39]
(iv) these structures receive information from the ventral visual pathway [40–42], which provides the input necessary for summing the relative salience of an object.

Expression of salience in the oculomotor network

There are two mechanisms by which the salience map operates that yield measurable consequences in behavior: salience effects (defined by the task) and inhibition of return. Over the last decade, much has been learned about how salience and inhibition of return are reflected in neural activity through the use of visual search and the non-predictive cue–target task (see Box 1 for interpretations of the neural correlates described here).

Salience and inhibition of return in visual search

The visual search task has provided an effective way to explore the neural correlates of salience and inhibition of return. The majority of the work we describe has originated from Schall’s laboratory [29,43] whose members have monitored neural activity in the frontal eye fields as monkeys performed an oddball search task (Figure 1a). The typical pattern of neural activity obtained during this task is illustrated in Figure 1c: the target is differentiated from the distractor in the recurrent epoch (see Box 1). No differentiation between the target and the distractors is seen in the first feedforward sweep. The same neural correlate is witnessed for targets defined by color or motion [44] and, importantly, the conjunction of color and shape as well [18] (see Box 2). This outcome is consistent with the

Figure 1. Visual search. (a) Examples of different visual search displays and (b) the corresponding changes in performance across set size. (Adapted with permission from [8].) (c) Example of the activity from a single frontal-eye-field neuron during distinct feature search. The blue line shows neural activity when the target appeared within the response field of the neuron. The red line shows neural activity when the distractor appeared within the response field of the neuron. The arrow indicates the time of target discrimination. (Adapted with permission from [43].) (d) Example of the activity from another single neuron monitored during similar feature search. The orange arrow illustrates the time of target discrimination for the distinct feature task for comparison. (Adapted with permission from [44].) (e) Behavioral and neural correlates associated with inhibition of return (IOR). (Adapted with permission from [45].)
Box 1. Neural signals in the oculomotor network

The oculomotor network is spread across several regions of the brain, including the frontal eye fields, the lateral intraparietal area, the superior colliculus, and the brainstem reticular formation (Figure Ia). The frontal eye fields and superior colliculus are crucial nodes in this network: both receive convergent input from many cortical areas and project directly to the brainstem reticular formation. Both structures contain three types of neurons, each identified by the particular events that it registers. Visual neurons register the appearance of a visual stimulus in their receptive field. Motor neurons register the initiation of a saccadic eye movement to a particular locus in space. Visuomotor neurons register the appearance of a visual stimulus and the initiation of a saccadic eye movement. (This classification scheme is over simplified, see [42,71,78]).

As described in the main text, the neural correlates of attentional selection are associated with a change in the neural representation of the target. This sensory representation can be divided into two epochs [46]: the feedforward sweep of visual input corresponds to the initial registration of a visual stimulus by the neuron (Figure Ib). Recurrent processing corresponds to a second epoch of neural activity that is thought to originate from widespread interactions across visual areas in the brain.

![Figure I](https://www.sciencedirect.com)

Figure I. (a) Regions of the brain involved in the generation of eye movements and shifts of attention in the human (left) and macaque (right). (b) Changes in neural activity associated with feedforward, recurrent, and motor processing. Target-aligned and saccade-aligned refer to the alignment of neural activity to the appearance of the target or the initiation of a saccade, respectively.

Salience and inhibition of return revealed through the cue–target task

The basic logic underlying the non-predictive, cue–target task is different than that of visual search: rather than exploring the properties of a salient object in a crowded display, the cue–target task reveals the consequence of a salient event across time [3,33,34,47]. Despite this different way of probing attentional selection, the cue target task is thought to measure the same underlying processes as visual search [7].

As illustrated in Figure 2a, the observer begins each trial by maintaining fixation upon a central dot. Then, an abrupt flash appears in the visual periphery (the cue). The cue is extinguished and, after a delay, a target appears to which the observers generate a response. There are two crucial manipulations that reveal the consequences of the salient cue: the positioning of the target with respect the cue (same location versus opposite locations) and the interval between the onset of the cue and the target. The consequence of these two factors on the performance of monkeys is illustrated in Figure 2b. Like human observers, monkey saccadic reaction times are faster when the cue and target appear at the same location at short cue–target intervals and slower at longer cue–target intervals (see [48] for a direct human and monkey comparison). This change in performance is thought to reflect the initial capture of attention by the salient cue followed by inhibition of return [3,33,34,47].

Both behavioral effects are associated with noteworthy changes in neural activity in the superior colliculus [20,22,48–50]. The capture of attention (Figure 2c, left) is associated with a stronger neural representation of the target when the cue and target appeared at the same location. By contrast, inhibition of return (Figure 2c, right)
is associated with a weaker neural representation of the target.

For the cue–target task, a close relationship exists between the neural representation of the target and behavior [22]. Focusing first on the capture of attention (Figure 2d), performance at the 50 ms cue–target interval was inconsistent across sessions: in some sessions, monkeys responded faster when the cue and target appeared at the same location, whereas in other sessions they responded more slowly. Contrasting across the sessions in which the capture of attention was or was not obtained in behavior revealed that the neural data followed suit: a strong target-related response was observed in the sessions yielding a same location advantage in behavior and a weak target-related response was found in the sessions yielding an opposite location advantage [22].

Inhibition of return, seen at longer cue–target intervals, also corresponds closely to changes in neural activity [20,22,48–50]. Importantly, the causal role between the weak target-related response and inhibition of return has been established. Dorris and colleagues [49] observed that the neural activity registering the target was weak even though the neurons appeared more excitable after the appearance of the cue (i.e. higher activity after the cue; Figure 2c, right). To see whether this weak target-related signal was responsible for longer reaction time, they replaced the appearance of a visual target with a brief train of microstimulation in the superior colliculus to evoke a saccade. As seen in Figure 2e, faster reaction times were obtained when the visual target was replaced with mild electrical stimulation, in contrast to the slower reaction times obtained when the visual target guided the action. Therefore, it is the diminished salience of the target that is responsible for inhibition of return when using this cue–target task (see also Box 2).

In summary, salience and inhibition of return are represented as a change in the feedforward representation of the target when using the cue–target task: the capture of attention is associated with a strong target-related response and inhibition of return is associated with a weak target-related response.

**Stages of processing within the salience network**

Consider one outcome of the evidence presented thus far: the neural correlates of salience and inhibition of return differ across tasks. For visual search, salience and inhibition of return are represented as a change in the amount of time it takes to discriminate the target from the distractor, which occurs within the recurrent epoch. For the cue–target task, salience and inhibition of return are represented as a change in the initial representation of the visual target, which occurs within the feedforward epoch.

Consider also that these distinct neural correlates appear to reflect differences across tasks, not differences across brain structures. Oddball visual search produces similar neural correlates in the frontal eye fields [43], the lateral intraparietal area [51], and the superior colliculus [27]. Although significant differences across studies make comparisons of the cue–target task tenuous, similar changes in neural activity have been observed in the superior colliculus [20,22,48–50,52], the lateral intraparietal area [53], and other structures [22]. This synopsis implies that the salience map may be the property of a network.

When probing the properties of a network, it is difficult to assess ‘where’ a cognitive behavior originates because many regions receive similar inputs and are reciprocally interconnected. Thus, the neural expression of the cognitive behavior will be similar across brain areas. An alternative way to ask the question ‘where’ a cognitive behavior originates is to assess at which level of processing these neural correlates emerge. The organization of the superior colliculus allows us to make such comparisons directly (Figure 3a). Its superficial layers receive visual input from early stages in the sensory-to-motor processing path: the retina, the primary visual cortex, and areas V2 and V3 [22]. Its intermediate layers receive input from later stages in the sensory-to-motor processing path: the lateral intraparietal area, the frontal eye fields, the dorsal lateral prefrontal cortex, and the inferotemporal cortex [22]. Thus, by comparing the neural correlates in these two subregions of the superior colliculus, it is possible to assess whether the input to the superficial layers or the intermediate layers contains the information necessary to reveal these neural correlates of salience and inhibition of return.

For visual search (Figure 3b), neurons in the intermediate layers of the superior colliculus discriminate the target...
from the distractors, but the same pattern is not observed in the superficial layers [27]. Thus, for visual search, salience and inhibition of return originate late in sensory processing. For the cue–target task (Figure 3c), strong activity representing the capture of attention (top) is seen in the intermediate layers, but the same pattern is not observed in the superficial layers [22]. By contrast, the weak signal representing inhibition of return (bottom) is seen in both layers [22]. Thus, for the cue–target task, salience is a property originating late in sensory processing, whereas inhibition of return originates early.

In summary, the neural correlates of salience require input from later stations in the visual hierarchy, which means that the salience map is not a summary of visual processing occurring at early stages of the visual hierarchy. Albeit contrary to the basic construction of the salience map, this observation is consistent with many studies showing that complex visual objects can be perceptually distinct in visual search [11,54] (see also Box 2). By contrast, the neural correlates of inhibition of return can be viewed at early and late stages of the visual hierarchy, depending on the task. This observation is consistent with studies showing that different variants of inhibition of return exist [55].

The priority map: combining salience and relevance
Salience and the salience map refer to bottom-up processes in attentional selection – neither the relevance of an object nor the goals of observers play any part in this conceptual framework. Yet the terms salience and relevance are often treated as synonyms in the neurophysiological literature [18,19,21,23,24,29,30]. This discrepancy might reflect...
the ‘top-down’ knowledge that neurophysiologists bring to the issue – the relevance of an object influences how it is processed in oculomotor structures and elsewhere [56–59]. This has obvious consequences for the feasibility of the salience map because bottom-up and top-down sources of input converge to produce an amalgamated representation of ‘priority’ [10,60].

Even though salience and relevance affect the same structures, the unique contributions of both sources of input can be teased apart. In a recent study [20], monkeys performed two versions of the cue–target task: the non-predictive version, described above, and a predictive version, in which the cue accurately identified the upcoming location of the target 75% of the time. The same neurons were monitored in both conditions, allowing the unique contributions of salience and relevance to the neural signal to be determined.

The predictive cue produced clear changes in the monkeys’ behavior: the monkeys responded faster when the cue and target appeared at the same location for predictive cues than for non-predictive cues, especially when the cue–target interval was short (Figure 4a). The neural representation of the target yielded a similar pattern: its representation was stronger for predictive cues, in particular at short cue–target intervals. The reason for this change is shown in Figure 4b. Following the feedforward registration of the predictive cue, the neural activity was elevated, this elevated activity was maintained across the cue–target interval, and augmented the registration of the target. Thus, salience and relevance yield distinct neural signals – salience is reflected in the initial registration of the target, and relevance is reflected in the elevated activity following the predictive cue.

At present, a conceptually similar analysis cannot be conducted for visual search because foreknowledge of the upcoming location of the target has not been manipulated directly. Foreknowledge of the feature identifying the target has been manipulated, though. Bichot and Schall [45] presented the target in the same color for 10 trials before switching it to another color. This manipulation allowed the target to be identified faster across trials in behavior and allowed the target to be discriminated sooner in neural activity. Therefore, advance knowledge of the feature identifying the target acts on the same neural correlates as salience and inhibition of return in visual search.
Taken together, the oculomotor system appears to represent the priority of objects – the combined representation of an object’s bottom-up distinctiveness and its relevance to observers. Manipulations of spatial attention produce separate correlates of salience and relevance, as found through the cue–target task. By contrast, manipulations of feature-based attention do not yield discernable signals of salience and relevance; instead, the oculomotor system registers the summary of these two processes. This suggests that the attentional control settings involved in enhancing the features of a desired object are processed outside of the oculomotor network, yet facilitate target discrimination in the priority map.

**Relationship between priority and action**

How we choose the next object of attention seems to be closely related to how we choose the next target of a saccade [67], as shown through behavioral evidence [68,69], functional imaging and neuropsychological investigations [35,70]. These correspondences raise the question—what is the relationship between selective attention and oculomotor action?

Computational and psychological theories treat the salience map and action as separate processes [4]. In the brain, however, the neural representations of priority and action are found in the same neurons. Although this convergence suggests that a direct relationship between these processes might exist, the relationship between attention and oculomotor planning remains unclear.

In some circumstances, visual information can be used to trigger action directly. This is seen through ‘express saccades’ – the ability of a visual target to trigger a saccade.
immediately when the superior colliculus is in a state of high excitability [71]. Moreover, slower reaction times linked to inhibition of return originate from the weak registration of the target, as shown through electrical stimulation studies [49].

In other circumstances, visual signals appear to be dissociated from action. Juan and his colleagues reported that attending to a target does not affect saccadic planning ([72], but see also [73]). In addition, two distinct classes of visuomotor neurons have been observed in visual search in the frontal eye fields [74] and the superior colliculus [27]. Discriminatory neurons identify the target in search array, but bear little resemblance to ultimate motor act. Non-discriminatory neurons register the appearance of a visual stimulus, do not identify the target in the search array, and bear a close correspondence to the ensuing saccade. Thus, in some instances, a direct relationship exists between visual signals and action, whereas, in other instances, it does not (see Box 3).

Conclusions
The salience map is a concept upon which different disciplines have converged: psychologists and computational scientists have implemented the salience map in their models of attentional selection. In this article, our primary goal was to review the neurophysiological evidence taken to support the salience map. Our secondary goal was to assess if and how evidence from neurophysiology should modify our theoretical perspectives.

With regards to the first goal, we have provided ample evidence indicating that the oculomotor network shares important features with the theoretical salience map and yields patterns of neural activity that are consistent with its functioning – the neural representation of the visual target is enhanced when it is salient and suppressed under conditions of inhibition of return.

With regards to our second goal, neurophysiological investigations also reveal that the concept of the salience map must be broadened to include top-down influences. We recommend the term priority map [60] to properly reflect the combined roles of salience and relevance in this selection process.

The neurophysiological evidence has additional implications. (1) The neural expression of salience and inhibition of return change across tasks, which indicates that we are not tapping into the same processes with different tasks, but different processes with similar manifestations in behavior. (2) For a given task, salience and inhibition are not separate mechanisms, but represent a continuum of neural processing. (3) The priority map appears to be the property of a network, forcing us to reconsider the suggestions that different structures have fundamentally different roles in attentional selection.

Despite its limitations, the salience map has been a successful theoretical construct, from which our models will develop further [1,4,75–77] (see also Box 3). It has stimulated a great deal of research in cognitive, computational, and neurophysiological sciences. Finally, it is one of those rare concepts that fosters direct comparisons between psychological and computational theory and neurophysiology. It is through these cross-disciplinary interactions that our knowledge of the neural basis of cognitive function will blossom.

References
2 James, W. (1890) Principles of Psychology, Dover Publications
29 Thompson, K.G. and Bichot, N.P. (2005) A visual salience map in the primate frontal eye field. Prog. Brain Res. 147, 251