Cognitive Neuroscience
Section 6

There are many aspects to attention. It can be *controlled*. It can be *focused* on a particular sensory modality or item. It can be *divided*. It can *set* a perceptual system. It has a finite *capacity*. It can determine what items enter memory and it can itself be influenced by memory.

The single most characteristic feature of attention is its *selectivity*. It selects one item for perception to the exclusion of others, or selects one motor act to the exclusion of others.

Why is this selectivity necessary? It has been explained in neural terms as being necessary to conserve certain limited “resources”.

Such resources could be critical neuronal populations in the cortex that would be overwhelmed with processing tasks if attentional mechanisms did not allocate them to one specific task at a time.
Recast in terms of cognit activation, the explanation for attentional selectivity may be that the activation of multiple cognits at the same time could cause massive interference by activation of associated cognits. Attention would then be necessary to select only a limited number of cognits for activation at a time.

As Fuster states, “The role of attention is to select one of those networks at a time and to keep it active for as long as it serves a cognitive function or the attainment of a behavioral goal.”
Biological roots of attention

Any organism must process a continuous flow of sensory information and produce a continuous stream of motor output in order to adapt to the environment in a way that is conducive to its survival.

In lower organisms the only feedback from effector systems to sensory systems is external, i.e. through the external environment. Hence, their behavior consists of simple reactions to sensory stimuli.
Higher organisms also have *internal feedback* mechanisms from effector to sensory systems, allowing a greater degree of behavioral control.

Internal feedback provides a means for effector systems to regulate sensory systems to “set” some of them for improved receptivity and analysis of sensory input ("improved" meaning more consistent with adaptive behavior) and to suppress processing in others.

Internal feedback is a form of corollary discharge, which allows “tuning” of sensory systems to produce more adaptive behavior through enhanced sensory analysis.

For example:
1) a visual stimulus is initially processed hierarchically in a feedforward manner, activating multiple visual and frontal executive areas at all hierarchical levels.
2) higher-order visual areas provide feedback to lower areas for perception
3) executive areas may initiate movements that enhance visual processing, e.g. moving the eyes to foveate the stimulus for improved acuity.
4) executive areas may also send *internal feedback* signals to the visual system to directly enhance sensory processing.
5) this feedback represents a kind of *attentional control* that provides the visual system with information from the frontal hierarchy, e.g. about the behavioral relevance of the stimulus.
Attentional control operates to both facilitate and attenuate sensory and motor processes. This control may operate by a combination of excitation and inhibition. Excitation and inhibition interact at all levels of the central nervous system.
Sensory example: *lateral inhibition* (on-center; off-surround) enhances spatial and temporal contrast of LGN cells – useful for *edge detection* in cortex.

Motor example: *reciprocal innervation* of extensor & flexor muscles – motoneurons in the spinal cord innervate extensor & flexor muscles; when the extensor is contracted, its muscle spindles send afferent sensory signals back to the spinal cord; afferents synapse on extensor motoneuron to maintain contraction by positive feedback & synapse on inhibitory interneurons that synapse on flexor motoneuron; inhibition allows cooperative action of the two antagonistic muscles.

Figure 6.2. Reciprocal interplay of excitation and inhibition in sensory and motor systems. A: Reactions of a cell in the lateral geniculate body to light in different parts of its receptive field: the cell is excited by central illumination and inhibited by peripheral (annular) illumination; thus, the illumination of either the center or the periphery inhibits input from the other, enhancing contrast. B: Reciprocal spinal innervation of flexor and extensor muscles facilitating extension of the knee. From Kuffler and Nicoll (1976), slightly modified, with permission.
The same cooperative duality of excitation and inhibition seen in sensory and motor systems is utilized in attention:

1) enhancement of processing in selected cognits (i.e. foci of attention or targets)
2) suppression of processing in competing cognits (i.e. distractors)

In hierarchical systems, with both divergence and convergence from one level to the next, there may be critical nodes that could become bottlenecks to the flow of activity without there being control of their inputs. In other words, selection may be critical to prevent such bottleneck nodes from becoming overwhelmed by an excess of inputs.

Attentional processing can be explained within the cortical system of cognitive networks without need for specialized structures dedicated to attention as a specialized function.
Perceptual attention

Proper functioning of the cerebral cortex requires arousal by the “continuous inflow of nonspecific activating influences from several structures of the brain stem.”

Attention is not arousal, but requires a minimal level of arousal.

The mesencephalic reticular formation is an important source of diffuse excitatory cholinergic input to the cortex.

This system increases its activity in arousal from sleep to wakefulness, and is tonically active in the maintenance of vigilance and general alertness.

Emotion is known to influence attention. Influences on perceptual attention from mood and affect probably come about through inputs to cortical perceptual network hierarchies that originate in the limbic system. The amygdala plays an important role in this, but anterior thalamic nuclei, entorhinal & cingulate cortical areas also are involved.

The selective aspect of attention can be explained by processes restricted to the neocortex.
The “control” of perceptual attention is a balance between bottom-up and top-down influences.

**Bottom-up attention**

Bottom-up attention is thought to originate at the lowest levels of sensory processing and to be determined by properties of sensory stimuli, such as **saliency** and **novelty**. That is, stimuli that are salient or novel will be processed with greater bottom-up attention. The classical example of bottom-up attentional processing based on saliency in vision is **figure-ground separation**, or **pop-out**.

Bottom-up attentional control was studied extensively by Julesz, who used the name **preattentive processing**.

According to Julesz, preattentive processing:

1) is fast
2) operates in parallel
3) is automatic (i.e. does not require consciousness)

Bottom-up processing of salient stimuli may be an elaboration in the cortex of the simple center-surround antagonism for contrast enhancement in the LGN discussed above.
Top-down attention

Top-down influences occur during the processing of perceptual information as the selective modulation of cortical networks in perceptual hierarchies by those in the executive hierarchy or those higher in the perceptual hierarchy.

Top-down influences in perceptual attention promote perceptual matching. Remember the ART model:

a) Sensory inputs gain access to sets of higher-level networks representing different perceptual categories.

b) Each categorical network performs a matching process with the sensory input signal.

c) The categorical networks compete, and the one with the closest match generates a top-down expectation signal that is sent to lower-level networks.

This process is a form of top-down attention because it selects certain categories for further processing, and blocks others.

The executive hierarchy may also be a source of top-down attention. Executive cognits may selectively “prime” specific perceptual cognits (at different levels) that are consistent with ongoing behavioral performance (as part of the perception-action cycle).
Evidence for top-down attentional modulation

A. Unit recording in monkey inferior temporal cortex:

(1) In one experiment (Moran & Desimone 1985), attention to a spot of light in the receptive field (rf) of an IT cell:
   a. enhances the response (i.e. firing rate) of the cell to that spot in the rf
   b. suppresses the response to unattended spots in the rf

Therefore, both the inclusionary and exclusionary aspects of attention take place within the rf of the same IT neuron.
(2) In another experiment (Fuster 1990), monkeys were required to attend selectively to different features (color or symbol) of a complex visual stimulus (gray symbol on colored disk). Depending on the symbol, the monkey had to memorize either the symbol or the color, and use the memorized information after a delay to complete the task.

![Graph showing neuronal responses](image)

Some IT neurons responded indiscriminately to all stimuli. Others responded selectively to one color or one symbol, i.e. they were symbol-selective (15%) or color-selective (21%). Some of the color-selective cells were color-attentive, meaning that they responded significantly more to their preferred color (e.g. green or red) when it was relevant (i.e. had to be retained) than when it was not.
Thus, when the monkey was required to first interpret the symbol in order to decide that a color (e.g., green) was relevant, the (green) color-attentive cells showed a stronger differential response to color than otherwise. E.g., the amount by which the response to green was above that to red was enhanced when the green color was attended.

This suggests that the color-attentive IT neurons received top-down attentional modulation from higher cortical levels in order to attend and put in memory their appropriate color.

(3) Tomita et al (1999) recorded the responses of IT neurons to visual stimuli using the posterior-split-brain paradigm. They transected the posterior corpus callosum & anterior commissure, sparing the anterior corpus callosum, which interconnects the prefrontal cortices. Because of this preparation, IT neurons do not receive bottom-up visual inputs from stimuli presented in the ipsilateral visual hemifield. Any influence on IT from ipsilateral stimuli must cross from contralateral to ipsilateral prefrontal cortices and then descend from ipsilateral prefrontal cortex.
The monkeys were first trained to memorize paired visual associates. Then, the cue stimulus was presented in the ipsilateral hemifield and the choice stimulus in the contralateral hemifield.

Neurons in ipsilateral IT showed delayed activation by the cue stimulus, indicating that they were activated by a top-down signal. Behavioral performance was severely impaired if this top-down signal was blocked.
B. Human ERP studies:

(1) Attention to the location or quality of a stimulus enhances ERP components following the stimulus.
(2) Attentional enhancement in vision is largest over posterior visual cortex.
(3) It occurs with long enough latency to suggest top-down control.

C. Human clinical studies:

Studies of patients with lesions involving prefrontal cortex suggest that top-down attentional control originates in this region. These patients show deficits in performing tasks requiring them to attend to one or more stimuli to correctly perform the task.

Specifically, patients with large lateral prefrontal cortex lesions have difficulty:
1. attending to changes in the sensory environment
2. shifting attention from 1 aspect of the sensory environment to another

They also show lowered attentional enhancement of ERP components.
These results suggest that in controlling the executive sequences required for task performance, the prefrontal cortex also controls the perceptual input that it receives from posterior cortical areas.

We see here that prefrontal damage produces symptoms similar to the classical neglect syndrome described as resulting from posterior parietal damage. How can we reconcile these observations?
The *hemispatial neglect syndrome* is a neuropsychological condition in which, after damage to one hemisphere of the brain, a deficit in attention to and awareness of one side of space is observed. It is defined by the inability for a person to process and perceive stimuli on one side of the body or environment that is not due to a lack of sensation. It results most commonly from brain injury to the right cerebral hemisphere, most commonly to the right posterior parietal cortex.
There are 3 main cortical regions of the right hemisphere where damage produces neglect:
1) Inferior Parietal Lobe (IPL)
2) Inferior and Middle Frontal Gyri (MFG, IFG)
3) Temporo-Parietal Junction (TPJ)

Figure 4. Cortical right hemisphere brain regions associated with neglect include the angular (Ang) and supramarginal (Smg) gyri of the inferior parietal lobe (IPL), the temporo-parietal junction (TPJ), and the inferior (IFG) and middle frontal (MFG) gyri. Additionally, the diagram also shows the superior parietal lobe (SPL) and intraparietal sulcus (IPS).
Mesulam has proposed a distributed network for spatial attention that includes some of these same areas.

Figure 4. A large-scale distributed network for spatial attention.
D. Human imaging studies

In humans performing visual attention tasks (Kastner et al 1999), prefrontal and inferior temporal cortical areas are jointly activated. This suggests that the prefrontal cortex functionally interacts with IT in visual attention.
Working memory

The term *working memory* originated in cognitive psychology to refer to the temporary memory of sensory stimuli used by people during cognitive operations.

It now has a more general meaning as *active memory* used for short-term perceptuo-motor processing in humans or animals.

Fuster sees working memory as a form of “attention focused on the internal representation of a recent event for a pending action”.

**Neural evidence for working memory:**

Neurons in the prefrontal cortex of monkeys are found that show elevated firing during the delay period of a delayed match-to-sample task in comparison to the periods between trials.

**The delayed match-to-sample task:**

a) a sample stimulus is briefly presented, followed by the delay period.

b) after the delay, the monkey is presented with 2 or more test stimuli, and is required to respond in a way that indicates which one is a match to the sample.

c) the motor act depends, or is *contingent*, on the stimulus information held in memory.

d) the monkey receives a reward for correct performance.
Properties of delay-period prefrontal neuron activation:
  1) it is strictly dependent on the task requirement to perform a motor act contingent on some information kept in working memory
  2) it is not induced by expectancy of the reward
  3) it is correlated with the monkey’s ability to remember the information
  4) it can be obliterated by distraction

Interpretation of these findings:
  a) prefrontal neuron activation reflects sustained attention to the sample stimulus after it has disappeared from view.
  b) these cells are part of a prefrontal cognit that represents the properties of the stimulus within the framework of task-related behavior.

Goldman-Rakic was the first neurobiologist to identify prefrontal cells related to working memory. Based on anatomical pathway tracing, she proposed that working memory is maintained in a distributed network connecting prefrontal cortex with posterior areas. This proposal is similar to Mesulam’s proposal of a distributed network for spatial attention that was based on human lesion studies. Both proposals include prefrontal and posterior parietal cortex.
The organization of working memory in large-scale distributed networks is supported by physiological studies:

1) visual memory cells have been found in inferior temporal cortex
2) spatial memory cells have been found in posterior parietal cortex
3) tactile memory cells have been found in somatosensory cortex

Conclusion: In executing a memory task, executive aspects of the task activate networks in prefrontal cortex and perceptual aspects activate networks in different posterior cortical areas depending on the modality of the sensory components involved.

In human neuroimaging studies of working memory:

a) the lateral prefrontal cortex is active
b) sensory association cortex in the parietal and inferior temporal cortex is active
c) the anterior cingulate cortex (part of Mesulam’s spatial attention network) is also often active

These studies show that working memory and attention are closely related:

a) working memory consists of the joint activation of a distributed network of perceptual and executive memory
b) selective attention to items of working memory consists of sustained and selective activation of parts of that network
A cooling study: Fuster (1985) trained monkeys to perform a delayed match-to-sample task for color: presentation of sample color disc, followed by a 10-20 sec delay period, followed by 2 or more discs of which one is a match to the sample, followed by choice of the match, followed by reward for correct choices.

These monkeys had implants that allowed cooling of lateral prefrontal and inferior temporal cortices.

While the task was being performed, one of the 2 areas was cooled and units were recorded from the other.

Results of the cooling study:
a) Bilateral cooling of either area induced performance deficits in the working memory task, but not for performance of the task without delay.
b) Many neurons in the un-cooled cortex showed a reversible decline in their ability to distinguish colors, i.e. their role in working memory was diminished. This effect of cooling in one cortex on performance in the other may be explained as the disruption of reentrant excitation between the two cortices.

Conclusion from the cooling study: a distributed network maintains visual working memory in this task. It consists of an executive component in prefrontal cortex and a perceptual component in inferior temporal cortex. Both are necessary for correct task performance.
Executive attention

In the control of action, neural resources must be selectively allocated, as in the control of perception. This requires *executive attention*.

Within the context of the perception-action cycle, it is obvious that attention must control executive as well as perceptual processes. “The selective allocation of executive cognits and effector networks is essential for purposive action.”

Since the prefrontal cortex is the highest and most integrative stage in the action hierarchy, it follows that it exerts the highest level of executive attentional control.

What has often been called executive control is just another term for executive attention.

Damage to lateral prefrontal cortex leads to a set of symptoms that have been characterized as the *disexecutive syndrome* (Baddeley 1986). This syndrome consists of the inability to formulate, initiate, and execute plans of action.

It can be attributed to a failure of activation of cognits that store action schema, as well as those necessary for the initiation of action, and the maintenance of the course of action toward a goal.
Executive attention is needed to select from alternatives that exist about the course of action to follow at every moment. It arises from the interaction of executive cognits and perceptual cognits.

In Fuster’s view, the mirror neurons of premotor cortex (discussed in Section 3), represent evidence for executive attention. The cognits of which these neurons are a part must be responsible for attention to action because they are activated during execution of an action and also by observing another individual execute that same action.
From neuroimaging studies, three major regions of prefrontal cortex are implicated in executive attention:

1) anterior cingulate cortex

Activated by tasks that require high motivation or resolution of conflict. In the Stroop task, the subject must name the color meaning of a word that appears in a different color, e.g. the word “red” appearing in green. Here conflict is set up between the color in which the word appears and meaning of the word itself. The cingulate cortex is involved in resolving the conflict in this task.
2) lateral prefrontal cortex
Activated by:
   a) shift of set or gaze
   b) integration of information across time

3) orbital frontal cortex
Activated in tasks requiring suppression of interference -- OFC plays a role in the exclusionary aspect of attention.

All three regions may be activated to differing degrees in tasks that require executive attention.
Set and expectancy

As stated above, the lateral prefrontal cortex is involved in tasks requiring the integration of contingencies over time.

Temporal integration is necessary when the information required for behavioral performance is discontinuous in time, i.e. the items of information are available only at certain times, with gaps in between.

The lateral prefrontal cortex is involved in mediating contingencies between elements of information that are separated in time.

The mediation of cross-temporal contingencies is a crucial function for complex behavior that allows information that was only available at some time in the past to be used as if it were available at the present. Potentially, the availability of past information may have been distributed over many previous times, and all the information is available together at one time. This capability gives species that possess it an advantage in competition with other species that do not.

In the mediation of cross-temporal contingency, the lateral prefrontal cortex makes use of two complementary functions: working memory and preparatory set.
(a) Working memory is attention directed to the internal representation of previously encountered sensory information. It is retrospective (i.e. directed to the past).

(b) Preparatory set is the priming (i.e. partial pre-activation) of perceptual and motor cognits for expected events. It is prospective (i.e. directed to the future). It is also called attention directed to the future.

Preparatory set is attention focused on expected events and on the anticipated consequences of present events and actions.

This ability to anticipate future consequences of present activity lies at the heart of the ability to formulate plans of action. Lesions of lateral prefrontal cortex disrupt this ability.
Neural correlates of set:

1. The *Contingent Negative Variation (CNV)* is a slow surface-negative potential that is maximal over frontal cortex during the period between S1 and S2, where the contingency of S2 on S1 has been learned. Because it is related to the expectancy of S2 brought about by the occurrence of S1, the CNV is also called an expectancy wave. The size of the CNV increases with time during the S1-S2 period. Therefore, it is larger when this period is longer.

2. The *Lateralized Readiness Potential (LRP)* is also a slow surface-negative potential like the CNV.

The LRP occurs when a subject is about to make a planned motor response. It is often studied in a paradigm where the subject receives a warning stimulus (WS), then has the opportunity to prepare a motor response (RS) during the delay period, then makes the RS in response to an imperative stimulus (IS).

The LRP occurs during the delay period of this paradigm. It is maximal over SMA, and is larger contralateral to the RS when the RS is unilateral.
3. Neurons related to set have been found in lateral prefrontal cortex in close proximity to working memory cells. They increase their discharge as the time of an impending action approaches. Discharge accelerates as a function of the probability with which the animal can precisely predict the motor act to be performed.

The relation between frontal EEG slow waves (e.g., CNV & LRP) and unit activity in set is explained by the *excitability regulation theory* of Rockstroh et al (1989):

a) The excitability of prefrontal cortex is increased by depolarization of the apical dendrites of pyramidal neurons – by synapses from input fibers from the *nonspecific thalamic nuclei*.

b) Depolarization of pyramidal cell populations in frontal cortex gives rise to: (a) increased surface negativity in the EEG; and (b) increased firing rates of the pyramidal cells.

Summary: preparatory set complements working memory in temporal integration. For temporal integration to take place, information must be transferred from neurons representing a signal in the recent past (working memory cells) to neurons that prepare the consequent response to that signal in the near future (set cells).
Execution and monitoring

Executive attention is intimately related to the actual execution of action.

Both consist in the selective allocation of components of the motor system for a specific action.

Therefore, it is not surprising that systems for the control of executive attention highly overlap those for execution of action: the same cortical areas that control executive attention are also involved in controlling gaze, head position, and body postures – actions involved in orienting behavior – actions that promote executive attention.

Executive attention systems are involved in **visual spatial attention**.

1) They control the oculomotor system to control the direction of gaze.
2) By controlling the direction of gaze, they restrict the content of visual perception.
3) In keeping with the perception-action cycle, the control of gaze by executive attention areas is to a large degree based on perceptual analysis of the environment.
Although the frontal cortex is organized hierarchically, it does not follow that participation of frontal areas in execution of movements follows a rigid top-down sequential order. There is evidence (e.g., Kalaska et al 1998) that different frontal areas are active in parallel, controlling different movement parameters, during preparation for, and execution of, movements.

Executive attention is also highly distributed in space and time for planning and control of action.

Automatic, well-rehearsed actions may be performed largely in parallel at lower levels of the motor hierarchy, and without executive attention.

Actions requiring high levels of executive attention, however, require the temporal integrative function of prefrontal cortex, and consequently a greater reliance on serial processing.
The term *monitoring* refers to attention directed to sensory input specifically for the guidance of action. Monitoring is a “tool” of executive attention, used by the prefrontal cortex to control actions, both as they occur and in preparation for prospective events.

**Monitoring reflects the role of prefrontal cortex in executive attention:**
1) The prefrontal cortex exerts both inclusionary and exclusionary control on perceptual and motor systems.
2) The prefrontal cortex receives perceptual signals that modulate its control of action.
3) Perceptual signals to prefrontal cortex may:
   a. result from processing of sensory inputs by perceptual cortical areas
   b. come from those areas without sensory inputs
   c. arise as a consequence of movement in the form of proprioceptive feedback
   d. be generated in response to efferent copies of current motor action received as corollary discharge from the motor system
4) The exclusionary control is needed to protect goal-directed actions from interference caused by distracting stimuli and incompatible schema; it is thought to originate in orbital frontal cortex (OFC).