Perception is Cognition: Selective Attention Mechanisms in Visual Processing

Matthew P. Burruss

1. Introduction

In many ways, perception is cognition. How we visualize the world around us guides our decisions, actions, and beliefs. As we interact with the world, our knowledge then influences how and what we perceive. One key component in this cycle of perception and learning is visual selective attention. Interestingly, visual selective attention is a mechanism that, despite its everyday use, is challenging to specifically define. In many ways, visual selective attention is an emergent mechanism that arises from the whole cognitive system working together. On the other hand, it could be an autonomous, causal mechanism that acts independently of other subsystems in the brain. This paper will take the approach that human visual selective attention system is best defined as a tightly integrated emergent system possessing causal mechanisms that would be necessary for replication in a computational model. Therefore, developments in advancing our understanding of selective visual attention both computationally and theoretically will provide questions related to human cognition at a higher level due to the close and natural link between perception and cognition.

The paper is organized to first look at how the causal properties meet the emergent properties by defining the neural mechanisms of human selective attention and the various ideas and evidences that come together to form the underlying psychological theories (*Section 2).* The emergent properties will then be explained through high-level cognitive experiments that have shown coupling between selective attention and other cognitive phenomenon (*Section 3)*. Finally, four computational models that incorporate attention mechanisms are presented and compared to what we know about human selective attention (*Section 4).*

2. Neural Mechanisms for Selective Visual Attention

The phenomenon of visual selective attention in humans is based upon two fundamental properties. First, we have a limited capacity for processing visual information which leads to a competition among the parts of a visual input. Second, we apply selective, top-down biases to filter the visual information while being affected by bottom-up biases independent to the task at hand but related to our learned knowledge base. These two properties lead to the generalized *biased competition model* of selective visual attention which results in the collection of information that is relevant to behavior (Desimone & Duncan 1995). The following sections highlight the anatomical mechanisms that lead to these basic properties and current theories unifying these findings.

2.1 Attention Networks and Selective Biases

The human visual selective attention network is organized into two primary streams or processing pathways that both begin at the primary visual cortex or V1 (Figure 1, Desimone & Duncan 1995). The *ventral stream* leads to the inferior temporal cortex and is relevant in object recognition and the discovery of behaviorally relevant stimuli (Posner & Peterson 1990). The *dorsal stream* leads to the posterior parietal cortex and is important in spatial recognition and visuomotor activities. Naturally, the question arises as to how these two streams converge to produce a coherent response.

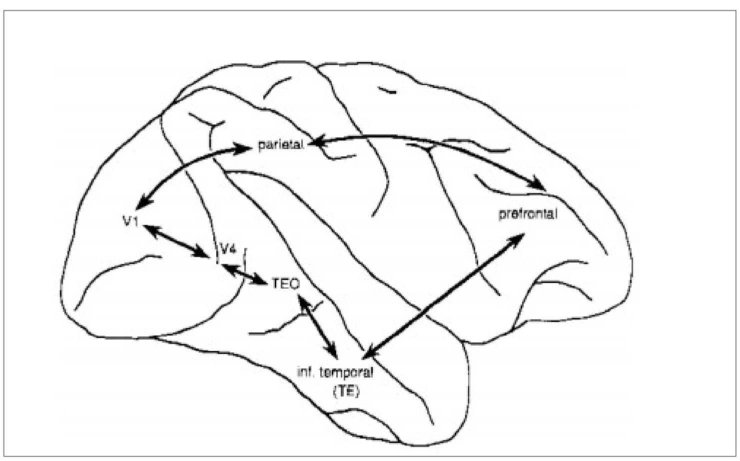

Figure 1. The ventral stream (bottom) and dorsal stream (top) begin at the primary visual cortex (V1) and converge at the prefrontal cortex due to the effects of biases and memory.

It is widely believed that both bottom-up and top-down biases guide the convergence of the two streams. The ventral stream uses bottom-up biases which detect inhomogeneities in the stimulus whereas the dorsal stream is guided by both bottom-up and top-down biases (Posner & Peterson 1990). For example, a bottom-up bias could be a circle in a field of squares and a top-down bias could be a particular stimuli that is task-specific or relevant to a necessary behavioral response (Figure 2). Research has shown that while both approaches are necessary, humans dynamically switch between top-down and bottom-up approaches (Schneider et al. 2012). Therefore, attention along the ventral and dorsal streams can be thought of as influenced by the weighted response of these top-down and bottom-up mechanisms.

Data is also represented differently along these streams. As the visual input passes successive neurons, the complexity of the neuronal properties increase. For example, the cells in V1 largely function as spatiotemporal energy filters but in V2 they respond to illusory contours in the visual stimulus (Heydt et al. 1984). The receptive fields of a neuron, which is a critical processing resource for visual tasks, also increases as one progresses along the ventral and dorsal streams (Mangun 1995). As a paradoxical consequence, more information is processed by later neurons. The following section will explain why this is the case by considering the effects of memory on attention.

2.2 Filtering and Memory: Are we sure we want that?

As the visual input progresses from V1 to other areas of the brain, top-down filtering techniques are applied at successive pathways. These top-down filtering techniques are responsible for answering what defines an object in space. However, at later stages, the primary role of these biases is to determine what information to ignore as the receptive fields increase.

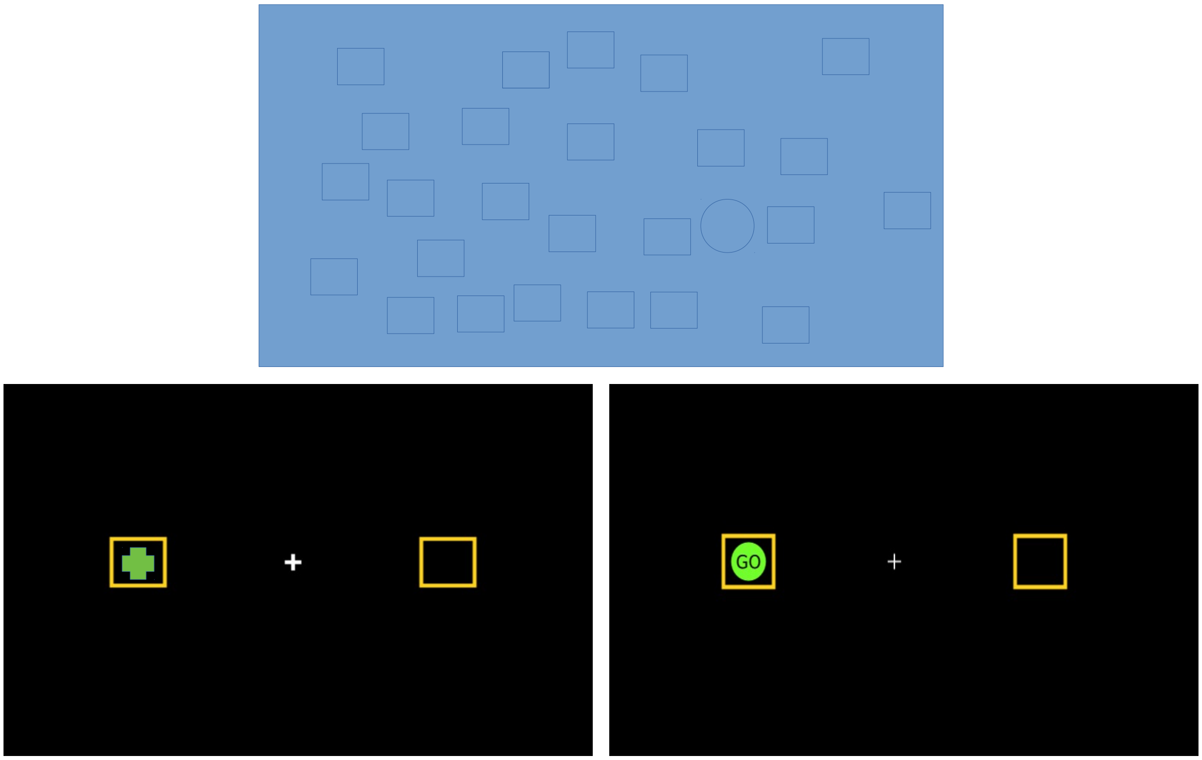

Figure 2. A spatial image affected primarily by bottom-up biases that detect the inhomogeneous shape (top). A spatial cue experiment where the spatial cue (bottom left) is provided, and the subject is asked to quickly determine the box containing the target (bottom right) in which mainly top-down biases are in effect. (bottom figure screenshotted from psytoolkit.org/experiment-library/experime nt_cueing.html)

Mangun 1995 showed that filtering of unwanted information is a two stage process that utilizes memory primarily in the second stage by inhibiting the activation of certain neurons. Moran & Desimone (1985) showed that the first stage works over small ranges where spatial filters are more influential, and the second stage works in larger spatial ranges over the IT cortex where working memory has a greater effect. Convergence of the various visual subsystems therefore is a combination of top-down and bottom-up biases applied in the two attention networks as well as the effect of learned knowledge-bases. Working memory assists in determining what to exclude at later stages along these pathways and where to attend to next in the visual stimuli.

2.3 Integrated Competition Hypothesis & Theory of Visual Selection

Based on the neural mechanisms of attention, two related theories explain how the visual attention network converges to produce a response and how this response is related to our knowledge base in memory.

The first theory is the integrated competition hypothesis (ICH) which presents three guiding principles for how these systems converge based on top-down biases (Duncan et al. 1997). First, the subsystems of the brain are activated in parallel by the visual input. The attention networks then plays an active role in integration and synchronicity that results in convergence towards a response. Finally, task-specific selection is primed by a top-down bias which can exist in memory and/or come directly from the visual stimulus. For example, if a person is told to look for a red object in a field of blue objects, their attention networks will be primed to compete for objects that match that representation. One may notice that bottom-up biases seem to be excluded from ICH; however these biases are independent of the cognitive task at hand and are learned through experience (Duncan et al. 1997). Another consequence of the hypothesis is that attention is an emergent behavior that results from parallel processes where the response is the result of a race of various signals. *Section 4.1* will describe a mathematical framework that uses this idea to model visual attention (Logan 1996).

The second concept is the theory of visual selection which describes how cognition affects perception (Duncan & Humphrey 1989). First, a stage of perceptual description produces a structured representation of the input. Then the input descriptors are matched with an internal attention template of information in LTM which is brought to short term memory (STM). The attention template attaches short-term descriptions to the input which can then be further processed by the attention network. Consequently, attention templates can model the dynamic effect of top-down and bottom-up biases as well as hierarchical representations of attention; therefore, the theory of visual selection is useful in deciding how to construct a computational model of selective visual attention. For example, the cognitive architecture CHREST uses a mechanism similar to attentional templates to incorporate memory into the model of visual attention (*Section 4.2,* Lane et al. 2008).

3. Experimental Evidence of High Level Congitive Effects

It has long been studied that expert chess players are able to accurately recall the configuration of a chess board after a brief visual stimulus (De Groot 1946). Along with an extreme familiarity with the domain, this ability is made possible by expertly scanning a chess board configuration. What then allows experts to have accurate recall ability and a fast response time in attention dependent tasks?

Kahneman & Treisman (1984) describes two experimental paradigms, selective filtering experiments and selective set experiments, to explore attention-related questions such as these. In selective filtering experiments, a subject selects a target from a large stimulus set and typically the experimenters measure accuracy. In selective set experiments, a subject selects a target from a small set of stimuli and the experimenters typically measure response time. The interpretation of the experiments is dependent on the experimental paradigm as selective filtering experiments tend to expose the relationship between memory and attention more so than selective set experiments which expose the automatic mechanisms of attention.

The results of attention-based experiments have shown that the ability to selectively pay attention to areas of a visual stimuli improves one’s capacity to recall details from long-term memory, increases reaction time for detection of targets, and improves object discrimination (Lang et al. 2014, Posner et al. 1978, Steinman et al. 1995). This section examines famous experiments in the field of selective attention and explores how the underlying mechanisms described in the previous section come together to construct these phenomenon.

3.1 Object Recognition and Memory Recall

Spatial cueing can have a significant affect on object recognition. However, attention mechanisms involve more than stimulus-based spatial cues. A famous selective experiment performed by Sperling (1960) found that when subjects were presented 12 letters for 50 ms, they only correctly reported 4-5 items on average (Figure 2). Sperling then cued a random, specific location after the display had disappeared, finding that subjects were able to consistently and accurately report all items from the randomly cued location. As the onset of this cue increased, the subject's accuracy decreased, suggesting that the stimulus had left working memory.

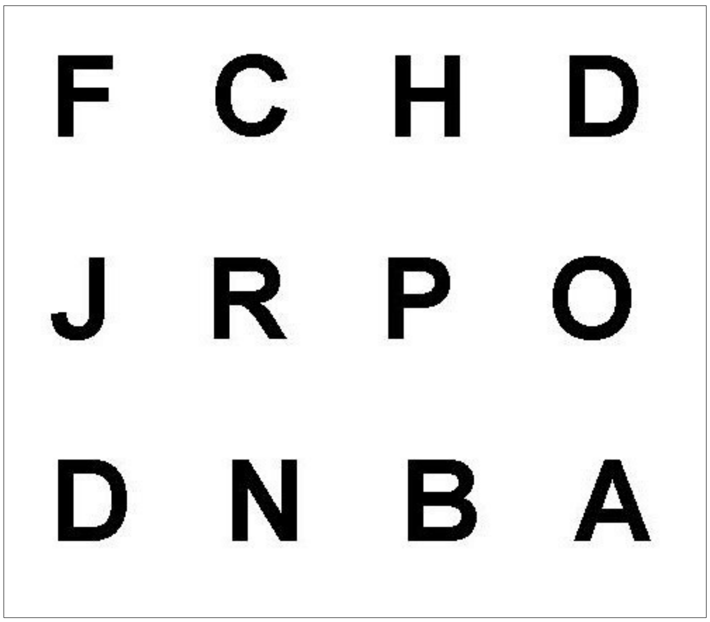

Figure 3. Sperling (1960) showed individuals a screen similar to the one above for 50 ms before asking the subjects to recall as many letters as possible.

Sperling (1960) showed that visual selective attention can improve memory recall; however attention is also affected by memory capacity. Watson & Kramer 1999 show that individuals with low working memory require more time to adjust their attention than individuals with high working memory. The results of these two studies therefore suggest that in a computational paradigm, the properties of working memory, including its rate of decay and capacity, affect the human visual selective attention mechanism.

One theory of how these attention events are retrieved from memory is through ‘visual indexing’ (Pylyshyn 1989). Visual indexing theory argues that the visual system is able to reference multiple objects in parallel based on bottom-up processes. Later, these objects can be indexed in memory based on these references. The CHREST cognitive architecture uses a modified version of visual indexing that also takes into account working memory and top-down biases in generating the indices (Lane et al 2008). This is a necessary step as later research has uncovered that object recognition in selective attention mechanisms is dependent on the object and not just the space the object occupies (Holmes & Horrax 1919, Posner et al. 1978).

3.2 Perceptual Load and Priming Effects

Sperling (1960) showed that cued locations enhance visual processing and memory recall in that zone. Steinman et al. 1995 expanded this to see how stimuli outside of this attention zone might be affected. The experiments found a positive attentional focus 12.8 degrees off of the focal center. Object discrimination and recall within the attention zone increases while processing outside the zone is inhibited. Furthermore, the strength of the stimulus used for the cue and priming effects can change the size of this focal zone..

Priming and interference can also affect the sensitivity to specific stimuli. For example, semantically priming subjects with similar words such as “doctor” can enhance the subsequent recognition of the word “nurse” (Meyer & Schavaneveldt 1971). Experiments using interference have been used to explain how and why humans are able to process distractors like negative primers. In a selective set experiment, Driver and Tipper (1989) had subjects count red items and ignore black items (Figure 4). If the red items were numbers, the subjects performed poorly due to interference; however if the black items were numbers, the subjects did not encounter interference. Thus, Driver and Tipper (1989) ruled that when subjects filtered out interferences early, there was little effect on response time.

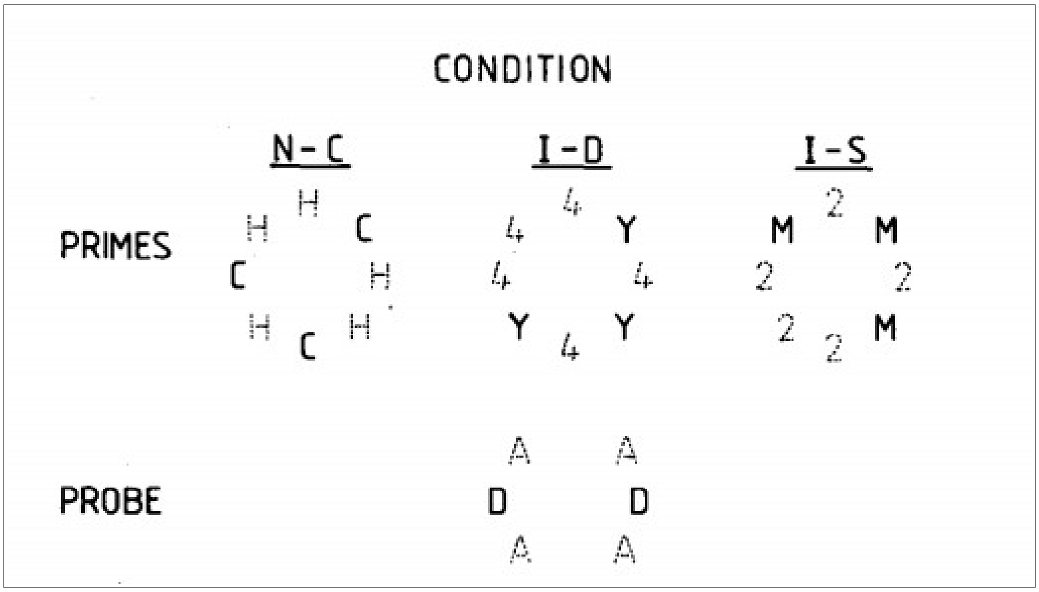

Figure 4. To test their hypothesis, Driver and Tipper (1989) designed three experimental scenarios (N-C = neutral control, I-D interference-different, and I-S interference-same). Solid items represent red items and dashed items represent black items (Figure from Driver and Tipper 1989).

The experiments above explain two important concepts. Firstly, the limited attended area in a visual stimuli uses the majority of visual processing resources while processing outside this area is inhibited. Secondly, selective attention is affected by priming effects held in memory and task-specific interferences that are largely stimulus based.

3.3 Object and Spaced Based Attention

While the previous experiments focused on attention among disparate objects, the visual selective attention mechanisms must also be able to create perceptual groups to attend to whole objects. At a high-level, visual selective attention can group items based on two principles. S*paced-based* attention selects regions of space independent of the features within that space whereas o*bject-based* attention selects objects and features independent of the space they occupy. Research has found that these two attention mechanisms share common neural mechanisms and therefore have joint relevance in human selective attention (Fink et al. 1997).

Space-based attention is a well studied mechanism of object grouping in visual selection tasks based on spatial cues and other bottom-up inhomogeneities (Heydt et al. 1984, Duncan et al. 1997). For example, due to spatial cues humans respond faster to stimuli in expected areas of the visual input than stimuli in unexpected areas (Posner et al. 1978, Sperling 1960). These findings have also shown that sensitivity to changes in this expected region attenuate over time. However, space-based mechanisms are not the only factors guiding object-grouping. For example, Tipper et al. 1991 showed that inhibition to previously attended objects was independent of the object's location. That is, if an attended object was moved, a subject still displayed a decreased sensitivity to changes in that object. Object-based attention is therefore influenced by higher-level cognitive functions like object recognition, as well as task-specific and independent biases and priming.

Additional support for object-based attention comes from study of Balint syndrome which affects the posterior parietal lobe, an important area of the attention network (Moreaud 2003). *Simultanagnosia* is a symptom of Balint syndrome in which a patient is unable to perceive more than one object at a time despite normal visual processing. Holmes & Horrax (1919) showed that patients could not determine if two parallel lines had the same length; however they could identify simple shapes if the lines were connected to form a square. Thus, Balint syndrome exposes an inability to attend to multiple objects in a visual stimuli as the result of damage to the attention network.

4. Computational Models of Selective Visual Attention

The following section describes four computational models (CTVA, CHREST, agent-based, and MORSEL) which attempt to model the neural mechanisms and high-level cognitive effects of human selective attention.

4.1 CODE Theory of Visual Attention (CTVA)

The COntour DEtector (CODE) theory of visual attention (CTVA) provides a mathematical framework for unifying object-based and space-based attention by clustering perceptual objects and regions of space (CODE) and defining a method for higher level categorization (TVA) (Logan 1996).

CODE theory developed by Oeffelen and Vos (1983) represents a visual scene as an analog representation of an object’s location and as an abstract representation of object groupings. Bottom-up processes guide the analog representations while both top-down processes and bottom-up processes generate object groupings (Figure 5). The top-down biases introduce a threshold from which related objects can be grouped. As the threshold increases, large groupings are broken into smaller ones. In this manner, CODE can create a hierarchy of objects by adjusting the priming of the top-down bias.

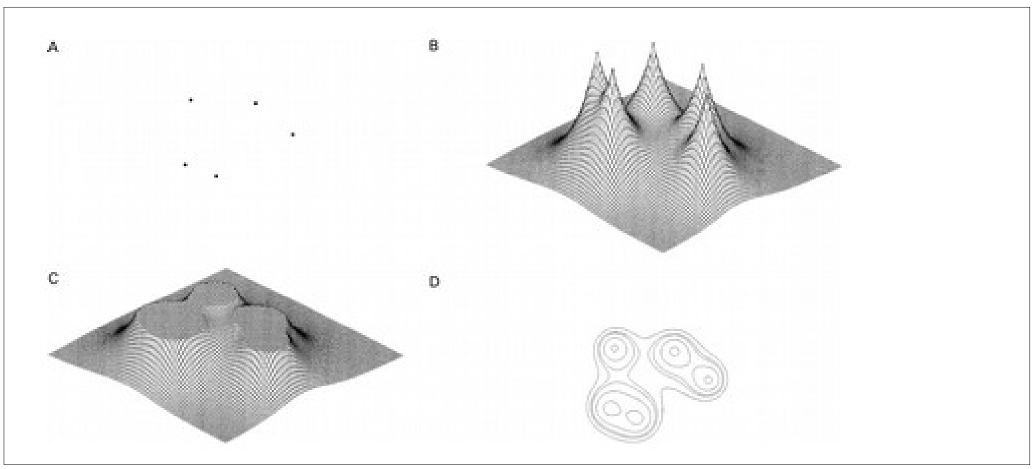

Figure 5. Dot patterns of items (a), code surface (b), the results of thresholding (c), and contour mapping of the grouping of the dots (d). In CTVA, the CODE surface (b) is how space is represented and the groupings (d) represent an object. The threshold allows for the creation of hierarchical representations and TVA is used to make a categorization based on the groupings (Figure from Logan 1996).

However among these various objects, CODE cannot achieve within-object discrimination. Logan 1996 uses Bundesen’s (1990) formulation of the theory of visual attention (TVA) which models what object humans select within a visual input based on a computational race. TVA uses perceptual features (color, shape, etc) to create a linking parameter which captures the evidence that a grouping of features belong to a specific category.

The race towards a a specific categorization is based on this linking parameter and can be affected by attention weights towards a given category, representing, for example, an individual being primed to look for a particular colored object. Furthermore, Bundensen (1990) proposed using TVA to predict response time by interpreting the result of TVA as a rate parameter in a exponential distribution. Thus, the output of TVA for the various categorizations can be seen as the mean processing time of a visual scene, minus the time to perceive and generate a motor response. Thus, CTVA can be used in experiments modeling human response time in attention-related tasks.

A major disadvantage of CTVA is that CODE assumes that objects are points in space and is therefore unable to deal with objects that overlap or extend in space due to motion. Proximity is also the only grouping principle utilized by TVA. Other forms of groupings may exist, such as common fate where two items may be moving with similar velocities (e.g. a flock of birds).

4.2 CHREST Cognitive Architecture of Visual Expertise

CTVA looks at attention through a mathematical lens whereas CHREST presents a cognitive architecture of human visual expertise (Lane et al. 2008). Attention mechanisms in CHREST are encoded as a unique set of search heuristics that guide the recollection of chunks from memory. Chunks are pieces of information acquired through experience that are directly affected by bottom-up and top-down biases and encoded symbolically in memory.

The CHREST architecture first perceives a visual input through a limited field of view from which features are extracted. The contents of STM in addition to any domain-specific knowledge and spatially relevant features are used to refocus on a new segment of the visual input. CHREST can also search for features in long-term memory (LTM) and place features in short-term memory (STM), thus resembling the theory of visual selection.

CHREST's attention mechanism is based on a *perception-learning-perception* cycle. A simulated eye is used to retrieve information from the visual input and is guided by a set of heuristics. First, CHREST selects a fixated item from the center of the field of view. It stores the item and its location in short-term memory until a termination condition is met. The termination condition is when there is no next item to fixate upon or it revisits an already fixated item. Upon termination, CHREST learns the pattern generated in STM.

What guides the eye movement is a set of domain independent and domain dependent heuristics. An example of a domain independent heuristic is attending to an unobserved portion of the visual input or a random position on the periphery. In Chess, a domain dependent heuristics may involve focusing on important pieces like the safety of the king. One important heuristic to discuss is selecting a new item to fixate on based on information in LTM. This heuristic attempts to discover the largest related chunk in LTM and then constructs a larger chunk based on contents found in STM.

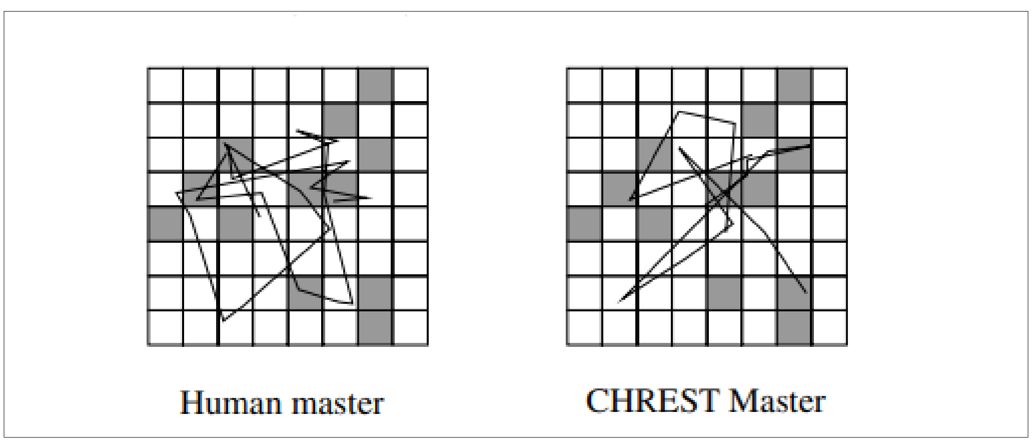

Figure 6. An example of an expert chess player's eye movement across a board configuration and the simulated eye movement of CHREST using the search heuristics. The gray scales are important squares (Figure from Lane et al. 2008).

As a consequence of this heuristic, CHREST creates larger chunks as it explores the domain. Thus, CHREST is able to mimic the transition from novice to experts through better memory recall through selective attention (Figure 6). Furthermore, the symbolic encoding of chunks allows attention to affect learning and learning to then affect attention.

4.3 Belief-Desire-Intention Agent: Agent-Based Approach

Previous computational approaches have used bottom-up and top-down biases to influence selective attention. Duncan et al. 1997 describes various factors that can cause a feature in an visual stimuli to stand out, such as its temporal relevance, spatial-context, and uniqueness. These biases are low-level constructs of the input; however, higher-level biases such as emotions, surprise, and the congruency of beliefs are well known catalyzers of attention (Kahnmen 1973). Macedo (2012) provides a agent-based model that attempts to select information based on these higher-level biases.

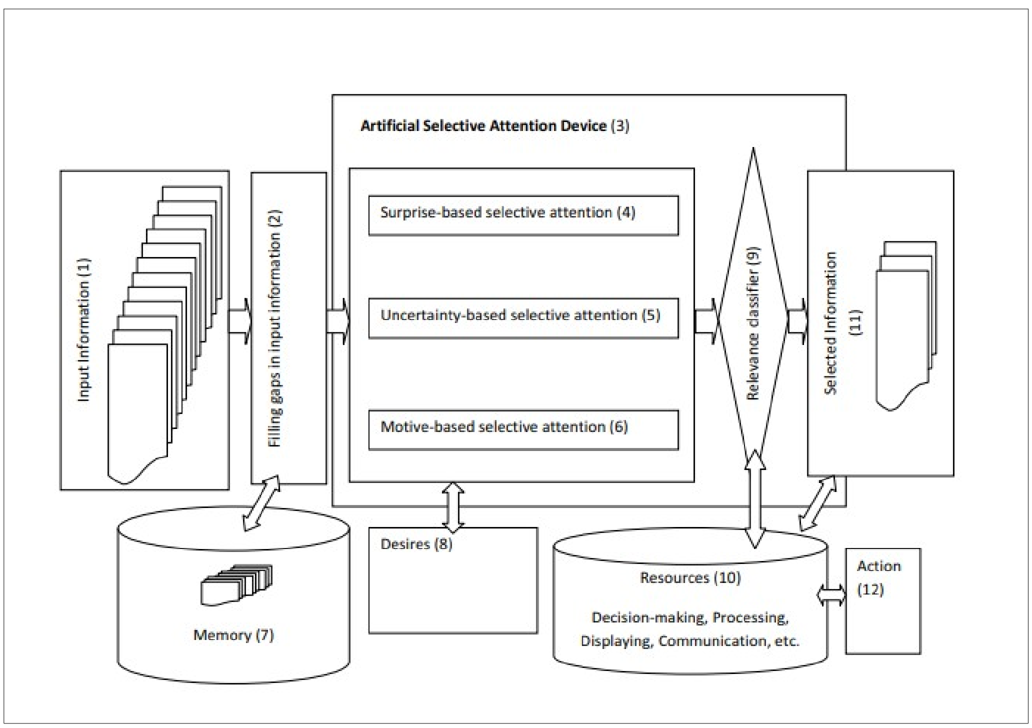

Figure 7. The belief-desire-agent architecture (Macedo 2012).

Figure 7 shows the modules of the agent-based architecture where expectations taken from the knowledge base in addition to the input are used to construct the world state. The model then assigns a relevance classification to each event in the world based on external goals, surprise, uncertainty, and motives. An action is selected based on available resources. Within this framework, surprise is the expectation of the given event compared to the most likely event in the world. Uncertainty is the information gained or the difference between the prior uncertainty and future uncertainty. Finally, motive is a domain specific score which captures the desire to see a particular event.

The paper proposes simple thresholds for each of the three attention mechanisms. The relevance classifier produces a YES or NO response to the question “is this event interesting?” or formally if a particular attention mechanism score exceeds its respective threshold. The paper does not discuss the decision-making process and leaves that formulation up to a particular domain and task. Despite limited scope and rather rigid mechanisms, Macedo (2012) presents a framework for modeling surprise, uncertainty, and motive. These higher-level top-down biases should be interpreted as areas of future study or possible extensions to more advanced selective attention models.

4.4 MORSEL: Connectionist Model of Visual Attention to Model Neglect Dyslexia

MORSEL is a connectionist computational model of attention. The model was designed to analyze visual objects in parallel and to mimic perceptual errors of humans due to distractors and interferences. Mozer & Behrmann (1990) used an intentionally damaged version of MORSEL to model neglect dyslexia. Neglect dyslexia is a neuropsychological syndrome which causes patients to misread words through letter omission, addition, or substitution (Moore & Demeyere 2019). For example, a patient might read BOY as TOY or BEACH as REACH. MORSEL has three main functional components used in replicating this behavior (Figure 8).

The first component of MORSEL is the attention mechanism (AM) that guides the model and prevents it from processing too much information. The AM receives visual input as well as high-level processing input. The second component is BLIRNET which creates location invariant representations of visually presented words or letters in parallel. The final component is the pull-out net which is used to clean the noisy perceptual data presented by BLIRNET and to ensure semantic coherence.

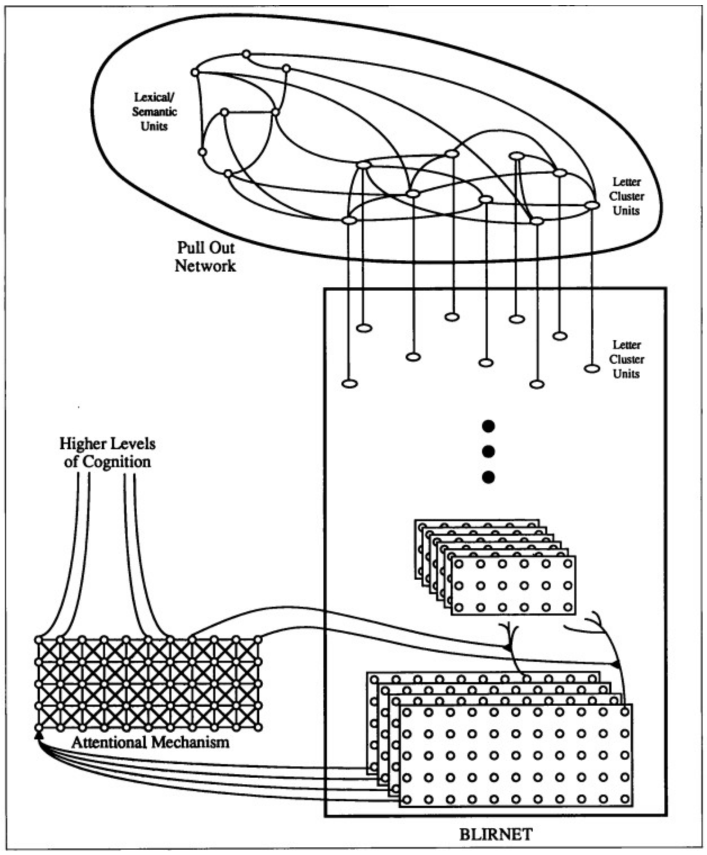

Figure 8. A sketch of MORSEL architecture used to model Neglect dyslexia. (Mozer & Behrmann 1990).

Mozer and Behrman (1990) describe the four properties of MORSEL that allow it to account for neglect dyslexia. First, attention is placed on new stimuli early in processing. Second, attention selects an item which for MORSEL is a bundle of dense features in a given space. Thirdly, the AM gates the flow of activity in the object recognition system (BLIRNET) allowing for damage to be modeled. Finally, the clean-up mechanism recovers meaningful information from BLIRNET.

To mimic neglect dyslexia, Mozer and Behrman damage the bottom-up mechanisms of the AM, affecting the probability that features in the attended space will be detected by the AM. Experiments with MORSEL found that when presented with a pair of words, not only would a damage to the AM cause a directional bias towards selecting one word over another, but that within that word the letters experienced a similar bias in the later stages of the model. Thus, affecting the bottom-up processing of the AM not only affects lower-level attention mechanisms (which area of the stimulus are processed) but also higher-order cognitive functions (which area of the attended stimulus are selected) which is a spectacular insight showing the link between selective attention and cognition.

A disadvantage is the model can only take into account view centered instances of neglect dyslexia; however, there are other frames of reference patients experience, such as their body, eyes, or head (Moore & Demeyere 2019). Still, MORSEL shows the effectiveness of using modeling to gain insight into disabilities like neglect dyslexia.

5. Discussion

In their formulation of CTVA, Lane et al. 2008 presents 5 key questions that face any computational theory of visual spatial attention. The questions ask 1) how space is represented, 2) what defines an object in that space, 3) what is the shape of the attention zone, 4) how does selection occur within the focus of attention, and 5) how does selection between objects occur. In the discussion below, the agent-based model is omitted due to its incomplete nature; although useful insights from the model will be mentioned.

For CTVA, space is represented from the bottom-up by a CODE surface and from the top-down as perceptual groups or objects categorized by TVA. The size of the attention zone is determined by the threshold which forms perceptual groups. Selection among the group is based on Bundesen’s (1990) TVA model of selection; however, selection between objects is not well specified. A major advantage of CTVA is that it can be used to model changes in response time (RT) due to various bottom-up and top-down biases. It also has the fewest restrictions on the representation of low-level bottom-up biases in comparison to the other architectures and also invites both space-based and object-based grouping effects.

Unlike CTVA, CHREST's formulation of attention provides a feedback mechanism resembling how humans begin to develop expertise in a given domain. For CHREST, space is represented as the relationship of objects in a 2D surface. CHREST uses heuristics to capture the top-down and bottom-up biases as well as higher-level symbolic representations in memory. In CHREST, an object is the formed chunks that result from a complete perception-learning-perception cycle. The attention zone is determined by the simulated eye. Selection within this zone and between objects is based on knowledge-base in LTM as well as independent space-based heuristics that guide perception across the stimuli.

In MORSEL space is a 2D frame and objects are the visually presented words and letters. The attention zone is the total area covered by the AM while selection between objects is determined by healthy areas in the AM. MORSEL imposes the most restrictions on the input in order to focus on the selection of objects to model neglect dyslexia. Selection among objects, such as letters within a word, is due to the activation response in the pull-out network.

By now it should be obvious that memory is linked to attention mechanisms; therefore, it is necessary to extend Lane et al. 2008's questions to ask to what extent working memory is included. CHREST, MORSEL, and the agent-based model have concepts of working memory yet could be advanced by incorporating the transiency and capacity of working memory (Sperling 1960, Watson & Kramer 1999). Also while CTVA and CHREST involve the use of top-down and bottom-up biases, they fail to incorporate recent research which has shown the effects of these to be a weighted combination (Schneider et al. 2012). However, CTVA's ability to represent space and object-based attention is an advancement in more accurately modeling human attention and possibly modeling object-based attention experiments like Tipper’s 1991 object-based inhibition experiments. Furthermore, the agent-based model presents calculations that could be used to model higher-order attention template features which could be incorporated by models like CHREST to achieve a more realistic memory model.

Another crucial theory discussed in selective attention was the integrated competition hypothesis (ICH) and the theory of visual selection. The agent-based model and CTVA include the concept of limited visual resources which MORSEL and CHREST both lack. However, only CHREST parallels the attentional templates described by the theory of visual selection in the use of chunks stored in LTM and brought to STM. Furthermore, CHREST’s perception-learning-perception learning cycle closely resembles the emergent behavior of attention found in many high-level experiments (Sperling 1960, Meyer & Schavaneveldt 1971, Tipper 1991). MORSEL on the other hand is useful in that limited visual resources could be modeled by damaging the network, like when modeling neglect dyslexia.

The area for computational improvement that requires the most future consideration is computationally filtering out information at later stages based on the task and the developed knowledge-base. To accomplish such a task, a model would have to provide an inhibitory signal to a neuron which is based on information in working memory as well as priming and interference effects. However, more research is required to better understand the neural mechanisms of this filtration and the role of working memory. An experimental framework to explore such questions would likely require the use of negative primers and distractions to determine not only what information is filtered out but at what stage in the neural pipeline.

6. Conclusion

Visual attention is one of the most well experienced phenomenon in our lives, but upon further investigation, it is clear that understanding and replicating its mechanisms in computational models is surprisingly advanced. Upon analyzing the neural mechanisms of the attention network, the causal effects of bottom-up and top-down stimulus-based biases appear at early stages. At later stages experience-based top-down biases such as priming and task-specific interference begin to affect the attention mechanism. These findings indicate that attention mechanisms while causal in nature are also emergent as an individual gains expertise in a visual task and builds a knowledge base in memory. The emergent mechanism of visual attention invites the study of high level cognitive experiments and provides leeway in representing attention computationally while causal mechanisms of attention like spatial filters are likely needed to accurately represent the underlying mechanisms. Most important in this analysis of the emergent and causal properties of visual selective attention is the understanding that attention and cognition are finely linked though not inseparable; therefore a better understanding of human selective attention in visual tasks is a step closer in better understanding human cognition at a high level.

7. References

Bundesen, C. (1990). A theory of visual attention. *Psychological review*, *97*(4), 523.

De Groot, A. D. (1946). *Thought and choice in chess* (Vol. 4). Walter de Gruyter GmbH & Co KG.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review*, *96*(3), 433.

Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current opinion in neurobiology*, *7*(2), 255-261.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193-222.

Driver, J., & Tipper, S. P. (1989). On the nonselectivity of selective seeing: Contrasts between interference and priming in selective attention. Journal of Experimental Psychology: Human Perception and Performance, 15, 304–314.

Fink, G. R., Dolan, R. J., Halligan, P. W., Marshall, J. C., & Frith, C. D. (1997). Space-based and object-based visual attention: shared and specific neural domains. *Brain: a journal of neurology*, *120*(11), 2013-2028.

Holmes, G., & Horrax, G. (1919). Vol. I APRIL. 1919 No. 4 DISTURBANCES OF SPATIAL ORIENTATION AND VISUAL ATTENTION, WITH LOSS OF STEREOSCOPIC VISION. *Archives of Neurology and Psychiatry*, *1*(4), 385.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Treisman, A. M. (1984). Changing views of attention and automaticity. In R. Parasuraman, & D. R. Davies (Eds.). Varieties of attention. Orlando, FL: Academic Press.

Lane, P. C., Gobet, F., & Smith, R. L. (2008, May). Attention mechanisms in the CHREST cognitive architecture. In *International Workshop on Attention in Cognitive Systems* (pp. 183-196). Springer, Berlin, Heidelberg.

Logan, G. D. (1996). The CODE theory of visual attention: an integration of space-based and object-based attention. *Psychological review*, *103*(4), 603.

Macedo, L. (2013, September). Arguments for a computational model for forms of selective attention based on cognitive and affective feelings. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (pp. 103-108). IEEE.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*(1), 4-18.

Meyer, D. E., & Schavaneveldt, R. W. (1971). Facilitation in recognising pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 90, 227–234

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782

Moreaud, O. (2003). Balint syndrome. *Archives of neurology*, *60*(9), 1329-1331.

Moore, M. J., & Demeyere, N. (2019). Neglect dyslexia as a word-centred impairment: A single case study. *Cortex*, *119*, 543-554.

Mozer, M. C., & Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience*, *2*(2), 96-123.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual review of neuroscience*, *13*(1), 25-42.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. *Modes of perceiving and processing information*, *137*(158), 2.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65-97.

Schneider, D., Beste, C., & Wascher, E. (2012). On the time course of bottom-up and top-down processes in selective visual attention: An EEG study. *Psychophysiology*, *49*(11), 1660-1671.

Steinman, B. A., Steinman, S. B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a center—surround organization. *Vision research*, *35*(13), 1859-1869.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology*, *43*(2), 289-298.

Von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*(4654), 1260-1262.

van Oeffelen, M. P., & Vos, P. G. (1983). An algorithm for pattern description on the level of relative proximity. Pattern Recognition, 16, 341–348.

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics*, *61*(1), 31-49.