Neuron
Perspective

CellPress

From Behavior to Neural Dynamics: An Integrated Theory of Attention

Timothy J. Buschman[1,*] **and Sabine Kastner**[1,*]
[1]Department of Psychology, Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA
*Correspondence: tbuschma@princeton.edu (T.J.B.), skastner@princeton.edu (S.K.)
http://dx.doi.org/10.1016/j.neuron.2015.09.017

The brain has a limited capacity and therefore needs mechanisms to selectively enhance the information most relevant to one's current behavior. We refer to these mechanisms as "attention." Attention acts by increasing the strength of selected neural representations and preferentially routing them through the brain's large-scale network. This is a critical component of cognition and therefore has been a central topic in cognitive neuroscience. Here we review a diverse literature that has studied attention at the level of behavior, networks, circuits, and neurons. We then integrate these disparate results into a unified theory of attention.

Introduction

Over 125 years ago, William James defined attention as the "taking possession by the mind…of one out of what seem simultaneously possible objects or trains of thought" (James, 1890). James' intuitive understanding of attention is remarkably close to our modern definition: attention is the selective prioritization of the neural representations that are most relevant to one's current behavioral goals. Such prioritization is necessary because the brain is a limited capacity information system. Representations of external stimuli and internal thoughts compete for access to these limited processing resources, and attention helps to resolve that competition in favor of the information that is currently task relevant.

Attention research has been central to the fields of cognitive neuroscience, psychology, and systems neurophysiology. This has led to the discovery of a large number of attention effects at each of these levels of observation. In the first three sections, we briefly review this literature, highlighting key insights at the behavioral, network, and neuronal levels. Our goal for this review is to integrate these disparate findings into a single unified framework, which we outline in the fourth section.

We should note that we will largely constrain our review to visual attention, as it has been the best studied. We acknowledge the importance of extending our understanding to other sensory modalities and to interactions between modalities, and we hope the knowledge gained from understanding visual attention will reveal principles of neural processing that may be fundamental to cognition more generally.

Furthermore, even though attention is often studied in isolation, a mechanism that prioritizes task-relevant information will likely interface with many cognitive domains such as action control and decision making, motivation and emotions, memories at different timescales, and awareness. We will review our current knowledge of some of these interactions in the last section. Understanding the interaction of selective attention with other cognitive domains will ultimately lay the foundation for reaching a cohesive understanding of the general principles of cognition and their associated neural mechanisms (Nobre and Kastner, 2014).

Behavioral Effects—Building Blocks and Shifting Concepts

Classical Attention Paradigms

The two most commonly used paradigms to study visual attention are visual spatial orienting (Posner et al., 1980) and visual search (Treisman and Gelade, 1980).

In spatial orienting tasks, subjects are instructed by a predictive cue to direct attention to a particular spatial location where they must detect or discriminate a target stimulus. The classic finding is that subjects benefit from the cue as they respond faster and more accurately to stimuli occurring at the cued location than to stimuli occurring at other locations. This facilitation comes at the expense of other objects in the visual environment, reflecting the competitive nature of attention.

While orienting tasks typically involve only a single target stimulus, visual search tasks more closely relate to our everyday experience, where we typically face cluttered scenes. In search tasks, subjects are given an array of stimuli and asked to find a particular target stimulus defined by one or more features in the array (e.g., find the green "T" in an array of green and blue "T's" and "L's"; see Figure 1A). Hence, in visual search, the selection process is informed by features of the target (i.e., feature-based attention), which then guides spatial attention.

Performance in visual search tasks is affected by how many features the target shares with other stimuli in the array. If the target has a unique feature, such as being a different color than the distracters, the search is completed quickly and effortlessly, regardless of the number of elements in the array. This phenomenon is known as "pop-out" or efficient (parallel) search. However, just by changing the distractors in the search array, the search for the same target can be made much more difficult. For example, if the target is defined by a conjunction of features that each are shared by distracters (as in Figure 1A), search time increases as a function of the number of elements in the array. This is known as inefficient search, and the increase in search times is thought to reflect a serial target search, which is mediated by a spatial "spotlight" mechanism that can shift from location to location about every 50 ms (Buschman and Miller, 2009; Wolfe et al., 2011). However, under some circumstances, only a subset of the array needs to be searched. Simple features, such

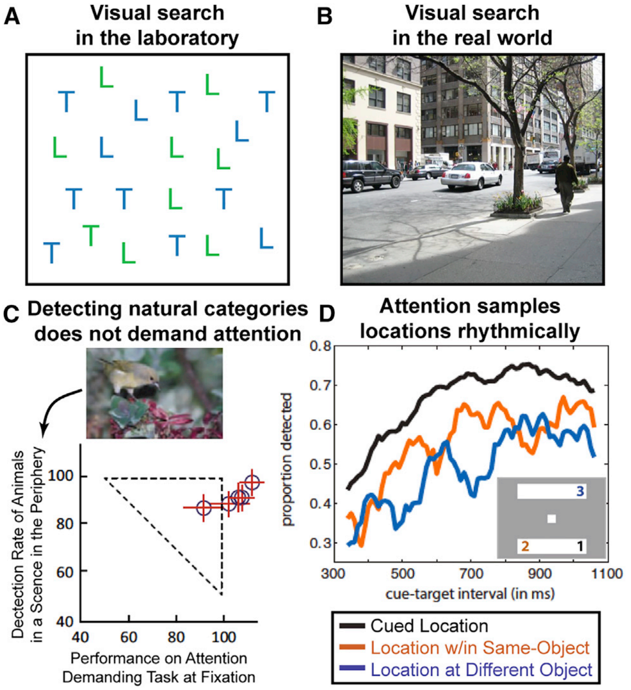

Figure 1. Behavioral Studies

(A) Visual search in artificial displays versus in real-world scenes. Detecting the presence of a green T (conjunction search) is effortful and time-consuming, such that reaction times increase as a function of display items.

(B) In contrast, detecting the presence of categorical object information such as "people" or "cars" in real-world scenes requires only a single glance, despite the large number and variety of distracter objects.

(C) The detection of animals or vehicles in natural scenes does not require focused spatial attention. In a dual-task paradigm, subjects performed a central discrimination task, while detecting animals in scenes presented in the periphery. Performance is normalized to a condition when only a single task was performed. Performance in the peripheral animal detection task was only mildly impaired by simultaneously performing the central discrimination task. Adapted from Li et al. (2002).

(D) Selective attention has rhythmic properties. Subjects detected the dimming of a part of a rectangular object at a spatially cued location (black line; location 1 in the two-object display depicted as an example), at an uncued location of the same object than the cued location (orange line; location 2), or at an uncued location of a different object than the cued location (blue line; location 3). Accuracy is plotted as a function of the cue-target interval revealing the following rhythmic properties: at the cued location, detection performance fluctuated at 8 Hz, whereas at the same- and different-object locations a characteristic anti-phase relationship of a 4-Hz rhythm was observed. Adapted from Fiebelkorn et al. (2013).

as color, can be used to guide the search to just those elements that share a particular target feature (Wolfe et al., 1989). Search difficulty also depends on the similarity of the target to the distracters and to the dissimilarity of the distracters to each other (Duncan and Humphreys, 1989).

The results of studies using classical attention paradigms have shaped our current theoretical concepts and have been foundational for investigations at the neural level that we will review below. However, attention mechanisms have evolved to function in real-world scenarios. Recently, there have been a growing number of studies that have asked whether the knowledge and concepts gained from simplified laboratory conditions translate to more ecologically relevant situations.

Real-World Visual Search

An important first step to investigate attentional prioritization under more naturalistic conditions has been to study the selection of categorical object information from natural scenes (for an in-depth review, see Peelen and Kastner, 2014). In daily life, we select meaningful objects from meaningful scenes such as looking for cars when crossing a street. What would be the behavioral prediction for detecting a car in the scene of Figure 1B on the basis of classic search paradigms? Typical scenes contain dozens of distracter objects with highly variable appearance, and there is not one feature that uniquely defines a target. On the basis of classical attention theories, one would predict a long response time reflecting a particularly inefficient search. However, the opposite is the case. The detection of familiar object categories in scenes is extremely rapid (Thorpe et al., 1996), and the search is highly efficient—adding additional items to a scene has little cost (Wolfe et al., 2011). Furthermore, one can accurately perform such real-world search tasks while simultaneously performing a second, attention-demanding task at fixation (Figure 1C; Li et al., 2002). This suggests that real-world search of object categories does not require focused spatial attention.

Neuroimaging studies in humans have begun to investigate the neural basis of real-world search by having subjects detect the presence of objects from a target category in briefly presented photographs (Peelen and Kastner, 2011; Peelen et al., 2009) or short movie segments (Çukur et al., 2013). It was found that the pattern of neural activity in object-selective cortex evoked by the scenes fully depended on task relevance: target objects embedded in natural scenes were only represented when one was actively searching for them. Responses in many parts of the brain increased with the appearance of a stimulus in the target category, or a semantically similar category, suggesting that category-based attention may have widespread influences on brain activity. Together, these results provide neural evidence that the attentional selection mechanism that biases the processing of scenes acts at the level of natural categories. Future work is needed to extend our traditional concepts of attention to incorporate mechanisms that are optimized for naturalistic conditions. The key to this will be the development of appropriate paradigms in animal models in order to study the underlying neural mechanisms in greater detail.

Rhythmic Properties of Selective Attention

Classic attention theories (Posner et al., 1980; Treisman and Gelade, 1980) propose a unique and indivisible spotlight of attention that highlights a selected item. To process an entire scene, this spotlight was thought to be continuously moving from location to location, shifting at a rate of approximately 20 Hz (Wolfe et al., 2011). Previous studies suggested that this shifting may be regular, moving the spotlight of attention in a rhythmic fashion around a visual scene (Buschman and Miller, 2009). Surprisingly, recent evidence shows that even when this spotlight is sustained at one location, it is not static, but rather appears to flash rhythmically. Using electroencephalogram (EEG), Busch and VanRullen (2010) demonstrated that the detection of a visual target at threshold was systematically related to the phase of an ongoing theta oscillation (~7 Hz). This phase-behavior relationship was contingent on the allocation of attentional resources following a cue and was absent at other locations in the visual