Attention and Cognitive Control in Humans, Animals and Intelligent Systems

2009-09-08

Thomas Hellström

UMINF-09.16

ISSN-0348-0542

Department of Computing Science

Umeå University

Umeå

Sweden

Abstract

While often being associated only with visual cognition, attention is gradually becoming relevant for other sensor modalities and also for modeling other stages of cognitive information processing than perception. From having been a topic primarily studied by psychologists, the area has become increasingly interesting also for ethologists, neuro scientists and roboticists.

This paper summarizes important research results, and gives an overview of common bottom-up and top-down models for visual attention. Several examples of mechanisms related to attention are given and put into context. Models for executive attention a.k.a. cognitive control are also explained and put into relation with perceptual attention. A number of experiments and results related to the function of pre-frontal cortex are reported and discussed. Besides giving overviews of the areas and important results, the paper tries to formulate a unified view and generalization of the many, seemingly different, mechanisms denoted attention and cognitive control. As an example of how perceptual attention and cognitive control often are intertwined, a model inspired by Miller and Cohen (1994) is finally adopted to describe the rapid eye movements called saccades.

1. Introduction

Cognitive control was “a major ingredient of psychological consideration” already in the late 1800s and early 1900s (Hommel, Ridderinkhof, & Theeuwes 2002). Furthermore, *will* and *attention* were already then important concepts when describing personal goals and interests within cognitive psychology. Within neuroscience and cognitive science the study of attention has increased markedly over the last decade. The research reported in this paper will mainly refer to this period.

While often being associated with visual cognition, attention is gradually becoming relevant for other sensor modalities and also for modeling other stages of cognitive information processing than perception. From having been a topic primarily studied by psychologists, the area has become increasingly interesting for ethologists, neuro scientists and roboticists.

We will in this paper describe some common models that are used to explain visual attention and also executive attention a.k.a. cognitive control. Besides giving overviews of the areas and important results, we will try to formulate unified views and generalization of the many, seemingly different, mechanisms denoted attention and how it connects to cognitive control.

Section 2 gives a general introduction to the concept of attention, followed by a thorough overview of common models and issues related to visual attention in Section 3. Section 4 contains conclusions regarding perceptual attention and Section 5 introduces cognitive control in depth. Section 6 discusses the connection between perceptual attention and cognitive control, converging at a model that describes the rapid eye movements called saccades.

2. Attention

Attention is a common term in psychology and cognitive science. However, it is often used in very different contexts and meanings such as in expressions like

- “The smell caught my attention”
- “I didn’t see you, I was paying attention to the lecture”
- “Pay attention to how you twist your ankle when you hit the ball”
- “Pay attention to how you spend your money”
- “I don’t remember even crossing the street. I must not have been paying attention”

We are probably talking about many different things, and to make the difference clearer, the following division into three main categories may be useful:

- Perceptual attention, for example:
  - Focusing on a part of the visual field
  - Hearing one conversation in a noisy environment, the “cocktail party effect”: when we are talking with a friend in a crowded party, we still can hear and understand what people say even if the room is very noisy. We are able to ignore what another nearby person is saying. Then if someone on the other side of the room calls out our name, we immediately react to that particular sound
  - Being more attentive to the distance to the car in front of your own than to your cell phone conversation
- Action related attention, for example:
  - Focusing on one task (e.g. how to hit a ball)
  - Maintaining behavioral goals (e.g. how to spend money)
  - Guiding behaviors depending on the situation (e.g. guiding eye movements)
- Mental attention, for example:
  - Mind-wandering or "spontaneous thought”
  - Being conscious about your actions (e.g. remembering having crossed the street)

Definitions of new concepts often serve a purpose of high-lighting important aspects and components of the concept in question. Two definitions of *attention* are:

- “Seeing the world around you is like drinking from a fire hose. The flood of information that enters the eyes could easily overwhelm the capacity of the visual system… [Visual attention] allows selective processing of the information relevant to current goals” (Kanwisher, & Downing, 1988).
- “…the capacity to select and enhance limited aspects of currently processed information, while suppressing the remaining aspects” (De Pisapia, Repov, Braver, 2007).

The first definition talks about reduction of visual data and also adds the aspect of controlling attention depending on the current goal. The second definition is more general and covers not only sensory data but also action related and mental concepts. In this way, so called *executive attention* (a.k.a. *cognitive control*) is included.

3. Visual attention

When we look at and interact with the physical world, the focus of what we are attending to constantly changes. A lot of research in attention has been and still is focused on this visual form of perceptual attention. A number of interesting mechanisms control how the focus is set and how information is selected, both in general situations and when performing a defined task involving processing of visual information. As an illustration of the complexity of these mechanisms, consider the following two pictures:

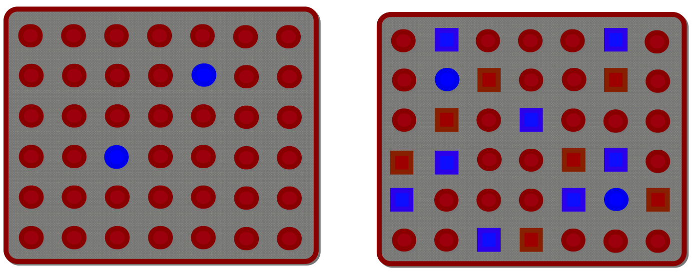


Figure 1. Illustration of two types of visual attention mechanisms. Left: the blue circles "Pop-out" with low effort. Right: A sequential search is necessary since a combination of features is necessary to extract the blue circles.

The blue circles in the left picture are easy to identify since they "Pop-out" in a seemingly automatic fashion. The right picture is more complicated. To identify the blue circles we have to "manually" scan all candidates in a time consuming process.

Humans and other animals do not look at a scene in a steady way. Instead, the eyes move around locating interesting parts of the scene. The reason for this is believed to be the fact that the central part of the retina has higher resolution at the fovea. Even if the human eyes have a very wide field of view, we have to move the eyes such that the important objects are centered on the retina. Normally, we move our eyes 3–5 times per second (that is, 150,000 to 250,000 times every day) (Itti & Koch, 2001). These motions are to a large extent automatic and unconscious and are denoted saccades. Saccades have a typical duration 20-200 ms and a length up to 90 degrees. Newborns follow simple rules like looking for edges (Haith, 1980). However, in other cases it has been shown that the motion depends on the task. Typical saccade movements for test objects who have different tasks when being exposed to a picture are shown in Figure 2. The motion patterns clearly reflect the task, such that important information automatically is gathered.

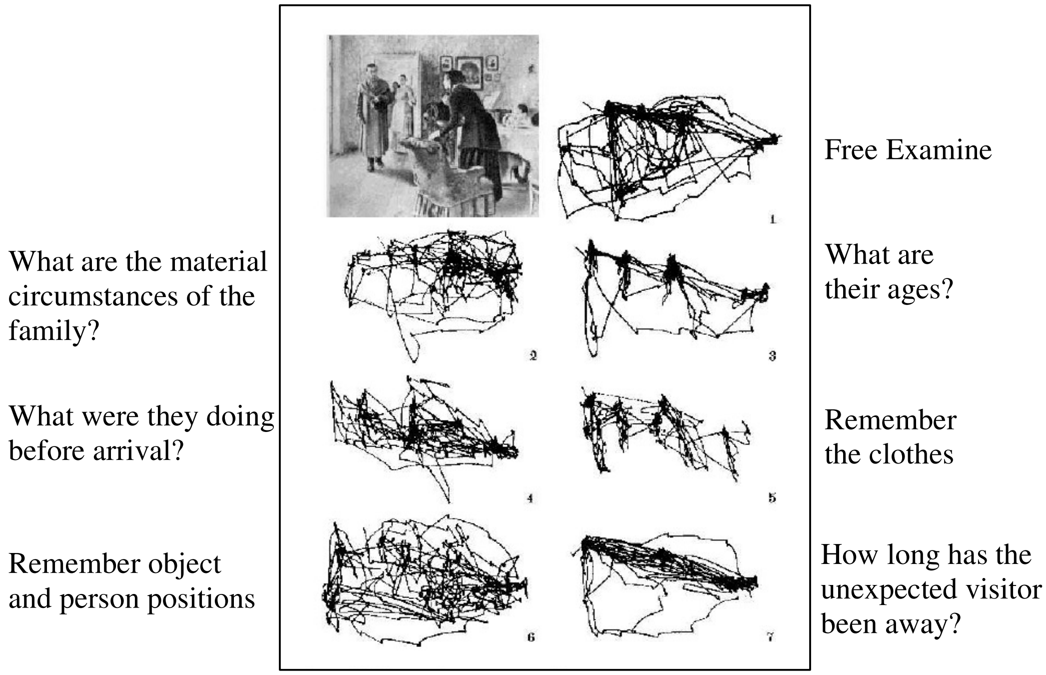


Figure 2. Saccade movements when looking at a picture with a specific given task Picture from (Yarbus, 1967)**.**

There is a general consensus (De Pisapia, Repov, & Braver, 2007) that attention is a core component in visual perceptual processing. Research in visual attention has also identified advanced innate mechanisms supporting visual attention. Human infants less than a month have been shown to prefer patterned stimuli, 3-D objects, high contrast objects, and face-like objects (Slater, 1998). Attention is also higher for moving objects (Vinter, 1986) than for static parts of the scene. Possible explanations for this kind of phenomenon will be given in the remainder of this section.

We will start with four views on why evolution has equipped us with mechanisms for visual attention at all (De Pisapia, Repov, & Braver, 2007):

1. Our processing capabilities are limited and the optic nerve handles 10^7-10^8 bits of data each second. There is obviously a clear need to remove some information
2. Not all information is relevant. Hence there is a need to detect and select the most informative and relevant aspects of all information
3. Information is not coherent. Hence there is a need to integrate visual features into a unified view. Selected features may in this way be bound together into higher-level object representations.
4. Attention is an emergent phenomenon in a network with inhibitory competition

It is clear that the four views can also be applied to other kinds of sensor modalities than vision.

Computational models of visual attention have been developed for a number of reasons. They can be used for testing existing theories and experiments of visual attention. They can also be used to explain certain neuropsychological deficits, for instance *attentional neglect* (Cohen, 1994). Models may also contribute to the understanding of neurophysiological or neurobiological effects, for example salient locations in the visual field (Itti and Koch, 2000). Finally, an important reason to develop models of visual attention is to advance development of computer vision, intelligent systems and robotics. We will describe two well-known models, both inspired by the Feature Integration Theory by Treisman and Gelade (1980).

3. 1 Koch & Ullman's model of visual attention

One early and very influential model for visual attention was developed by Koch and Ullman (1985). It is a bottom-up model that builds from extracted features that are computed at a number of different spatial scales. Local and global measures of deviations from mean values are computed for all features. In this way a red dot attracts more attention than a uniform red field, and a single red dot attracts more attention when it is the only red dot in the visual field. The complete computational scheme is described in Figure 3. The resulting saliency map is finally used to extract winners by a Winner-take-all (WTA) mechanism. The result can be fed to motor behaviors changing the focus of attention (FOA) to the selected location, or to further perceptual processing. An inhibition of return function suppresses already selected locations in the saliency map such that a sequential search of the scene can take place. This computational model was fully implemented by Itti, Koch, and Niebur (1998). The pictures in Figure 4 show how the FOA jumps between different parts of the picture, in an order determined by the computed saliency values.

While some attentional effects can be explained by bottom-up models such as the Koch & Ullman model many observed phenomena remain unexplained. One such case is called Conjunctive search and has already been illustrated in Figure 1. The basic observation is that conjunctions of selection criteria are much harder to handle than one single criterion. While a search for blue circles among red circles can be done in constant time, finding blue circles among red and blue circles and squares requires active search: "An attentional spotlight moving from item to item in the display" (Treisman, 1988). The required time is typically proportional to the number of distractors in the image. Wolfe addressed this search process in the Guided Search model described below.

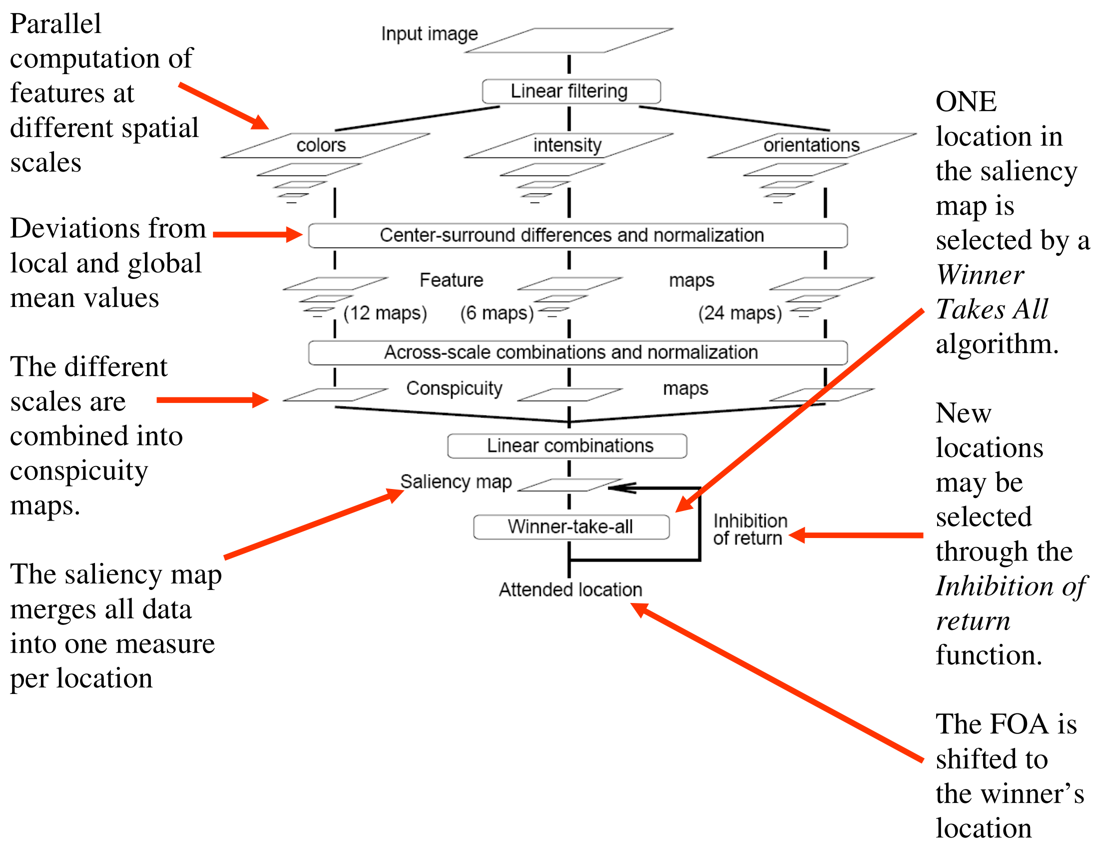


Figure 3. Koch and Ullman's model of visual attention. The attended locations will be selected in sequence through the Inhibition of return function.

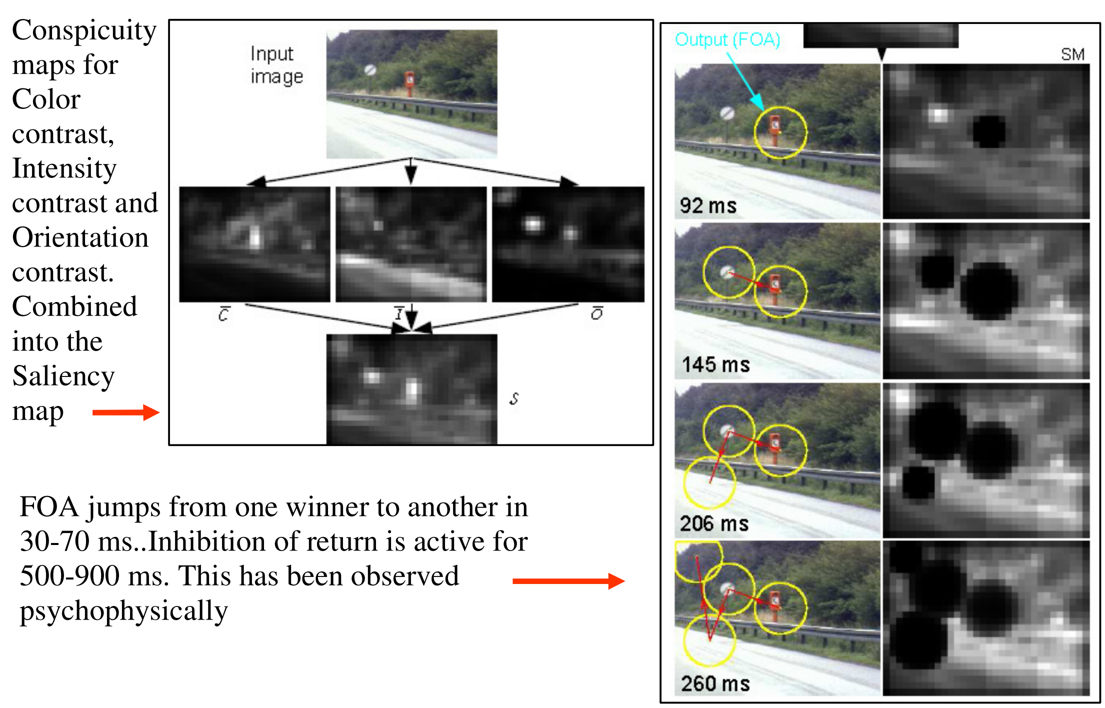


Figure 4. Implementation of Koch's and Ullman's model. Picture adopted from (Itti, Koch, & Niebur, 1998).

3. 2 Wolfe's "Guided Search 2.0"

The model by Wolfe (1994) combines Bottom-up and Top-down views of human visual attention and visual search. In this model, attention is controlled not only bottom-up by sensory data, but also top-down by high level goals. Visual stimuli are processed by a variety of low-level feature detectors which produce individual feature maps. These are weighted and summed into a single activation map corresponding to the Saliency map in the Koch and Ullman model. Peaks in the activation map are regarded as areas of interest. High-level cognitive systems can influence the process by modifying the weights. See Figure 5 for a schematic overview. The model explains some of the conjunctive search effects described above and was successfully implemented in the Kismet robot (Breazeal & Scassellati, 2001), see Figure 6. An overview of related research and suggested extensions to the basic model are given by Mozer & Baldwin (2008).

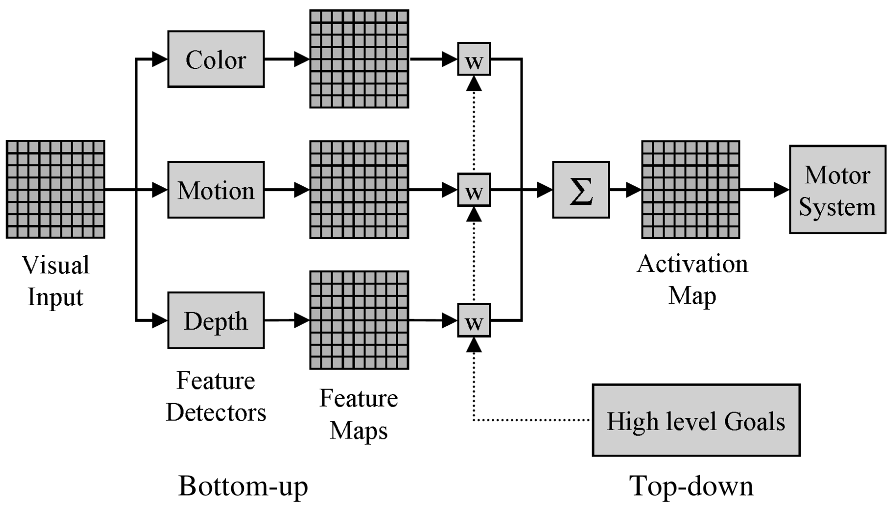


Figure 5. Wolfe's model for visual attention includes Top-down control such that High level goals may change the criteria for attention. Picture from (Scassellati,2001).

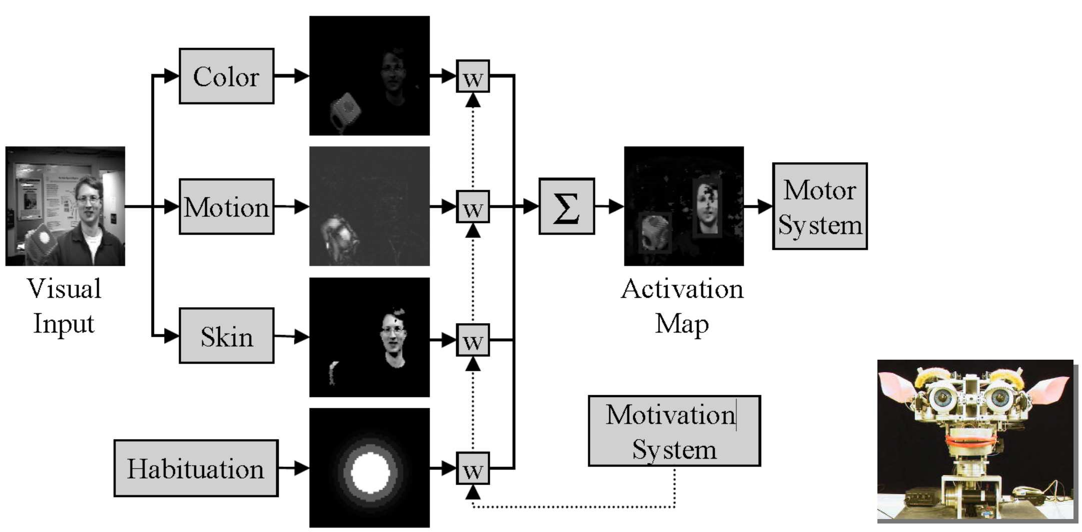


Figure 6. Overview of the attention system in the Kismet robot. Picture from (Scassellati, 2001).

3.3 Other models of visual attention

Attention may so far look like a purely perceptual mechanism that performs pre-processing and/or data selection to simplify data analysis for high-level cognitive functions. But already the models for visual attention described above contain elements of action, namely the motor system that receives the attention information and moves the eyes accordingly. In the light of recent theories of sensory-motor coordination (Pfeifer & Scheier, 1997), the division of sensing and action becomes even more dubious. Sensing and acting is believed to be intertwined in general, so there is no reason to assume that the two concepts should be separated in the attention mechanisms.

Balkenius (2000) draws further consequences from this connection and proposes a model in which attention mechanisms are seen as actions and, as such, are subject to learning just as other action oriented behaviors may be. He uses *habituation* and *instrumental conditioning* as learning methods. The resulting model could for example be used to construct a robot that is able to learn where to look to gain the most information about its environment.

3.4 More visual attention mechanisms

Some attentional effects can be explained by bottom-up models such as the Koch and Ullman model or by combined Bottom-up/Top-down models such as Wolfe's Guided Search model. However, many behaviors of the human visual attention system remain unexplained, especially when attention depends on the specific task or goal of the ongoing activity. The task dependent saccade movements described in the beginning of Section 3 is one example of this. A few other attention related mechanisms are described below.

Inattentional blindness

Experiments by Simons and Chabris (1999) suggest that we sometimes fail to perceive clearly visible and salient unexpected objects in our environment when engaged in a task. Figure 7 illustrates experiments in which observers of video tapes were told to count the total number of passes made with the basket ball on the recorded video. After some while of this action, either of two unexpected events occurred: a person holding an umbrella walked from left to right or a person wearing a gorilla costume walked through the action in the same way. In both cases, the unexpected event lasted 5 s, and the players continued their actions during and after the event.

Somewhat surprisingly, 46% of 192 tested observers failed to notice the unexpected event. The results clearly show how we perceive and remember only those objects and details that receive focused attention. Surprisingly, spatial proximity of the critical unattended object to attended locations does not appear to affect detection, suggesting that observers attend to objects and events, not spatial positions.

Figure 7. Observers focusing on counting the total number of passes with the basket ball often fail to observe seemingly salient, but unexpected, events. Pictures from (Simons and Chabris, 1999).

Attentional blink

The attentional cost of attending to one visual stimulus may lead to impairments in identifying a second stimulus presented within 500 ms of the first. This phenomenon is known as the attentional blink. A typical experiment is to ask subjects to identify both a number and a letter that are rapidly and sequentially presented on a display. Videos illustrating inattentional blindness, attentional blink, and other visual awareness phenomena are available for sale at http://www.viscog.com/.

Change blindness

Change blindness refers to a striking difficulty to see large changes that normally would be seen easily. If the pictures in Figure 8 are repeatedly shown with a short blank screen inserted between the two pictures, the difference is very hard to notice. The explanation is that the onset of the blank screen disturbs the local motion signals caused by a change. This crashes the automatic system that normally draws attention to changes in a picture. Instead, attention is controlled by slower, higher-level search mechanisms which sequentially check object by object, until the object that is changing is found. Change blindness and other related phenomena have been studied by Simons and Rensink (2005).

Figure 8. Pictures demonstrating change blindness. Extracted from videos available at *http://www.psych.ubc.ca/~rensink/flicker*

4. So what is Attention so far?

The discussion and the given examples have so far focused on attention in connection to visual perceptual processing. A few examples have also been given for other sensor modalities. We will now try to extract the common characteristics of what the concept attention usually refers to and really means in terms of information processing. Attentional mechanisms "assign attention" to parts of sensory data that are judged important. The process can be divided into two parts: Data extraction and utilization.

Data extraction may involve both attenuation and selection components. It may be a mapping from a high-dimensional sensor data space to a low-dimensional space, i.e. a removal of irrelevant dimensions in data, for instance attention to a small part of the visual scene, or a focus on blue circles in a picture. It may also be selection of temporal chunks from a sensory data stream, such as in the "cocktail party effect" where the attention mechanism blocks the high-level perception of background noise until the listener's name is recognized. Another example would be the detection of a moving object in the visual field which activates the visual perception.

The result may be utilized in different ways. It may be integrated into the sensing mechanism, such as in saccading, where the image location with highest attention is fed to the motor control of the eyes such that focus of attention is changed to that new location (Balkenius, 1999). The extracted data may also be used as input by other cognitive functions such as the reflex behavior that flexes the neck to protect the head if a ball is thrown at you. The assigned attention may also simply be brought up to the conscious level, memorized and be part of the basis for general decision making. Often, these two ways of utilizing the extracted data are mixed. For example, saccading works as a continuous active-sensing process in which the eyes move such that other cognitive processes are fed with relevant data.

Attention mechanisms are controlled in two major ways (both are usually active simultaneously):

- Bottom-up (exogenous) control is driven by properties in sensor data. Visual attention typically works with properties like color, intensity, orientation, edges, motion, contrasts, and faces.
- Top-down (endogenous) control is driven by goals and intentions supplied by higher level cognitive functions.

A unified view is to the see bottom-up mechanisms as "general purpose" settings for top-down mechanisms governed by arbitrary or default goals.

5. Executive Attention and Cognitive Control

There is a general consensus in psychology that there is an executive system that controls our thoughts and actions to produce coherent behavior. This function is most often referred to as cognitive control (Miller, 2000; Miller & Cohen, 2001; Freedman et al., 2003; Wallis, Anderson & Miller, 2001; Miller, Erickson & Desimone, 1996) but also executive function, or executive attention. We will in the following mostly use the term cognitive control. However, the name executive attention refers to an interesting view where general brain patterns are seen as sensory data subject to an attention process. Hence, execute attention extends the attention concept discussed above to cover more than selection and activation of sensory information.

There is no exact definition cognitive control. Typical descriptions involve maintaining behavioral goals, and using these goals as a basis for choosing what aspects of the environment to attend to and which action to select. One early definition is by Minsky (1985): *"consciously manipulate thoughts and behaviors using attention to deal with conflicting goals and demands"*

5.1 A model of Cognitive Control

One model of cognitive control by Miller and Cohen (2001) is illustrated in Figure 9. Cognitive control is here seen as the ability of the brain to take control over reflexive[1] reactions and behaviors. The flow from Sensors into Behaviors and out to Actuators illustrates such bottom-up reflexive behaviors. Some of these are hardcoded by evolution or through early development. For instance, heat sensors in the fingers automatically trigger a reflexive behavior that rapidly moves the arm if strong heat is sensed. In addition to these "hardwired" behaviors, our ability to learn allows us to form new long-term memories, skills and behaviors. Still, such a system would only allow us to react to the environment in a very non flexible way. To be pro-active we need to be able to make plans to achieve goals. This is handled by the top-down acting *Executive Functions* that select among pre-existing pathways in long-term memory, activate goal-relevant pathways and suppress irrelevant ones. This means flexibility: Instead of responding mechanically in the same way to the same sensory information, we may change our response by simply changing our goals. It also makes it possible to activate goal-relevant mechanisms in the absence of any external sensory input. In Section 5.2, a detailed example of how the model works is given.

The activation and suppression of pathways can be seen as a way to assign high attention to a specific pattern at the same time as all other patterns are given low attention, as illustrated in Figure 10. The dog associates the sound from the human with negative reward and the smell of the bone and the view of the interesting object with positive reward. The cognitive control mechanism determines which behavior to execute: approach, escape, attack, or ignore by combined activation and suppression of the available alternative behaviors.

[1] In this context, the term reflexive behaviors refer to "hardwired" stimulus-response mappings that directly connect sensors to actuators.

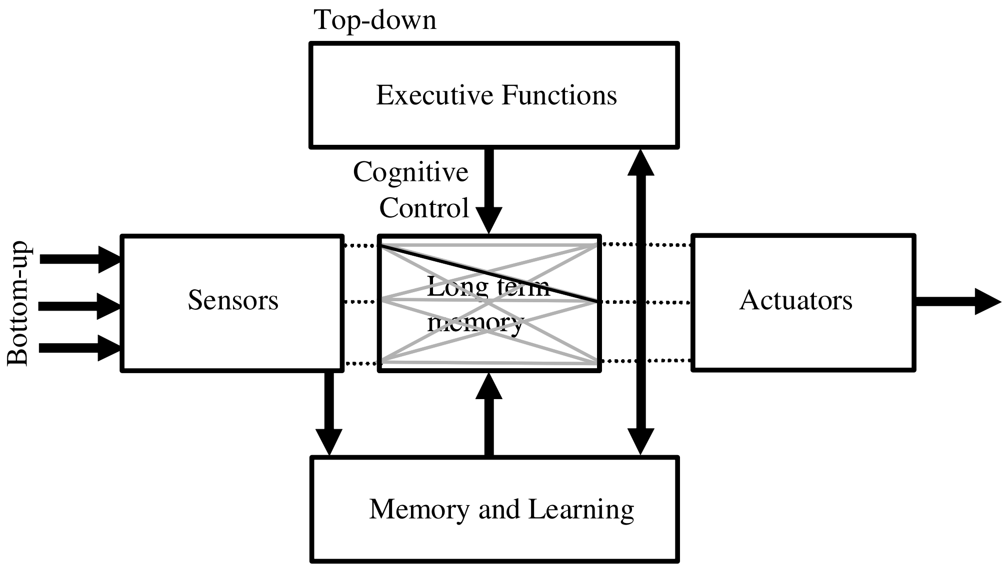


Figure 9. Model of Cognitive control proposed by Miller and Cohen (2001). Executive functions suppress and select patterns in Long term memory.

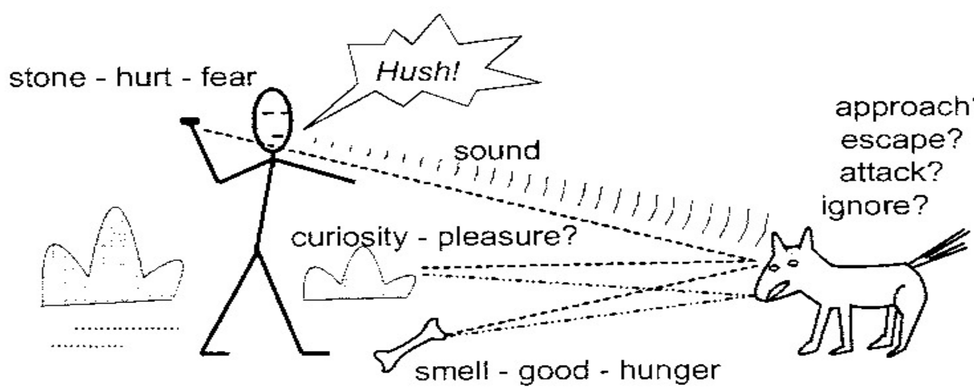


Figure 10. Cognitive control may be seen as a process of selection and suppression of patterns. Picture from (Haikonen, 2003).

Executive functions are believed to modulate many different types of cognitive activities in the brain. Some examples are (Miller, 2003):

- Generate plans and monitor task progress
- Focus on task related information
- Maintain and update goal information
- Inhibit distractions to current tasks
- Shift between different levels of cognition ranging from routine actions to complex deliberation
- Learn new responses in novel situations

5.2 Pre frontal cortex

Pre frontal cortex (PFC) is believed to be a center for executive functions in the brain of humans and other animals (Miller & Cohen, 2001). The size of PFC is correlated to the intelligence of the animals, see Figure 11. The PFC is interconnected with virtually the entire cerebral cortex with the exception of primary sensory and motor areas. This supports the view that the PFC is responsible for coordination of higher-level processes disconnected from direct interaction with the world. To accomplish this, a number of advanced competences are necessary, for example abilities to handle

- Perceptual categories
    - What is a Cat and what is a Dog?
- Concepts
    - Quantity (numerosity)
    - General and abstract
- Rules
    - Concrete: *Stop if red light*!
    - Abstract: Rules not tied to any particular stimulus or behavioural response. Example: “If the second shown picture matches the first then *Hold lever down*. Else, *Release lever*!”.

There is much evidence that the PFC actually has these capabilities. Supporting experiments are reported in the next section.

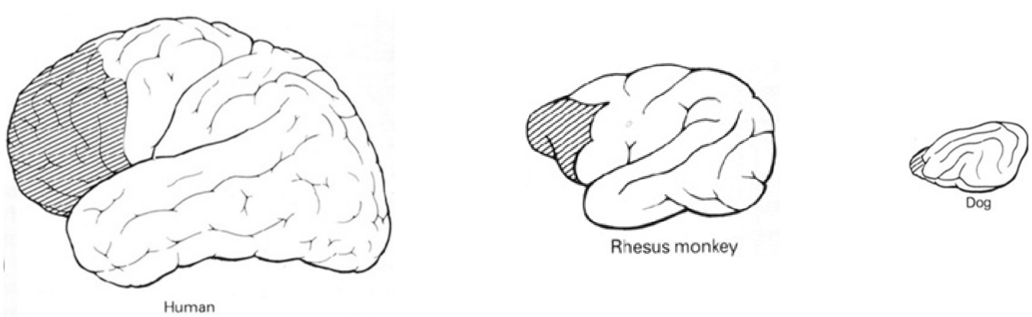


Figure 11. The size of the brain and the pre frontal cortex (PFC) for a human, rhesus monkey and dog. The size of PFC is believed to be correlated to intelligence. Pictures from (Miller, 2003).

The previously described model (see Figure 9) of cognitive control, assigns PFC as the center of executive functions. Figure 12 illustrates the proposed mechanism with an example from (Miller, 2003). Symbolic high-level units include a cue (*Phone rings*) and two possible actions. If you are *At home* you are supposed to *Answer* and if you are a *Guest*, you are supposed to *Don’t answer*. PFC communicates with most parts of cerebral cortex and is consequently modeled with access to all relevant units. During learning, or even by chance, you may sometimes execute the correct action. This is illustrated by thick lines in the figures. The *Phone* rings and you *Answer*. A reward system triggers a learning mechanism in PFC such that the neurons that were active around the time when the reward arrived get connected. This is shown as dotted lines. This result in a chain of connections that activate the *Answer* action whenever the *Phone rings* and you are *At home*. In a similar way, when you are a *Guest*, the correct action when *Phone rings* may be found by chance or by learning. The active neurons may also in this case be affected by the reward system, as illustrated in Figure 13. The signals from PFC can be seen as bias in the selection of action when the *Phone rings*. Eventually connections may also get formed outside PFC as illustrated in Figure 14. In this way responses become more reflexive and PFC gets less involved. However PFC is still important as a mechanism for suppressing inappropriate reflexes. In Figure 15, two reflex patterns are illustrated. One corresponding to the *Guest* case, and one to the *At home* case. Since the latter case receives much more reward, a reflex directly connecting *Phone rings* with *Answer*ing is triggered even when being a *Guest*. However, the *Don't answer* reflex is reinforced by the PFC pattern previously developed (see Figure 14) such that the correct response is executed.

In general, the model says that activity patterns in PFC represent internal goals and provide bias signals for other brain patterns. In this way the PFC guides pathways between inputs, internal states, and outputs. The theory is supported in studies of humans and monkeys with PFC damages.

The same general ideas were used by Cohen, Dunbar and McClelland (1990) in a specific model that explains the classic Stroop test. In this test (Stroop, 1935) the test subjects were given sequences of color names, such as "green, red, blue, purple". Some sequences were written with all black ink and some with colored ink, however in colors not matching the individual words, such as **Blue Purple Red Green Purple Green**. If the task was to read the words, the reading ability (reaction time and accuracy) was not affected by the colors. However, if the task was to name the ink color, the written word confused the reading. The explanation is that regular word reading is automatic (overlearned) while the color naming requires attention. Cognitive control is needed to bias processing of this weaker pathway. The theory was modeled with a feed-forward neural network that illustrated how attention (focus on colors or focus on words) could be used as an additional input to strengthen one of several possible pathways. The model is illustrated in Figure 16. The network was trained with input and output patterns and converged to the weights and biases shown in the figure.

Gaps in time between associated stimuli and/or responses are inevitable. To still be able to associate and connect neurons, *sustained activity* in the PFC neurons is essential (Fuster, 1985). This means that neurons are able to sustain their activity for several seconds even in the absence of further stimulation. In addition to enabling the formation of association between entities separated in time, this also allows task rules to be maintained until the task is completed.

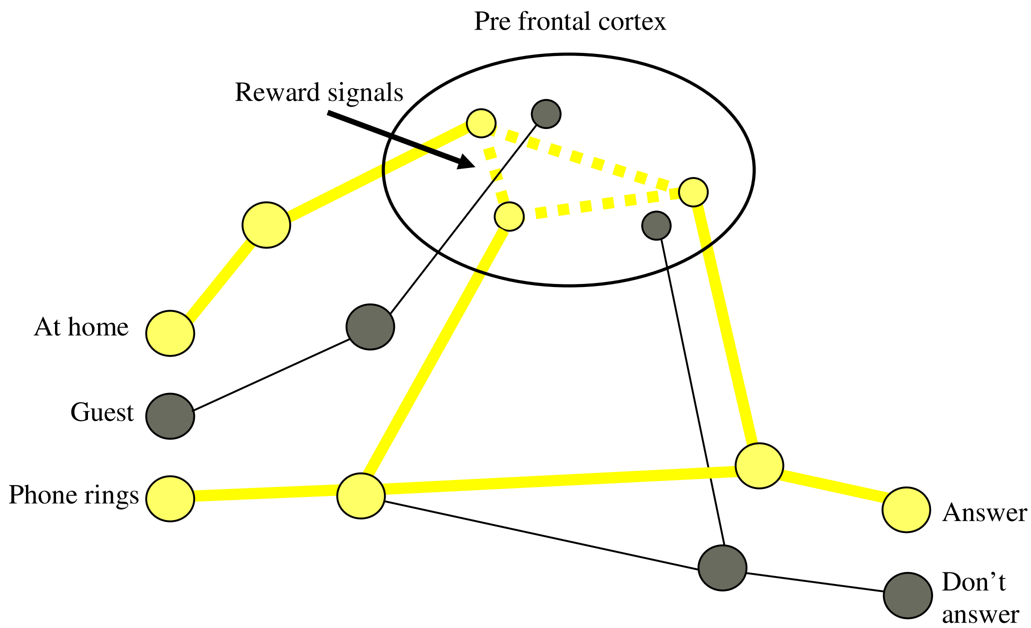


Figure 12. If the *Phone rings w*hen you are *At home*, and you by chance *Answer*, reward signals create associative patterns in PFC (dotted lines) such that the path going to *Answer* will be automatically activated when you are *At home*.

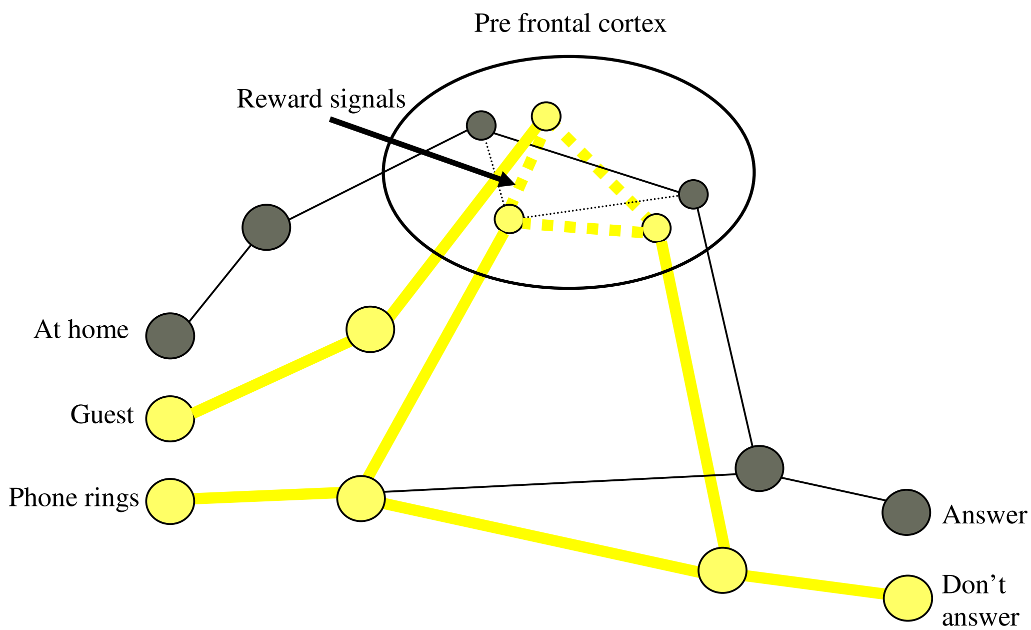


Figure 13. If the *Phone rings w*hen you are a *Guest*, and you by chance *Answer*, the path from *Phone rings* to *Don't answer* is reinforced. The PFC pattern from Figure 12 is here permanent but not active since *At home* is not active. Hence, the *Don't answer* action is chosen.

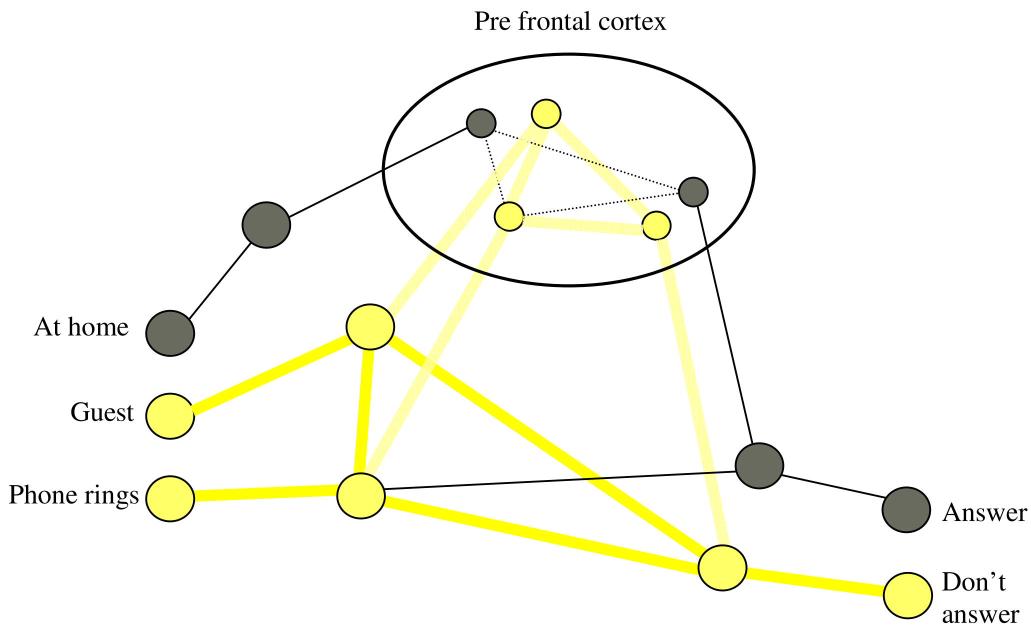


Figure 14. Eventually connections may get formed also outside PFC. The figure shows this for the *Guest* case. In this way responses become more reflexive and PFC gets less involved.

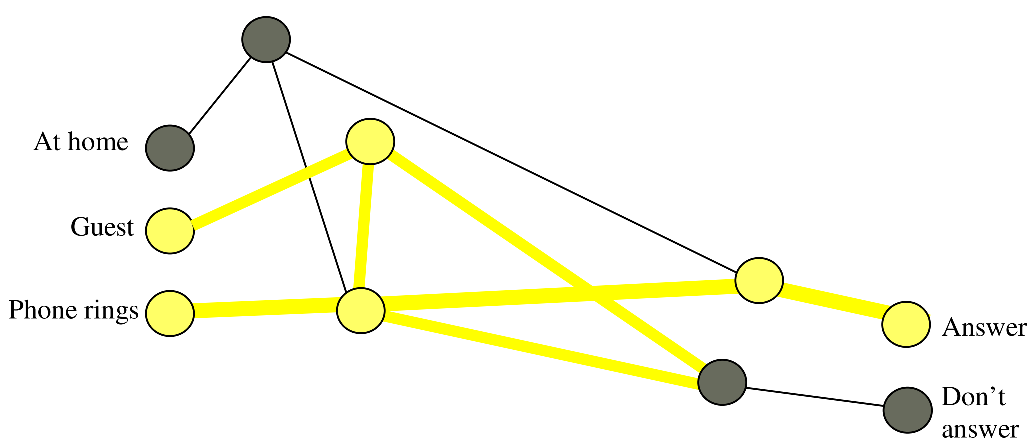


Figure 15. A direct pathway between *Phone rings* and *Answer* might become more strongly established than the other pathway since you are more often at home. As a result, you develop an unconditional reflex to answer the phone (even when you are a *Guest*).

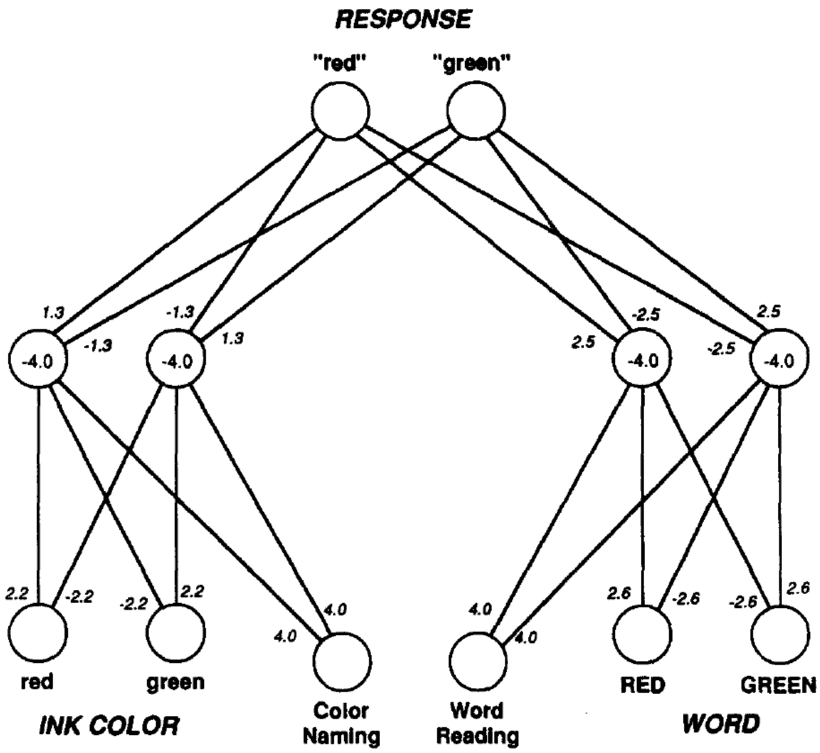
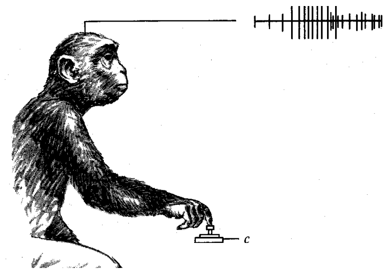


Figure 16. Neural network model of the Stroop test (picture from Cohen Dunbar & McClelland, 1990). Top-down cognitive control introduces bias such that either one of the tasks *Color Naming* and *Word Reading* is executed.

5.3 Experiments related to Executive functions



A research group lead by Earl Miller at MIT (http://www.millerlab.org/) have conducted a number of experiments (Miller, 2000; Miller, 2003; Miller & Cohen, 2001; Freedman et al., 2003; Wallis, Anderson, & Miller, 2001; Miller, Erickson, Desimone, 1996; Miller, Freedman & Wallis, 2003) related to cognitive control and executive function.

Figure 17. Picture from (Miller, 2003).

In a series of experiments, the group trained monkeys on tasks designed to isolate cognitive operations connected to executive control. During task execution, they recorded neuron activity from groups of single neurons. The experiments show that high level concepts such as categories and numerosity are represented by single neurons or small groups of neurons.

A "Dog Neuron" in the PFC

In one experiment (Miller, 2003), the monkeys were exposed to artificial pictures of animals that looked like mixtures of dogs and cats, see Figure 18. The mixture was controlled such that the amount of "dog" and "cat" could be varied continuously from a pure "cat" to a pure "dog". The firing rate of a selected neuron is illustrated in Figure 19. Each curve represents response versus time for exposure to one specific picture (see the annotations). After approx 1500ms, the neuron has "decided" what to think and shows a clear separation between the dog and cat pictures.

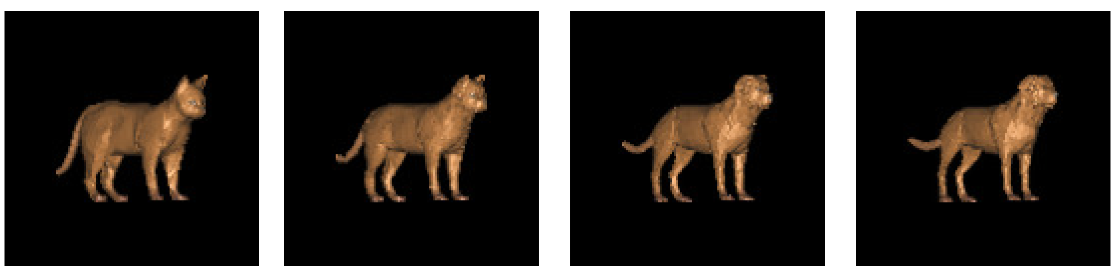


Figure 18. Morphed pictures showing varying mixtures of dog and cat. Picture from (Miller, 2003).

This neuron clearly distinguished between pictures of cats and dogs, but did not distinguish between different levels of catness and dogness. This means that the neuron is encoding information about category membership and not low-level physical properties.

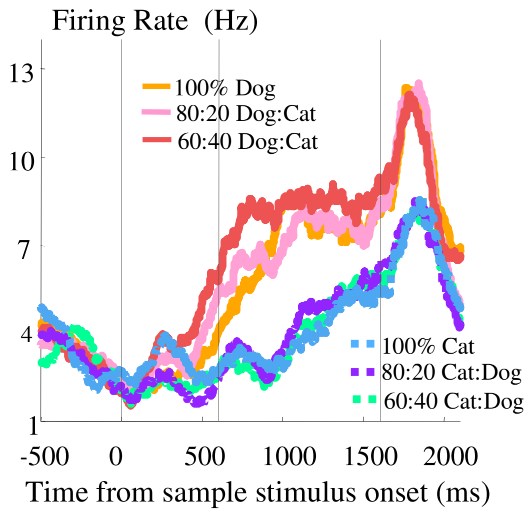


Figure 19. Neuron response for the "Dog neuron" when the monkey is exposed to pictures with varying mixtures of cat and dog. Picture from (Miller, 2003).

Quantity (numerosity)

In this experiment (Miller, 2003) the test subjects were shown pictures with varying numbers of dots, see Figure 20 to the right. The diagram to the left of the figure shows the firing activity for a neuron that showed the highest activity to pictures with one dot. After approx 1500ms, the neuron has "decided" on firing or not. The neurons representing other numbers showed a progressive drop-off in activity as numbers increased. The behavior is independent on the exact layout of the pictures. The experiment shows how the neurons represent the number concept rather than specific sensory data. The result fulfils the demands for Numerosity (Miller, 2003):

- Preservation of numerical order – numbers are not independent categories.
- Numerical Distance Effect – discrimination between numbers improve with increasing distance between them.
- Numerical Magnitude Effect – discrimination between numbers of equal numerical distance is increasingly difficult as their size increases (e.g., 1 and 2 are easier to tell apart than 5 and 6).

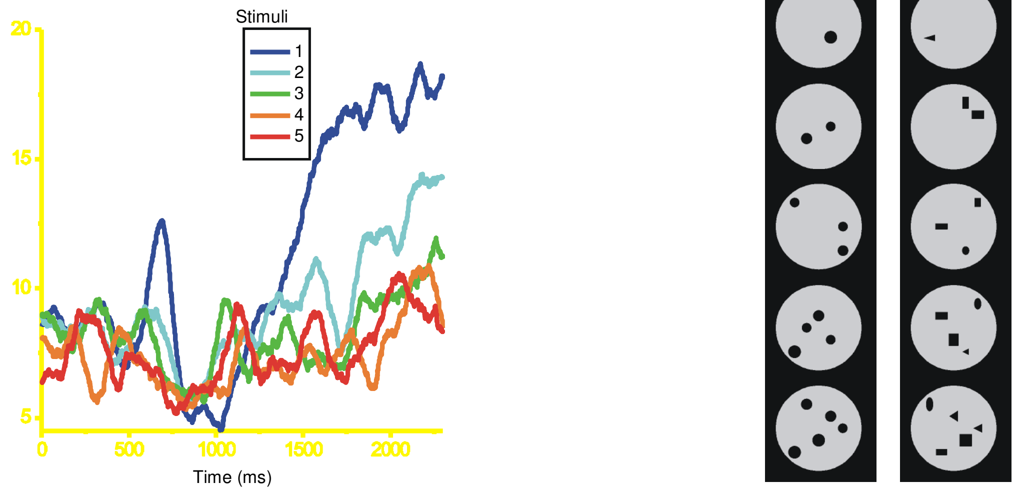


Figure 20.. Firing rates for neurons representing numbers when the test person is exposed to one of the top circles to the right. The "one" neuron stabilizes as winner after approx. 1500 ms. Picture from (Miller, 2003).

Abstract rules

In this experiment (Wallis, Anderson & Miller, 2001), illustrated in Figure 21, monkeys were trained to switch between two rules: *Match* and *Non-match*. Monkeys saw a sample picture followed by a delay and then a single test picture. For the match rule, the monkeys were given a reward if they released a lever when the pictures were the same, and held it if pictures were different. For the non-match rule, the rule was reversed. The trained monkeys could perform the task the first time they saw a new set of pictures. I.e.: The behavior did not depend on fixed stimulus-response associations, but on an "understanding" of the underlying rule.

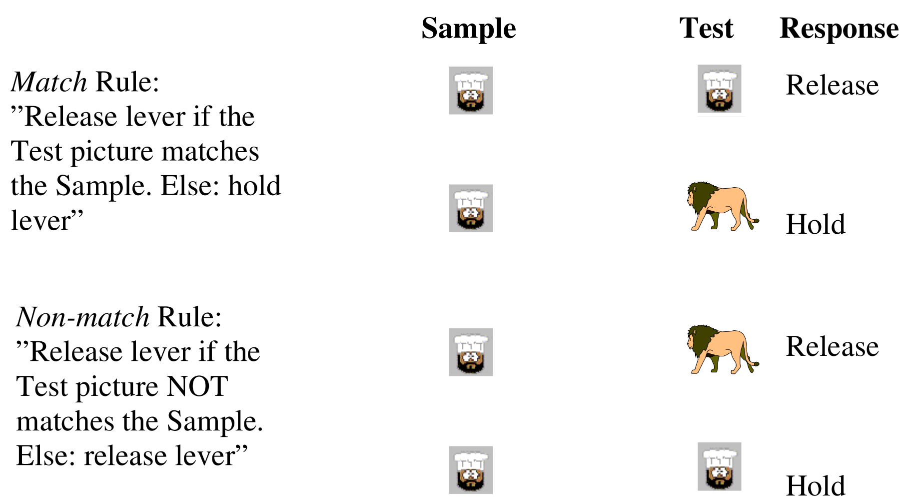


Figure 21. Layout of experiment to show how abstract rules can be formed in the pre frontal cortex of monkey brains. Picture adopted from (Miller, 2003).

6. Conclusions

Attention can be seen as mechanisms responsible for selection/attenuation of brain patterns of various kinds. The three main categories outlined in the introduction of the article; perceptual, action related, and mental attention may be described in a common framework performing selection and rejection of brain patterns. Perceptual attention works with competing sensor dimensions and features. Action based attention works with competing actions, and motor schemas. Mental attention works with other types of memory items not directly connected to sensor or motor areas. While these memory items are clearly different in many respects, they are all neuron patterns that may be subject to cognitive control, i.e. various selection and rejection mechanisms.

The reported experiments related to pre frontal cortex (PFC) show that cognitive control has a number of high-level concepts at its disposal as result of advanced data processing: Classification, Numerosity and Advanced abstract rule handling all yield high-level information that can be part of the cognitive control in PFC, for instance as input data to executive functions.

7. A Cognitive Control Model of Saccading

To illustrate how perceptual attention and cognitive control may be intertwined, saccading will be described with a model inspired by Miller's and Cohen's model of cognitive control introduced in Section 4.1. In Figure 22, a bottom-up visual perception module extracts salient features, objects, and locations from the stream of vision data. The results are placed in memory where also a representation of the high-level goal, to report the age of all persons in the picture, resides. The executive function performs the saccade control by enabling motor patterns that are appropriate for scanning the picture for age-related information (such as looking for faces). For this it needs access both to the high-level goal and to the high-level percepts being already computed and placed in memory. The enabling of motor patterns is accompanied by disabling other competing patterns, such as the ones corresponding to other high-level goals and to the default eye motion pattern. The enabled behavior will move the eyes to locations where age-related information may be extracted by *Other cognitive functions*.

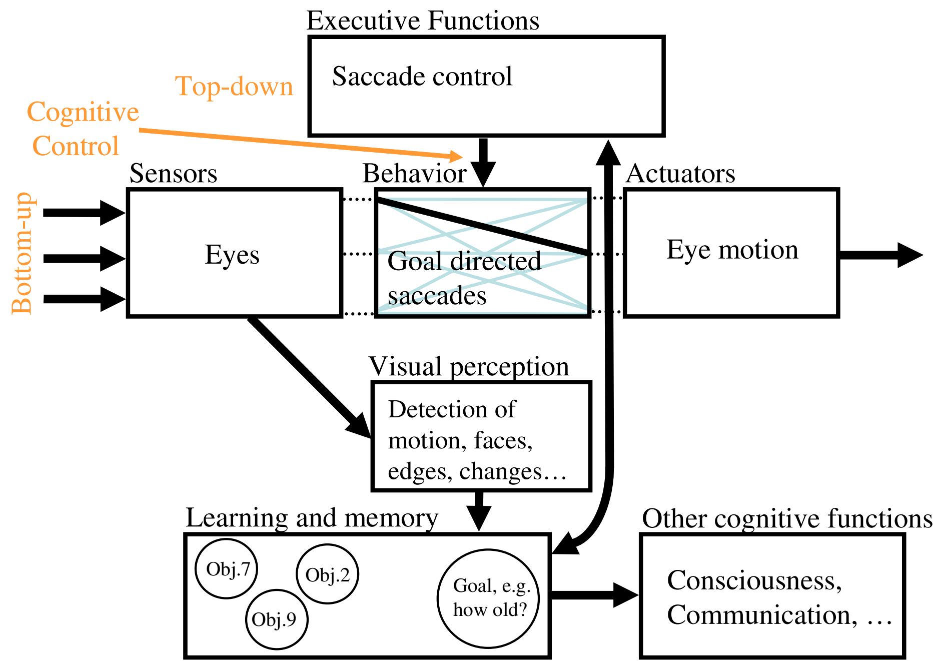


Figure 22. Model of saccading using Miller's and Cohen's (2001) model of cognitive control. The goal specification is used by the execute function to select an eye motion schema appropriate for the set goal.

References

Balkenius, C. (2000). Attention, habituation and conditioning: Toward a computational model. Cognitive Science Quarterly, 1(2), pp. 171-204.

Breazeal, C., & Scassellati, B. (2001). Infant-like Social Interactions between a Robot and a Human Caretaker, Adaptive Behavior.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. Psychological Review, 97, pp. 332-361.

De Pisapia, N., Repov, G., Braver, T.S. (2007). Computational models of attention and cognitive control. The Cambridge Handbook of Computational Psychology (Edited by Ron Sun), Chapter 15, pp. 422-450.

Freedman, D.J., Riesenhuber, M., Poggio, T., & Miller, E.K. (2003). J. Neuroscience, 23:5235-5246.

Fuster, J. M. (1985). The prefrontal cortex, mediator of crosstemporal contingencies. Hum. Neurobiol. 4, 169–179.

Haikonen, P. (2003). The Cognitive Approach to Conscious Machine, Imprint Academic.

Haith, M. (1980). Rules That Babies Look by. Hillsdale, NJ: Erlbaum.

Hellström, T. (2009). Imitation and Priming in Humans, Animals and Intelligent Systems, Technical report Department of Computing Science, Umeå University, Umeå, Sweden.

Hommel, B.,Ridderinkhof, K. R., Theeuwes, J. (2002). Cognitive control of attention and action: Issues and trends, Psychological Research 66: 215–219.

Itti, L., Koch, C. (2001). Computational Modelling of Visual Attention, Nature Reviews Neuroscience, Vol. 2, No. 3, pp. 194-203.

Itti, L., Koch, C., Niebur, E. (1998). A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp. 1254-1259, Nov.

Kanwisher N, Downing P. (1988). Separating the wheat from the chaff, Science, 282(5386):57-58.

Kawamura, K., Gordon, S., Ratanaswasd, P., Garber, C., & Erdemir, E. (2007). Implementation of Cognitive Control for Robots, Proc. of the 4th COE Workshop on Human Adaptive Mechatronics (HAM), pp. 41-54, Tokyo Denki University, Japan, March 2-3.

Koch, C., Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. Human Neurobiology, 4 (4). pp. 219-27.

Miller, E.K. (2000). The prefrontal cortex and cognitive control. Nature Reviews Neuroscience, 1:59-65

Miller, E.K. (2003). Cognitive Control, Fundamental of Brain and Mind Lecture Series, Mass. Institute of Technology, June 11-13.

Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annual Review of Neuroscience, 24:167-202.

Miller, E.K., Erickson, C.A., Desimone, R. (1996). "Neural mechanisms of visual working memory in prefrontal cortex of the macaque", Journal of Neuroscience. 16 (16), pp. 5154-5167.

Miller E.K., Freedman D.J., Wallis J.D. (2002). The Prefrontal Cortex: Categories, Concepts, and Cognition, Philosophical Transactions of the Royal Society London: Biological Sciences 357, pp. 1123-1136.

Miller E.K., Freedman D.J., Wallis J.D. (2003). The Prefrontal Cortex: Categories, Concepts, and Cognition. In: The Physiology of Cognitive Processes, Parker A., Derrington A., & Blakemore C. (eds.), Oxford University Press, pp 252-273.

Minsky, M. (1985). The Society of Mind, NY: Simon & Schuster.

Mozer, M. C., & Baldwin, D. S. (2008). Experience-guided search: A theory of attentional control. In J. Platt, D. Koller, & Y. Singer (Eds.), Advances in Neural Information Processing Systems 20 (pp. 1033-1040). Cambridge, MA: MIT Press

Pfeifer, R., Scheier, C. (1997), Sensory-motor coordination: The metaphor and beyond, Robotics and Autonomous Systems, 20(2-4): 157-178.

Revel, A., & Nadel, J. (2007). How to build an imitator. In K. Dautenhahn & C. Nehaniv (Eds.), Imitation and social learning in robots, humans and animals: Behavioural, social and communicative dimensions, pp. 279-300. Cambridge: Cambridge University Press.

Scassellati, B. (2001). Foundations for a Theory of Mind for a Humanoid Robot, 2001, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, PhD Thesis.

Simons, D.J., & Chabris, C.F. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events, Perception, volume 28, pp. 1059–1074.

Simons, D.J., & Rensink R.A. (2005). Change blindness: Past, present, and future, Trends in Cognitive Sciences, Vol.9 No.1.

Slater, A. (1998). The competent infant:i innate organization and early learning in infant visual perception. In Slater A. (ed.), Perceptual Development. Hove: Psychology Press, 105-30.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions.
Journal of Experimental Psychology, 18, pp. 643-662.
Treisman, A. (1988). Features and objects: The fourteenth Barlett memorial lecture, The Quarterly Journal of Experimental Psychology 40A: 201–237.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention, Cognitive Psychology 12, pp. 97–136.

Vinter, A. (1986). The role of movement in eliciting early imitations. Child Development, 57:66–71.

Wallis, J.D., Anderson, K.C., & Miller, E.K. (2001). Single neurons in the prefrontal cortex encode abstract rules, Nature, 411:953-956.

Wolfe, J.M. (1994). Guided Search 2.0: A Revised Model of Visual Search, Psychonomic Bulletin & Review, 1(2): 202-238.

Yarbus, A. (1967). Eye movements and vision, New York: Plenum Press.