

Visual Search: How Your Mind Finds Targets in the Crowd

Authored by
mohammad looti

June 16, 2026

RECOMMENDED CITATION

mohammad looti (2026). *Visual Search: How Your Mind Finds Targets in the Crowd*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=38073>

Visual search is a type of perceptual task requiring attention that typically involves an active scan of the visual environment for a particular object or feature (the target) among other objects or features (the distractors). Visual search can take place with or without eye movements. The ability to consciously locate an object or target amongst a complex array of stimuli has been extensively studied over the past 40 years. Practical examples of using visual search can be seen in everyday life, such as when one is picking out a product on a supermarket shelf, when animals are searching for food amongst piles of leaves, when trying to find your friend in a large crowd of people, or simply when playing visual search games such as Where's Wally? Many visual search paradigms have used eye movement as a means to measure the degree of attention given to stimuli. However, vast research to date suggests that eye movements move independently of attention, and therefore are not a reliable method to examine the role of attention. Much previous literature on visual search uses reaction time in order to measure the time it takes to detect the target amongst its distractors. An example of this could be a green square (the target) amongst a set of red circles (the distractors).

Search Types

Feature Search

Feature search (also known as "disjunctive" or "efficient" search) is a visual search process that focuses on identifying a previously requested target amongst distractors that differ from the target by a unique visual feature such as color, shape, orientation, or size. An example of a feature search task is asking a participant to identify a white square (target) surrounded by black squares (distractors). In this type of visual search, the distractors are characterized by the same visual features. The efficiency of feature search in regards to reaction time (RT) and accuracy depends on the "pop out" effect, bottom-up processing, and parallel processing. However, the efficiency of feature search is unaffected by the number of distractors present. The "pop out" effect is an element of feature search that characterizes the target's ability to stand out from surrounding distractors due to its unique feature. Bottom-up processing, which is the processing of information that depends on input from the environment, explains how one utilizes feature detectors to process characteristics of the stimuli and differentiate a target from its distractors. This draw of visual attention towards the target due to bottom-up processes is known as "saliency." Lastly, parallel processing is the mechanism that then allows one's feature detectors to work simultaneously in identifying the target.

Conjunction Search

Conjunction search (also known as inefficient or serial search) is a visual search process that focuses on identifying a previously requested target surrounded by distractors possessing one or

more common visual features with the target itself. An example of a conjunction search task is having a person identify a red X (target) amongst distractors composed of black Xs (same shape) and red Os (same color). Unlike feature search, conjunction search involves distractors (or groups of distractors) that may differ from each other but exhibit at least one common feature with the target. The efficiency of conjunction search in regards to reaction time(RT) and accuracy is dependent on the distractor-ratio and the number of distractors present. As the distractors represent the differing individual features of the target more equally amongst themselves(distractor-ratio effect), reaction time(RT) increases and accuracy decreases. As the number of distractors present increases, the reaction time(RT) increases and the accuracy decreases. However, with practice the original reaction time(RT) restraints of conjunction search tend to show improvement. In the early stages of processing, conjunction search utilizes bottom-up processes to identify pre-specified features amongst the stimuli. These processes are then overtaken by a more serial process of consciously evaluating the indicated features of the stimuli in order to properly allocate one's focal spatial attention towards the stimulus that most accurately represents the target. In many cases, top-down processing affects conjunction search by eliminating stimuli that are incongruent with one's previous knowledge of the target-description, which in the end allows for more efficient identification of the target. An example of the effect of top-down processes on a conjunction search task is when searching for a red 'K' among red 'Cs' and black 'Ks', individuals ignore the black letters and focus on the remaining red letters in order to decrease the set size of possible targets and, therefore, more efficiently identify their target.

Real World Visual Search

In everyday situations, people are most commonly searching their visual fields for targets that are familiar to them. When it comes to searching for familiar stimuli, top-down processing allows one to more efficiently identify targets with greater complexity than can be represented in a feature or conjunction search task. In a study done to analyze the reverse-letter effect, which is the idea that identifying the asymmetric letter amongst symmetric letters is more efficient than its reciprocal, researchers concluded that individuals more efficiently recognize an asymmetric letter amongst symmetric letters due to top-down processes. Top-down processes allowed study participants to access prior knowledge regarding shape recognition of the letter N and quickly eliminate the stimuli that matched their knowledge. In the real world, one must use his prior knowledge everyday in order to accurately and efficiently locate his phone, keys, etc. amongst a much more complex array of distractors. While bottom-up processes may come into play when identifying objects that are not as familiar to a person, overall top-down processing highly influences visual searches that occur in everyday life.

Reaction Time Slope

It is also possible to measure the role of attention within visual search experiments by calculating the slope of reaction time over the number of distractors present. Generally, when high levels of attention are required when looking at a complex array of stimuli (conjunction search), the slope increases as the reaction times increase. For simple visual search tasks (feature search), the slope decreases due to reaction times being fast and requiring less attention.

Visual Orienting and Attention



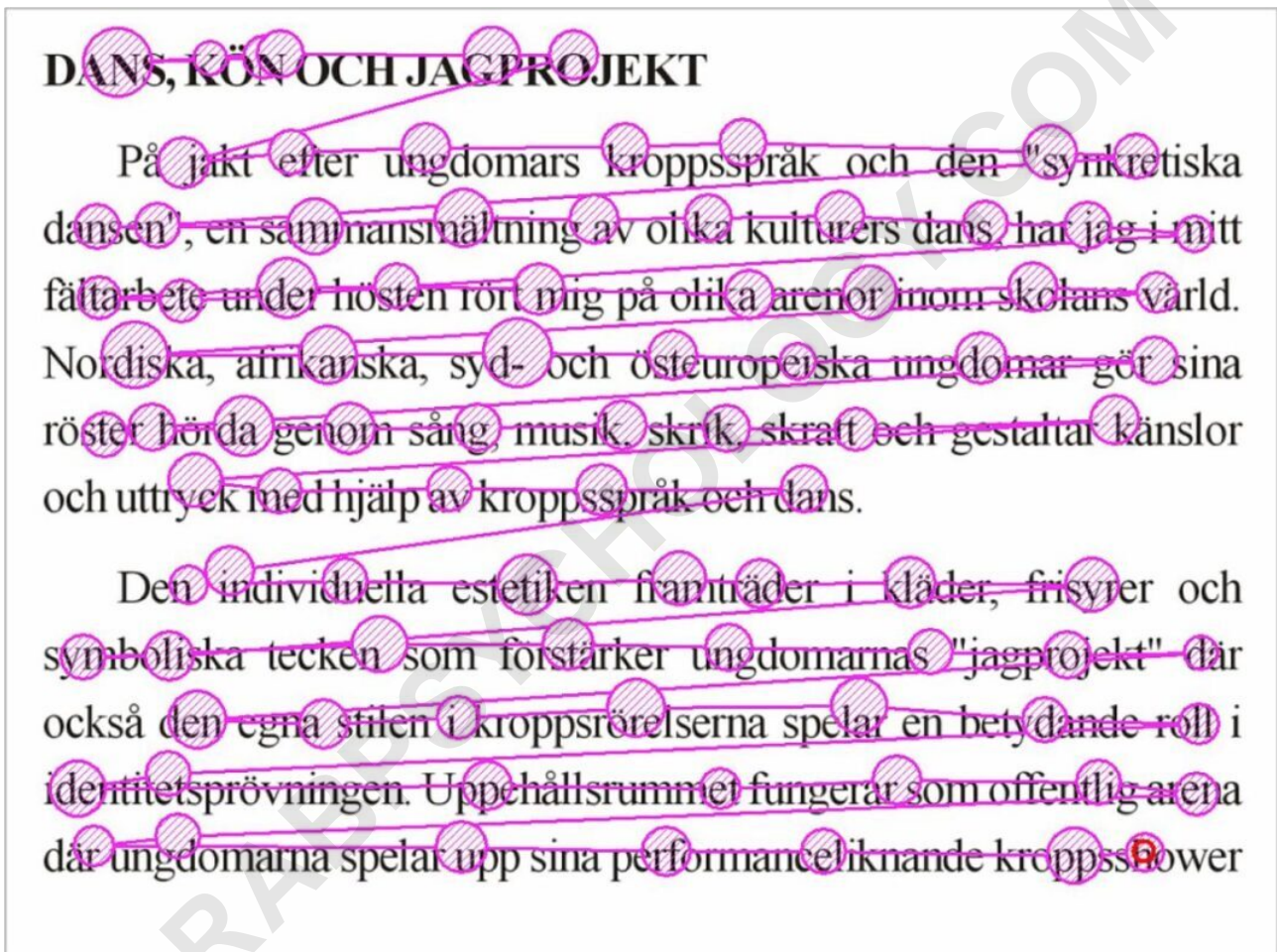
A photograph that simulates Foveation

One obvious way to select visual information is to turn towards it, also known as visual orienting. This may be a movement of the head and/or eyes towards the visual stimulus, called a saccade. Through a process called foveation, the eyes fixate on the object of interest, making the image of the visual stimulus fall on the fovea of the eye, the central part of the retina with the sharpest visual acuity.

There are two types of orienting:

Exogenous orienting is the involuntary and automatic movement that occurs to direct one's visual attention toward a sudden disruption in his peripheral vision field. Attention is therefore externally guided by a stimulus, resulting in a reflexive saccade.

Endogenous orienting is the voluntary movement that occurs in order for one to focus his visual attention on a goal-driven stimulus. Thus, the focus of attention of the perceiver can be manipulated by the demands of a task. A scanning saccade is triggered endogenously for the purpose of exploring the visual environment.



A plot of the saccades made while reading text. The plot shows the path of eye movements and the size of the circles represents the time spent at any one location.

Visual search relies primarily on endogenous orienting because participants have the goal to detect the presence or absence of a specific target object in an array of other distracting objects.

Visual orienting does not necessarily require overt movement, though. It has been shown that people can covertly (without eye movement) shift attention to peripheral stimuli. In the 1970s, it was found that the firing rate of cells in the parietal lobe of monkeys increased in response to

stimuli in the receptive field when they attended to peripheral stimuli, even when no eye movements were allowed. These findings indicate that attention plays a critical role in understanding visual search.

Subsequently, competing theories of attention have come to dominate visual search discourse. The environment contains a vast amount of information. We are limited in the amount of information we are able to process at any one time, so it is therefore necessary that we have mechanisms by which extraneous stimuli can be filtered and only relevant information attended to. In the study of attention, psychologists distinguish between preattentive and attentional processes. Preattentive processes are evenly distributed across all input signals, forming a kind of "low-level" attention. Attentional processes are more selective and can only be applied to specific preattentive input. A large part of the current debate in visual search theory centres on selective attention and what the visual system is capable of achieving without focal attention.

Theory

Feature Integration Theory (FIT)

A popular explanation for the different reaction times of feature and conjunction searches is the feature integration theory (FIT), introduced by Treisman and Gelade in 1980. This theory proposes that certain visual features are registered early, automatically, and are coded rapidly in parallel across the visual field using preattentive processes. Experiments show that these features include luminance, colour, orientation, motion direction, and velocity, as well as some simple aspects of form. For example, a red X can be quickly found among any number of black Xs and Os because the red X has the discriminative feature of colour and will "pop out." In contrast, this theory also suggests that in order to integrate two or more visual features belonging to the same object, a later process involving integration of information from different brain areas is needed and is coded serially using focal attention. For example, when locating an orange square among blue squares and orange triangles, neither the colour feature "orange" nor the shape feature "square" is sufficient to locate the search target. Instead, one must integrate information of both colour and shape to locate the target.

Evidence that attention and thus later visual processing is needed to integrate two or more features of the same object is shown by the occurrence of illusory conjunctions, or when features do not combine correctly. For example, if a display of a green X and a red O are flashed on a screen so briefly that the later visual process of a serial search with focal attention cannot occur, the observer may report seeing a red X and a green O.

The FIT is a dichotomy because of the distinction between its two stages: the preattentive and attentive stages. Preattentive processes are those performed in the first stage of the FIT model, in

which the simplest features of the object are being analyzed, such as color, size, and arrangement. The second attentive stage of the model incorporates cross-dimensional processing, and the actual identification of an object is done and information about the target object is put together. This theory has not always been what it is today; there have been disagreements and problems with its proposals that have allowed the theory to be amended and altered over time, and this criticism and revision has allowed it to become more accurate in its description of visual search. There have been disagreements over whether or not there is a clear distinction between feature detection and other searches that use a master map accounting for multiple dimensions in order to search for an object. Some psychologists support the idea that feature integration is completely separate from this type of master map search, whereas many others have decided that feature integration incorporates this use of a master map in order to locate an object in multiple dimensions.

The FIT also explains that there is a distinction between the brain's processes that are being used in a parallel versus a focal attention task. Chan and Hayward have conducted multiple experiments supporting this idea by demonstrating the role of dimensions in visual search. While exploring whether or not focal attention can reduce the costs caused by dimension-switching in visual search, they explained that the results collected supported the mechanisms of the feature integration theory in comparison to other search-based approaches. They discovered that single dimensions allow for a much more efficient search regardless of the size of the area being searched, but once more dimensions are added it is much more difficult to efficiently search, and the bigger the area being searched the longer it takes for one to find the target.

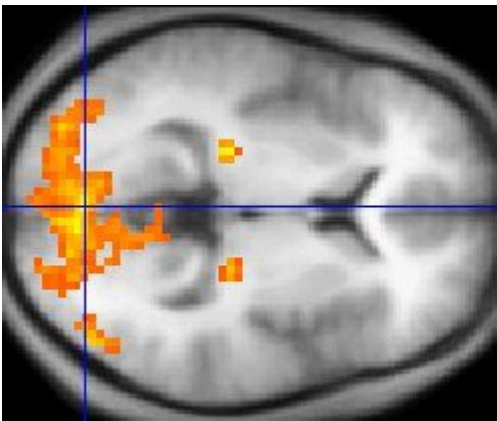
Guided Search Model

A second main function of preattentive processes is to direct focal attention to the most "promising" information in the visual field. There are two ways in which these processes can be used to direct attention: bottom-up activation (which is stimulus-driven) and top-down activation (which is user-driven). In the guided search model by Jeremy Wolfe, information from top-down and bottom-up processing of the stimulus is used to create a ranking of items in order of their attentional priority. In a visual search, attention will be directed to the item with the highest priority. If that item is rejected, then attention will move on to the next item and the next, and so forth. The guided search theory follows that of parallel search processing.

An activation map is a representation of visual space in which the level of activation at a location reflects the likelihood that the location contains a target. This likelihood is based on preattentive, featural information of the perceiver. According to the guided search model, the initial processing of basic features produces an activation map, with every item in the visual display having its own level of activation. Attention is demanded based on peaks of activation in the activation map in a search for the target. Visual search can proceed efficiently or inefficiently. During efficient search,

performance is unaffected by the number of distractor items. The reaction time functions are flat, and the search is assumed to be a parallel search. Thus, in the guided search model, a search is efficient if the target generates the highest, or one of the highest activation peaks. For example, suppose someone is searching for red, horizontal targets. Feature processing would activate all red objects and all horizontal objects. Attention is then directed to items depending on their level of activation, starting with those most activated. This explains why search times are longer when distractors share one or more features with the target stimuli. In contrast, during inefficient search, the reaction time to identify the target increases linearly with the number of distractor items present. According to the guided search model, this is because the peak generated by the target is not one of the highest.

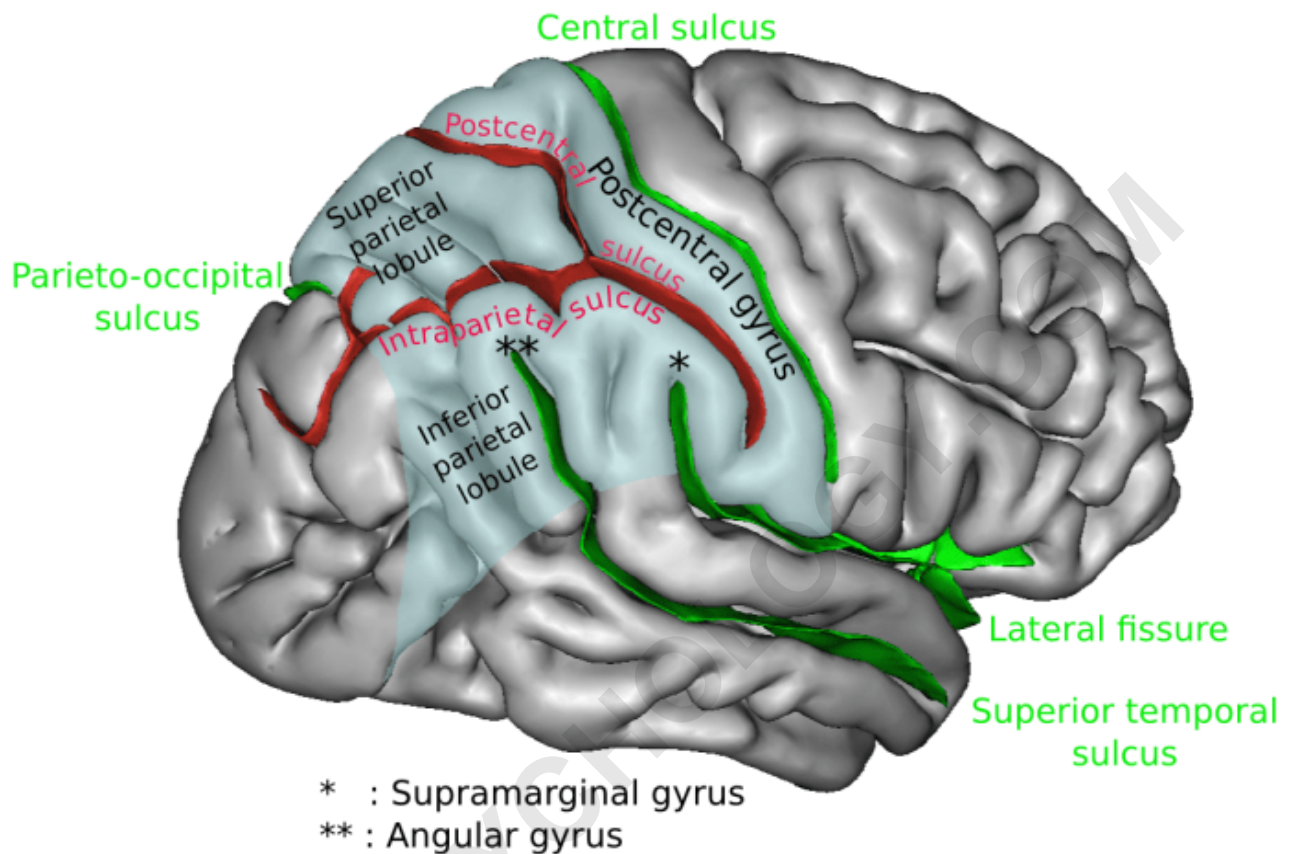
Biological Basis



A pseudo-color image showing activation of the primary visual cortex during a perceptual task using functional magnetic resonance imaging (fMRI)

During visual search experiments the posterior parietal cortex has elicited much activation during functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) experiments for inefficient conjunction search, which has also been confirmed through lesion studies. Patients with lesions to the posterior parietal cortex show low accuracy and very slow reaction times during a conjunction search task but have intact feature search remaining to the ipsilesional (the same side of the body as the lesion) side of space. Ashbridge, Walsh, and Cowey in (1997) demonstrated that during the application of transcranial magnetic stimulation (TMS) to the right parietal cortex, conjunction search was impaired by 100 milliseconds after stimulus onset. This was not found during feature search. Nobre, Coull, Walsh and Frith (2003) identified using functional magnetic resonance imaging (fMRI) that the intraparietal sulcus located in the superior parietal cortex was activated specifically to feature search and the binding of individual perceptual features as

opposed to conjunction search. Conversely, the authors further identify that for conjunction search, the superior parietal lobe and the right angular gyrus elicit bilaterally during fMRI experiments.



Visual search primarily activates areas of the parietal lobe.

In contrast, Leonards, Sunaert, Vam Hecke and Orban (2000) identified that significant activation is seen during fMRI experiments in the superior frontal sulcus primarily for conjunction search. This research hypothesises that activation in this region may in fact reflect working memory for holding and maintaining stimulus information in mind in order to identify the target. Furthermore, significant frontal activation including the ventrolateral prefrontal cortex bilaterally and the right dorsolateral prefrontal cortex were seen during positron emission tomography for attentional spatial representations during visual search. The same regions associated with spatial attention in the parietal cortex coincide with the regions associated with feature search. Furthermore, the frontal eye field (FEF) located bilaterally in the prefrontal cortex, plays a critical role in saccadic eye movements and the control of visual attention.

Moreover, research into monkeys and single cell recording found that the superior colliculus is involved in the selection of the target during visual search as well as the initiation of movements.

Conversely, it also suggested that activation in the superior colliculus results from disengaging attention, ensuring that the next stimulus can be internally represented. The ability to directly attend to a particular stimuli during visual search experiments has been linked to the pulvinar nucleus (located in the midbrain) while inhibiting attention to unattended stimuli. Conversely, Bender and Butter (1987) found that during testing on monkeys, no involvement of the pulvinar nucleus was identified during visual search tasks.

Evolution

There is a variety of speculation about the origin and evolution of visual search in humans. It has been shown that during visual exploration of complex natural scenes, both humans and nonhuman primates make highly stereotyped eye movements. Furthermore, chimpanzees have demonstrated improved performance in visual searches for upright human or dog faces, suggesting that visual search (particularly where the target is a face) is not peculiar to humans and that it may be a primal trait. Research has suggested that effective visual search may have developed as a necessary skill for survival, where being adept at detecting threats and identifying food was essential.



Henri Rousseau, Jungle with Lion

The importance of evolutionarily relevant threat stimuli was demonstrated in a study by LoBue and DeLoache (2008) in which children (and adults) were able to detect snakes more rapidly than other targets amongst distractor stimuli.

Given that the environment in which humans live has changed significantly over time, questions arise as to whether the purpose of visual search is falling away, or whether humans have adapted it to identify new salient targets. Research into the relevance of visual search in modern society has included identifying target nutritional information on product labels, identifying salient features while driving and manipulating consumer shopping habits using different shelf display characteristics. Another modern application of visual search has been the development of artificial visual search engines, such as Google Goggles.

Face Recognition

Over the past few decades there have been vast amounts of research into face recognition, specifying that faces endure specialized processing within a region called the fusiform face area (FFA) located in the mid fusiform gyrus in the temporal lobe. Debates are ongoing whether both faces and objects are detected and processed in different systems and whether both have category specific regions for recognition and identification. Much research to date focuses on the accuracy of the detection and the time taken to detect the face in a complex visual search array. When faces are displayed in isolation, upright faces are processed faster and more accurately than inverted faces, but this effect was observed in non-face objects as well. When faces are to be detected among inverted or jumbled faces, reaction times for intact and upright faces increase as the number of distractors within the array is increased. Hence, it is argued that the 'pop out' theory defined in feature search is not applicable in the recognition of faces in such visual search paradigm. Conversely, the opposite effect has been argued and within a natural environmental scene, the 'pop out' effect of the face is significantly shown. This could be due to evolutionary developments as the need to be able to identify faces that appear threatening to the individual or group is deemed critical in the survival of the fittest. More recently, it was found that faces can be efficiently detected in a visual search paradigm, if the distractors are non-face objects, however it is debated whether this apparent 'pop out' effect is driven by a high-level mechanism or by low-level confounding features. Furthermore, patients with developmental prosopagnosia, suffering from impaired face identification, generally detect faces normally, suggesting that visual search for faces is facilitated by mechanisms other than the face-identification circuits of the fusiform face area.

Patients with forms of dementia can also have deficits in facial recognition and the ability to recognize human emotions in the face. In a meta-analysis of nineteen different studies comparing normal adults with dementia patients in their abilities to recognize facial emotions, the patients with frontotemporal dementia were seen to have a lower ability to recognize many different emotions. These patients were much less accurate than the control participants (and even in comparison with Alzheimer's patients) in recognizing negative emotions, but were not significantly impaired in recognizing happiness. Anger and disgust in particular were the most difficult for the dementia patients to recognize.

Face recognition is a complex process that has many more factors that can affect one's recognition abilities. Other aspects to be considered include race and culture and their effects on one's ability to recognize faces. Some factors such as the other race effect can influence one's ability to recognize and remember faces. There are so many factors, both environmental and individually internal, that can affect this task that it can be difficult to isolate and study each and every idea.

Considerations

Ageing

Research indicates that performance in conjunctive visual search tasks significantly improves during childhood and declines in later life. More specifically, young adults have been shown to have faster reaction times on conjunctive visual search tasks than both children and older adults, but their reaction times were similar for feature visual search tasks. This suggests that there is something about the process of integrating visual features or serial searching that is difficult for children and older adults, but not for young adults. Studies have suggested numerous mechanisms involved in this difficulty in children, including peripheral visual acuity, eye movement ability, ability of attentional focal movement, and the ability to divide visual attention among multiple objects.

Studies have suggested similar mechanisms in the difficulty for older adults, such as age related optical changes that influence peripheral acuity, the ability to move attention over the visual field, the ability to disengage attention, and the ability to ignore distractors.

A study by Lorenzo-López et al. (2008) provides neurological evidence for the fact that older adults have slower reaction times during conjunctive searches compared to young adults. Event-related potentials (ERPs) showed longer latencies and lower amplitudes in older subjects than young adults at the P3 component, which is related to activity of the parietal lobes. This suggests the involvement of the parietal lobe function with an age-related decline in the speed of visual search tasks. Results also showed that older adults, when compared to young adults, had significantly less activity in the anterior cingulate cortex and many limbic and occipitotemporal regions that are involved in performing visual search tasks.

Alzheimer's Disease

Research has found that people with Alzheimer's disease (AD) are significantly impaired overall in visual search tasks. Surprisingly, AD sufferers manifest enhanced spatial cueing, but this benefit is only obtained for cues with high spatial precision. Abnormal visual attention may underlie certain visuospatial difficulties in patients with (AD). People with AD have hypometabolism and neuropathology in the parietal cortex, and given the role of parietal function for visual attention, patients with AD may have hemispatial neglect, which may result in difficulty with disengaging attention in visual search.

An experiment conducted by Tales et al. (2000) investigated the ability of patients with AD to perform various types of efficient visual search tasks. Their results showed that search rates on the "pop-out" tasks were similar for both AD and control groups, however, people with AD searched significantly slower compared to the control group on the conjunction task. One interpretation of these results is that the visual system of AD patients has a problem with feature binding, such that it is unable to communicate efficiently the different feature descriptions for the stimulus. Binding of features is thought to be mediated by areas in the temporal and parietal cortex, and these areas are known to be affected by AD-related pathology.

Another possibility for the impairment of people with AD on conjunction searches is that there may be some damage to general attentional mechanisms in AD, and therefore any attention-related task will be affected, including visual search.

Tales et al. (2000) detected a double dissociation with their experimental results on AD and visual search. Earlier work was carried out on patients with Parkinson's disease (PD) concerning the impairment patients with PD have on visual search tasks. In those studies, evidence was found of impairment in PD patients on the "pop-out" task, but no evidence was found on the impairment of the conjunction task. As discussed, AD patients show the exact opposite of these results: normal performance was seen on the "pop-out" task, but impairment was found on the conjunction task. This double dissociation provides evidence that PD and AD affect the visual pathway in different ways, and that the pop-out task and the conjunction task are differentially processed within that pathway.

Autism

Studies have consistently shown that autistic individuals performed better and with lower reaction times in feature and conjunctive visual search tasks than matched controls without autism. Several explanations for these observations have been suggested. One possibility is that people with autism have enhanced perceptual capacity. This means that autistic individuals are able to process larger amounts of perceptual information, allowing for superior parallel processing and hence faster target location. Second, autistic individuals show superior performance in discrimination tasks between similar stimuli and therefore may have an enhanced ability to differentiate between items in the visual search display. A third suggestion is that autistic individuals may have stronger top-down target excitation processing and stronger distractor inhibition processing than controls. Keehn et al. (2008) used an event-related functional magnetic resonance imaging design to study the neurofunctional correlates of visual search in autistic children and matched controls of typically developing children. Autistic children showed superior search efficiency and increased neural activation patterns in the frontal, parietal, and occipital lobes when compared to the typically developing children. Thus, autistic individuals' superior performance on visual search tasks may be due to enhanced discrimination of items on the display, which is associated with occipital activity, and increased top-down shifts of visual attention, which is associated with the frontal and parietal areas.

Consumer Psychology

In the past decade, there has been extensive research into how companies can maximise sales using psychological techniques derived from visual search to determine how products should be positioned on shelves. Pieters and Warlop (1999) used eye tracking devices to assess saccades

and fixations of consumers while they visually scanned/searched an array of products on a supermarket shelf. Their research suggests that consumers specifically direct their attention to products with eye-catching properties such as shape, colour or brand name. This effect is due to a pressured visual search where eye movements accelerate and saccades minimise, thus resulting in the consumer's quickly choosing a product with a 'pop out' effect. This study suggests that efficient search is primarily used, concluding that consumers do not focus on items that share very similar features. The more distinct or maximally visually different a product is from surrounding products, the more likely the consumer is to notice it. Janiszewski (1998) discussed two types of consumer search. One search type is goal directed search taking place when somebody uses stored knowledge of the product in order to make a purchase choice. The second is exploratory search. This occurs when the consumer has minimal previous knowledge about how to choose a product. It was found that for exploratory search, individuals would pay less attention to products that were placed in visually competitive areas such as the middle of the shelf at an optimal viewing height. This was primarily due to the competition in attention meaning that less information was maintained in visual working memory for these products.