# Which came first, Attention or Perception?

Consider a visual search task in which an observer is presented with an array of 15 green letters and 1 red letter and must report whether the red letter is a T. A task such as this can generally be performed quite easily but the recall of which letters were found on the green cards is found to be quite poor, which indicates that attention was indeed focused on the one red item.

Usually, attention is not attracted to a single location in the image, but rather to an extended region that is likely to constitute an object or a part of an object. It is only after the attention system has selected a region of the visual field that objects are identified by the recognition system. This makes object based attention a chicken and egg problem that is rather difficult to disentangle.

To get to the depths, we question, why, exactly, is the observer unable to report the green letters? One possibility is that attention operates at a very early stage such that the observer literally does not see the green items. Another possibility is that low level, feature based processing occurs for all items, but integrated object representations are formed only for attended items, making it impossible for the observer to report anything more than the basic features of the green items. A third possibility is that every item in the array is fully identified, but only attended items are stored in working memory so that they can be reported.

How can these three models of attention be empirically distinguished?

In an experiment conducted by Ford et. at. subjects performed a visual search task with arrays of 12 squares, 10 of which were gray and 2 of which were colored, and they were required to report whether a particular color was present in each array. In some trial blocks subjects performed this task alone, and in other trial blocks this task was combined with a central task in which subjects had to report whether a letter presented at fixation was a consonant or a vowel. The letter was degraded with random visual noise. The subjects were told that the central task was to be considered the primary task and that they should devote all of their attention to this task. The ERPs from this experiment are shown in the fig. A small but consistent N2pc (ERP component linked to selective attention) component was observed when the visual search task was performed in isolation. When the central task was added, however, the N2pc component was completely eliminated. This result indicates that the perceptual-level attentional mechanism is not necessary for accurate feature detection.
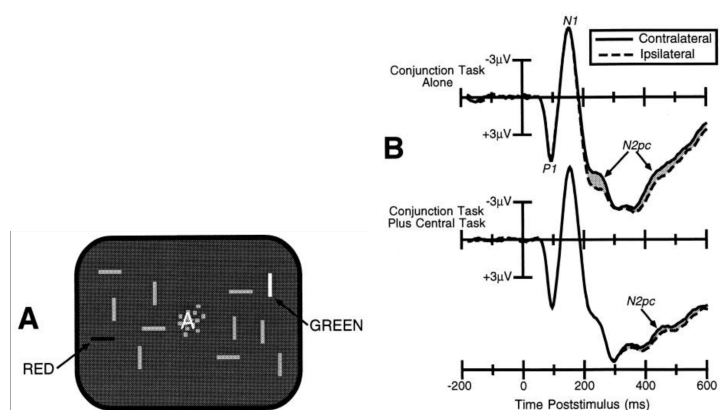
Hence low-level, feature-based processing occurs first for all items. This extracts features for orientations, intensities, colors etc and builds up a saliency maps which give the relevant locations for the further process. Attention is directed to these relevant locations and acts as a spotlight on these locations. I choose to use the term 'location' as I believe the distinction between objects and locations blurs in many ways, for example if the shape of a spatial spotlight is allowed to deform around an object.

Original image   Saliency map   Attended regions   Not attended regions
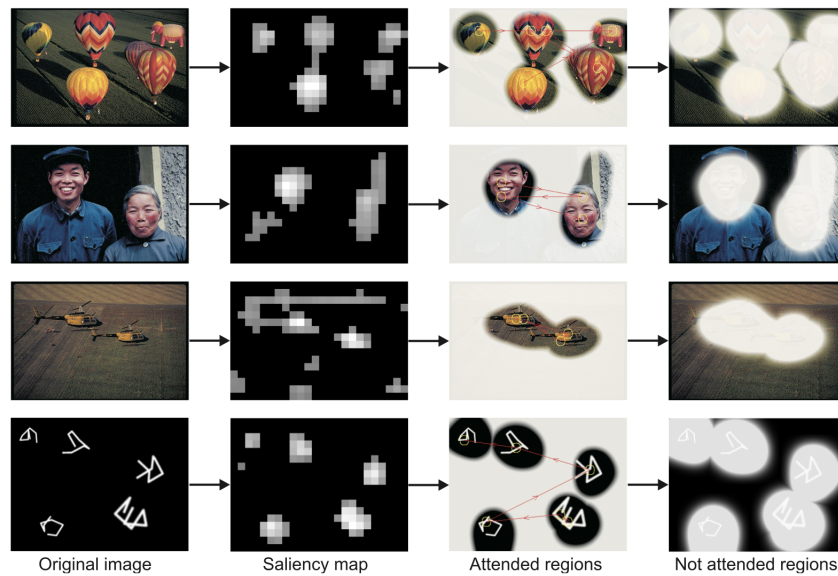
Hence instead of a circular problem, there is some level of perceptual processing that occurs before and helps in attentional selection. As for the relationship between object and space based based attention, I believe attentional selection is first directed to a spatial location by preliminary perceptual processes and then from the specific spatial location an object contour is selected. My opinion on the debate that is concered about object vs spatial selection, various examples stated by Scholl such as the handgame and the ballgame experiment, same object advantage experiment are convincing enough that even though initially attention is directed to a spatial location, there is definite selection of object contour later.
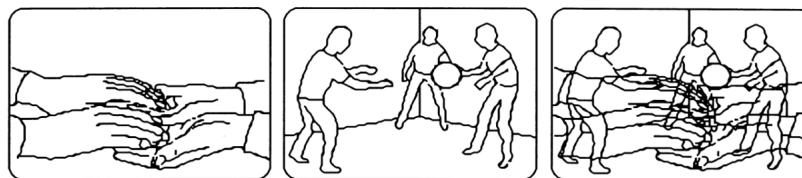

Fig. 1. Sample displays from Neisser and Becklen (1975). The two scenes – the 'hand game' and the 'ballgame' – are superimposed, and subjects are then induced to attend to only one of them, for example to count the number of times the hands clap each other. In this case, subjects fail to perceive incredible sustained events which occur in the other scene, despite the superimposition. See text for details.
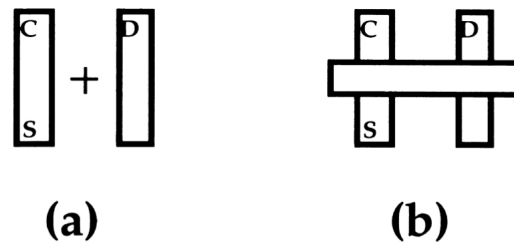
Fig. 3. Stimuli from various experiments used to demonstrate 'same-object advantages' in the automatic spread of attention: (a) Egly, Driver, and Rafal (1994); (b) Moore, Yantis, and Vaughan (1998). In each case 'C' indicates the cued location, 'S' indicates a same-object target location, and 'D' indicates a different-object target location. See text for details. Note that the Moore et al. study actually used a slightly different task. (Adapted from Egly, Driver, and Rafal (1994) and Moore et al. (1998).)

References:

1. Steve Luck, Michael Ford : "On the role of selective attention in visualperception"

2. Dirk Walther, Laurent Itti, Maximilian Riesenhuber, Tomaso Poggio, and Christof Koch : "Attentional Selection for Object Recognition { a Gentle Way"

3. Brian J. Scholl : "Objects and attention: the state of the art"