Fundamental Components of Attention

Eric I. Knudsen

Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305-5125; email: eknudsen@stanford.edu

Key Words
working memory, stimulus selection, prefrontal cortex, parietal cortex, superior colliculus, vision

Abstract
A mechanistic understanding of attention is necessary for the elucidation of the neurobiological basis of conscious experience. This chapter presents a framework for thinking about attention that facilitates the analysis of this cognitive process in terms of underlying neural mechanisms. Four processes are fundamental to attention: working memory, top-down sensitivity control, competitive selection, and automatic bottom-up filtering for salient stimuli. Each process makes a distinct and essential contribution to attention. Voluntary control of attention involves the first three processes (working memory, top-down sensitivity control, and competitive selection) operating in a recurrent loop. Recent results from neurobiological research on attention are discussed within this framework.
INTRODUCTION

To behave adaptively in a complex world, an animal must select, from the wealth of information available to it, the information that is most relevant at any point in time. This information is then evaluated in working memory, where it can be analyzed in detail, decisions about that information can be made, and plans for action can be elaborated. The mechanisms of attention are responsible for selecting the information that gains access to working memory.

Four component processes are fundamental to attention: (a) working memory, (b) competitive selection, (c) top-down sensitivity control, and (d) filtering for stimuli that are likely to be behaviorally important (salience filters). Working memory is a highly dynamic form of memory that operates over periods of seconds and temporarily stores selected information for detailed analysis (Baddeley 2003). Competitive selection is the process that determines which information gains access to working memory (Desimone & Duncan 1995). Top-down sensitivity control is a process that regulates the relative signal strengths of the different information channels that compete for access to working memory (Egeth & Yantis 1997). Salience filters automatically enhance responses to stimuli that are infrequent in space or time or are of instinctive or learned biological importance (Koch & Ullman 1985).

The engagement of these processes leads directly to increased behavioral sensitivity and shortened response latencies (the traditional metrics of attention) as well as to the cognitive benefits that we associate with attention.

The past decade has witnessed an enormous surge in neurophysiological research on attention. Nearly all this research has focused on various phenomena that are associated with attention. One major goal, however, is to understand the neural mechanisms that underlie attention. Progress toward this goal would be facilitated by a model that accounts for the phenomena of attention in terms of neurobiological components. This chapter establishes the framework for such a model and discusses recent results within the context of this framework.

The proposed framework for attention is shown in Figure 1, a framework inspired by the models of Desimone & Duncan (1995) and Miller & Cohen (2001). The central nervous system contains information about the world, about stored memories, and about the internal state of the animal. At any point in time, the information that gains access to working memory is selected by a competitive process from this repertoire of information on the basis of its relative signal strength. Signal strength reflects the combined effects of the quality of the encoded information, top-down bias signals, and bottom-up salience filters. The information with the greatest signal strength enters the circuitry for working memory and competes with existing information for control of working memory. The information that controls working memory also directs top-down bias signals that modulate the signal strengths of relevant ascending representations, forming a recurrent loop that underlies voluntary attention.

WORKING MEMORY

Working memory is a special form of memory with extraordinary capabilities (Baddeley 2003). Working memory holds a limited
Figure 1

Functional components of attention. Processes that contribute to attention are shown in red. Information about the world (green ellipse) is transduced by the nervous system and is processed by salience filters that respond differentially to infrequent or important stimuli (bottom-up). Neural representations in various hierarchies encode information about the world, movements, memories, the animal's emotional state, etc. A competitive process selects the representation with the highest signal strength for entry into the circuitry that underlies working memory. Working memory can direct top-down bias signals that modulate the sensitivity of representations that are being processed in working memory. The selection process can also direct top-down bias signals that reflect the result of the competitive selection. Working memory and competitive selection direct eye movements and other orienting behaviors that modify the effects of the world on the animal’s nervous system. Corollary discharges associated with gaze control modulate sensitivity control. Voluntary attention involves working memory, top-down sensitivity control, and competitive selection operating as a recurrent loop (dark arrows).
amount of information for periods of seconds while the information is evaluated and manipulated in a uniquely powerful and flexible fashion on the basis of the animal’s internal state and stored memories. Working memory itself comprises competitive processes, and multiple types of information may compete for full control of the circuitry underlying working memory at any moment in time. The degree to which one type of information gains full control of working memory reflects the relative strengths of the competing representations. The information that is held in working memory serves as the basis for decisions and the planning of complex behaviors (Genovesio et al. 2006, Yoshida & Ishii 2006) and, most importantly for this discussion, controls top-down signals that modulate the sensitivity of neural representations that contribute to that information (Miller & Cohen 2001).

Working memory and attention are inextricably inter-related. When an animal attends to an object, information associated with that object enters working memory. Conversely, information in working memory is information that is associated with objects to which an animal has attended (LaBar et al. 1999). Thus, working memory represents the objects of attention.

The capacity of working memory to manipulate information is limited at any one time to a single domain (e.g., verbal, mathematical, visuospatial). The portions of the brain that participate in working memory depend on the information being processed. For example, functional imaging studies on humans show that verbal working memory tasks activate the ventrolateral prefrontal cortex (PFC) and language areas in the temporal and inferior parietal cortex on the left side (Schumacher et al. 1996). In contrast, visuospatial working memory tasks activate the dorsolateral PFC, inferior parietal cortex on the right side, and high-order visual areas in the occipital cortex (Smith et al. 1996).

The PFC is one area of the brain that is activated consistently in working memory tasks. Clinical reports in humans and lesion studies in monkeys confirm a central role of the PFC in working memory (Miller & Cohen 2001). These studies indicate that lesions in the PFC cause general deficits in working memory, with no apparent deficits in sensory discrimination or motor performance (Diamond & Goldman-Rakic 1989, Duncan et al. 1996, Vendrell et al. 1995). In contrast, although lesions in other areas of the brain can also lead to deficits in working memory, the effect of these lesions is specific for the sensory or motor information represented in these areas and is accompanied by corresponding sensory or motor deficits.

The data suggest that working memory is a function that is usually distributed widely in the brain, with the PFC acting as an executive controller. During working memory tasks, the PFC engages with cortical and subcortical regions that process sensory information, motor information, information about internal state, or stored memories, depending on the task at hand (Baddeley 2003, Constantinidis & Wang 2004). The extensive, reciprocal anatomical connections of the PFC with most cortical and many subcortical regions are consistent with this view (Miller & Cohen 2001).

Neurophysiological studies in the monkey PFC have revealed a neural correlate of working memory, referred to as “delay-period” or “persistent” activity (Funahashi et al. 1989, Fuster & Alexander 1971). In monkeys trained to remember a target stimulus for brief periods of time, neurons in the PFC not only respond to the target while it is presented, but also continue to discharge for many seconds after the target has disappeared and until the animal reports the target with a response (Figure 2). This persistent activity exhibits many of the properties of working memory. It is tuned for the stimulus parameters upon which the monkey must render a decision (Figure 2b; best versus worst sample), tuning that presumably reflects the involvement of various sensory, motor, limbic, or memory areas in working memory (Miller et al. 1996, Suzuki et al. 1997). The persistent activity is
modulated according to decisions made by the animal regarding the stimulus (Kim & Shadlen 1999). Moreover, the activity persists until the animal reports the target stimulus, even though the target may be followed by a series of “distracter” stimuli that are behaviorally irrelevant (di Pellegrino & Wise 1993, Fuster 1995, Miller et al. 1996). An exceptionally high proportion of neurons in the PFC exhibits delay-period activity (Miller & Cohen 2001). Together, these data support the central role of the PFC in working memory.

All types of information about the world and the organism are processed in working memory. However, the spatial location of a stimulus in the world is a feature that is analyzed in a special pathway, reflecting the importance of location as a parameter for making decisions and planning goal-oriented behaviors. The processing in working memory of
information according to location is referred to as spatial working memory.

Data from functional imaging studies demonstrate that spatial working memory tasks consistently activate two major cortical areas: the dorsolateral PFC and the posterior parietal cortex (PPC) (Curtis 2006). These two structures are strongly interconnected by reciprocal pathways (Schwartz & Goldman-Rakic 1984). Both structures contain neurons with delay-period activity that is tuned for the location of stimuli in space (Constantinidis & Wang 2004), and lesions of either structure interfere with the monkey’s ability to plan responses using the remembered locations of stimuli (Chafee & Goldman-Rakic 2000).

As discussed above, the properties of the PFC demonstrate its critical role in working memory, including spatial working memory (Constantinidis & Wang 2004). In contrast, the functional properties and organization of the PPC, which are discussed later, indicate that this region is likely to be more involved in the processes of competitive selection and top-down sensitivity control than in working memory (Colby & Goldberg 1999). Persistent activity in the PPC is less prevalent and more susceptible to interruption by distracting stimuli (Powell & Goldberg 2000) than is the persistent activity in the PFC (Fuster 1995, Miller et al. 1996). Instead of robustly representing a behaviorally relevant target that is stored in working memory, activity in the PPC represents the relative salience of all stimuli (Bisley & Goldberg 2003) as well as the goal locations of movements that a monkey intends to make (Andersen et al. 2004, Batista & Andersen 2001, Ipata et al. 2006).

**TOP-DOWN SENSITIVITY CONTROL**

In the context of attention, not only does working memory accept, store, and manipulate information, but it also generates signals that improve the quality of the information that it processes (Miller & Cohen 2001, Miller & D’Esposito 2005). One mechanism for improving information quality is to direct orienting movements toward targets (Figure 1, gaze control). For example, by directing orienting movements of the eyes toward an object, working memory optimizes the resolution of visual information about the object (Andersen et al. 2004, Colby & Goldberg 1999). The same principle applies to orienting movements of other appendages (e.g., the hand) and other sensory systems (e.g., somatic sensation).

A second strategy for improving information quality is to modulate the sensitivity of neural circuits that represent the information (Figure 1, sensitivity control). This top-down mechanism can improve the signal-to-noise in all domains of information processing: sensory, motor, internal state, and memory. Experiments employing functional imaging in humans have revealed the vast extent of brain areas that can be modulated by attention-related bias signals. Along with consistent activation of the PFC, attention tasks are capable of enhancing activity in many regions of the neocortex, limbic cortex, basal ganglia, pulvinar nucleus, superior colliculus, and cerebellum. The regions of the brain that are activated during a task depend specifically on what domain of information is being attended (Corbetta et al. 1991, Mesulam 1999, Muller et al. 2006, Shomstein & Behrmann 2006, Shomstein & Yantis 2004, Summerfield et al. 2006). These kinds of experiments have not, however, distinguished increased activity due to information being processed in working memory (persistent activity) from increased responses due to top-down regulation of neural sensitivity.

The effects of top-down bias signals have been observed neurophysiologically in monkeys trained to discriminate among sensory stimuli. When monkeys make choices on the basis of the properties of a stimulus, the responses of neurons that represent the stimulus increase compared with when the same stimulus is presented but is behaviorally irrelevant.
(Desimone & Duncan 1995). Such increases in responsiveness have been observed at many levels in various information processing hierarchies (Maunsell & Cook 2002, McAdams & Reid 2005, McAlonan et al. 2006). The largest and most consistent increases are observed at higher levels in the hierarchies owing, at least in part, to the fact that activity at these higher levels reflects increases that have occurred at all lower levels.

Top-down modulations of neural responsiveness are precise for the features upon which judgments will be made (Desimone & Duncan 1995, Maunsell & Treue 2006), a precision that distinguishes attention-related modulations from general arousal. Consequently, these top-down modulations improve the signal-to-noise of the encoded information: Only neurons with receptive fields that contain the stimulus and that are tuned for the parameter values of the attended stimulus exhibit an increase in sensitivity. In contrast, neurons tuned for different stimulus parameters often exhibit a decrease in sensitivity (Chelazzi et al. 1993, Reynolds & Desimone 2003, Treue & Martinez Trujillo 1999). The inhibition of neurons that are not tuned for the stimulus suggests that top-down bias signals activate local inhibitory circuitry, as well as excitatory circuitry. This balanced influence on excitatory and inhibitory circuitry increases responsiveness while maintaining sharp feature tuning (Shu et al. 2003).

The effect of the top-down bias signals on the responses of individual neurons is multiplicative in high-order visual areas tuned for line orientation (V4; Williford & Maunsell 2006) or the direction of stimulus motion (MT; Martinez-Trujillo & Treue 2002); but see Reynolds et al. 2000: In both V4 and MT, attention causes neural responses to increase proportionately more as the stimulus more closely matches the orientation or direction tuning of the neuron. In addition, the effectiveness of the top-down signals is graded with the difficulty of the task. For example, when a monkey must discriminate among line orientations and the difference between possible orientations is small, the attention-related increases in neural responses to a target are greater than when the difference between possible orientations is large (Boudreau et al. 2006, Spitzer et al. 1988). Thus, the effectiveness of multiplicative bias signals increases as the demand for resolution increases.

Top-down modulations of neural responsiveness can be precise not only for features but also in their timing (Khayat et al. 2006, Motter 1994). Neurons that represent an attended stimulus may exhibit elevated spike rates that decline rapidly once a monkey has made its decision but before it has made its response (Ghose & Maunsell 2002). This match of elevated discharge rates with the period of decision indicates that top-down bias signals modulate rapidly (within tens of ms) and that they increase differentially during the decision process.

Increases in neuronal sensitivity caused by the task relevance of stimuli have been documented in circuits at many levels of processing, from the thalamus and primary sensory cortex to the PFC (Khayat et al. 2006, McAlonan et al. 2006, Miller & D’Esposito 2005). These increases have been consistently interpreted as reflecting solely attention-related processes. Maunsell (2004) has cautioned, however, that in many behavioral paradigms, response increases may include the effects of value judgments associated with the stimuli (top-down signals) that may influence attention, but may act independently of attention processes. Neurons in the LIP increase the strength of their responses to visual stimuli depending on the magnitude and probability of the reward associated with the stimulus (Platt & Glimcher 1997, Sugrue et al. 2005), and the paradigms used to study the effects of reward are essentially the same as those that researchers have used to study attention-related effects. The effects of reward on neural responses are likely to increase at higher levels in processing hierarchies, and the effects should be strong particularly in circuits that underlie working
memory, where the importance of information is the critical factor in determining what information is maintained. In the future, behavioral tasks that seek to analyze attention-related processes will need to differentiate between these two potential effects by manipulating reward and attention load independently.

**BOTTOM-UP SALIENCE FILTERS**

Information does not need to be modulated by top-down bias signals to gain access to working memory (Egeth & Yantis 1997, James 1890). Certain properties of the world can evoke exceptionally strong neural responses that may win access to working memory (Itti & Koch 2001, Remington et al. 1992). Stimulus-driven access to working memory, commonly referred to as bottom-up attention, reflects the effects of salience filters (Figure 1) at many levels in the central nervous system that select for properties of stimuli that are likely to be important. Typically, salient stimuli occur infrequently in space or time, for example, a sudden sound, a flash of light, or a red dot in a field of green dots. Salience filters may also select for stimuli of instinctive (e.g., looming stimuli) or learned (e.g., voice of a parent) biological importance. The nervous system responds automatically to such salient stimuli with unusually strong responses and/or with responses distributed across large populations of neurons.

A variety of neural mechanisms give rise to salience filters. Mechanisms of adaptation can create filters for stimuli that occur infrequently in time. Adaptation mechanisms, both intrinsic to cells as well as those generated by network dynamics, cause neurons that respond strongly at first to reduce their responses or to cease responding entirely to sustained or repeated stimuli. Network connectivity can create filters for stimuli that occur infrequently in space. For example, networks containing widespread lateral inhibition, particularly divisive (shunting) inhibition, can detect isolated stimuli.

The unusually strong neural activation that results from these filters gives the representations of salient stimuli an advantage in the competition for access to working memory. Such stimuli are perceived as “popping out” from the scene (Egeth & Yantis 1997). In some cases, the advantage conferred on the representation of a salient stimulus is sufficiently great that the representation wins the competition for working memory, even while working memory is engaged in processing other kinds of information (Egeth & Yantis 1997, James 1890).

Neural signals representing salient stimuli may influence working memory momentarily, for a period of less than a few hundred ms (Bisley & Goldberg 2003). Once the information enters working memory, its importance can be evaluated and compared with the importance of other information already being processed in working memory (Baddeley 2003). The information that is deemed to be of greatest importance maintains control of working memory and serves as the basis for subsequent top-down sensitivity control (Miller & Cohen 2001, Miller & D’Esposito 2005).

Unexpected or highly salient stimuli can trigger top-down modulations of sensitivity and orienting behaviors even before the neural activity representing the stimulus enters working memory. Although the information associated with unexpected or highly salient stimuli also enters working memory (and, therefore, is attended), during the first brief period of time just after stimulus onset, competitive selection, sensitivity control, and gaze control operate independently of working memory (Figure 1; arrows from competitive selection to sensitivity and gaze control). Highly salient stimuli can begin to modulate the sensitivity of ascending circuits and can trigger eye saccades within 120 ms of stimulus onset. Such short latency saccades (“express saccades”) are mediated by the superior colliculus (and not by the motor cortex), they
occur only in response to salient stimuli, and they never occur when judgments about properties of the stimulus must be made to select a correct endpoint for the saccade (McPeek & Keller 2004, Schiller et al. 1987). The latter property indicates that express saccades are initiated before they can be guided by working memory processes.

**SPACE-SPECIFIC SENSITIVITY CONTROL**

Of all features, stimulus location is arguably of the most fundamental importance. The behavioral implications of a stimulus are profoundly affected by the location of the stimulus relative to the animal. Stimulus location also serves as a powerful filter for selecting information for intensive analysis in working memory and for guiding goal-directed behaviors (Andersen et al. 1997, Colby & Goldberg 1999, Maunsell & Treue 2006). For these reasons, just as working memory employs top-down bias signals to improve information quality for other features (often referred to as feature attention), it also employs space-specific bias signals that improve the localization and representation of stimuli (referred to as spatial attention) (Figure 3).

The effects of space-specific bias signals on neural responsiveness have been demonstrated in monkeys trained to attend a cued location. Behaviorally, animals increase their sensitivity and decrease their response latency to target stimuli presented at the cued location (Desimone & Duncan 1995). At the same time, neurons at high levels in the visual pathway increase their discharge rates to stimuli that appear at the cued location (Maunsell & Treue 2006, Treue & Maunsell 1999).

The effect of space-specific bias signals is most dramatic when the target and a distracter stimulus are both located within the receptive field of a high-order neuron. In this case, attention can sharpen the spatial tuning of the neuron to the cued location, diminishing or eliminating the suppressive effects of the distracter so that the neuron represents the cued stimulus almost exclusively (Everling et al. 2002, Luck et al. 1997, Treue & Maunsell 1999, Womelsdorf et al. 2006a). At the same time, neurons with receptive fields that overlap spatially with the cued location, but that are tuned for stimulus parameters that differ from those of the target, exhibit a decrease in sensitivity. These push-pull effects of top-down bias signals increase the signal-to-noise of the representation. These effects indicate that top-down bias signals activate local excitatory as well as local inhibitory neurons in the network.

A fundamental problem with using “location” as the basis for regulating response sensitivity is that location is represented in a large number of different frames of reference in the nervous system. Working memory often carries out analyses and makes decisions in frames of reference that are abstract, linked neither to sensory surfaces nor to motor workspaces (Baddeley 2003). In contrast, top-down bias signals exert their influence on visual cortical areas that represent information in retinocentric frames of reference (Van Essen 1985). Thus, whenever working memory processes information in a reference frame that is not retinocentric, it must translate spatial information from that reference frame into a retinocentric frame before these bias signals can be sent to appropriate portions of visual cortical areas.

The translation of spatial information from one reference frame into another for the generation of space-specific bias signals may be one function of the PPC (Figure 3). The PPC receives input from all sensory modalities as well as movement-related corollary discharges and proprioceptive feedback (Andersen et al. 1997). The PPC is organized into functional areas that represent information relevant to different kinds of movements. For example, the lateral intraparietal (LIP) area represents information relevant to eye saccades; the medial intraparietal area represents information relevant to arm movements; and the ventral intraparietal area represents information relevant to the
Figure 3
Schema for top-down sensitivity control. **Blue arrows:** Bias signals that regulate neural responsiveness. **Green arrows:** bottom-up information filtered for salience in the superior colliculus and in visual cortical areas. Top-down bias signals from working memory that modulate the representations of nonspatial aspects of information are transmitted directly to the cognate representations (upper right). Top-down bias signals from working memory that modulate the representation of information on the basis of object location are transmitted to the posterior parietal cortex (PPC) where they are represented in various reference frames. Retinocentric bias signals are transmitted from working memory also to the frontal eye fields (FEF). Retinocentric bias signals are distributed from the PPC and FEF to retinocentric sensory areas. Bias signals for modulating representations in other reference frames are transmitted from the PPC to the cognate representations. Bottom-up bias signals combine with top-down bias signals in all these structures.

head (Colby & Goldberg 1999). These various representations of spatial information require transformations from sensory frames of reference into frames of reference appropriate for different brain regions. Transformations are accomplished by integrating sensory spatial information with body or limb position information. For example, by integrating
eye-centered (retinocentric) visual information with eye position information, the LIP represents visual information relative to the head (head-centered coordinates) (Andersen et al. 2004). Within the various functional areas of the PPC, sensory information is transformed into a number of different coordinate frames (Figure 3; arrow to nonretinocentric space representations.

The information about the relative positions of the eyes, head, limbs, and body could also be used by the PPC to translate high-order spatial representations from working memory into frames of reference that are appropriate for top-down control of sensory processing areas. For example, the LIP might combine spatial information from working memory that is represented in an egocentric frame of reference with eye-position and head-position information and, thereby, translate the spatial information into a retinocentric frame of reference. Once spatial bias signals are represented in a retinocentric frame of reference, they can be distributed to retinocentric sensory representations to regulate their sensitivity (Figure 3; dashed box).

The distribution of retinocentric spatial bias signals may be carried out by both the LIP and the forebrain gaze control area, the frontal eye fields (FEF) (Figure 3). The FEF is reciprocally connected with the LIP and the PPC and mediates voluntary control of gaze direction (Schiller et al. 1987, Stanton et al. 1995). The role of the FEF in controlling orienting eye movements has been explored extensively in the past, but only recently has its role in distributing top-down bias signals become appreciated (Awh et al. 2006, Moore et al. 2003).

Psychophysicists were the first to discover the tight linkages that exist between gaze control and spatial attention (Rizzolatti et al. 1987). They found, for example, that each time we make a saccadic eye movement to a new location, our sensitivity to stimuli at that location increases tens of ms before the eyes move (Shepherd et al. 1986). Thus, orienting eye movements and spatial attention are functionally linked (although separable) in the brain.

The tight linkage between gaze control and spatial attention has been demonstrated directly by applying electrical microstimulation to gaze control areas in the monkey (Moore & Fallah 2004). Monkeys were trained to monitor a cued location in space without moving the eyes (covert attention tasks). When tested behaviorally, they exhibited increased sensitivity to luminance changes of stimuli specifically at the cued location. An electrode for microstimulating the cortex was then placed in the FEF. When high current levels (50–150 μA) were delivered through the electrode, the eyes made a fixed-vector saccadic movement to a new location, defined in retinocentric coordinates as the movement field for the site. When the same FEF site was stimulated with low current levels, below the level required to evoke eye movements, monkeys demonstrated an increase in behavioral sensitivity to stimuli located specifically in the movement field of the microstimulation site, as though their attention had been directed to that location by the focal activation of the FEF. Investigators have reported analogous results for the effects of low-level microstimulation of the superior colliculus on behavioral detection of visual motion or stimulus change (Cavanaugh et al. 2006, Cavanaugh & Wurtz 2004, Muller et al. 2005).

The same kind of weak electrical microstimulation of the FEF also causes top-down bias signals to be distributed to retinotopically matched areas in V4 (Armstrong et al. 2006, Moore & Armstrong 2003). When the activity of neurons in V4 was recorded during FEF microstimulation, neurons with receptive fields that contained the movement field of the FEF site, and were tuned for the properties of the stimulus, increased their responsiveness to visual stimulation (Figure 4). The response increases mimicked the response increases that occur in V4 when monkeys are cued to attend a location (Reynolds et al. 1999).
Figure 4

Effect of FEF microstimulation on the responses of V4 neurons in the monkey. Microstimulation in the FEF and single-unit recordings in the V4 were carried out while the monkey fixated a central spot. For each FEF site, the current threshold for inducing an eye saccade and the direction and magnitude of the saccades (movement field) were measured. For V4 neurons, the location of the visual receptive field and the tuning for stimulus bar orientation were measured, and the preferred and nonpreferred bar orientations were determined. (a) Microstimulation and recording sites are shown on a lateral view of the monkey brain. The horizontal bars indicate the timing of the appearance of the visual stimulus in the V4 receptive field and the timing of the 50 ms FEF microstimulation. The peristimulus-time histogram shows responses of a single V4 neuron to a preferred stimulus with (red) and without (black) low-level FEF microstimulation; microstimulation current levels were always well below the threshold for inducing eye movement. Unit responses to the visual stimulus were enhanced immediately after FEF microstimulation. (b) Summary data of the average change in the responses of V4 neurons ($n = 33$) that was caused by FEF microstimulation. The data compare responses to preferred (open bars) and nonpreferred (solid bars) stimulus orientations. Error bars indicate the standard error of the mean. When the movement field for the FEF site (red arrow) aligned with the V4 receptive field (dashed circle), FEF microstimulation induced increases in V4 responses that were much greater for the preferred than for the nonpreferred stimulus. The effect increased dramatically when a distracter stimulus was present in the visual field (middle pair of bars). When the movement field of the FEF site did not align with the V4 receptive field (right pair of bars), FEF microstimulation decreased responses to the preferred stimulus. FEF microstimulation altered V4 activity only when a stimulus was present in a unit’s receptive field, and the direction of the effect depended specifically on the mutual alignment of the FEF movement field with the V4 receptive field. The data are from Moore & Armstrong (2003).

The role of the forebrain gaze control area as a distributor of top-down space-specific bias signals holds across species and across sensory modalities (Winkowski & Knudsen 2006). The FEF equivalent in the avian brain is the arcopallial gaze field (AGF). In barn owls, electrical microstimulation applied to the AGF evokes fixed-vector eye and head saccades. Low-level microstimulation causes space-specific increases in the responsiveness of auditory neurons in the optic tectum (the avian equivalent of the mammalian superior colliculus), similar to the effects of FEF microstimulation on visual responses in V4 (Armstrong et al. 2006).

The bias signals elicited by FEF or AGF microstimulation are precise in both time and space. The increase in neural sensitivity that results from microstimulation typically lasts 100 ms after the end of stimulation (although...
the effect lasts longer in Figure 4a) (Moore et al. 2003, Winkowski & Knudsen 2006), about the same duration as the increase in behavioral performance following FEF microstimulation (Moore & Fallah 2004). In addition, AGF microstimulation sharpens the spatial tuning of auditory neurons in the optic tectum according to the spatial location represented at the AGF stimulation site. Analogously, FEF microstimulation enhances the responsiveness of V4 neurons only for stimuli located in the portion of the receptive field that corresponds to the movement field of the FEF stimulation site (Armstrong et al. 2006). This sharpening of spatial tuning is analogous to the spatial sharpening of receptive fields that researchers have observed in high-order visual areas during spatial attention tasks (Everling et al. 2002, Luck et al. 1997, Treue & Maunsell 1999, Womelsdorf et al. 2006a). Because visual receptive fields in most high-order visual areas are larger than movement fields in the FEF (Van Essen 1985), FEF microstimulation is likely to sharpen visual spatial tuning in these areas, although this idea remains to be tested.

Single-unit recordings in the FEF indicate that the neurons that encode top-down bias signals are different from the neurons that exclusively encode eye saccades (Thompson et al. 2005). In monkeys trained to suppress saccades to specific targets, the responses of visual FEF neurons that encode the target are strongly enhanced while movement-related FEF neurons that encode eye saccades to that location are suppressed. Visual activity in the FEF represents salient stimuli and could be distributed to other visual areas to act as bias signals for modulating sensory responsiveness in a space-specific manner (Figure 3). The data from this experiment are consistent with a common origin for retinocentric bias signals and eye movement control signals, but these signals separate at the level of the FEF, which can distribute retinocentric bias signals even without commanding an eye movement.

The FEF heavily projects back to the LIP, which, in turn, projects to a wide range of cortical and subcortical sensory areas (Moore et al. 2003, Schall et al. 1995, Stanton et al. 1995). Hence, the effects of FEF microstimulation on space-specific modulations of neuronal responsiveness may well be mediated by the LIP. However, the FEF also projects extensively to these sensory areas and could, therefore, convey these bias signals directly. Moreover, spatial information from working memory that is already represented in a retinocentric reference frame could pass directly to the FEF and, from there, directly to visual cortical areas, without being processed in the LIP. Indeed, these pathways may act in parallel (Figure 3). To determine whether either the FEF or LIP is essential for mediating top-down spatial bias signals, the effects of pharmacological inactivation of the FEF or LIP could be studied in animals trained to direct spatial attention based on nonspatial cues (e.g., green cue = attend to the right).

COMPETITIVE SELECTION

The selection of information for entry into working memory is a highly competitive process (Desimone & Duncan 1995). Information about the external world, from memory stores, and about the animal’s internal state is processed extensively and automatically in parallel hierarchies of networks in the central nervous system. Competition for representation occurs at many levels in these hierarchies. The competition compares signal strengths that result from the combined effects of the quality of the encoded information, modulations by top-down bias signals, and the influences of bottom-up salience filters (Figure 1). The competition at each level helps to eliminate the effects of distracting stimuli and to select the most salient stimulus in a given parameter space. At low levels in a hierarchy, the competition occurs within the representations of basic stimulus parameters (for example, stimulus location or sound frequency). At higher levels, the competition can occur among neurons tuned for higher-order features (for example, shapes or types of
A final competition takes place at the interface with working memory, where different domains of information (for example, vision, audition, or somatic sensation) compete for entry into working memory networks (Baddeley 2003).

Information on which the selection of stimuli for working memory is based (i.e., the relative salience of stimuli across the visual field) is represented in the LIP (Bisley & Goldberg 2003). In monkeys trained to discriminate visual targets in the presence of visual distracters, neurons in the LIP encode both the target and the distracter with elevated discharge rates that reflect the effects of top-down bias signals and bottom-up salience filters, respectively. Moreover, the relative spike rates that represent the target versus the distracter change over the course of a trial. Spike rates are high for the target before a distracter (a flashed dot) is presented. Immediately after the distracter is flashed, spike rates are higher for the distracter than for the target. Then, within 300 ms following the distracter, spike rates are again higher for the target. The changes in the relative spike rates to the target and to the distracter correlate with changes in the monkey’s behavioral performance in discriminating the target stimulus. These data are consistent with the proposition that the LIP contains a representation of relative stimulus salience across the entire visual field and that the relative level of activity within this population predicts the information that will gain access to the circuitry of working memory.

Thus, in addition to its proposed role in translating top-down bias signals into various reference frames (discussed previously), the LIP also appears to contribute importantly to competitive selection. These functions are mutually compatible; indeed, competitive selection should act on networks that are modulated by both top-down and bottom-up mechanisms. Although the circuits that mediate top-down sensitivity control and competitive selection may overlap, they are not the same. Top-down bias signals originate in a different network and modulate sensitivity to specific information. In contrast, competitive selection reflects a computation that is intrinsic to a network, a competition for representation that is based on the relative strength of activity (salience) across the entire network.

The competitions that contribute to the selection process take place at various hierarchical levels. These competitions have a special requirement: They must compare response strengths to multiple, simultaneous stimuli and select the strongest responses, whereas the information represented by each region of neural activation in a network is preserved (i.e., neural activity representing different stimuli must not be combined or averaged to arrive at a single solution). This demand (to compare response strengths without altering information) indicates a special class of winner-take-all process. A neural network that could mediate such a competition involves a special type of inhibitory neuron that receives input from a restricted portion of a network and extends inhibitory connections throughout the entire network. Unlike typical inhibitory neurons that operate locally for such purposes as contrast enhancement, regulation of excitability, or spike synchronization, these neurons would establish mutual inhibition, and therefore competition, among all channels in a representation. A winner-take-all competition suggests that the inhibition is nonlinear (Lee et al. 1999). The gain of the network could be increased with the addition of positive recurrent connections (Brody et al. 2003, Major & Tank 2004, Shu et al. 2003). A network with these properties could mediate the final selection of information for entry into working memory as a competition among different networks.

Inhibitory circuits that exhibit this unusual architecture have been described in the retina and in the avian superior colliculus, called the optic tectum (Famiglietti 1992, Stafford & Dacey 1997, Wang et al. 2004). The circuit in the optic tectum is of particular interest because this structure participates in stimulus selection (McPeek & Keller 2004). The
optic tectum represents the locations of visual, auditory, or somatosensory stimuli as a topographic map of space. The responses of tectal neurons are suppressed by a special class of inhibitory neuron that resides in the nucleus isthmi pars magnocellularis (Imc) (Wang et al. 2004). Each Imc neuron is excited by input from a discrete location in the tectal space map and projects back with inhibitory input to the entire space map, except to the location from which it received its excitatory input. In addition, Imc neurons also inhibit cholinergic modulatory neurons that project to those same regions of the optic tectum and may provide local positive recurrent input (Wang et al. 2006). Although the unusual anatomy of this circuit suggests that it could mediate a winner-take-all, competitive selection for stimulus location, the function of this circuit has yet to be determined.

Although stimulus selection is usually dominated by cortical networks, subcortical structures exert a powerful influence on the selection process. This is true particularly for the superior colliculus. As mentioned previously, changes in gaze direction that are mediated by eye saccades cause momentary shifts in spatial attention to stimuli located at the target of the impending eye saccade (Rizzolatti et al. 1987, Shepherd et al. 1986). This implies that changes in gaze direction are accompanied by neural signals that cause the representation of the stimulus selected for the next eye saccade, to win the competition for entry into the circuitry for working memory as an eye saccade occurs. Corollary discharges associated with eye saccades occur in the superior colliculus and propagate to the FEF, via the mediodorsal thalamic nucleus (Figure 3). In the FEF, these corollary discharges shift the locations of visual receptive fields tens of ms before each eye saccade so that FEF neurons represent stimuli at the future locations of their receptive fields (Sommer & Wurtz 2006). This influence of the superior colliculus on the FEF likely contributes to stabilization of the visual world during eye movements. Similar predictive shifts of visual receptive fields occur in the LIP before each eye saccade (Duhamel et al. 1992, Umeno & Goldberg 1997). The same corollary discharges, when they impinge on neurons with foveal receptive fields (which represent the target of an impending eye saccade) may also act as bias signals that increase the responsiveness of these neurons (Figure 1; arrow from gaze control to sensitivity control). The differentially increased responses of these neurons could confer a momentary competitive advantage on the representation of the target for an impending eye saccade. Thus, corollary discharges from the superior colliculus could confer stimulus selection during eye saccades.

In addition, the superior colliculus contributes to stimulus selection in certain visual discrimination tasks (McPeek & Keller 2004). In lower mammals, and even more so in nonmammalian vertebrates, the superior colliculus plays a major role in form vision (Stein & Meredith 1993). In primates, however, the role of the superior colliculus in form vision has been largely usurped by the visual cortex (Van Essen 1985). This finding makes especially noteworthy the demonstration by McPeek & Keller (2004) that the superior colliculus contributes to visual target selection. In this experiment, monkeys were trained to make an eye saccade to the oddball-colored dot in a four-dot display (e.g., to the red dot among three green dots). Both the color and location of the four dots were represented in the visual cortex. In contrast, only the locations of the dots were represented in the superior colliculus because collicular neurons are not selective for color. Nevertheless, when the superior colliculus was focally inactivated so that the representation of the target (red) stimulus was suppressed in the colliculus (but was still present in the cortex) the monkey no longer discriminated the oddball color, but instead selected each of the 4 dots with approximately equal probability. The conditions of this experiment (rapid responses to flashed stimuli) optimized the influence of the superior colliculus relative to that of the visual cortex because the colliculus is differentially
involved in short latency saccades, and collicular neurons are particularly responsive to salient (flashed) stimuli (Schiller et al. 1987, Stein & Meredith 1993). Nevertheless, the results demonstrate a strong influence of the superior colliculus on target selection.

These results imply that information from the superior colliculus influences the selection process in parallel with information from the visual cortex (Figure 3). Information from the superior colliculus reaches the FEF via the mediodorsal thalamic nucleus (Sommer & Wurtz 2006) and the LIP via the pulvinar nucleus, a thalamic nucleus known to play a critical role in spatial attention (Robinson & Petersen 1992). In addition, the FEF and the LIP are heavily interconnected. These pathways provide the superior colliculus with access to representations of stimulus salience in the FEF and LIP, and these pathways operate in parallel with those that originate in the visual cortex.

NEURAL CORRELATES OF STIMULUS SELECTION

One difficulty in studying the neural mechanisms that select information for working memory is identifying when information is, indeed, being gated into working memory. The problem is that access to working memory may depend not on the absolute spike rates of competing neurons, but rather on their relative spike rates, as proposed for the LIP by Bisley & Goldberg (2003). Therefore, unless the spike rates of all competing neurons are monitored simultaneously, it may be impossible to determine which neurons are providing input to working memory at any point in time.

A potential solution to this problem is the observation that neurons can exhibit a distinctive temporal discharge pattern when the information they encode gains access to working memory. Single-unit studies in monkeys, as well as electroencephalographic studies in humans, report that when an animal attends to a particular target stimulus, neurons that represent the target in high-order sensory areas, in the PFC, and in the PPC exhibit synchronized discharges with a periodicity of 40–70 Hz, referred to as gamma frequencies (Bauer et al. 2006, Bichot et al. 2005, Steinmetz et al. 2000, Taylor et al. 2005, Tiitinen et al. 1993). The association of attentional selection of stimuli with oscillations specifically in the gamma-band is controversial, although it has been replicated using a variety of attention tasks. In a recent study (Womelsdorf et al. 2006b), for example, monkeys were trained to detect a small change in the color of a target at a cued location in the visual field in the presence of a distracting stimulus at another location. Single-unit recordings in V4 demonstrated that when the monkey attended the target, units tuned for the target stimulus increased their discharge rates and synchronized their spikes with the local field potential, which oscillated at gamma frequencies. A comparison of the increase in discharge rates with the increase in synchronization showed that synchronization was more sensitive than discharge rate as an indicator of behavioral performance.

Oscillations at gamma frequencies occur in a wide range of networks under various conditions (Gray et al. 1989, Lee 2003, Liu & Newsome 2006). Clearly, they are not specific for attention. However, it is tempting to hypothesize that when the synchronization of unit activity at gamma frequencies increases dramatically during attention tasks, the synchronized activity represents information that is entering the circuitry for working memory.

SUMMARY POINTS

The conceptual framework presented here proposes that attention reflects the combined contributions of four distinct processes: working memory, competitive selection, top-down sensitivity control, and automatic filtering for salient stimuli. Attention selects the
information that gains access to the circuitry for working memory. Access to working memory is determined by the relative signal strengths of competing representations of information. Signal strength is modulated automatically by bottom-up salience filters and is modulated top-down by bias signals that are controlled by working memory and by corollary discharges that accompany gaze changes. Voluntary control of attention is mediated by a recurrent loop comprising working memory, top-down sensitivity control, and competitive selection.

Information is evaluated and decisions are made in working memory. According to the proposed framework, attention does not identify targets; working memory does. In addition, attention is not “deployed” but rather is an ongoing competition among information processing hierarchies vying for access to working memory. What is “deployed” are top-down bias signals based on decisions made in working memory. Top-down bias signals can selectively enhance representations of certain information so that that information continues to have a high probability of gaining entry into working memory. Eye movements, along with other orienting movements, are also guided by decisions made in working memory and serve, together with top-down bias signals, to improve the resolution of information provided to working memory.

FUTURE ISSUES

The framework for attention proposed in this review is intended to act as a heuristic tool to facilitate the study of neural mechanisms underlying attention. By identifying the key functional components of attention, this framework allows studies of basic neural mechanisms to be interpreted in the broader context of attention. For example, studies on the short-term maintenance of information by persistent activity may provide insight to the mechanisms of working memory. Such studies are being carried out on a wide range of preparations, from the entorhinal cortex to the brainstem and spinal cord (Constantinidis & Wang 2004, Fransen et al. 2006, Major & Tank 2004). Mechanisms that could underlie competitive selection can be explored in networks that perform winner-take-all computations on their inputs. The mechanisms of top-down sensitivity control might be elucidated by studying networks that exhibit spatially and temporally precise regulation of neuronal sensitivity. Finally, a mechanistic understanding of salience filters may result from examining intrinsic cellular and network mechanisms of adaptation.

A further benefit of this conceptual framework is in interpreting the symptoms of disease or dysfunction. Many disorders affect attention, but they do so in different ways. Different manifestations of attention disorders indicate that the components of attention, particularly working memory, competitive selection, and top-down sensitivity control, are differentially affected by disorders. For example, prominent among the symptoms of schizophrenias is the inability to ignore irrelevant or imagined stimuli (Phillips & Silverstein 2003), suggesting a particular problem with mechanisms of competitive selection either within, or for, working memory. In contrast, attention deficit disorder frequently includes an inability to retain information in working memory and/or an inability to maintain attention on a specific task (Biederman & Faraone 2005), suggesting problems with working memory and top-down sensitivity control, respectively. These different components of attention are mediated by different, although potentially overlapping,
sets of neural mechanisms. Therefore, the development and selection of optimal ther-
apies for ameliorating such disorders of attention require that we both greatly expand
our knowledge of the neural mechanisms that underlie attention and diagnose the symp-
toms of attention disorders precisely and in the context of this knowledge. Hopefully
the framework for attention presented here will be useful in this regard.

ACKNOWLEDGMENTS

I thank K. Armstrong, M. Goldberg, and W. Newsome for reviewing the manuscript and
P. Knudsen for figure preparation. This work was supported by grants from the National
Institutes of Deafness and Other Communication Disorders.

LITERATURE CITED

transformations in the posterior parietal portex. In The Cognitive Neurosciences III, ed. MS
Andersen RA, Snyder LH, Bradley DC, Xing J. 1997. Multimodal representation of space
20:303–30
Armstrong KM, Fitzgerald JK, Moore T. 2006. Changes in visual receptive fields with micro-
Awh E, Armstrong KM, Moore T. 2006. Visual and oculomotor selection: links, causes and
4:829–39
Batista AP, Andersen RA. 2001. The parietal reach region codes the next planned movement
in a sequential reach task. J. Neurophysiol. 85:539–44
Bauer M, Oostenveld R, Peeters M, Fries P. 2006. Tactile spatial attention enhances gamma-
band activity in somatosensory cortex and reduces low-frequency activity in parieto-
Bichot NP, Rossi AF, Desimone R. 2005. Parallel and serial neural mechanisms for visual
Bisley JW, Goldberg ME. 2003. Neuronal activity in the lateral intraparietal area and spatial
attention. Science 299:81–86
Boudreau CE, Williford TH, Maunsell JH. 2006. Effects of task difficulty and target likelihood
J. Neurosci. 24:11236–43
Chafee MV, Goldman-Rakic PS. 2000. Inactivation of parietal and prefrontal cortex re-
83:1550–66
Chelazzi L, Miller EK, Duncan J, Desimone R. 1993. A neural basis for visual search in inferior


Contents

Information Processing in the Primate Retina: Circuitry and Coding
  G.D. Field and E.J. Chichilnisky .......................................................... 1

Orbitofrontal Cortex and Its Contribution to Decision-Making
  Jonathan D. Wallis .................................................................................. 31

Fundamental Components of Attention
  Eric I. Knudsen ..................................................................................... 57

Anatomical and Physiological Plasticity of Dendritic Spines
  Veronica A. Alvarez and Bernardo L. Sabatini ............................................ 79

Visual Perception and Memory: A New View of Medial Temporal Lobe Function in Primates and Rodents
  Elisabeth A. Murray, Timothy J. Busey, and Lisa M. Saksida ................. 99

The Medial Temporal Lobe and Recognition Memory
  H. Eichenbaum, A.P. Yonelinas, and C. Ranganath .................................. 123

Why Is Wallerian Degeneration in the CNS So Slow?
  Mauricio E. Vargas and Ben A. Barres .................................................. 153

The Head Direction Signal: Origins and Sensory-Motor Integration
  Jeffrey S. Taube ..................................................................................... 181

Peripheral Regeneration
  Zu-Lin Chen, Wei-Ming Yu, and Sidney Strickland ............................ 209

Neuron-Glial Interactions in Blood-Brain Barrier Formation
  Swati Banerjee and Manzoor A. Bhat .................................................... 235

Multiple Dopamine Functions at Different Time Courses
  Wolfram Schultz .................................................................................... 259

Ventral Tegmental Area Neurons in Learned Appetitive Behavior and Positive Reinforcement
  Howard L. Fields, Gregory O. Hjelmstad, Elyssa B. Margolis, and Saleem M. Nicola  .......................................................... 289
Copper and Iron Disorders of the Brain
Erik Madsen and Jonathan D. Gitlin .......................................................... 317

The Micromachinery of Mechanotransduction in Hair Cells
Melissa A. Vollrath, Kelvin Y. Kwan, and David P. Corey .......................... 339

Neurobiology of Feeding and Energy Expenditure
Qian Gao and Tamás L. Horvath ............................................................... 367

Mechanisms that Regulate Establishment, Maintenance, and
Remodeling of Dendritic Fields
Jay Z. Parrish, Kazuo Emoto, Michael D. Kim, and Yuh Nung Jan ............. 399

Dynamic Aspects of CNS Synapse Formation
A. Kimberley McAllister .............................................................................. 425

Adhesion Molecules in the Nervous System: Structural Insights into
Function and Diversity
Lawrence Shapiro, James Love, and David R. Colman .............................. 451

Development of Neural Systems for Reading
Bradley L. Schlaggar and Bruce D. McCandliss ........................................ 475

Molecular Architecture of Smell and Taste in Drosophila
Leslie B. Vosshall and Reinhard F. Stocker .............................................. 505

The Neural Basis of Decision Making
Joshua I. Gold and Michael N. Shadlen .................................................... 535

Trinucleotide Repeat Disorders
Harry T. Orr and Huda Y. Zoghbi ............................................................... 575

Indexes

Cumulative Index of Contributing Authors, Volumes 21–30 .................... 623
Cumulative Index of Chapter Titles, Volumes 21–30 .............................. 627

Errata

An online log of corrections to Annual Review of Neuroscience chapters (if any, 1997
to the present) may be found at http://neuro.annualreviews.org/