Cognitive influences on attention

In 1880, William James famously defined attention as taking possession by the mind of one out of several simultaneously possible objects or train of thoughts. The modern study of attention continues to work within this broad definition. It is generally accepted that an individual is only aware of a small fraction of the information provided to the brain by the sensory systems. Attention is the name given to the process which governs which material enters awareness and which does not. The present article focuses on how cognitive factors—goals and expectations of an observer—influence visual attention.

Whether a given stimulus is attended depends both on its inherent salience and the state of the observer. The interplay between these two factors is exemplified by the following two situations: (1) Imagine searching for a friend in a crowd. As one searches through the crowd, attention may be captured by elements that are inherently salient, for instance, a person in a red coat among people in black coats or by dynamic cues such as a person running. However, where one looks in the crowd also depends on one's knowledge. So, if the friend is known to be wearing a black coat and a blue hat, attention may be less likely to be captured by an otherwise salient red coat, but perhaps be misdirected to a blue coat as one searches for anything blue in the crowd. (2) Imagine walking by a golf course and worrying about being hit by a golf ball. Although it is likely that a quickly moving white thing in one's visual field will be noticed under most circumstances, the act of walking by a golf course and worrying about being hit by a golf ball may lower the threshold for detecting quickly moving white things. Directing one's attention to blue things because blue is currently relevant and increasing sensitivity to detecting moving white things are instances of *cognitive influences on attention*. Traditional accounts of attention have placed little emphasis on such influences in comparison to factors which were thought to automatically capture attention (e.g., a red thing among black things). However, it can be argued that the very purpose of visual attention is the selection of information most relevant to a present goal. Hence, the effects of goals, expectations, recent history, and even emotions on visual attention have become an area of very active research.

Studying attention: behavioral methods

A commonly used paradigm for studying visual attention is the eponymously named Posner cuing task. The basic version of the task requires subjects to press a button anytime they detect a small circle (the target) while looking at the center of a screen without moving their eyes. Prior to the appearance of the target, a light (the cue) flashes on the left or the right side of the screen. The location of the cue either coincides with the location of the subsequently appearing target (valid trials) or does not (invalid trials). The basic finding is that reaction times are shorter on valid trials than on uncued trials, and are slowest on invalid trials. The interpretation is that attention is automatically "deployed" to the cued region. Targets that appear in the attended region are processed faster than targets appearing in an unattended region (which require an attentional shift from the previously cued region).

In addition to flashing lights, the cues can be symbolic such as right or left arrows presented in the center of the screen. Classic studies from the late 1970's showed that flashing lights (also called exogenous or peripheral cues) elicit attentional shifts even when they do not predict the location of the target. In contrast, arrow cues (also called endogenous or central cues) only produce shifts of attention when they are predictive (e.g., 80% of the time the arrow predicts the position of the target). These results have been interpreted to mean that, unlike flashing lights which automatically capture attention, central cues need to be cognitively interpreted and will shift attention only if subjects have a reason to process them. These classic findings led to a dichotomy between automatic (stimulus-driven, bottom-up, exogenous) attentional processes and controlled (cognitive, top-down, endogenous) processes.

Controversies

Recent studies have argued against this dichotomy in favor of a view in which (1) learned associations determine whether nonpredictive endogenous cues elicit attentional shifts, and (2) highly salient cues can fail to capture attention if they conflict with the viewer's goals. For example, studies from the laboratory of Alan Kingstone have shown that pictures of eyes elicit attentional shifts in the direction of their gaze even when the direction does not predict the target. Similarly, nonpredictive arrows and even printed words like "up" and "down" elicit attentional shifts. Conversely, whether a traditional exogenous cue like a unique color (e.g., a patch of red among greens) captures attention, appears to depend on how the viewer is processing the scene. Consider performance on a task developed by Jan Theeuwes to study the degree to which various visual properties automatically capture attention. In this task, participants are presented with shapes arranged on an imaginary circle around a central fixation point. The goal is to report whether a line that appears in a target shape is, for example, vertically oriented. In a basic version of this task, the target is defined by its unique shape (e.g., it is the only diamond among circles). On distractor trials, the display appears with one of the non-target shapes in a different color from the rest. Because color is irrelevant to the task, greater reaction times on distractor trials indicate that the unique color automatically captured attention. Based on such findings, Theeuwes and colleagues argued that, uniquely colored or shaped objects (singletons) automatically capture attention. Howard Egeth and colleagues challenged this conclusion by showing that whether unique objects capture attention depends on the processing mode of the viewer. If the viewer is in a “singleton detection mode,” tuned to detect unique objects, attention is broadly focused and is indeed captured by task-irrelevant singletons. However, if the viewer is specifically looking for a certain feature such as a diamond shape (“feature detection mode”) then salient, but task-irrelevant distractors do not capture attention. Nevertheless, there do exist properties that capture attention regardless of task relevance or processing mode. One such property is “sudden onset.” A suddenly appearing object generally captures attention. However, when the object is task-irrelevant, attention is disengaged quite quickly (typically in less than 100 ms).

Although there is now wide agreement that the viewers’ goals can affect which objects or features are attended, the locus of these effects remains highly controversial. Does having a goal like “look for the red things” change the priority of redness, but not affect visual processing? Or, does the goal actually change how red things are represented throughout the visual system? Traditional accounts have denied the latter claim. For example, Zenon Pylyshyn has argued for the existence of an early vision system—a modular system that is encapsulated from information outside vision such as the observer’s knowledge and goals, and is thus “cognitive impenetrable.” Support for the claim that attention changes basic visual processes has come from behavioral studies showing that attended objects are actually perceived as more salient (e.g., brighter), from electrophysiological studies on non-human animals, and neuroimaging studies on humans.

Effects of Goals on Visual Processing and Attention: Evidence from Electrophysiology and Neuroimaging

Electrophysiological studies on behaving animals have allowed researchers to isolate the influence of higher-level influences on attention from the processing that reflects the physical properties of the stimuli. For instance, because neurons in the primary visual cortex (V1) fire most to bars with a certain orientation, a researcher can compare the firing of the neuron to a vertical bar when it is task-relevant versus irrelevant. Any difference in neuron's firing rate to a particularly tuned bar between the first and second task reflects the demands of the task because the physical stimulus is identical in both cases. Such studies generally show that V1 neurons fire more vigorously when their preferred orientation is behaviorally relevant.

An immediate implication of such findings is that responses of neurons even in V1—the first part of cortex to receive visual input—reflect not simply the physical characteristics of a stimulus, but also the cognitive goals of the observer. More recent findings have shown that primary sensory neurons have two types of receptive fields (RFs). The first is the so-called classical RF and corresponds roughly to what is observed in anesthetized animals and is the initial response of a neuron in an awake animal. The classical RF of a V1 neuron is a line segment of a certain orientation projected into a very specific part of the visual field. Within a short time period (often under 50 ms, and sometimes as short as 2 ms), the classical RF is modulated by higher-level information including the goals of the observer, the visual context, and the organization of the scene— producing the non-classical RF. While the classical RF of a V1 neuron includes only positional and orientation information, the non-classical RF includes information such as whether the bar is part of a figure, the background, or an object boundary and whether the figure that the segment is a part of is behaviorally relevant.

In similar studies measuring firing rates of V4 neurons (sensitive to color properties of a stimulus), attentional capture of task-irrelevant color singletons is reflected in high firing rates which peak at ~120 ms after stimulus onset. When the task requires the monkey to ignore the color singleton, one can observe neural responses to a task-relevant color continue to remain at a high level, while responses to task-irrelevant color singletons become down-modulated after ~75 ms.

Is early visual neural activity immune to cognitive influences? In electrophysiological studies, it has been found that when a task is performed repeatedly such as attending to a vertical bar for numerous consecutive trials, the classical RF may disappear entirely—the neurons’ response being immediately modulated by the current task. Recent neuroimaging work in humans confirms the conclusion that activity in anatomically early visual areas is permeable to cognitive influences. For instance, when human observers are trained to associate cues with either color or attention, presenting the cue alone modulates activity in visual areas (fusiform gyrus for color; lingual gyrus for location) and, crucially, the amount of modulation strongly predicts performance on the upcoming target-detection trials. This suggests that early visual processing can be tuned by goals and expectations.

Recall that attention is thought to be closely linked to awareness (one notable exception is the phenomenon of blindsight). A claim that a red circle among black circles automatically captures attention generally means that one becomes aware of the red circle even if one's goal is to avoid it. Would showing that some early neural activity evoked by the red circle is impermeable to top-down influences be evidence that automatic awareness of the red circle is directly subserved by those early neural activity? Recent work suggests that such a conclusion is unwarranted. Rather, it appears there may be no awareness without top-down modulation of early visual representations. Much of this evidence has come from studies relying on Event Related Potentials (ERPs) Transcranial Magnetic Stimulation (TMS). Both methods rely on the earlier time-course of bottom-up versus the later time-course of top-down processes to map their respective contributions. In one study, subjects detected visual figures which were briefly presented and then concealed with a pattern mask. By correlating subjects' performance with electrical potential measured by electrodes on the scalp, J. Fahrenfort and colleagues showed that bottom-up activity, which peaked at ~120 ms after stimulus onset was not correlated with conscious perception, while top-down (recurrent) activity peaking later (160 ms) was. Thus, it appears that top-down modulation in anatomically "early" visual areas (e.g., V1) by higher-level regions (e.g., prefrontal cortex) is necessary for visual awareness. Further evidence for the causal role of recurrent processing comes from TMS studies in which a high-intensity magnetic field is briefly applied to a selected region of a subject's scalp as he or she performs a task. This pulse creates a temporary disruption in neural processing. Several recent studies have shown that disrupting feedback activity in early visual areas disrupts awareness.

Conclusion

The study of attention has classically focused on the physical characteristics that determine whether stimuli are attended. Recent studies have shifted the focus to cognitive factors such as expectations and goals of the viewer. These studies show that neural activity causally linked to awareness is deeply permeated by cognitive factors. Although highly controversial, one conclusion is that it may be impossible to fully study visual attention by separating observers from their goals and environments.

Gary Lupyan

University of Pennsylvania

See also

Attention and Consciousness; Attention: Physiological; Object Perception; Illusory (Non-Veridical) Perception; Neural Representation/Coding

Further Reading

Fahrenfort, J., Scholte, H., & Lamme, V. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, *19*(9), 1488-1497.

Foxe, J., & Simpson, G. (2002). Flow of activation from V1 to frontal cortex in humans - A framework for defining "early" visual processing. *Experimental Brain Research*, *142*(1), 139-150.

Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, *1080*(1), 63-72. doi: 10.1016/j.brainres.2005.09.068.

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nat Rev Neurosci*, *1*(2), 91-100. doi: 10.1038/35039043.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! it is time to take a look at the real world. *Current Directions in Psychological Science*, *12*(5), 176-180. doi: 10.1111/1467-8721.01255.

Lamme, V., & Roelfsema, P. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571-579.

Lamy, D., & Egeth, H. E. (2003). Attentional capture in singleton-detection and feature-search modes. *Journal of Experimental Psychology. Human Perception and Performance*, *29*(5), 1003-20. doi: 10.1037/0096-1523.29.5.1003.

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology-General*, *109*(2), 160-174.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341-+.

Theeuwes, J., & Van der Burg, E. (2007). The role of spatial and nonspatial information in visual selection. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(6), 1335-51. doi: 10.1037/0096-1523.33.6.1335.